AOS 6.8

Acropolis Advanced Administration Guide

May 20, 2024

Contents

OVERVIEW

AOS is the operating system of the Nutanix Controller VM–the VM that runs in the hypervisor to provide Nutanix-specific functionality. It provides core functionality used by workloads and services running on the platform. It contains several data services and features for data protection, space efficiency, scalability, automated data tiering, and security.

AOS is a back-end service that allows for workload and resource management, provisioning, and operations. Its goal is to abstract the facilitating resource (for example, hypervisor, on-prem infrastructure, and cloud based infrastructure) and give workload the ability to seamlessly move between hypervisors, cloud providers, and platforms.

The *AOS Advanced Administration Guide* provides an introduction to AOS and its key feature. It covers advanced topics and tasks that you can do within the system, and references that describe the tasks.

This information is for experienced Windows or Linux system administrators who are familiar with virtualization.

Introduction to AOS

AOS is the base operating system, the so-called **data plane** that packages (encapsulates) the run time of storage, compute, security, and network. It is an artificial intelligence core hardened operating system. AOS runs on top of the Nutanix native hypervisor (AHV), or other hypervisors such as the VMware vSphere ESXi and Microsoft Hyper-V. AOS creates Nutanix Acropolis clusters on the hypervisor to control and virtualize storage of the node as storage pool, container, and Volumes using direct path I/O PCI pass through mechanism.

AOS is installed as a Controller Virtual Machine (CVM) atop a hypervisor to manage everything in a Nutanix cluster.

AOS provides data services and consists of three foundational components–the Distributed Storage Fabric (DSF), the App Mobility Fabric (AMF), and AHV.

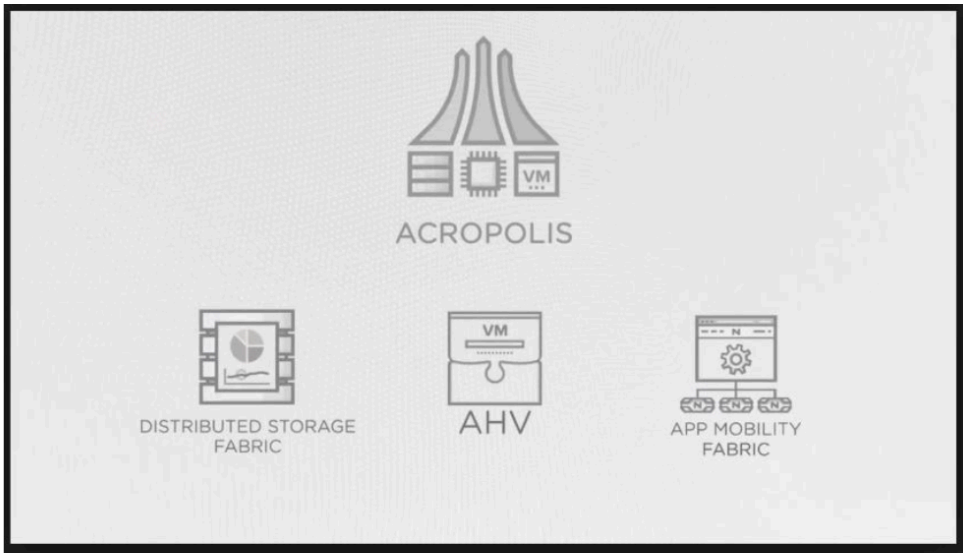


Figure 1: AOS Components

- **Distributed Storage Fabric (DSF)**

  DSF pools the flash and hard disk drive storage across the cluster and presents it as a datastore to the hypervisor. DSF exposes various storage systems (SMB, NFS, and SCSI) with no single point of failure. The datastore is like a centralized storage to the hypervisor but the I/O is managed locally and provides high performance.

  DSF consists of the following:

  - A Storage Pool, which is a logical pool of physical devices that include SSD and HDD devices for the cluster. Storage pool can span multiple Nutanix nodes and expand with the cluster. In most configurations, only one storage pool is necessary.
  - A Container, which is the logical segmentation of storage pools and consists of a group of VMs or vDisk. Container is important because feature like redundancy factor is configured at container level and applied at VM or file level.
  - A vDisk, which is a file larger than 512 KB on the DSF including VMDK and VM hard disks. There are no artificial limits on the size of vDisk and the theoretical maximum size of vDisk is 9 exabytes (*1 exabyte - 1 billion gigabyte). vDisks are logically composed of vBlocks.
  - A vBlock, which is a 1 MB chunk of virtual address space on a vDisk. Each vBlock is mapped to an extent in the vDisk.
  - An Extent, which is a 1 MB piece of large contiguous data that consists of a number of contiguous blocks.
  - An Extent Group, which is a 1 MB or 4 MB piece of physically contiguous stored data. This data is stored as a file on the storage device owned by the CVM. Extents are dynamically distributed among extent groups to provide data striping across nodes and disks to improve performance.

    > Note: The 1 MB and 4 MB extent groups are for dedup and non-dedup data correspondingly.

  DSF divides user data at a fundamental level into Extents. It can compress, ratio-code, snapshot, or dedupe this user data. Compression process can reduce the size of the Extent from 1 MB to a few kilobytes.

  Data Extents can move around. Data that is accessed frequently (Hot Data) is moved to the SSD tier in the DSF and data that is not accessed as frequently (Cold Data) is moved to the HDD. This DSF capability is called Intelligent Tiering.

  Extents are also stored on the nodes on which the guest VM is running.

  DSF has enterprise grade features like Performance Accelerations, Capacity Optimization, Data Protection, and Disaster Recovery.

  - Performance Accelerations. DSF uses the following key capabilities for performance acceleration:

    - Intelligent Tiering.

      DSF continuously and automatically monitors data access patterns and then optimizes data placement. DSF moves data intelligently between the SSD and HDD tiers to provide for optimal performance without requiring an administrator.

    - Data Locality.

      This capability refers to the storage of VM data on the node where the VM is running. It ensures that the read I/O does not have to go through the network. Data locality optimizes the performance and reduces the network congestion. When the VM is moved from one node to another using either vMotion, lab migration, or due to an HA event, the migrated VM data is also moved to ensure data locality.

    - Automatic Disk Balancing.

      This capability ensures data is distributed uniformly across the cluster. Any node in the Nutanix cluster can use storage resources across the cluster. This ensures manual rebalancing is unnecessary. Automatic disk balancing reacts to changing workloads and once utilization reaches self threshold, disk balancing keeps it uniform among the nodes.

- Capacity Optimization. DSF provides deduplication, compression, and erasure coding for Storage optimization.

    - Deduplication.

      There are two types of deduplication – Performance tier and Post process map reduce. The performance tier deduplication removes the duplicate data inline with the content cache to reduce the footprint of the applications working set. The post-process deduplication reduces repetitive data in the capacity tier to increase the effective storage capacity of a cluster.

    - Compression.

      This capability consists of an inline and post process compression. These are intelligently determined based on sequential or random access patterns to ensure optimal performance.

    - Erasure Coding.

      Provides resilience and increases usable capacity by up to 75%. Erasure coding encodes a strip of data block on different nodes and calculates parity. If a node or disk fails, parity is used to find and calculate the missing data blocks.

- Data Protection. It is integrated at the VM level and ensures continuous availability of data. Depending on the recovery time objective (RTO) and recovery point objectives (RPO), data protection features include Time Stream and Cloud Connect for minor incidents, and Async, NearSync, and Sync replication for Major incidents.

  Data protection can create a limited local metadata based on local Snapshots with VM for application level consistency. Since this snapshot is based on metadata, data protection requires minimum disk overhead and ensures high performance recovery.

- Disaster Recovery. The Nutanix data recovery and replication capabilities are built on snapshot technology. VM snapshots can be asynchronously replicated or backed-up to another data center based on user defined schedule. Replication topologies are flexible and bi-directional. Disaster recovery can be one-to-one, one-to-many, and many-to-many deployments. During replication, data is compressed and replicated at the sub block level for maximum efficiency and lower WAN bandwidth consumption.

  Nutanix offers Metro availability for critical workloads that require zero Recovery Point Objective (RPO) and near zero Recovery Time Objective (RTO) to ensure continuous data availability across separate sites in a metro.

  Administrators can set up metro availability bi-directionally between two sites connected over a metro area network. This requires a round trip latency of less than 5 milliseconds. Data is written synchronously to both sides and is always available to the applications in the event of a site failure or when a site undergoes maintenance. VMs can be non-destructively migrated between sites for planned maintenance events and other needs.

- **App Mobility Fabric (AMF)**

  This is a collection of technologies that allows applications and data to move freely between runtime environments. AMF is an open environment capable of delivering powerful virtual machine (VM) placement, VM migration, VM conversion, cross hypervisor High Availability, and integrated disaster recovery.

  AMF supports most virtualized applications and provides a more seamless path to containers and hybrid cloud computing.

  - VM Placement and Migration.

    This comes from Acropolis Dynamic Scheduling (ADS) and is automatically available on every AHV cluster. ADS makes the initial placement decisions for VMs using CPU, memory and storage data points and continuously monitors these data points to make migration decisions. ADS also watches for any anomalies and makes migration decisions to avoid hotspots. Using machine learning, ADS adjusts threshold overtime from the initial fixed values to achieve efficiency without sacrificing performance.

  - Hypervisor Conversion.

    With a few clicks, the admin can convert an ESXi cluster to an AHV cluster. All the VMs running on the ESXi hypervisor are automatically converted to AHV. This VM conversion is host-independent and the VM downtime is less than five minutes.

  - Cross Hypervisor Disaster Recovery (High Availability).

    Allows VMs to be migrated from one hypervisor to another. For example, from ESXi to AHV or the other way round. This is done using Protection Domains by taking and replicating snapshots and then recovering the VMs from the snapshots.

- **AHV**

  While the DSF supports traditional hypervisors such as VMware vSphere and Microsoft Hyper-V, AOS also includes its own native hypervisor – The Acropolis Hypervisor (AHV), based on the proven Linux KVM hypervisor. It provides enhanced security, self healing capabilities based on SaltStack, and enterprise-grade VM management. AHV delivers an overall better user experience at a lower TCO. It is the first hypervisor to plug into the AMF.

AOS can also be **control or management plane** when AHV is the hypervisor. Prism is the control plane of Nutanix and provides infrastructure management for virtual environments running on AOS. Prism provides simplicity to infrastructure management and streamlines time-consuming IT tasks. Prism includes:

- One-click software upgrades for more efficient maintenance,
- One-click insight for detailed capacity trend analysis and planning, and
- One-click troubleshooting for rapid issue identification and resolution.

For an administrator, Nutanix Prism provides the following:

- Convergence of storage, compute and virtualization resources into a unified system to provide an end-to-end view of all workflows – something difficult to achieve with legacy three-tier architectures.
- Advanced machine learning technology with built-in heuristics and business intelligence to easily and quickly mine large volumes of system data, and generate actionable insights for enhancing all aspects of infrastructure performance.
- True consumer-grade user experience with sophisticated search technology that makes management tasks elegantly simple and intuitive, with no need for specialized training.

For more information, see Prism Web Console Guide.

Nutanix AOS Architecture

Nutanix Acropolis (AOS) does not rely on traditional Storage Area Network (SAN) or Network Attached Storage (NAS) interconnects. AOS combines high capacity storage and server compute (CPU and RAM) into a single platform building block. Each building block delivers a unified, scalable, shared-nothing architecture with no single points of failure.

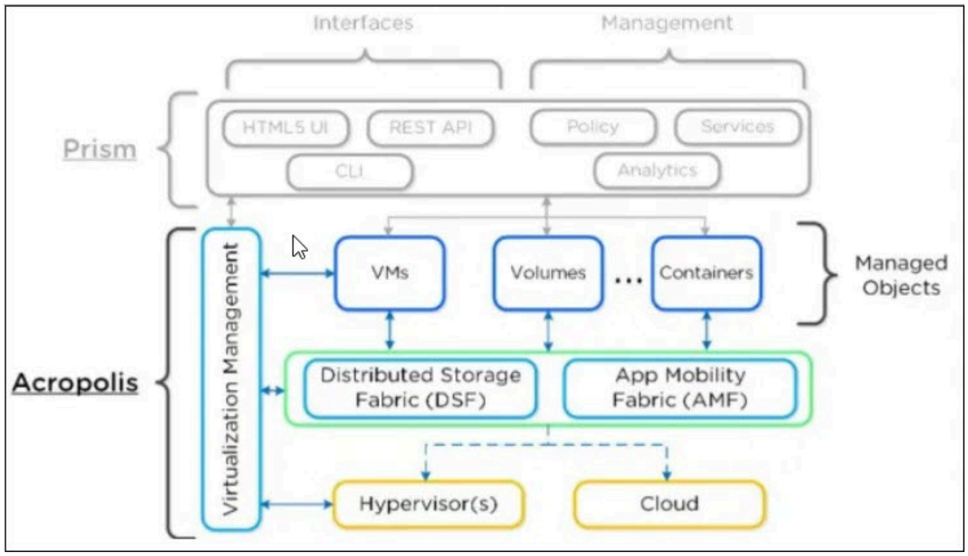


Figure 2: Nutanix AOS Architecture

The Nutanix solution does not require SAN constructs, such as LUNs, RAID groups, or expensive storage switches. All storage management is VM-centric, and I/O is optimized at the VM virtual disk level. The software solution runs on nodes from a variety of manufacturers that are either all-flash for optimal performance, or a hybrid combination of SSD and HDD that provides a combination of performance and additional capacity.

The Distributed Storage Fabric (DSF) automatically tiers data across the cluster to different classes of storage devices using intelligent data placement algorithms. The algorithm ensures the most frequently used data is available in memory or in flash on the node local to the VM.

Key Components

Nutanix Acropolis has five key components that make it a complete solution for delivering any infrastructure service:

Built-in AHV Virtualization

Nutanix AOS includes AHV, its own native virtualization solution, and additionally supports the virtualization solutions of VMware ESXi, and Microsoft Hyper-V. Nutanix AHV is a comprehensive enterprise virtualization solution tightly integrated into AOS and is provided with no additional license cost. AHV delivers the features required to run enterprise applications, for example:

- Combined VM Operations and Performance Monitoring via Nutanix Prism.
- Backup, Disaster Recovery, Host and VM High Availability.
- Dynamic Scheduling (Intelligent placement and resource contention avoidance).

- Broad Ecosystem Support (Certified Citrix Ready, Microsoft Validated via SVVP).

Platform Services

Nutanix AOS delivers a comprehensive set of software-defined platform services so that IT organizations can consolidate all their workloads on the Nutanix platform and manage them centrally.

This include the following:

- VM-centric storage to support almost any virtualized application.
- Container Centric storage with persistent storage support for Kubernetes and Docker.
- With Nutanix Volumes, a scale-out storage solution where every Controller VM (CVM) in a cluster.
- Nutanix Files, a file storage solution for unstructured data such as large-scale home directories, user profiles and more.

Enterprise Storage Capabilities

Nutanix AOS employs MapReduce technology to deliver highly distributed Enterprise grade storage to ensure no single points of failure and negligible impact to real-time performance. Nutanix Enterprise Storage capability include, but are not limited to:

- Performance acceleration capabilities such as caching, data tiers, and data locality.
- Storage optimization technologies, such as Deduplication, Compression and Erasure Coding.
- Data-at-Rest Encryption, supporting both Self-encrypting drives (SED) with KMS and Software-only Encryption.
- Data protection technologies to support snapshots to local, remote and cloud based sites.
- Disaster Recovery features including synchronous and asynchronous replication.

Networking Services

Nutanix AOS provides a comprehensive set of services to visualize the network, automate common network operations and, in the near future, secure the network through native services and partner integration. These services include, but are not limited to:

- Application-centric visualization of the physical and virtual network topology to instantly diagnose and fix common networking issues.
- Open APIs that enable network devices and services such as top-of-rack switches, application delivery controllers (ADC) and firewalls to automatically adapt based on application lifecycle events.

AOS Lifecycle Management

The life cycle manager (LCM) tracks software and firmware versions of the various components in a cluster. LCM allows you to view information about the current inventory and update the versions as needed.

LCM supports software and firmware updates for all platforms that use Nutanix software. For more information, see *Nutanix Compatibility and Software Interoperability* in the Acropolis Family Release notes.

You can access the LCM framework using the Prism interface. All communication between the cluster and LCM modules go through the LCM framework. To view the LCM dashboard, select LCM from the pull-down list on the left of the main menu. For more information, see Life Cycle Manager Guide.

Documentation References

Refer to the guides in this table for additional information required for AOS administration.

Guide	Purpose
Getting Started Guide	This guide provides you with all the required summary to get the Nutanix Acropolis system up and running.
Nutanix Rack Mounting Guide	This guide provides information about mounting a block based on various models.
Field Installation Guide	This guide provides information about using the Foundation for deploying a node and creating a cluster automatically. It allows you to configure a pre-imaged node, or image a node with AOS and a hypervisor of your choice. It also allows you to form a cluster out of nodes whose hypervisor and AOS versions are the same, with or without reimaging.
Prism Web-Console Guide	This guide provides the Prism Element workflows and additional AOS configurations using Prism Element. It also describes the legacy 1-click upgrade method, using the **Upgrade Software** option, in Prism Element to perform software upgrades.
AHV Administration Guide	This guide provides information about managing nodes with AHV. AHV is the native hypervisor of the Nutanix solution. It represents a unique approach to virtualization that offers powerful virtualization capabilities such as core VM operations, live migration, VM high availability, and virtual network management.
vSphere Administration Guide for Acropolis (using vSphere HTML Client)	This guide describes how to configure and manage the Nutanix cluster in vSphere.
Hyper-V Administration for Acropolis	This guide describes how to configure and manage the Nutanix cluster in Hyper-V.
Data Protection and Recovery with Prism Element	This guide describes the concepts and procedures for configuring data protection using protection domains. You can configure data protection using protection domains in Prism Element (in other words, by signing in to an individual cluster through its web console).

Guide	Purpose
Getting Started with Nutanix Community Edition	This guide provides information about the Nutanix Community Edition. This is a free version of Nutanix AOS, which powers the Nutanix enterprise cloud platform. The Community Edition of AOS is designed for people interested in test driving its main features on their own test hardware and infrastructure. As stated in the end user license agreement, Community Edition is intended for internal business operations and non-production use only.
Nutanix Cluster Check Guide	This guide provides information about Nutanix Cluster Check (NCC), which is a cluster-resident software that diagnoses cluster health and identifies configurations qualified and recommended by Nutanix. NCC continuously and proactively runs hundreds of checks and takes the needed action towards issue resolution. Depending on the issue discovered, NCC raises an alert or automatically creates Nutanix Support cases. NCC can be run provided that the individual nodes are up, regardless of cluster state.

REMOTE CONSOLE IP ADDRESS CONFIGURATION

The Intelligent Platform Management Interface (IPMI) is a standardized interface used to manage a host and monitor its operation. To enable remote access to the console of each host, you must configure the IPMI settings within BIOS.

The Nutanix cluster provides a Java application to remotely view the console of each node, or host server. You can use this console to configure additional IP addresses in the cluster.

The procedure for configuring the remote console IP address is slightly different for each hardware platform.

Refer to the Third Party Platform section to know more about configuring the non-OEM platforms.

Configuring the Remote Console IP Address (BIOS)

This procedure provides information about configuring IPMI settings in BIOS and enable access to the console of each host.

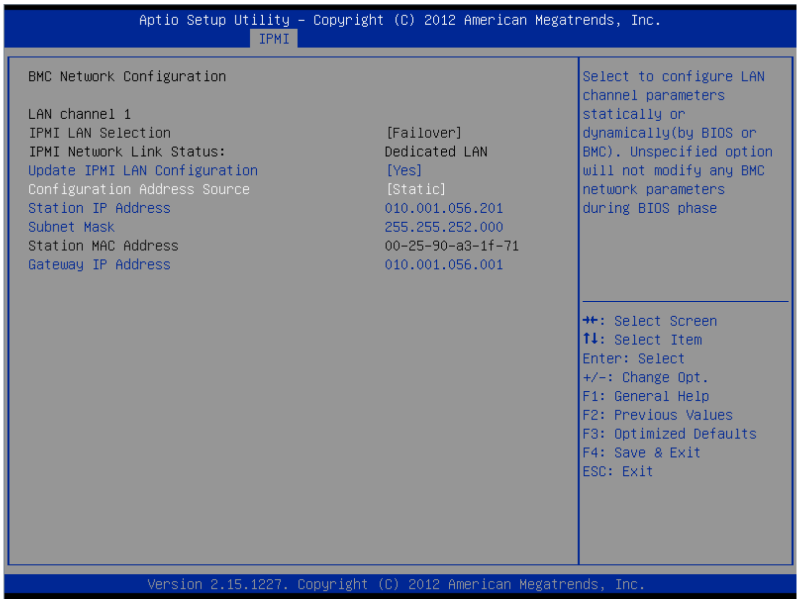
About this task

To configure the IPMI settings in BIOS to enable remote access to the console of each host, do the following:

Procedure

1. Connect a keyboard and monitor to a node in the Nutanix block.
2. Restart the node and press **Delete** to enter the BIOS setup utility.
   There is a limited amount of time to enter BIOS before the host completes the restart process.
3. Press the right arrow key to select the **IPMI** tab.
4. Press the down arrow key until **BMC Network Configuration** is highlighted and then press **Enter**.
5. Press down the arrow key until **Update IPMI LAN Configuration** is highlighted and press **Enter** to select **Yes**.
6. Select **Configuration Address Source** and press **Enter**.
7. Select **Static** and press **Enter**.

8. Assign the **Station IP Address**, **Subnet Mask**, and **Router IP Address**.



```
Aptio Setup Utility – Copyright (C) 2012 American Megatrends, Inc.
IPMI

BMC Network Configuration

LAN channel 1
IPMI LAN Selection                          [Failover]
IPMI Network Link Status:                   Dedicated LAN
Update IPMI LAN Configuration               [Yes]
Configuration Address Source                [Static]
Station IP Address                          010.001.056.201
Subnet Mask                                 255.255.252.000
Station MAC Address                         00-25-90-a3-1f-71
Gateway IP Address                          010.001.056.001

Select to configure LAN
channel parameters
statically or
dynamically(by BIOS or
BMC). Unspecified option
will not modify any BMC
network parameters
during BIOS phase

→←: Select Screen
↑↓: Select Item
Enter: Select
+/-: Change Opt.
F1: General Help
F2: Previous Values
F3: Optimized Defaults
F4: Save & Exit
ESC: Exit

Version 2.15.1227. Copyright (C) 2012 American Megatrends, Inc.
```

9. Review the BIOS settings and press **F4** to save the configuration changes and exit the BIOS setup utility.
   The node restarts.

10. Log on to any Controller VM in the cluster, and run the following command to perform rolling restart for Genesis on all the Controller VMs in the cluster.

    ```
    nutanix@cvm$ allssh genesis restart
    ```

    > Note: If you are reconfiguring IPMI address on a node because you have replaced the motherboard, restart Genesis on the Controller VM only for that node using the following command:
    >
    > ```
    > nutanix@cvm$ genesis restart
    > ```

    If the restart is successful, output similar to the following is displayed.

    ```
    Stopping Genesis pids [1933, 30217, 30218, 30219, 30241]
    Genesis started on pids [30378, 30379, 30380, 30381, 30403]
    ```

Configuring the Remote Console IP Address (Command Line)

This procedure provides information about configuring the remote console IP address in your cluster using the CLI script.

About this task

You can configure the management interface from the hypervisor host on the same node.

Perform the following steps once from each hypervisor host in the cluster where you want to change the management network configuration.

Procedure

1. Log on to the hypervisor host with SSH (AHV or vSphere) or remote desktop connection (Hyper-V).

2. Set the networking parameters.

   - For AHV

     ```
     root@ahv# ipmitool -U ADMIN -P ADMIN lan set 1 ipsrc static
     root@ahv# ipmitool -U ADMIN -P ADMIN lan set 1 ipaddr mgmt_interface_ip_addr
     root@ahv# ipmitool -U ADMIN -P ADMIN lan set 1 netmask mgmt_interface_subnet_addr
     root@ahv# ipmitool -U ADMIN -P ADMIN lan set 1 defgw ipaddr mgmt_interface_gateway
     ```

   - For vSphere

     ```
     root@esx# /ipmitool -U ADMIN -P ADMIN lan set 1 ipsrc static
     root@esx# /ipmitool -U ADMIN -P ADMIN lan set 1 ipaddr mgmt_interface_ip_addr
     root@esx# /ipmitool -U ADMIN -P ADMIN lan set 1 netmask mgmt_interface_subnet_addr
     root@esx# /ipmitool -U ADMIN -P ADMIN lan set 1 defgw ipaddr mgmt_interface_gateway
     ```

   - For Hyper-V

     ```
     > ipmiutil lan -e -I mgmt_interface_ip_addr -G mgmt_interface_gateway
      -S mgmt_interface_subnet_addr -U ADMIN -P ADMIN
     ```

   Replace:

   - *`mgmt_interface_ip_addr`* with the new IP address for the remote console.
   - *`mgmt_interface_gateway`* with the gateway IP address.
   - *`mgmt_interface_subnet_addr`* with the subnet mask for the new IP address.

3. Show current settings.

   - For AHV

     ```
     root@ahv# ipmitool -v -U ADMIN -P ADMIN lan print 1
     ```

   - For vSphere

     ```
     root@esx# /ipmitool -v -U ADMIN -P ADMIN lan print 1
     ```

   - For Hyper-V

     ```
     > ipmiutil lan -r -U ADMIN -P ADMIN
     ```

   Confirm that the parameters are set to the correct values.

4. Log on to any Controller VM in the cluster, and run the following command to perform rolling restart for Genesis on all the Controller VMs in the cluster.

   ```
   nutanix@cvm$ allssh genesis restart
   ```

   > Note: If you are reconfiguring IPMI address on a node because you have replaced the motherboard, restart Genesis on the Controller VM only for that node using the following command:
   >
   > ```
   > nutanix@cvm$ genesis restart
   > ```

   If the restart is successful, output similar to the following is displayed.

   ```
   Stopping Genesis pids [1933, 30217, 30218, 30219, 30241]
   Genesis started on pids [30378, 30379, 30380, 30381, 30403]
   ```

Changing the Controller VM IP Addresses in your Nutanix Cluster (CLI Script)

This procedure provides information about configuring the Controller VM IP address in your cluster using the CLI script.

Before you begin

> Note: This procedure is applicable only if you have already created a cluster and want to change the IP addresses of the CVM, hypervisor host, IPMI address schema into a new infrastructure.

For more information and cautions about the impact of changing the cluster IP addresses, see Modifying Cluster Details in the *Prism Web Console Guide*.

To change the Controller VM IP address, you must run the external IP address reconfiguration script (`external_ip_reconfig`).

You can use the external IP address reconfiguration script in the following scenarios:

- Change the IP addresses of the CVMs in the same subnet.
- Change the IP addresses of the CVMs to a new or different subnet.

  In this scenario, the external IP address reconfiguration script works successfully if the new subnet is configured with the required switches and the CVMs can communicate with each other in the new subnet.
- Change the IP addresses of the CVMs to a new or different subnet if you are moving the cluster to a new physical location.

  In this scenario, the external IP address reconfiguration script works successfully if the CVMs can still communicate with each other in the old subnet.

> Caution: Do not use the external IP address reconfiguration script (`external_ip_reconfig`) in the following scenarios:
>
> - If you are changing the IP address of the CVM to a subnet that is non-routable to the current subnet. For this scenario, contact Nutanix Support for assistance.
> - If you are using the network segmentation feature on your cluster and you want to change the IP addresses of the backplane (eth2) interface. For instructions about how to change the IP addresses of the backplane (eth2) interface, see Reconfiguring the Backplane Network in the *Security Guide* .

Ensure that the following prerequisites are met before you change the Controller VM IP address:

- Before you decide to change the CVM, hypervisor host, and IPMI IP addresses, consider the possibility of incorporating the existing IP address schema into the new infrastructure by reconfiguring your routers and switches instead of Nutanix nodes and CVMs. If that is not possible and you must change the IP addresses of CVMs and hypervisor hosts, proceed with the procedure described in this document.
- Guest VM downtime is necessary for this change, because the Nutanix cluster must be in a stopped state. Therefore, plan the guest VM downtime accordingly.

- Verify if your cluster is using the network segmentation feature.

  ```
  nutanix@cvm$ network_segment_status
  ```

  Note the following if you are using the network segmentation feature.

  - The network segmentation feature enables the backplane network for CVMs in your cluster (eth2 interface). The backplane network is always a non-routable subnet and/or VLAN that is distinct from the one which is used by the external interfaces (eth0) of your CVMs and the management network on your hypervisor. Typically, you do not need to change the IP addresses of the backplane interface (eth2) if you are updating the CVM or host IP addresses.
  - If you have enabled network segmentation on your cluster, check to make sure that the VLAN and subnet in-use by the backplane network is still going to be valid once you move to the new IP scheme. If not, change the subnet or VLAN.

    For information about the AOS version and instructions to disable the network segmentation feature before you change the CVM and host IP addresses, see Disabling Network Segmentation on an AHV Cluster section in *Security Guide*. After you have updated the CVM and host IP addresses by following the steps outlined later in this document, you can proceed to re-enable network segmentation. Follow the instructions in the *Security Guide*, to designate the new VLAN or subnet for the backplane network.
  - Use the following CLI commands if network segmentation is enabled on the cluster and you wish to change both the external CVM eth0 (Management) and internal CVM eth2 (backplane) IP addresses:

    - Reconfigure the external CVM IP address (eth0).

      ```
      nutanix@cvm$ external_ip_reconfig
      ```
    - Reconfigure the internal CVM IP address (eth2).

      ```
      nutanix@cvm$ backplane_ip_reconfig
      ```
- If you have configured remote sites for data protection, either wait until any ongoing replications are complete or stop them. After you successfully reconfigure the IP addresses, update the reconfigured IP addresses at the remote sites before you resume the replications. For information about changing the Controller VM IP addresses of metro availability clusters, see Changing the Controller VM IP Addresses in Metro Availability Clusters.
- Nutanix recommends that you prepare a spreadsheet that includes the existing and new CVM, hypervisor host, and IPMI IP addresses, subnet masks, default gateway, and cluster virtual IP addresses and VLANs (download the IP Address Change Worksheet Template).

- You can change the virtual IP address of the cluster either before or after you change the CVM IP address. The virtual IP address of the cluster is required to configure certain data protection features.

  > Caution: All the features that use the cluster virtual IP address are impacted if you change that address. For more information, see Virtual IP Address Impact in the *Prism Web Console Guide*.

  Use the following steps to change the virtual IP address of the cluster.

  1. Clear the existing virtual IP address of the cluster.

     ```
     nutanix@cvm$ ncli cluster clear-external-ip-address
     ```

  2. Set a new virtual IP address for the cluster.

     ```
     nutanix@cvm$ ncli cluster set-external-ip-address
     external-ip-address=insert_new_external_ip_address
     logon-name=admin password=prism_admin_user_password
     ```

     - Replace ***`insert_new_external_ip_address`*** with the new virtual IP address for the cluster.
     - Replace ***`prism_admin_user_password`*** with password of the Prism admin account.

- You can change the iSCSI data services IP address of the cluster while you are changing the virtual IP address of the cluster. This IP address is used by Nutanix Volumes and other data services applications.

  > Caution: For certain features, changing the external data services IP address can result in unavailable storage or other issues. The features in question include Volumes, Calm, Leap, Karbon, Objects, and Files. For more information, see KB 8216 and iSCSI Data Services IP Address Impact in the *Prism Web Console Guide*.

- Ensure that the cluster NTP and DNS servers are reachable from the new Controller VM IP addresses. If you are using different NTP and DNS servers, remove the existing NTP and DNS servers from the cluster configuration and add the new ones. If you do not know the new addresses, remove the existing NTP and DNS servers before cluster reconfiguration and add the new ones afterwards.

Web Console	• In the gear icon pull-down list, click **Name Servers**. • In the gear icon pull-down list, click **NTP Servers**.
nCLI	• `ncli> cluster remove-from-name-servers servers="name_servers"` • `ncli>cluster add-to-name-servers servers="name_servers"` • `ncli> cluster remove-from-ntp-servers servers="ntp_servers"` • `ncli>cluster add-to-ntp-servers servers="ntp_servers"` Replace: • *`"name_servers"`* with IP addresses of the servers that you want to remove from the list of name servers. • *`"ntp_servers"`* with IP address of the NTP server that you want to remove from the list of NTP servers.

- Log on to a Controller VM in the cluster and check that all hosts are part of the metadata store.

  ```
  nutanix@cvm$ ncli host ls | grep "Metadata store status"
  ```

  For every host in the cluster, `Metadata store enabled on the node` is displayed.

  > Warning: If `Node marked to be removed from metadata store` is displayed, do not proceed with the IP address reconfiguration, and contact Nutanix Support to resolve the issue.

> Warning: If you are using distributed switches in your ESXi clusters, migrate the distributed switches to standard switches before you perform any IP address reconfiguration procedures that involve changing the management VMkernel port of the ESXi host to a different distributed port group that has a different VLAN.

About this task

Perform the following steps to change the IP addresses on a Nutanix cluster:

1. Check the health of the cluster infrastructure and resiliency. For more information, see the *Before you begin* section of this document.
2. Stop the cluster.
3. Change the VLAN and NIC Teaming configurations as necessary.

   > Note: Check the connectivity between CVMs and hosts, that is all the hosts must be reachable from all the CVMs and vice versa before you perform step 4. If any CVM or host is not reachable, contact Nutanix Support for assistance.
4. Change the CVM IP addresses by using the external_ip_reconfig script.
5. Change the hypervisor host IP addresses if necessary.
6. Restart the CVMs.
7. Perform the initial series of validation steps.
8. Start the cluster.
9. Perform the final series of validation steps.
10. Change the IPMI IP addresses if necessary.

The external IP address reconfiguration script performs the following tasks:

1. Checks if cluster is stopped.
2. Puts the cluster in reconfiguration mode.
3. Restarts Genesis.
4. Prompts you to type the new netmask, gateway, and external IP addresses, and updates them.
5. Updates the IP addresses of the Zookeeper hosts.

Perform the following procedure to change the Controller VM IP addresses

> Warning: If you are changing the Controller VM IP addresses to another subnet, network, IP address range, or VLAN, you must also change the hypervisor management IP addresses to the same subnet, network, IP address range, or VLAN.
>
> If you have configured a data services IP address for guest VMs that use ISCSI volumes and you are changing the IP addresses of the CVMs to a different subnet, you must change the data services IP address to that subnet. After you change the data services IP address, update the guest VM ISCSI client configuration with the new data services IP address.

For instructions about how to change the IP address of an AHV host, see Changing the IP Address of on AHV Host in the *AHV Administration Guide*.

For instructions about how to change the IP address of an ESXi host, see Changing a Host IP Address in the *vSphere Administration Guide for Acropolis*.

For instructions about how to change the IP address of a Hyper-V host, see Changing a Host IP Address in the *Hyper-V Administration for Acropolis* guide.

Procedure

1. Log on to the hypervisor using SSH (AHV or vSphere), remote desktop connection (Hyper-V), or the IPMI remote console.

   If you are unable to reach the IPMI IP addresses, reconfigure by using the BIOS or hypervisor command line.

   For using BIOS, see Configuring the Remote Console IP Address (BIOS).

   For using the hypervisor command line, see Configuring the Remote Console IP Address (Command Line).

2. Log on to any Controller VM in the cluster.

   - For AHV or vSphere—Log on to the controller VM in the cluster.

     ```
     root@host# ssh nutanix@cvm_ip_address
     ```

   - For Hyper-V—Log on to the controller VM in the cluster.

     ```
     > ssh nutanix@cvm_ip_address
     ```

   Replace `cvm_ip_address` with the IP address of the Controller VM.

   Accept the host authenticity warning if prompted, and enter the password.

3. Stop the Nutanix cluster.

   ```
   nutanix@cvm$ cluster stop
   ```

   > Warning: This step affects the operation of a Nutanix cluster. Schedule a down time before performing this step.

   If you are using VLAN tags on your CVMs and on the management network for your hypervisors and you want to change the VLAN tags, then stop the cluster and make the changes mentioned in the following guides.

   For information about assigning VLANs to hosts and the Controller VM, see the indicated documentation:

   - For AHV—See Assigning an AHV Host to a VLAN and Assigning the Controller VM to a VLAN in the *AHV Administration Guide*.
   - For ESXi—See Configuring Host Networking (Management Network) in the *vSphere Administration Guide for Acropolis* for instructions about tagging a VLAN on an ESXi host by using DCUI.

     > Note: If you are relocating the cluster to a new site, the **`external_ip_reconfig`** script works only if all the CVMs are up and accessible with their old IP addresses. Otherwise, contact Nutanix Support to manually change the IP addresses.

   After you have stopped the cluster, shut down the CVMs and hosts and move the cluster. Proceed with step 4 only after you power on the cluster at the desired site and you have confirmed that all CVMs and hosts can SSH to one another. As a best practice, ensure that the out-of-band management Remote Console (IPMI, iDRAC, and ILO) is accessible on each node before you proceed further.

Verify that upstream networking is configured to support the changes to the IP address schema .

For example, check the network load balancing or LACP configuration to verify that it supports the seamless transition from one IP address schema to another.

4. Run the external IP address reconfiguration script (external_ip_reconfig) from any one Controller VM in the cluster.

   ```
   nutanix@cvm$ external_ip_reconfig
   ```

5. Follow the prompts to type the new netmask, gateway, and external IP addresses. On successful completion of the script, the system displays a similar output:

   ```
   External IP reconfig finished successfully. Restart all the CVMs and start the cluster.
   ```

   > Note:
   >
   > `If the external_ip_reconfig script fails, re-run the script. If the script fails on the second attempt also, contact Nutanix Support.`

6. Restart each Controller VM in the cluster.

   ```
   nutanix@cvm$ sudo reboot
   ```

   > Note: If you have changed the CVMs to a new subnet, you must now update the IP addresses of hypervisor hosts to the new subnet. Change the hypervisor management IP address or IPMI IP address before you restart the Controller VMs.

   Enter the nutanix password if prompted.

7. After you turn on every CVM, log on to each CVM and verify if the IP address has been successfully changed.

   > Note: It can take up to 10 minutes for the CVMs to show the new IP addresses after they are turned on.

   > Note: If you see any of the old IP addresses in the following commands or the commands fail to run, stop and call Nutanix Support assistance.

   Run the following commands on every CVM in the cluster.

   a. Display the CVM IP addresses.

      ```
      nutanix@cvm$ svmips
      ```

   b. Display the hypervisor IP addresses.

      ```
      nutanix@cvm$ hostips
      ```

   c. From any one CVM in the cluster, verify that the following outputs show the new IP address scheme and that the Zookeeper IDs are mapped correctly.

      > Note: Never edit the following files manually. Contact Nutanix Support for assistance.

      ```
      nutanix@cvm$ allssh sort -k2 /etc/hosts
      nutanix@cvm$ allssh sort -k2 data/zookeeper_monitor/zk_server_config_file
      nutanix@cvm$ zeus_config_printer | grep -B 20 myid | egrep -i "myid|external_ip"
      ```

8. Start the Nutanix cluster.

   ```
   nutanix@cvm$ cluster start
   ```

   If the cluster starts properly, output similar to the following is displayed for each node in the cluster:

   ```
   CVM:host IP-Address Up
                                   Zeus   UP       [9935, 9980, 9981, 9994, 10015, 10037]
                              Scavenger   UP       [25880, 26061, 26062]
                                 Xmount   UP       [21170, 21208]
                       SysStatCollector   UP       [22272, 22330, 22331]
                              IkatProxy   UP       [23213, 23262]
                       IkatControlPlane   UP       [23487, 23565]
                          SSLTerminator   UP       [23490, 23620]
                         SecureFileSync   UP       [23496, 23645, 23646]
                                 Medusa   UP       [23912, 23944, 23945, 23946, 24176]
                     DynamicRingChanger   UP       [24314, 24404, 24405, 24558]
                                 Pithos   UP       [24317, 24555, 24556, 24593]
                             InsightsDB   UP       [24322, 24472, 24473, 24583]
                                 Athena   UP       [24329, 24504, 24505]
                                Mercury   UP       [24338, 24515, 24516, 24614]
                                 Mantle   UP       [24344, 24572, 24573, 24634]
                             VipMonitor   UP       [18387, 18464, 18465, 18466, 18474]
                               Stargate   UP       [24993, 25032]
                   InsightsDataTransfer   UP       [25258, 25348, 25349, 25388, 25391, 25393,
    25396]
                                  Ergon   UP       [25263, 25414, 25415]
                                Cerebro   UP       [25272, 25462, 25464, 25581]
                                Chronos   UP       [25281, 25488, 25489, 25547]
                                Curator   UP       [25294, 25528, 25529, 25585]
                                  Prism   UP       [25718, 25801, 25802, 25899, 25901, 25906,
    25941, 25942]
                                    CIM   UP       [25721, 25829, 25830, 25856]
                           AlertManager   UP       [25727, 25862, 25863, 25990]
                               Arithmos   UP       [25737, 25896, 25897, 26040]
                                Catalog   UP       [25749, 25989, 25991]
                              Acropolis   UP       [26011, 26118, 26119]
                                  Uhura   UP       [26037, 26165, 26166]
                                   Snmp   UP       [26057, 26214, 26215]
                      NutanixGuestTools   UP       [26105, 26282, 26283, 26299]
                             MinervaCVM   UP       [27343, 27465, 27466, 27730]
                          ClusterConfig   UP       [27358, 27509, 27510]
                               Aequitas   UP       [27368, 27567, 27568, 27600]
                            APLOSEngine   UP       [27399, 27580, 27581]
                                  APLOS   UP       [27853, 27946, 27947]
                                  Lazan   UP       [27865, 27997, 27999]
                                 Delphi   UP       [27880, 28058, 28060]
                                   Flow   UP       [27896, 28121, 28124]
                                Anduril   UP       [27913, 28143, 28145]
                                  XTrim   UP       [27956, 28171, 28172]
                          ClusterHealth   UP       [7102, 7103, 27995, 28209,28495, 28496,
    28503, 28510,
   28573, 28574, 28577, 28594, 28595, 28597, 28598, 28602, 28603, 28604, 28607, 28645, 28646,
    28648, 28792,
   28793, 28837, 28838, 28840, 28841, 28858, 28859, 29123, 29124, 29127, 29133, 29135, 29142,
    29146, 29150,
   29161, 29162, 29163, 29179, 29187, 29219, 29268, 29273]
   ```

Manually Configuring CVM IP Addresses

This procedure describes how to assign a static IP address to the Controller VM.

About this task

> Note: You can manually configure CVM IP address only when a cluster is not yet created.

> Note: Configure the Controller VM IP addresses by following Changing the Controller VM IP Addresses in your Nutanix Cluster (CLI Script) if you have already created a cluster. If you have not yet created a cluster, perform the procedure described in this section.

Perform the following steps to manually configure a static IP address to the Controller VM:

Procedure

1. Open the `ifcfg-eth0` file for editing.

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

2. Update the `NETMASK`, `IPADDR`, `BOOTPROTO`, `GATEWAY`, and `ONBOOT` parameters in the script.

   ```
   NETMASK=xxx.xxx.xxx.xxx
   IPADDR=xxx.xxx.xxx.xxx
   BOOTPROTO=none
   GATEWAY=xxx.xxx.xxx.xxx
   ONBOOT=Yes
   ```

   - `NETMASK`—Enter the desired netmask value.

     Replace *`xxx.xxx.xxx.xxx`* with the appropriate value.

   - `IPADDR`—Enter the appropriate static IP address (assigned by your IT department) for the Controller VM.
   - `BOOTPROTO`—Enter **none**.

     If you employ DHCP, change the value from `dhcp` to `none`. Only a static address is allowed; DHCP is not supported.

   - `GATEWAY`—Enter the IP address for your gateway.
   - `ONBOOT`—Enter **`Yes`**.

     > Important: The **`ONBOOT`** parameter is mandatory. Ensure that you set this parameter in the **`ifcfg-eth0`** script.

   > Note: Carefully check the file to ensure that there are no syntax errors, whitespace at the end of lines, or blank lines in the file.

3. Save the changes.

4. Run the following command to restart the network service:

   ```
   nutanix@cvm$ sudo systemctl restart network
   ```

   > Important: If the **`nutanix@cvm$ sudo systemctl restart network`** command does not work, the only option is to reboot the Controller VM.
   >
   > Run the following command to reboot the Controller VM:
   >
   > ```
   > nutanix@cvm$ sudo reboot
   > ```

CLUSTER MANAGEMENT

Although each host in a Nutanix cluster runs a hypervisor independent of other hosts in the cluster, some operations affect the entire cluster.

Most administrative functions of a Nutanix cluster can be performed through the web console (Prism), however, there are some management tasks that require access to the Controller VM (CVM) over SSH. Nutanix recommends restricting CVM SSH access with password or key authentication.

Refer to Controller VM Access and Nutanix User Access to Controller VM for information about how to access the Controller VM as an `admin` and `nutanix` user.

Refer to Performing Initial Configuration on a Cluster Using Prism Element for information about further configuring the cluster, once it is created, using Prism Element.

Performing Initial Configuration

About this task

Once the cluster is created it can be configured through the Prism Web Console (also known as Prism Element). A storage pool and a container are provisioned automatically when the cluster is created, but many other options require user input.

> Note: Prism Element is a built-in service on every Nutanix cluster. Each Nutanix cluster deployment has a unique Prism Element instance for local management. Multiple clusters are managed via Prism Central.

Managing a Nutanix cluster involves configuring and monitoring the entities within the cluster, including virtual machines, storage containers, and hardware components. You can manage a Nutanix cluster through a web-based management console or a command line interface (nCLI).

This topic introduces you to the Prism Web Console, which is a graphical user interface (GUI) that allows you to monitor cluster operations and perform a variety of configuration tasks. For more information, see Web Console Overview in the *Prism Web Console Guide*.

Use the following steps to perform the initial configuration using Prism Web Console:

Procedure

1. Specify the Timezone of the cluster.

   For more information, see Creating a VM (AHV) in the *Prism Web Console Guide*.

2. Specify an outgoing SMTP server.

   For more information, see Configuring an SMTP Server in the *Prism Web Console Guide*.

3. Enable the Remote Support Tunnel if the site security policy allows Nutanix customer support to access the cluster.

   For more information, see Controlling Remote Connections in the *Prism Web Console Guide*.

   > Caution: Failing to enable remote support prevents Nutanix Support from directly addressing cluster issues. Nutanix recommends that all customers send the Pulse data at minimum because it allows proactive support of customer issues.

4. Enable Pulse if the site security policy allows Nutanix Support to collect cluster status information.

   For more information, see Configuring Pulse in the *Prism Web Console Guide*.

   This information is used by Nutanix Support to send automated hardware failure alerts, and diagnose potential problems and assist proactively.

   Pulse securely transmits system-level diagnostic data to the Insights platform, enabling predictive health and context-aware support automation workflows. Pulse is enabled on a per-cluster basis and it is packaged as a part of NCC . This ensures that no matter which version of Nutanix AOS or hypervisor vendor is in use, customers always benefit from the latest Insights enhancements to improve infrastructure health.

   > Note: Nutanix Pulse does not gather, nor communicate with any guest VM specific data, user data, metadata or any personally-identifiable information such as administrator credentials. No system-level data from any customer is ever shared with third parties.

5. Add a list of Alert Email recipients, or if the security policy does not allow it, disable alert emails.

   For more information, see Configuring Alert Emails in the *Prism Web Console Guide*.

6. Specify email recipients for specific alerts.

   For more information, see Configuring Alert Policies in the *Prism Web Console Guide*.

7. Enable network segmentation on the cluster. Network segmentation separates management traffic from backplane traffic for improved security.

   Network traffic can be segmented (or separated) within a Nutanix cluster for various functions or purposes. For example, backplane traffic can be separated from Management-Plane Traffic so as to allow for even greater available bandwidth for the backplane traffic. Further, as another example, DMZ related traffic could be isolated to specific host uplinks. For more information, see Securing Traffic Through Network Segmentation in the *Security Guide*.

8. Run the Life Cycle Manager (LCM) inventory to ensure the LCM framework has the updated software and firmware version of the entities in the cluster.

   For more information, see the Performing Inventory with the Life Cycle Manager section in *Life Cycle Manager Guide*.

9. Enable the automatic downloads of upgrade software packages for cluster components if the site security policy permits.

   For more information, see Software and Firmware Upgrades in the *Prism Web Console Guide*.

   > Note: To ensure that automatic download of updates can function, allow access to the following URLs through your firewall:
   >
   > - *.compute-*.amazonaws.com:80
   > - release-api.nutanix.com:80

10. License the cluster.

    For more information, see Licensing in the *Prism Web Console Guide*.

    > Note: Licensing is required for insights discoveries on SWO clusters.

11. For ESXi and Hyper-V clusters, add the hosts to the appropriate management interface.

    - vCenter—See vSphere Administration Guide for Acropolis.
    - SCVMM—See the Hyper-V Administration for Acropolis guide.

12. If you are using Microsoft Hyper-V hypervisor on HPE DX platform models, ensure that the software and drivers on Hyper-V are compatible with the firmware version installed on the nodes.

    For more information, see Deploying Drivers and Software on Hyper-V for HPE DX in the *Field Installation Guide*. This procedure to deploy software and drivers is to be carried out after cluster creation and before moving the nodes to production.

13. Verify that the cluster has passed the latest Nutanix Cluster Check (NCC) tests.

    Nutanix Cluster Check (NCC) is a framework of scripts that can help diagnose cluster health. You can run NCC as long as individual nodes are up, regardless of cluster state. The scripts run standard commands against the cluster or the nodes, depending on the type of information being retrieved.

    - Check the installed NCC version and update it if a recent version is available. For more information, see Software and Firmware Upgrades in the *Prism Web Console Guide*.
    - Install the new version of NCC (if you have detected a new version of NCC and have not installed it yet).

        - Log on to any Controller VM in the cluster and establish an SSH session.
        - Check the health of all the services.

          ```
          nutanix@cvm$ ncc health_checks run_all
          ```

          If the check reports a status other than PASS, resolve the reported issues before proceeding

          If you are unable to resolve the issues, contact Nutanix Support for assistance.

    - Configure email frequency to allow the cluster to check, run, and email reports at regular intervals as configured.

      For more information, see Configuring Alert Policies in the *Prism Web Console Guide*.

Controller VM Access

Although each host in a Nutanix cluster runs a hypervisor independent of other hosts in the cluster, some operations affect the entire cluster.

Most administrative functions of a Nutanix cluster can be performed through the web console (Prism), however, there are some management tasks that require access to the Controller VM (CVM) over SSH. Nutanix recommends restricting Controller VM (CVM) SSH access with password or key authentication.

This topic provides information about how to access the CVM as an `admin` user and `nutanix` user.

> Note: The direct SSH access to CVM and Prism Central VM (PCVM) is disabled for the **root** user.

admin User Access

Use the `admin` user access for all tasks and operations that you must perform on the controller VM. As an `admin` user with default credentials, you cannot access nCLI. You must change the default password before you can use nCLI. Nutanix recommends that you do not create additional CVM user accounts. Use the default accounts (`admin` or `nutanix`), or use **sudo** to elevate to the `root` account.

For more information about `admin` user access, see Admin User Access to Controller VM on page 26.

nutanix User Access

Nutanix strongly recommends that you do not use the `nutanix` user access unless the procedure (as provided in a Nutanix Knowledge Base article or user guide) specifically requires the use of the `nutanix` user access.

For more information about `nutanix` user access, see Nutanix User Access to Controller VM on page 28.

You can perform most administrative functions of a Nutanix cluster through the Prism web consoles or REST API. Nutanix recommends using these interfaces whenever possible and disabling Controller VM SSH access with password or key authentication. Some functions, however, require logging on to a Controller VM with SSH. Exercise caution whenever connecting directly to a Controller VM as it increases the risk of causing cluster issues.

> Warning: When you connect to a Controller VM with SSH, ensure that the SSH client does not import or change any locale settings. The Nutanix software is not localized, and running the commands with any locale other than en_US.UTF-8 can cause severe cluster issues.
>
> To check the locale used in an SSH session, run **`/usr/bin/locale`**. If any environment variables are set to anything other than en_US.UTF-8, reconnect with an SSH configuration that does not import or change any locale settings.

Admin User Access to Controller VM

You can access the Controller VM as the admin user (**`admin`** user name and password) with SSH. For security reasons, the password of the admin user must meet Controller VM Password Complexity Requirements. When you log on to the Controller VM as the admin user for the first time, you are prompted to change the default password.

See Controller VM Password Complexity Requirements to set a secure password.

After you have successfully changed the password, the new password is synchronized across all Controller VMs and interfaces (Prism web console, nCLI, and SSH).

> Note:
>
> - As an **`admin`** user, you cannot access nCLI by using the default credentials. If you are logging in as the **`admin`** user for the first time, you must log on through the Prism web console or SSH to the Controller VM. Also, you cannot change the default password of the **`admin`** user through nCLI. To change the default password of the **`admin`** user, you must log on through the Prism web console or SSH to the Controller VM.
> - When you make an attempt to log in to the Prism web console for the first time after you upgrade to AOS 5.1 from an earlier AOS version, you can use your existing **`admin`** user password to log in and then change the existing password (you are prompted) to adhere to the password complexity requirements. However, if you are logging in to the Controller VM with SSH for the first time after the upgrade as the **`admin`** user, you must use the default **`admin`** user password (Nutanix/4u) and then change the default password (you are prompted) to adhere to the Controller VM Password Complexity Requirements.
> - You cannot delete the **`admin`** user account.

- The default password expiration age for the `admin` user is 60 days. You can configure the minimum and maximum password expiration days based on your security requirement.
    - `nutanix@cvm$ sudo chage -M MAX-DAYS admin`
    - `nutanix@cvm$ sudo chage -m MIN-DAYS admin`

When you change the `admin` user password, you must update any applications and scripts using the `admin` user credentials for authentication. Nutanix recommends that you create a user assigned with the admin role instead of using the `admin` user for authentication. The Prism Element Web Console Guide describes authentication and roles.

Following are the default credentials to access a Controller VM.

Table 1: Controller VM Credentials

Interface	Target	User Name	Password
SSH client	Nutanix Controller VM	`admin`	Nutanix/4u
		`nutanix`	nutanix/4u
Prism web console	Nutanix Controller VM	`admin`	Nutanix/4u

Accessing the Controller VM Using the Admin User Account

About this task

Perform the following procedure to log on to the Controller VM by using the admin user with SSH for the first time.

Procedure

1. Log on to the Controller VM with SSH by using the management IP address of the Controller VM and the following credentials.

    - User name: `admin`
    - Password: `Nutanix/4u`

    You are now prompted to change the default password.

2. Respond to the prompts, providing the current and new `admin` user password.

    ```
    Changing password for admin.
    Old Password:
    New password:
    Retype new password:
    Password changed.
    ```

    See the requirements listed in Controller VM Password Complexity Requirements to set a secure password.

    For information about logging on to a Controller VM by using the `admin` user account through the Prism web console, see Logging Into The Web Console in the *Prism Element Web Console Guide*.

Nutanix User Access to Controller VM

You can access the Controller VM as the `nutanix` user (`nutanix` user name and password) with SSH. For security reasons, the password of the `nutanix` user must meet the Controller VM Password Complexity Requirements on page 29. When you log on to the Controller VM as the `nutanix` user for the first time, you are prompted to change the default password.

See Controller VM Password Complexity Requirements on page 29to set a secure password.

After you have successfully changed the password, the new password is synchronized across all Controller VMs and interfaces (Prism web console, nCLI, and SSH).

Note:

- As a `nutanix` user, you cannot access nCLI by using the default credentials. If you are logging in as the `nutanix` user for the first time, you must log on through the Prism web console or SSH to the Controller VM. Also, you cannot change the default password of the `nutanix` user through nCLI. To change the default password of the `nutanix` user, you must log on through the Prism web console or SSH to the Controller VM.
- When you make an attempt to log in to the Prism web console for the first time after you upgrade the AOS from an earlier AOS version, you can use your existing **nutanix** user password to log in and then change the existing password (you are prompted) to adhere to the password complexity requirements. However, if you are logging in to the Controller VM with SSH for the first time after the upgrade as the **nutanix** user, you must use the default `nutanix` user password (nutanix/4u) and then change the default password (you are prompted) to adhere to the Controller VM Password Complexity Requirements on page 29.
- You cannot delete the `nutanix` user account.
- For enhanced access restrictions for the `nutanix` user, consider enabling the **Cluster Lockdown** feature. For more information, see Controlling Cluster Access section in *Security Guide*.

Important: Nutanix does not recommend changing the password expiry setting for the `nutanix` user account. An expired `nutanix` user account can cause cluster stability issues.

When you change the `nutanix` user password, you must update any applications and scripts using the `nutanix` user credentials for authentication. Nutanix recommends that you create a user assigned with the nutanix role instead of using the `nutanix` user for authentication. The Prism Element Web Console Guide describes authentication and roles.

Following are the default credentials to access a Controller VM.

Table 2: Controller VM Credentials

Interface	Target	User Name	Password
SSH client	Nutanix Controller VM	`admin`	Nutanix/4u
		`nutanix`	nutanix/4u
Prism web console	Nutanix Controller VM	`admin`	Nutanix/4u

Accessing the Controller VM Using the Nutanix User Account

About this task

Perform the following procedure to log on to the Controller VM by using the nutanix user with SSH for the first time.

Procedure

1. Log on to the Controller VM with SSH by using the management IP address of the Controller VM and the following credentials.

    - User name: `nutanix`
    - Password: `nutanix/4u`

    You are now prompted to change the default password.

2. Respond to the prompts, providing the current and new `nutanix` user password.

    ```
    Changing password for nutanix.
    Old Password:
    New password:
    Retype new password:
    Password changed.
    ```

    See Controller VM Password Complexity Requirements on page 29to set a secure password.

    For information about logging on to a Controller VM by using the nutanix user account through the Prism web console, see Logging Into The Web Console in the *Prism Element Web Console Guide*.

Controller VM Password Complexity Requirements

The password must meet the following complexity requirements:

- At least eight characters long.
- At least one lowercase letter.
- At least one uppercase letter.
- At least one number.
- At least one special character.

    > Note: Ensure that the following conditions are met for the special characters usage in the CVM password:
    >
    > - The special characters are appropriately used while setting up the CVM password. In some cases, for example when you use *! followed by a number* in the CVM password, it leads to a special meaning at the system end, and the system may replace it with a command from the bash history. In this case, you may generate a password string different from the actual password that you intend to set.
    > - The special character used in the CVM password are ASCII printable characters only. For information about ACSII printable characters, refer *ASCII printable characters (character code 32-127)* article on ASCII code website.

- At least four characters difference from the old password.

- Must not be among the last 5 passwords.
- Must not have more than 2 consecutive occurrences of a character.
- Must not be longer than 199 characters.

If a password for an account (CVM account) is entered five times unsuccessfully within a 15-minute period, the account is locked for 15 minutes.

Cluster Operations

This section describes how to manage cluster operations such as starting, stopping, destroying, and expanding a cluster.

- Stopping a Cluster on page 30
- Node Removal on page 31
- Starting a Nutanix Cluster on page 32
- Destroying a Cluster on page 33
- For information about cluster expansion, see Expanding a Cluster in the *Prism Web Console Guide*.

Stopping a Cluster

Before you begin

Shut down all guest virtual machines, including vCenter if it is running on the cluster. Do not shut down Nutanix Controller VMs.

> Note:
>
> - If you are running Files, stop Files before stopping your AOS cluster. This task stops all services provided by guest virtual machines and the Nutanix cluster.
> - If you are planning to stop your cluster that has metro availability configured, do not stop the cluster before performing some remedial actions. For more information, see Conditions for Implementing Data Protection (Metro Availability) in the *Prism Web Console Guide*.

(Hyper-V only) Stop the Hyper-V failover cluster by logging on to a Hyper-V host and running the `Stop-Cluster` PowerShell command.

About this task

> Note: This procedure stops all services provided by guest virtual machines, the Nutanix cluster, and the hypervisor host.

Procedure

1. Log on to a running CVM in the Nutanix cluster with SSH.
2. Verify all services are up on all Controller VMs.

```
nutanix@cvm$ cluster status
```

3. Stop the Nutanix cluster.

   ```
   nutanix@cvm$ cluster stop
   ```

   You must confirm by typing `I agree`.

   Wait to proceed until output similar to the following is displayed for every Controller VM in the cluster.

   ```
   CVM:host IP-Address
                                   Zeus   UP       [9935, 9980, 9981, 9994, 10015, 10037]
                              Scavenger   UP       [25880, 26061, 26062]
                                 Xmount   UP       [5556, 5593]
                       SysStatCollector DOWN       []
                              IkatProxy DOWN       []
                       IkatControlPlane DOWN       []
                          SSLTerminator DOWN       []
                         SecureFileSync DOWN       []
                                 Medusa DOWN       []
                     DynamicRingChanger DOWN       []
                                 Pithos DOWN       []
                             InsightsDB DOWN       []
                                 Athena DOWN       []
                                Mercury DOWN       []
                                 Mantle DOWN       []
                             VipMonitor   UP       [18387, 18464, 18465, 18466, 18474]
                               Stargate DOWN       []
                   InsightsDataTransfer DOWN       []
                                  Ergon DOWN       []
                                Cerebro DOWN       []
                                Chronos DOWN       []
                                Curator DOWN       []
                                  Prism DOWN       []
                                    CIM DOWN       []
                           AlertManager DOWN       []
                               Arithmos DOWN       []
                                Catalog DOWN       []
                              Acropolis DOWN       []
                                  Uhura DOWN       []
                                   Snmp DOWN       []
                      NutanixGuestTools DOWN       []
                             MinervaCVM DOWN       []
                          ClusterConfig DOWN       []
                               Aequitas DOWN       []
                            APLOSEngine DOWN       []
                                  APLOS DOWN       []
                                  Lazan DOWN       []
                                 Delphi DOWN       []
                                   Flow DOWN       []
                                Anduril DOWN       []
                                  XTrim DOWN       []
                          ClusterHealth DOWN       []
   ```

Node Removal

You may need to remove a node for various reasons such as replacement of a failed node or to deprecate an old node for cluster expansion. You can remove a node using the Prism web console or nCLI.

To remove a node (host) from the cluster, see Removing a Node in the *Prism Web Console Guide*.

Starting a Nutanix Cluster

About this task

This task describes the process of starting a cluster after you have either stopped or destroyed a cluster and have manually created the cluster again.

Perform the following steps to start a cluster:

Procedure

1. Log on to any CVM in the cluster using SSH.
2. Verify the Nutanix cluster.

   ```
   nutanix@cvm$ cluster start
   ```

   If the cluster is running properly, the output displays the status of all the applications on all the CVMs. All of them must display the status as UP. An output similar to the following is displayed for each node in the cluster:

   ```
   CVM:host IP-Address Up
                                   Zeus   UP       [9935, 9980, 9981, 9994, 10015, 10037]
                              Scavenger   UP       [25880, 26061, 26062]
                                 Xmount   UP       [21170, 21208]
                       SysStatCollector   UP       [22272, 22330, 22331]
                              IkatProxy   UP       [23213, 23262]
                       IkatControlPlane   UP       [23487, 23565]
                          SSLTerminator   UP       [23490, 23620]
                         SecureFileSync   UP       [23496, 23645, 23646]
                                 Medusa   UP       [23912, 23944, 23945, 23946, 24176]
                     DynamicRingChanger   UP       [24314, 24404, 24405, 24558]
                                 Pithos   UP       [24317, 24555, 24556, 24593]
                             InsightsDB   UP       [24322, 24472, 24473, 24583]
                                 Athena   UP       [24329, 24504, 24505]
                                Mercury   UP       [24338, 24515, 24516, 24614]
                                 Mantle   UP       [24344, 24572, 24573, 24634]
                             VipMonitor   UP       [18387, 18464, 18465, 18466, 18474]
                               Stargate   UP       [24993, 25032]
                   InsightsDataTransfer   UP       [25258, 25348, 25349, 25388, 25391, 25393,
    25396]
                                  Ergon   UP       [25263, 25414, 25415]
                                Cerebro   UP       [25272, 25462, 25464, 25581]
                                Chronos   UP       [25281, 25488, 25489, 25547]
                                Curator   UP       [25294, 25528, 25529, 25585]
                                  Prism   UP       [25718, 25801, 25802, 25899, 25901, 25906,
    25941, 25942]
                                    CIM   UP       [25721, 25829, 25830, 25856]
                           AlertManager   UP       [25727, 25862, 25863, 25990]
                               Arithmos   UP       [25737, 25896, 25897, 26040]
                                Catalog   UP       [25749, 25989, 25991]
                              Acropolis   UP       [26011, 26118, 26119]
                                  Uhura   UP       [26037, 26165, 26166]
                                   Snmp   UP       [26057, 26214, 26215]
                      NutanixGuestTools   UP       [26105, 26282, 26283, 26299]
                             MinervaCVM   UP       [27343, 27465, 27466, 27730]
                          ClusterConfig   UP       [27358, 27509, 27510]
                               Aequitas   UP       [27368, 27567, 27568, 27600]
                            APLOSEngine   UP       [27399, 27580, 27581]
                                  APLOS   UP       [27853, 27946, 27947]
                                  Lazan   UP       [27865, 27997, 27999]
                                 Delphi   UP       [27880, 28058, 28060]
                                   Flow   UP       [27896, 28121, 28124]
   ```

```
                              Anduril   UP       [27913, 28143, 28145]
                                XTrim   UP       [27956, 28171, 28172]
                        ClusterHealth   UP       [7102, 7103, 27995, 28209,28495, 28496,
 28503, 28510,
28573, 28574, 28577, 28594, 28595, 28597, 28598, 28602, 28603, 28604, 28607, 28645, 28646,
 28648, 28792,
28793, 28837, 28838, 28840, 28841, 28858, 28859, 29123, 29124, 29127, 29133, 29135, 29142,
 29146, 29150,
29161, 29162, 29163, 29179, 29187, 29219, 29268, 29273]
```

What to do next

After you have verified that the cluster is running, you can start guest VMs.

(Hyper-V only) If the Hyper-V failover cluster was stopped, start it by logging on to a Hyper-V host and running the `Start-Cluster` PowerShell command.

> Warning: By default, Nutanix clusters have redundancy factor 2, which means they can tolerate the failure of a single node or drive. Nutanix clusters with a configured option of redundancy factor 3 allow the Nutanix cluster to withstand the failure of two nodes or drives in different blocks.
>
> - Never shut down or restart multiple Controller VMs or hosts simultaneously.
> - Always run the **`cluster status`** command to verify that all Controller VMs are up before performing a Controller VM or host shutdown or restart.

Destroying a Cluster

This section provides information about how to destroy a cluster.

Before you begin

Ensure that you meet the following prerequisites before you destroy a cluster:

- Reclaim licenses from the cluster to be destroyed by following the unlicensing instructions in the License Manager Guide.

  > Note:
  >
  > - If you have destroyed the cluster and did not reclaim the existing licenses, contact Nutanix Support to reclaim the licenses.
  > - Reclaiming licenses is required to remove the cluster from the insights portal.

- If the cluster is registered with Prism Central (the multiple cluster manager VM), unregister the cluster before destroying it. For more information, see Register (Unregister) Cluster with Prism Central in the *Prism Web Console Guide*.
- If you are running Nutanix Files, stop Files before stopping your AOS cluster. This task stops all services provided by guest virtual machines and the Nutanix cluster.

About this task

> Warning: Destroying a cluster resets all nodes in the cluster to the factory configuration. All cluster configuration and guest VM data are unrecoverable after destroying the cluster. This action is not reversible and you must use this procedure discerningly.

> Note: You need admin user access to destroy a cluster.

Procedure

To destroy a cluster, perform the following steps:

1. Log on to any CVM in the cluster using SSH.
2. Power off all the VMs that are running on the hosts in the cluster.

    ```
    nutanix@cvm$ acli vm.off *
    ```

    Alternatively, you can log into the Web Console and power off all the VMs.

3. Stop the Nutanix cluster.

    ```
    nutanix@cvm$ cluster stop
    ```

    You must confirm by typing `I agree`.

    Wait to proceed until output similar to the following is displayed for every Controller VM in the cluster.

    ```
    CVM:host IP-Address
                                    Zeus   UP       [9935, 9980, 9981, 9994, 10015, 10037]
                               Scavenger   UP       [25880, 26061, 26062]
                                  Xmount   UP       [5556, 5593]
                        SysStatCollector DOWN       []
                               IkatProxy DOWN       []
                        IkatControlPlane DOWN       []
                           SSLTerminator DOWN       []
                          SecureFileSync DOWN       []
                                  Medusa DOWN       []
                      DynamicRingChanger DOWN       []
                                  Pithos DOWN       []
                              InsightsDB DOWN       []
                                  Athena DOWN       []
                                 Mercury DOWN       []
                                  Mantle DOWN       []
                              VipMonitor   UP       [18387, 18464, 18465, 18466, 18474]
                                Stargate DOWN       []
                    InsightsDataTransfer DOWN       []
                                   Ergon DOWN       []
                                 Cerebro DOWN       []
                                 Chronos DOWN       []
                                 Curator DOWN       []
                                   Prism DOWN       []
                                     CIM DOWN       []
                            AlertManager DOWN       []
                                Arithmos DOWN       []
                                 Catalog DOWN       []
                               Acropolis DOWN       []
                                   Uhura DOWN       []
                                    Snmp DOWN       []
                       NutanixGuestTools DOWN       []
                              MinervaCVM DOWN       []
                           ClusterConfig DOWN       []
                                Aequitas DOWN       []
                             APLOSEngine DOWN       []
                                   APLOS DOWN       []
                                   Lazan DOWN       []
                                  Delphi DOWN       []
                                    Flow DOWN       []
                                 Anduril DOWN       []
                                   XTrim DOWN       []
                           ClusterHealth DOWN       []
    ```

4. Destroy the cluster.

```
nutanix@cvm$ cluster destroy
```

> Caution: Performing this operation deletes all cluster and guest VM data in the cluster.

Follow the prompt to confirm destruction of the cluster.

Fingerprinting Existing vDisks

The vDisk manipulator utility fingerprints vDisks that existed in the cluster before deduplication was enabled.

Before you begin

The storage container must have fingerprint-on-write enabled.

Procedure

Run the vDisk manipulator utility from any Controller VM in the cluster.

» To fingerprint a particular vDisk:

```
nutanix@cvm$ vdisk_manipulator --operation="add_fingerprints" \
 --stats_only="false" --nfs_container_name="ctr_name" \
 --nfs_relative_file_path="vdisk_path"
```

- Replace *ctr_name* with the name of the storage container where the vDisk to fingerprint resides.
- Replace *vdisk_path* with the path of the vDisk to fingerprint relative to the storage container path (for example, `Win7-desktop11/Win7-desktop11-flat.vmdk`). You cannot specify multiple vDisks in this parameter.

» To fingerprint all vDisks in the cluster:

```
nutanix@cvm$ ncli vdisk list | grep "Name.*NFS" | awk -F: \
 '{print $4 ":" $5 ":" $6 ":" $7}' >> fingerprint.txt
nutanix@cvm$ for i in `cat fingerprint.txt`; do vdisk_manipulator --vdisk_name=$i \
 --operation="add_fingerprints" --stats_only=false; done
```

> Note: You can run `vdisk_manipulator` in a loop to fingerprint multiple vDisks, but run only one instance of `vdisk_manipulator` on each Controller VM at a time. Executing multiple instances on a Controller VM concurrently would generate significant load on the cluster.

IPv6 Enablement in a Cluster

Nutanix provides two ways to configure IPv6 on a cluster node. You can either use the `manage_ipv6` CLI command's interactive inputs to configure IPv6 on the CVM, or pass the `ips.json` file as an input.

> Note: The current IPv6 implementation is allowed only on a Prism Element (PE) cluster using AHV. Prism Central (PC) does not support IPv6. Before you enable IPv6 on a PE cluster, ensure that the PE is not registered to any PC instance.
>
> After you enable IPv6 on the PE cluster, you cannot register the PE cluster again to PC.

All external connections made to the Nutanix cluster support IPv6.

All inbound connection requests for the Cluster Virtual IP for Prism Element, Files CIFS and NFS clients, and DR of Files NFS and CIFS data support dual stack.

Considerations for IPv6 Enablement

The IPv6 enablement configuration is confined to:

- Cluster VIP—Supports dual stack and accessible by either IPv6 or IPv4 simultaneously without the need to reconfigure.
- FSVM CIFS and NFS client connections—Supports dual stack and accessible by either IPv6 or IPv4 simultaneously without the need to reconfigure.
- DR between clusters for Files data—The remote site connection supports either IPv4 or IPv6. Both are not supported at the same time for a single remote site.
- CVM and AHV access over SSH—CVM and AHV are accessible by both IPv4 address and IPv6 address simultaneously over SSH without the need to reconfigure.

IPv6 Architecture

The image below illustrates the IPv6 implementation on a cluster.

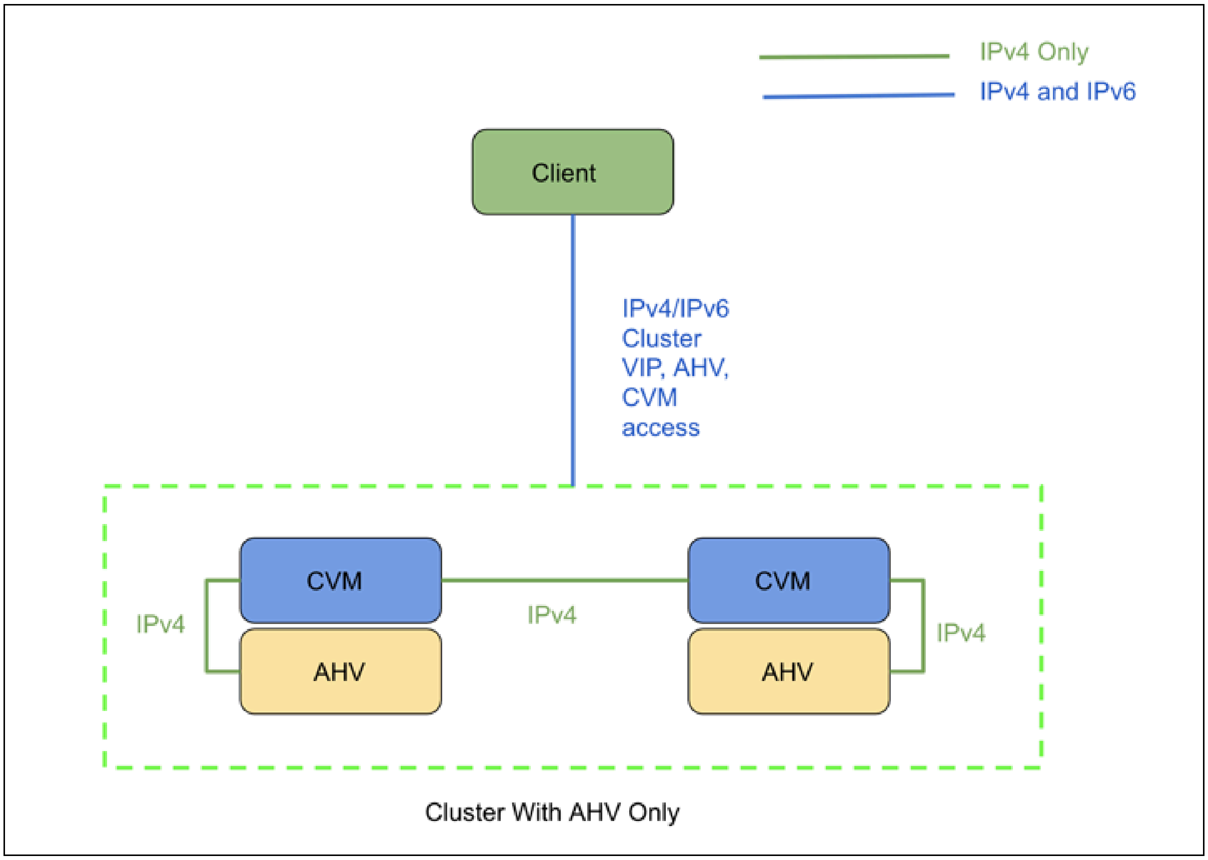


Figure 3: IPv6 Enablement Approach

IPv6 Configuration

Refer to the following sections to configure the IPv6 on a cluster:

- Configuring IPv6 Using Script
- Configure Virtual IPv6
- Adding IPv6 Node

- Configuring a Remote Site (Physical Cluster)

Configuring IPv6 Using ips.json File

About this task

Use the `ips.json` file to configure IPv6 addresses for all the existing nodes with IPv4 addresses in the cluster.

The `ips.json` file provides you an easier method to configure multiple nodes in a cluster instead of manually entering the IPv6 address for each node. The `ips.json` file also helps you to avoid cluster restart due to erroneous manual IPv6 configuration.

Procedure

1. Log on to the CVM.

2. Verify the IP addresses in the `ips.json` file.

   ```
   nutanix@cvm$ vim ips.json
   ```

   > Note: By default, IPv6 is not enabled on AHV. The **`manage_ipv6`** CLI command enables IPv6 on AHV.

   An output similar to the following is displayed:

   ```
   {
     "svmips": {
       "x.x.x.84": "2001:db8::",
       "x.x.x.87": "2001:db8::",
       "x.x.x.92": "2001:db8::"
     },
     "hostips": {
       "x.x.x.154": "2001:db8::",
       "x.x.x.83": "2001:db8::",
       "x.x.x.90": "2001:db8::"
     },
     "prefixlen": 32,
     "gateway": "fc00:0:0:10::1"
   }
   ```

3. Pass the `ips.json` file to configure the nodes with IPv6 addresses.

   ```
   nutanix@cvm$ manage_ipv6 -i ips.json configure
   ```

   This command takes two IPv6 addresses per node and assign them to the **`eth0`** Controller VM and **`br0`** AHV interface.

   An output similar to the following is displayed:

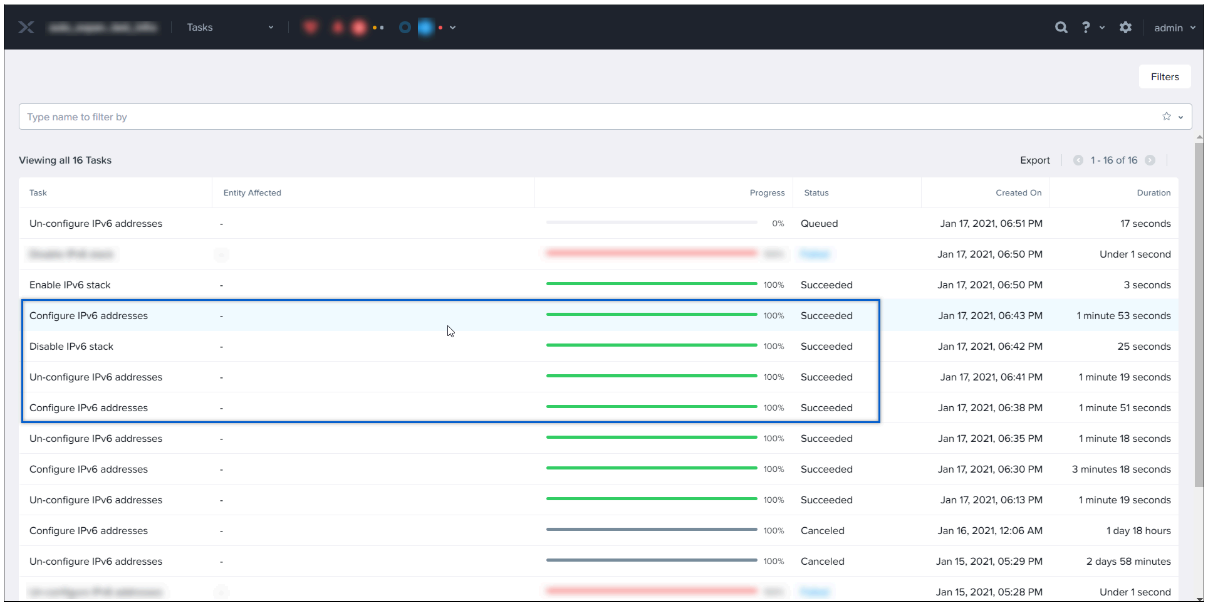
   ```
   [INFO] Reading IPv6 config from JSON file: ips.json
   [INFO] IPv6 config to configure: {
     "svmips": {
       "x.x.x.84": "2001:db8::",
       "x.x.x.87": "2001:db8::",
       "x.x.x.92": "2001:db8::"
     },
     "hostips": {
       "x.x.x.154": "2001:db8::",
       "x.x.x.83": "2001:db8::",
       "x.x.x.90": "2001:db8::"
     },
   ```

```
  "prefixlen": 32,
  "gateway": "fc00:0:0:10::1"
}
```

a. Enter **y** to apply IPv6 configuration. An output similar to the following is displayed:

   ```
   Proceed to apply above configuration? [Y/N]: y
   [INFO] 1. IPv6 enabled on all CVMs and hypervisors
   [INFO] 2. CVM and Hypervisor IPv6 addresses configured
   [INFO] 3. Stored IPv6 configuration in Zeus
   [INFO] 4. CVM and hypervisor firewall rules updated
   [INFO] 5. Necessary services have been restarted
   [INFO] Marked Ergon task 0dc5c0c8-09ff-44a5-8efd-f27d8a6cfec2 as kSucceeded
   [INFO] Action configure completed successfully
   Script output logged to /home/nutanix/data/logs/manage_ipv6.out
   ```

4. In the Prism Element web console, select the **Recent Task** icon in the taskbar and click **View All Tasks** to see the newly assigned IPv6 addresses on the nodes and other related task status.



Viewing all 16 Tasks

Task	Entity Affected	Progress	Status	Created On	Duration
Un-configure IPv6 addresses	-	0%	Queued	Jan 17, 2021, 06:51 PM	17 seconds
[redacted]	[redacted]	[redacted]	[redacted]	Jan 17, 2021, 06:50 PM	Under 1 second
Enable IPv6 stack	-	100%	Succeeded	Jan 17, 2021, 06:50 PM	3 seconds
Configure IPv6 addresses	-	100%	Succeeded	Jan 17, 2021, 06:43 PM	1 minute 53 seconds
Disable IPv6 stack	-	100%	Succeeded	Jan 17, 2021, 06:42 PM	25 seconds
Un-configure IPv6 addresses	-	100%	Succeeded	Jan 17, 2021, 06:41 PM	1 minute 19 seconds
Configure IPv6 addresses	-	100%	Succeeded	Jan 17, 2021, 06:38 PM	1 minute 51 seconds
Un-configure IPv6 addresses	-	100%	Succeeded	Jan 17, 2021, 06:35 PM	1 minute 18 seconds
Configure IPv6 addresses	-	100%	Succeeded	Jan 17, 2021, 06:30 PM	3 minutes 18 seconds
Un-configure IPv6 addresses	-	100%	Succeeded	Jan 17, 2021, 06:13 PM	1 minute 19 seconds
Configure IPv6 addresses	-	100%	Canceled	Jan 16, 2021, 12:06 AM	1 day 18 hours
Un-configure IPv6 addresses	-	100%	Canceled	Jan 15, 2021, 05:29 PM	2 days 58 minutes
[redacted]	[redacted]	[redacted]	[redacted]	Jan 15, 2021, 05:28 PM	Under 1 second

Figure 4: Cluster IPv6 Configuration Details

Refer to the following CLI commands for all the other possible IPv6 actions:

- Remove the IPs assigned to the nodes in the cluster using the `ips.json` script.

  ```
  nutanix@cvm$ manage_ipv6 unconfigure
  ```

  > Note: Remove the IPv6 VIP from the cluster if Prism Element is configured with VIP v6.

- Completely disable IPv6 addresses in the AOS. The user cannot configure IPv6 manually either. The script restarts all the services as required.

  ```
  nutanix@cvm$ manage_ipv6 disable
  ```

- Enable IPv6 addresses configuration in the AOS.

  ```
  nutanix@cvm$ manage_ipv6 enable
  ```

- Change the IP addresses assigned using the script.

  ```
  nutanix@cvm$ manage_ipv6 reconfigure
  ```

- View all the nodes' current IPv6 configuration in YAML format and any other irregularities, if present, in the cluster.

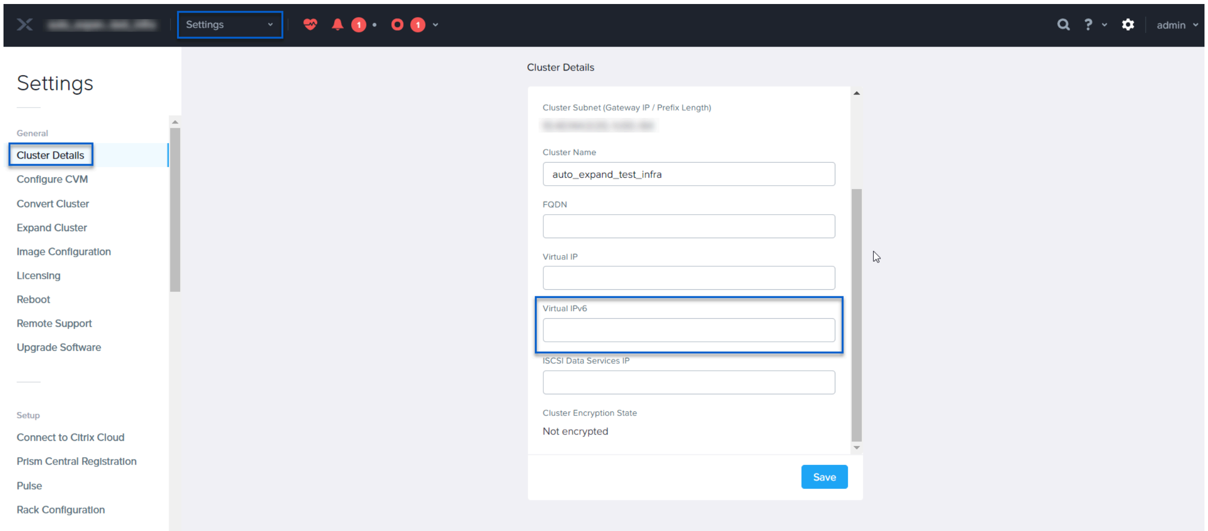
  ```
  nutanix@cvm$ manage_ipv6 show
  ```

5. To unconfigure an IPv6 address, manually remove the cluster Virtual IP address in Prism and run this command.

   ```
   manage_ipv6 unconfigure
   ```

Configure Virtual IPv6

After configuring IPv6 on the cluster using the `manage_ipv6` CLI command, assign a Virtual IPv6 using the Prism Element Web Console.



Cluster Details

Cluster Subnet (Gateway IP / Prefix Length)
[redacted]

Cluster Name
auto_expand_test_infra

FQDN

Virtual IP

Virtual IPv6

ISCSI Data Services IP

Cluster Encryption State
Not encrypted

Figure 5: Configuring Virtual IPv6 on a cluster

In the **Virtual IPv6** field, enter an IPv6 address that will be used as a virtual IPv6 for the cluster. A controller VM runs on each node and has its own IPv6 address, but this field sets a logical IPv6 address that always connects to one of the active Controller VM in the cluster (assuming at least one is active), which removes the need to enter the address of a specific Controller VM. The virtual IPv6 address is normally set during the initial cluster configuration (see the Field Installation Guide), but you can update the address at any time through this field.

For information on adding or modifying cluster parameters, see Modifying Cluster Details in the *Prism Web Console Guide*.

Name Server also supports IPv6. For more information about name servers, see Configuring Name Servers in the *Prism Web Console Guide*.

Adding IPv6 Node

A cluster is a collection of nodes. You can add new nodes to a cluster at any time after physically installing and connecting them to the network on the same subnet as the cluster. The cluster expansion process compares the AOS version on the existing and new nodes and performs any upgrades necessary for all nodes to have the same AOS version.

Before you begin

- Before attempting to add a node to the cluster, review the Prerequisites and Requirements in the *Prism Web Console Guide*. The process for adding a node varies depending on several factors, and this section covers specific considerations based on your AOS, hypervisor, encryption, and hardware configuration.
- Check the **Health Dashboard**. If any health checks are failing, resolve them to ensure that the cluster is healthy before adding any nodes. For more information, see Health Dashboard in the *Prism Web Console Guide*.
- Allow any current add node operations to complete.

- Ensure that all nodes are in the correct metadata state by checking the **Hardware Dashboard**. If any nodes show `Metadata store disabled on the node or Node is removed from metadata store`, enable the metadata store by clicking **Enable Metadata Store**. For more information, see Hardware Dashboard in the *Prism Web Console Guide*.

About this task

To add one or more nodes to an existing cluster (you can add multiple nodes at the same time), do the following:

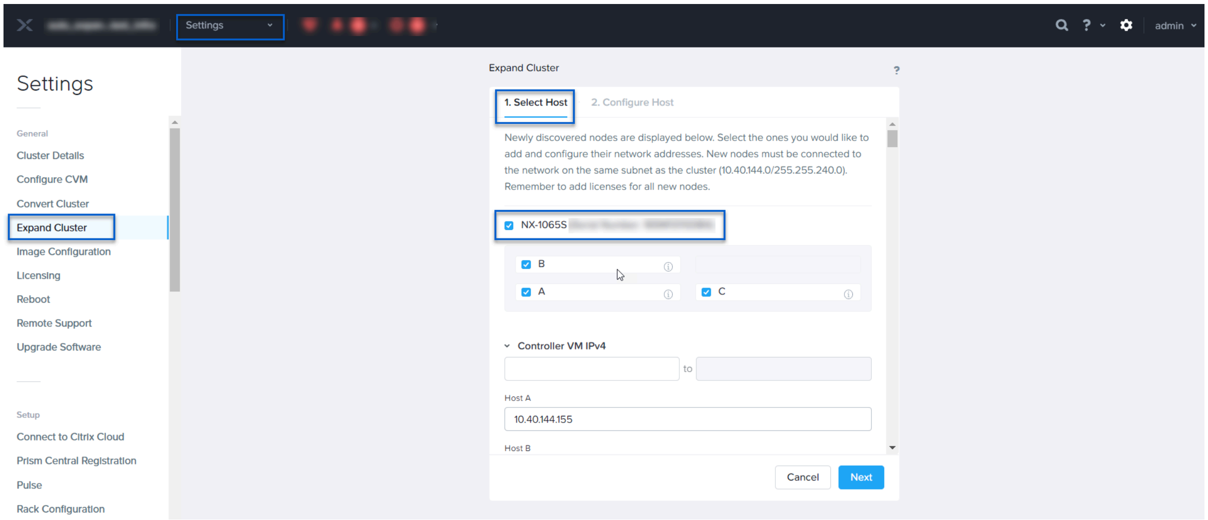
Procedure

1. Either click the gear icon in the main menu and then select **Expand Cluster** in the **Settings** page or go to the **Hardware Dashboard** and click the **Expand Cluster** button.

   The network searches for Nutanix nodes and then the **Expand Cluster** dialog box appears (on the **Select Host** screen) with a graphical list of the discovered blocks and nodes. Discovered blocks are blocks with one or more unassigned factory-prepared nodes (hypervisor and Controller VM installed) residing on the same subnet as the cluster. Discovery requires that IPv6 multicast packets are allowed through the physical switch.

   A lack of IPv6 multicast support prevents node discovery and successful cluster expansion.

2. Select the checkbox for each block that you want to add to the cluster. All nodes within the selected block are included automatically; Clear the checkbox of the nodes that are not required to be added to the cluster.



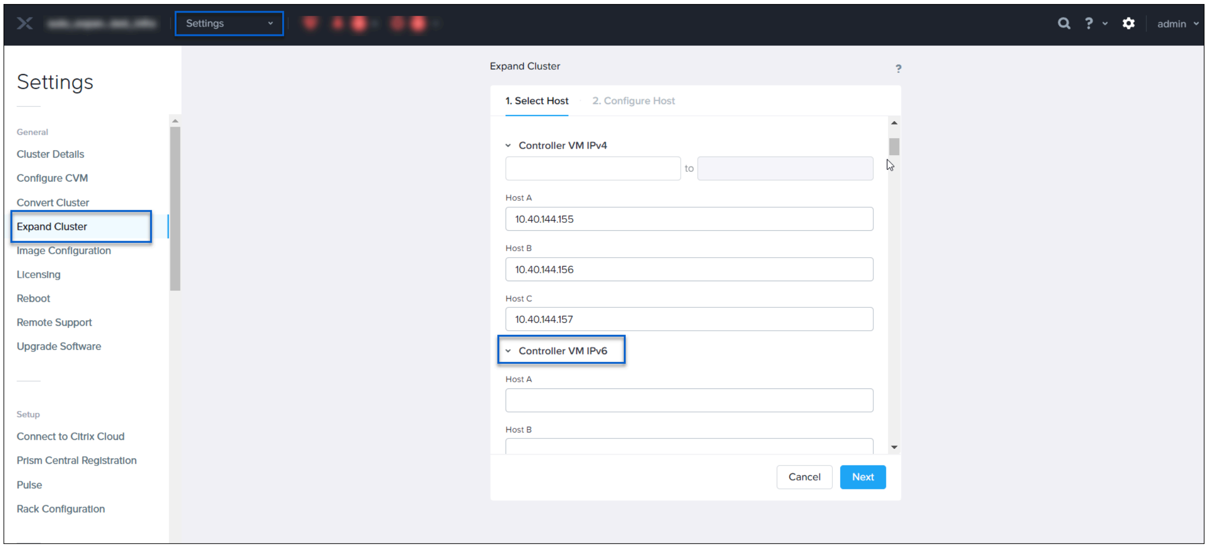
Newly discovered nodes are displayed below. Select the ones you would like to add and configure their network addresses. New nodes must be connected to the network on the same subnet as the cluster (10.40.144.0/255.255.240.0). Remember to add licenses for all new nodes.

Host A: 10.40.144.155

Figure 6: Adding an IPv6 node

When you select a block, more fields appear below the block diagram. A separate line for each node (host) in the block appears under each field name.

3. Do the following in the indicated fields for each checked block:



Controller VM IPv4

Host A: 10.40.144.155

Host B: 10.40.144.156

Host C: 10.40.144.157

Figure 7: Reviewing the Controller VMs IPv6 address

a. Host Name—Enter the name of the host.

   Enter just the host name, not the fully qualified domain name. Hyper-V requires hostname, but you can ignore this field if your clusters are running ESXi or AHV hypervisors.

b. Controller VM IPv4—Review the Controller VM IPv4 address assigned to each host and do one of the following:

   - If the address is correct, do nothing in this field.
   - If the address is not correct, either change the incorrect address or enter a starting address on the top line (for multiple hosts). The entered IPv4 address is assigned to the

Controller VM of the first host, and consecutive IPv4 addresses (sequentially from the entered address) are automatically assigned to the remaining hosts.

c. Controller VM IPv6—Review the Controller VM IPv6 address assigned to each host.

   If the address is not correct, either change the incorrect address or enter a starting address on the top line (for multiple hosts). The entered IPv6 address is assigned to the Controller VM of the first host, and the other IPv6 addresses are manually configured.

   > Note: This is an optional configuration and is required only if IPv6 is configured on the cluster.

d. Hypervisor IPv4—Repeat the previous step for this field.

   This field sets the hypervisor IPv4 addresses for all the hosts to be added.

e. Hypervisor IPv6—Repeat the previous step for this field.

   This field sets the hypervisor IPv6 addresses for all the hosts to be added.

   > Note: This is an optional configuration and is required only if IPv6 is configured on the cluster.

f. IPMI IPv4: Repeat the previous step for this field.

   This field sets the IPMI port IPv4 addresses for all the hosts to be added. An IPMI port is used for the hypervisor host console.

g. IPMI IPv6:Repeat the previous step for this field.

   This field sets the IPMI port IPv6 addresses for all the hosts to be added. An IPMI port is used for the hypervisor host console.

   > Note: This is an optional configuration and is required only if IPv6 is configured on the cluster.

   > Caution: This feature is for future use. Do not use tech preview features in production environments.

h. When all the node values are correct, click Next button (lower right).

   The network addresses are validated before continuing. If an issue is discovered, the problem addresses are highlighted in red. If there are no issues, the process moves to the Assign Rack screen with a message at the top when the hypervisor, AOS, or other relevant feature is incompatible with the cluster version.

For the remaining procedure to expand clusters, see Expanding a Cluster in the *Prism Web Console Guide*.

Configuring a Remote Site (Physical Cluster)

Before you begin

- For using network mapping, create the network connections and VLANs on both source and destination cluster. For more information about configuring network connections, see Network Configuration for VM Interfaces in the *Prism Web Console Guide*.
- Set up the network mapping on both the source and the destination clusters. For more information, see Network Mapping in the *Data Protection and Recovery with Prism Element Guide*.

About this task

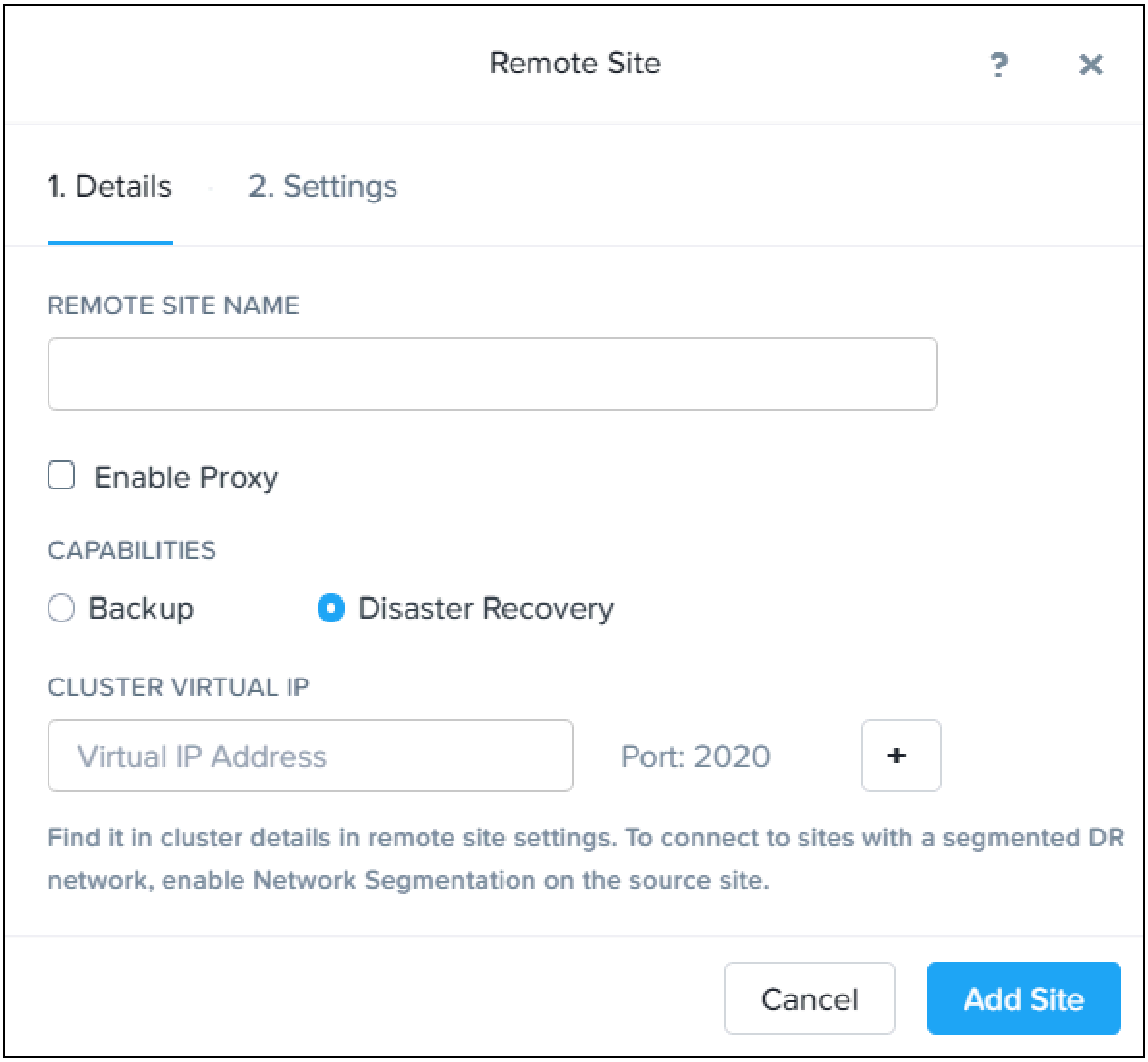
A remote site is the target location to store data replications for protected domains (see Configuring Data Protection with Asynchronous Replication in the *Data Protection and Recovery with Prism Element Guide*). The remote site can be either another physical cluster or a cluster located in a public cloud. To configure a remote physical cluster that can be used as a replication target, do the following:

> Note: Do not create multiple remote sites pointing to the single destination cluster. Otherwise, an alert will be generated.

Procedure

1. In the Data Protection dashboard (see Data Protection Dashboard in the Data Protection and Recovery with Prism Element Guide), click the **Remote Site** button and then select **Physical Cluster** from the pull-down list.

   The **Remote Site** dialog box appears.



Remote Site

1. Details 2. Settings

REMOTE SITE NAME

□ Enable Proxy

CAPABILITIES

○ Backup ● Disaster Recovery

CLUSTER VIRTUAL IP

Virtual IP Address Port: 2020

Find it in cluster details in remote site settings. To connect to sites with a segmented DR network, enable Network Segmentation on the source site.

Figure 8: Configuring Remote Site

2. Do the following in the indicated fields:

   a. Remote Site Name—Enter the remote site host name to create .

   > Note: This entity has the following naming restrictions:
   >
   > - The maximum length is 75 characters.
   > - Allowed characters are uppercase and lowercase standard Latin letters (A-Z and a-z), Simplified Chinese, decimal digits (0-9), dots (.), hyphens (-), and underscores (_).

   b. Enable Proxy—Check the box to enable a proxy.

   No proxy is used by default. Enabling this field allows the specified IP addresses—remote CVM IP addresses or remote Virtual IP address—to be used as a proxy to communicate with a Nutanix cluster on the remote site. In this case, the source cluster communicates with only one of the remote proxy IP addresses (remote CVM IP addresses or remote Virtual IP address), which forwards the requests to the appropriate remote CVMs. The proxy setting on the remote site limits the replication traffic to the defined destination remote site IP address (many to one) rather than to each individual CVM IP address in the destination cluster (many to many). The many-to-one replication approach can be used to simplify firewall rules configuration.

   > Note: It is recommended that you use a Virtual IP address for the proxy. Configuring a remote CVM for the proxy will make the CVM the single point of failure and break the entire communication if the CVM goes down.

   > Note: Network Address Translation (NAT) performed by any device in between the two Nutanix clusters is not currently supported.

   > Caution: Do not enable a proxy on remote sites that will be used with a metro availability Protection Domain.

   c. Backup—Select this option to enable backup (only) to this site.

   Backup allows the remote site to be used as a backup (replication) target. This means data can be backed up to this site and snapshots can be retrieved from the site to restore locally, but failover protection (that is, running failover VMs directly from the remote site) is not enabled.

   d. Disaster Recovery—Select this option to enable backup to and recovery from this site.

   Disaster recovery allows the remote site to be used both as a backup target and as a source for dynamic recovery. This means that failover VMs can be run directly from the remote site.

   > Note: During IPv6 configuration, the IPv6 addresses are stored in Zeus and the same is reflected in the IPS JSON output. When the disaster recovery fails, check for the **`Stored IPv6 configuration in Zeus`** in the output of the JSON script**.**

   e. Cluster Virtual IP

   - (Only on clusters without network segmentation for disaster recovery). Enter the Virtual IP (IPv4 or IPv6) address of the remote site cluster and port number , respectively, and then click the Add button. If you do not include a port number, a default port is used. The default port numbers are 2009 and 2020.

> Note: Ensure that the virtual IP address and the nodes in the remote cluster are in the same subnet.

- (Only on clusters with network segmentation for disaster recovery) Do the following:

    - In Segmented Subnet (Gateway IP/Prefix Length), enter the address of the subnet that is configured for isolating disaster recovery traffic.
    - In Port, enter the port number.
    - In Segmented Virtual IP, enter the virtual IP (IPv4 or IPv6) address of the disaster recovery service.

    > Note: Ensure that the virtual IP address and the nodes in the remote cluster are in the same subnet.

See Configuring a Remote Site (Physical Cluster) in the *Data Protection and Recovery with Prism Element Guide* for the remaining procedure of configuring a remote site

> Tip: Create a DR remote site:
>
> ```
> ncli> remote-site addresses=CVM_ipv6_ips name=Remote site name
> ```

LOGS

You can collect logs for Controller VMs, file server, hardware, alerts, and for the system. This topic provides more detailed information about common log files and AOS logs (such as Stargate, Cassandra, ZooKeeper, Genesis, Prism Gateway, ESXi, and Nutanix Calm logs).

Send Logs to Remote Syslog Server

The Nutanix command-line interface (nCLI) command **`rsyslog-config`** allows you to send logs from your Nutanix cluster to a remote syslog server. For more information about **`rsyslog-config`** command syntax, see the Command Reference guide.

Logs are forwarded from a Controller VM and they display the IP address of the Controller VM. The logs are forwarded from a Controller VM using either TCP or UDP. You can also forward logs to a remote syslog server by using Reliable Event Logging Protocol (RELP). To use RELP logging, ensure that you have installed **rsyslog-relp** on the remote syslog server.

> Note: You can use RELP logging only if the transport protocol is TCP.

After a remote syslog server is configured, it is enabled by default. (The Controller VM begins sending log messages once the syslog server is configured). **`rsyslog-config`** supports and can report messages from the following Nutanix modules.

> Note:
>
> For some modules, there is no change in the list of logs forwarded irrespective of the monitor logs setting.

Table 3: AOS Module Names for rsyslog-config

Logs are located in **`/home/nutanix/data/logs`** (except SYSLOG_MODULE).

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The `NUTANIX_LOG_LEVEL` log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
cassandra	system.log.`NUTANIX_LOG_LEVEL` dynamic_ring_changer.out dynamic_ring_changer.`NUTANIX_LOG_LEVEL`	cassandra_monitor.`NUTANIX_LOG_LEVEL` cassandra.out system.log.`NUTANIX_LOG_LEVEL` dynamic_ring_changer.out dynamic_ring_changer.`NUTANIX_LOG_LEVEL`
cerebro	cerebro.`NUTANIX_LOG_LEVEL` cerebro_cli.`NUTANIX_LOG_LEVEL`	cerebro.out cerebro.`NUTANIX_LOG_LEVEL` cerebro_cli.`NUTANIX_LOG_LEVEL`

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
curator	curator.***NUTANIX_LOG_LEVEL*** chronos_node_main.***NUTANIX_LOG_LEVEL*** chronos.out curator_cli.***NUTANIX_LOG_LEVEL***	curator.out curator.***NUTANIX_LOG_LEVEL*** ***NUTANIX_LOG_LEVEL*** chronos.out curator_cli.***NUTANIX_LOG_LEVEL***
genesis	genesis.out acropolis.out	genesis.out acropolis.out
	Forwards all task related to the *genesis* service to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The *`NUTANIX_LOG_LEVEL`* log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
prism	prism_gateway.log	prism_gateway.log prism_monitor.*`NUTANIX_LOG_LEVEL`* prism.out
stargate	stargate.*`NUTANIX_LOG_LEVEL`*	stargate.out stargate.*`NUTANIX_LOG_LEVEL`*

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***`NUTANIX_LOG_LEVEL`*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
zookeeper	zookeeper.out zeus_session_closer.***`NUTANIX_LOG_LEVEL`***	zookeeper_monitor.***`NUTANIX_LOG_LEVEL`*** zookeeper.out zeus_session_closer.***`NUTANIX_LOG_LEVEL`***
aplos	aplos.out aplos_engine.out	aplos.out aplos_engine.out
	Forwards all task related to the *aplos* service to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***`NUTANIX_LOG_LEVEL`*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
minerva_cvm	minerva_cvm.out minerva_cvm.log	minerva_cvm.out minerva_cvm.log
	Forwards all task related to the *minerva_cvm* service to a remote syslog server.	
uhura	uhura.out	uhura.out
	Forwards all task related to the *uhura* service to a remote syslog server.	
lazan	lazan.out solver.log	lazan.out solver.log

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
	Forwards all task related to the *lazan* service to a remote syslog server.	
anduril	anduril.out	anduril.out
	Forwards all task related to the *anduril* service to a remote syslog server.	
cluster_management	cluster_management_service.out	cluster_management_service.out
	Forwards all logs related to the *cluster_management* service to a remote syslog server.	
acropolis	acropolis.out	acropolis.out
	Forwards all logs related to the *acropolis* service to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
api_audit	api_audit.log api_audit_v3.log	api_audit.log api_audit_v3.log
	Forwards all logs related to the *api_audit* service to a remote syslog server.	
audit	consolidated_audit.log	consolidated_audit.log
	Forwards all logs related to the *audit* service to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
flow	hitCount1.log hitCount2.log hitCount3.log hitCount4.log	hitCount1.log hitCount2.log hitCount3.log hitCount4.log
	Forwards all logs related to the *flow* module to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
flow_service_logs	conntrack_stats_collector.log.***NUTANIX_LOG_LEVEL*** acropolis_ovs.log ovn-controller.log flow.out kafka.out cadmus.out microseg.out acropolis.out microsegmentation.out	conntrack_stats_collector.log.***NUTANIX_LOG_LEVEL*** acropolis_ovs.log ovn-controller.log flow.out kafka.out cadmus.out microseg.out acropolis.out microsegmentation.out

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***`NUTANIX_LOG_LEVEL`*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
	Forwards all logs related to the *flow_service_logs* module to a remote syslog server.	
flow_hitlogs	hitCount1.log hitCount2.log hitCount3.log hitCount4.log	hitCount1.log hitCount2.log hitCount3.log hitCount4.log
	Forwards all logs related to the *flow_hitlogs* module to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
lcm	genesis.out lcm_ops.out lcm_op.trace lcm_wget.log	genesis.out lcm_ops.out lcm_op.trace lcm_wget.log
	Forwards all logs related to the *lcm* service to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The `NUTANIX_LOG_LEVEL` log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
calm	jove.log hercules.log styx.log superevents.log iris.log helios.log algalon.log	jove.log hercules.log styx.log superevents.log iris.log helios.log algalon.log
	Forwards all logs related to the *calm* module to a remote syslog server.	

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
epsilon	jove.log narad.log achilles.log zaffi.log gozaffi_0.log indra_0.log indra_1.log karan_0.log karan_1.log arjun_0.log arjun_1.log gadarz_0.log	jove.log narad.log achilles.log zaffi.log gozaffi_0.log indra_0.log indra_1.log karan_0.log karan_1.log arjun_0.log arjun_1.log gadarz_0.log

Module name	Logs forwarded with monitor logs disabled	Logs forwarded with monitor logs enabled
SYSLOG_MODULE	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log	Controller VM: • /home/log/messages • /home/log/secure • /home/log/maillog • /home/log/cron • /home/log/spooler • /home/log/boot.log AHV host: • /var/log/messages • /var/log/secure • /var/log/maillog • /var/log/cron • /var/log/spooler • /var/log/boot.log Note: The additional monitor logs are also included.
	Note: • Forwards all the Controller VM logs that are stored in the following files to a remote syslog server. • Adding the SYSLOG_MODULE module to the rsyslog configuration configures rsyslog on compatible AHV hosts. A compatible host must be running an AHV release later than AHV-20160217.2. • The ***NUTANIX_LOG_LEVEL*** log file extension gets resolved at runtime depending upon the severity/level/priority that you provide to add modules for a rsyslog server. • These log files are supported on AHV, PE, and PC only .	
	Forwards all logs related to the *epsilon* service to a remote syslog server.	

Table 4: AOS Log Level Mapping to syslog Log Levels

AOS log levels	Contain information from these syslog log levels
INFO	DEBUG, INFO
WARNING	NOTICE, WARNING
ERROR	ERROR

AOS log levels	Contain information from these syslog log levels
FATAL	CRITICAL, ALERT, EMERGENCY

Ensure you enable module logs at the ERROR level, unless you require more information. If you enable more levels, the **`rsyslogd`** daemon sends more messages. For example, if you set the SYSLOG_MODULE level to INFO, your remote syslog server might receive a large quantity of operating system messages.

Limitations of Sending Logs to Remote Syslog Server

- You can only configure one rsyslog server; you cannot specify multiple servers.
- CPU usage might reach 10 percent when the **`rsyslogd`** daemon is initially enabled and starts processing existing logs. This is an expected condition on first use of an rsyslog implementation.

Configuring the Remote Syslog Server Settings

Before you begin

Install the Nutanix command-line interface (nCLI) and connect to a Controller VM in your cluster. For more information, see the Command Reference guide.

> Note: As the logs are forwarded from a Controller VM, the logs display the IP address of the Controller VM.

Procedure

1. As the remote syslog server is enabled by default, disable it while you configure settings.

   ```
   ncli> rsyslog-config set-status enable=false
   ```

2. Create a syslog server (which adds it to the cluster) and confirm it has been created.

   ```
   ncli> rsyslog-config add-server name=remote_server_name \
   relp-enabled={true | false} \
    ip-address=remote_ip_address port=port_num \
   network-protocol={tcp | udp}
   ncli> rsyslog-config ls-servers
   ```

   ```
    Name                          : remote_server_name
       IP Address                    : remote_ip_address
       Port                          : port_num
       Protocol                      : TCP or UDP
       Relp Enabled                  : true or false
   ```

   Replace:

   - *`remote_server_name`* with a descriptive name for the remote server receiving the specified messages.
   - *`remote_ip_address`* with the remote server's IP address.
   - *`port_num`* with destination port number on the remote server.
   - *`{true | false}`*, choose **`true`** to enable RELP and choose **`false`** to disable RELP.
   - *`{tcp | udp}`*, choose **`tcp`** or **`udp`** as the transport protocol.

3. Choose a module to forward log information from and specify the level of information to collect.

   ```
   ncli> rsyslog-config add-module server-name=remote_server_name
   ```

```
module-name=module level=loglevel include-monitor-logs={ false | true }
```

- Replace `module` with one of the following:
    - ACROPOLIS
    - AUDIT
    - CASSANDRA
    - CEREBRO
    - CURATOR
    - GENESIS
    - PRISM
    - STARGATE
    - SYSLOG_MODULE
    - ZOOKEEPER
    - UHURA
    - LAZAN
    - API_AUDIT
    - CALM
    - EPSILON
    - MINERVA_CVM
    - FLOW
    - FLOW_SERVICE_LOGS
    - LCM
    - APLOS
    - ANDURIL
    - CLUSTER_MANAGEMENT
    - FLOW_HITLOGS
- Replace `loglevel` with one of the following:
    - DEBUG
    - INFO
    - NOTICE
    - WARNING
    - ERROR
    - CRITICAL
    - ALERT
    - EMERGENCY

Enable module logs at the ERROR level unless you require more information.

- (Optional) Set `include-monitor-logs` to specify whether the monitor logs are sent. It is enabled (true) by default. If disabled (false), only certain logs are sent.

> Note: If enabled, the `include-monitor-logs` option sends all monitor logs, regardless of the level set by the `level=` parameter.

> Note: The rsyslog configuration is send to Prism Central, Prism Element, and AHV only if the module selected for export is applicable to them.

4. Configure additional modules if desired with `rsyslog-config add-module`.

5. Enable the server.

```
ncli> rsyslog-config set-status enable=true
```

Logs are now forwarded to the remote syslog server.

Common Log Files

Nutanix stores log files that contain cluster service events and errors in a common log root directory on the local filesystem of each CVM in a cluster. The logs contains details of all the relevant services required for cluster operation and monitoring.

The files in the common log area are further classified into different directories, depending on the type of information they contain.

> Note:
>
> - The timestamps for all Nutanix service logs are moved to UTC (in ISO 8601:2020-01-01 T00:00:00Z) from Prism version 5.18.
> - All operating system logs are not moved to UTC, hence Nutanix recommends that you set the server local time to UTC.

Nutanix Logs Root

The `/home/nutanix/data/logs` directory stores the Nutanix logs.

This location of the logs directory contains all the Nutanix process logs at the INFO, WARNING, ERROR and FATAL levels. It also contains the directories for the system stats (`sysstats`), and Cassandra system logs (`cassandra`).

The most recent FATAL log only contains the reason for the process to fail. More information can be found in the other types of logs by analyzing the entries leading up to the failure.

> Note: The symbolic link *`component_name`*`.[INFO|WARNING|ERROR|FATAL]` points to the most recent component log. For example:
>
> ```
> stargate.FATAL -> stargate.NTNX-12AM3K490006-2-
> CVM.nutanix.log.FATAL.201701-141913.30286
> ```

.FATAL Logs

If a component fails, it creates a log file named according to the following convention:

```
component-name.cvm-name.log.FATAL.date-timestamp
```

- `component-name` identifies to the component that created the file, such as Curator or Stargate.
- `cvm-name` identifies to the Controller VM that created the file.
- `date-timestamp` identifies the date and time when the first failure within that file occurred.

  Each failure creates a new .FATAL log file.

Log entries use the following format:

```
[IWEF]mmdd hh:mm:ss.uuuuuu threadid file:line] msg
```

The first character indicates whether the log entry is an Info, Warning, Error, or Fatal. The next four characters indicate the day on which the entry was made. For example, if an entry starts with F0820, it means that at some time on August 20th, the component had a failure.

> Tip: The cluster also creates .INFO and .WARNING log files for each component. Sometimes, the information you need is stored in one of these files.

Self-Monitoring (sysstats) Logs

The `/home/nutanix/data/logs/sysstats` directory stores the self-monitoring logs.

The node self-monitors itself by running several Linux tools every few minutes, including **`ping`**, **`iostat`**, **`sar`**, and **`df`**.

This directory contains the output for each of these commands, along with the corresponding timestamps.

/home/nutanix/data/logs/cassandra

The `/home/nutanix/data/logs/cassandra` directory stores the Cassandra metadata database logs. The Nutanix process that starts the Cassandra database (`cassandra_monitor`) logs to the `/home/nutanix/data/logs` directory. However, the most useful information relating to the Cassandra is found in the system.log* files located in the `/home/nutanix/data/logs/cassandra` directory.

This directory contains the output for each of these commands, along with the corresponding timestamps.

Consolidated Audit Logs

The `data/logs/consolidated_audit.log` directory stores the audit logs. The audit log allows you to view a list of actions performed across the clusters.

> Note: Audit logs with default values are generated when updates to VMs are initiated, either by Prism Central **Self Service** users or by using Nutanix v3 API calls for the first time.

Sample Audit Logs

Following are sample audit log outputs collected from Prism Web Console and Prism Central with admin and self service users .

Scenario 1: Creating a VM from Prism Web Console with Admin User

```
{"affectedEntityList":[{"entityType":"vm","name":"Test","uuid":"341a36a4-
bc1e-4f45-b829-41de340aa6f4"}],"alertUid":"VmCreateAudit","classificationList":
["UserAction"],"clientIp":"x.x.x.193","creationTimestampUsecs":"1625651019489357","defaultMsg":"Created
VM
Test","opEndTimestampUsecs":"1625651019485519","opStartTimestampUsecs":"1625651018983249","operationType":"
cb05-3882-ac1f6b161aaa","params":{"boot_device_order":"cdrom, disk,
network","disable_branding":"false","enable_cpu_passthrough":"false","enable_vga_console":"true","hardware_
a7b5-4b15456a757f","severity":"Audit","userName":"admin","uuid":"b2f4004f-5f22-438c-9eac-7bce0bdbd42b"}
```

Scenario 2: Powering on a VM from Prism Web Console with Self Service User (testuser)

```
{"affectedEntityList":[{"entityType":"vm","name":"Test","uuid":"341a36a4-bc1e-4f45-
b829-41de340aa6f4"},{"entityType":"host","name":"BayouBilly-3","uuid":"ed47be97-
ea80-4c79-bc32-749d5b9a5dac"}],"alertUid":"VmPowerOnAudit","classificationList":
["UserAction"],"clientIp":"x.x.x.193","creationTimestampUsecs":"1625651262934906","defaultMsg":"Powered
on VM
Test","opEndTimestampUsecs":"1625651262930860","opStartTimestampUsecs":"1625651258486149","operationType":"
cb05-3882-ac1f6b161aaa","params":
{"vm_name":"Test"},"recordType":"Audit","sessionId":"1adcde83-2dfe-437f-
af5b-3f532d7529c4","severity":"Audit","userName":"testuser","uuid":"b6634fdb-169c-453a-8e99-0da2623b6c50"}
```

Scenario 3: Creating a Category from Prism Central with Admin User

```
{"affectedEntityList":[{"entityType":"abac_category_key","name":"TEST","uuid":"673ba47b-
fd98-404c-af22-03db5af774ac"}],"alertUid":"CategoryKeyAudit","classificationList":
["UserAction"],"creationTimestampUsecs":"1625651133827490","defaultMsg":"Category key TEST has
been
created","opEndTimestampUsecs":"1625651133825892","opStartTimestampUsecs":"1625651133825892","operationType
b75e-688c5ce6cfe9","params":
{"category_key":"TEST"},"recordType":"Audit","severity":"Audit","userName":"admin","userUuid":"00000000-000
ad7b-fe63aa68fda7"}
```

Scenario 4: Tagging an Image to a Category with Admin User

```
{"affectedEntityList":[{"entityType":"image","name":"alpine-standard-3.14.0-
x86_64.iso","uuid":"390d8106-
f20b-4e5c-9bb6-2596b0f9fdf6"}],"alertUid":"CategoryAssignmentAudit","classificationList":
["UserAction"],"creationTimestampUsecs":"1625652191404317","defaultMsg":"Updated categories for
image alpine-standard-3.14.0-
x86_64.iso","opEndTimestampUsecs":"1625652191402891","opStartTimestampUsecs":"1625652191402891","operationTy
b75e-688c5ce6cfe9","params":
{"attached_categories":"TEST:Test","detached_categories":"","entity_name":"alpine-
standard-3.14.0-
x86_64.iso","entity_type":"image"},"recordType":"Audit","severity":"Audit","userName":"admin","userUuid":"0
a55d-dffa80ee2c03"}
```

Controller VM Log Files

The `/home/nutanix/data/logs` stores the Controller VMs logs.

Table 5: Controller VM Logs - Location: **`/home/nutanix/data/logs`**

Log	Contents	Frequency
alert_manager.[out, ERROR, FATAL, INFO, WARNING]	Alert manager process output.	
cassandra_monitor.[out, ERROR, FATAL, INFO]	Cassandra database monitor process output.	
catalina.out	Catalina/Tomcat for Prism process output.	

Log	Contents	Frequency
cerebro.[out, ERROR, FATAL]	DR and replication activity.	
check-cores.log	Core file processing.	every 1 min
check-fio	fio-status output.	every 1 hour
check-hardware.log	Power supply, fan speed, and DIMM temperature status.	every 1 min
check_intel.log	Intel PCIe-SSD status.	every 1 min
check-ip-connectivity.log	Network connectivity status to IPMI, hypervisor, and Controller VM of all nodes in the cluster.	every 1 min
chronos_node_main.[INFO, ERROR, FATAL, WARNING]	Write-ahead log (WAL) status	
connection_splicer.[out, ERROR, FATAL, INFO, WARNING]	Internal process connection status.	
cron_avahi_monitor.log	Avahi process status.	
cron_time_check.log	Check time difference across Controller VMs.	every 1 min
curator.[out, ERROR, FATAL, INFO, WARNING]	Metadata health and ILM activity.	
disk_usage.log	Disk and inode usage of all partitions on the Controller VM.	every 1 min
dynamic_ring_changer.[out, ERROR, FATAL]	Metadata migration across nodes activity.	
genesis.out	Nutanix software start process output.	
hyperint_monitor.[out, ERROR, FATAL, INFO, WARNING]	Hypervisor integration activity.	
pithos.[out, ERROR, FATAL, INFO, WARNING]	vDisk configuration activity.	
prism_gateway.[out, ERROR, FATAL, INFO]	Prism leader activity.	
prism_monitor.[out, ERROR, FATAL, INFO]	Prism (Web console, nCLI, REST API) monitor process output.	
scavenger.out	Log and core file clean-up status.	
send-email.log	E-mail alerts sent from the Controller VM.	every 1 min
snmp_manager.out	SNMP service logs.	

Log	Contents	Frequency
ssh_tunnel.log	Connect status to nsc.nutanix.com for the remote support tunnel.	
stargate.[out, ERROR, FATAL, INFO, WARNING]	NFS interface activity.	
stats_aggregator.[out, ERROR, FATAL, INFO]	Statistics aggregator process output.	
support-info.log	Daily automated support (ASUP) alerts.	
using-gflags.log	gflags status.	
zeus_config_printer.INFO	Contents of cluster configuration database.	
zookeeper_monitor.[out, ERROR, INFO]	Cluster configuration and cluster state activity.	

Table 6: Location: **`/home/nutanix/data/logs/cassandra`**

Log	Contents
system.log	Cassandra system activity.

Table 7: Location: **`/home/nutanix/data/logs/sysstats`**

Log	Contents	Frequency	Command
df.info	Mounted filesystems.	every 5 sec	df -h
disk_usage.INFO	Disk usage across disks.	every 1 hour	du
interrupts.INFO	CPU interrupts.	every 5 sec	
iostat.INFO	I/O activity for each physical disk.	every 5 sec	sudo iostat
iotop.INFO	Current I/O in real time.	every 5 sec	sudo iotop
lsof.INFO	List of top 10 processes or executables with most open file descriptors.	every 1 min	
meminfo.INFO	Memory usage.	every 5 sec	cat /proc/meminfo
metadata_disk_usage.INFO	Disk usage for metadata drives.	every 5 sec	
mpstat.INFO	CPU activities per CPU.	every 5 sec	mpstat
ntpq.INFO	NTP information.	every 1 min	ntpq -pn

Log	Contents	Frequency	Command
ping_gateway.INFO	Pings to the default gateway.	every 5 sec	ping
ping_hosts.INFO	Pings to all other Controller VMs.	every 1 min	ping
sar.INFO	Network bandwidth.	every 5 sec	sar -n DEV, -n EDEV
top.INFO	Real-time CPU and memory activity.	every 5 sec	top

Table 8: Location: **`/home/nutanix/data/serviceability/alerts`**

Log	Contents
`num`.processed	Alerts that have been processed.

Table 9: Location: **`/var/log`**

Log	Contents
dmesg	OS start messages.
kernel	OS kernel messages.
messages	OS messages after starting.

Correlating the FATAL log to the INFO file

This procedure provides useful techniques to discover relevant service events that lead up to a FATAL event and helps in troubleshooting and performing a root cause analysis.

About this task

When a process fails, the reason for the failure is recorded in the corresponding FATAL log. There are two ways to correlate this log with the INFO file to get more information:

Procedure

1. Search for the timestamp of the FATAL event in the corresponding INFO files.

    a. Determine the timestamp of the FATAL event.

    For example, the latest stargate.FATAL determines the exact timestamp.

    ```
    nutanix@cvm$ cat stargate.FATAL
    ```

    ```
    Log file created at: 2013/09/07 01:22:23
    Running on machine: NTNX-12AM3K490006-2-CVM
    Log line format: [IWEF]mmdd hh:mm:ss.uuuuuu threadid file:line] msg
    F0907 01:22:23.124495 10559 zeus.cc:1779] Timed out waiting for Zookeeper session
     establishment
    ```

b. Search for the timestamp in the corresponding INFO files.

   In this example, the timestamp is `F0907 01:22:23`, or September 7 at 1:22:23 AM.

c. Open the INFO file with `vi` and go to the bottom of the file (`Shift+G`).

```
nutanix@cvm$ grep "^F0907 01:22:23" stargate*INFO* |
cut -f1 -d:stargate.NTNX-12AM3K490006-2-CVM.nutanix.log.INFO.20130904-220129.7363
```

   This tells us that the relevant file to look at is `stargate.NTNX-12AM3K490006-2-CVM.nutanix.log.INFO.20130904-220129.7363`.

d. Analyze the log entries immediately before the FATAL event, especially any errors or warnings.

2. If a process fails repeatedly, it might be faster to do a long listing of the INFO files and select the one immediately preceding the current one. The current one would be the one referenced by the symbolic link.

   For example, in the output below, the last failure would be recorded in the file `stargate.NTNX-12AM3K490006-2-CVM.nutanix.log.INFO.20130904-220129.7363`.

```
ls -ltr stargate*INFO*
```

```
-rw-------. 1 nutanix nutanix 104857622 Sep  3 11:22 stargate.NTNX-12AM3K490006-2-
CVM.nutanix.log.INFO.20130902-004519.7363
-rw-------. 1 nutanix nutanix 104857624 Sep  4 22:01 stargate.NTNX-12AM3K490006-2-
CVM.nutanix.log.INFO.20130903-112250.7363
-rw-------. 1 nutanix nutanix  56791366 Sep  5 15:12 stargate.NTNX-12AM3K490006-2-
CVM.nutanix.log.INFO.20130904-220129.7363
lrwxrwxrwx. 1 nutanix nutanix        71 Sep  7 01:22 stargate.INFO ->
 stargate.NTNX-12AM3K490006-2-CVM.nutanix.log.INFO.20130907-012223.11357
-rw-------. 1 nutanix nutanix     68761 Sep  7 01:33 stargate.NTNX-12AM3K490006-2-
CVM.nutanix.log.INFO.20130907-012223.11357
```

> Tip: You can use the procedure above for the other types of files as well (WARNING and ERROR) in order to narrow the window of information. The INFO file provides all messages, WARNING provides only warning, error, and fatal-level messages, ERROR provides only error and fatal-level messages, and so on.

Stargate Logs

This section discusses common entries found in Stargate logs and what they mean. The Stargate logs are located at `/home/nutanix/data/logs/stargate.[INFO|WARNING|ERROR|FATAL]`.

Log Entry: `Watch dog fired`

```
F1001 16:20:49.306397 6630 stargate.cc:507] Watch dog fired
```

This message is generic and can happen for a variety of reasons. While Stargate is initializing, a watch dog process monitors it to ensure a successful startup process. If it has trouble connecting to other components (such as Zeus or Pithos) the watch dog process stops Stargate.

If Stargate is running, this indicates that the alarm handler thread is stuck for longer than 30 seconds. The stoppage could be due to a variety of reasons, such as problems connecting to Zeus or accessing the Cassandra database.

To analyze why the watch dog fired, first locate the relevant INFO file, and review the entries leading up to the failure.

Log Entry: `HTTP request timed out`

```
E0820 09:14:05.998002 15406 rpc_client.cc:559] Http request timed out
```

This message indicates that Stargate is unable to communicate with Medusa. This may be due to a network issue.

Analyze the ping logs and the Cassandra logs.

Log Entry: `CAS failure seen while updating metadata for egroup` *`egroupid`* **or** `Backend returns error 'CAS Error' for extent group id:` *`egroupid`*

```
W1001 16:22:34.496806 6938 vdisk_micro_egroup_fixer_op.cc:352]
CAS failure seen while updating metadata for egroup 1917333
```

This is a benign message and usually does not indicate a problem. This warning message means that another Cassandra node has already updated the database for the same key.

Log Entry: `Fail-fast after detecting hung stargate ops: Operation with id` *`opid`* `hung for 60secs`

```
F0712 14:19:13.088392 30295 stargate.cc:912] Fail-fast after detecting hung stargate ops:
 Operation with
id 3859757 hung for 60secs
```

This message indicates that Stargate restarted because an I/O operation took more than 60 seconds to complete.

To analyze why the I/O operation took more than 60 seconds, locate the relevant INFO file and review the entries leading up to the failure.

Log Entry: `Timed out waiting for Zookeeper session establishment`

```
F0907 01:22:23.124495 10559 zeus.cc:1779] Timed out waiting for Zookeeper session establishment
```

This message indicates that Stargate was unable to connect to Zookeeper.

Review the **`sysstats/ping_hosts.INFO`** log to determine if there were any network issues around that time.

Log Entry: `Too many attempts trying to access Medusa`

```
F0601 10:14:47.101438 2888 medusa_write_op.cc:85] Check failed: num_retries_ < 5 (5 vs. 5) Too
 many attempts
trying to access Medusa
```

This message indicates that Stargate had 5 failed attempts to connect to Medusa/Cassandra.

Review the Cassandra log (**`cassandra/system.log`**) to see why Cassandra was unavailable.

Log Entry:`multiget_slice() failed with error:` *`error_code`* `while reading n rows from` *`cassandra_keyspace`*

```
E1002 18:51:43.223825 24634 basic_medusa_op.cc:1461] multiget_slice() failed with error: 4
 while reading 1 rows
from 'medusa_nfsmap'. Retrying...
```

This message indicates that Stargate cannot connect to Medusa/Cassandra.

Review the Cassandra log (**`cassandra/system.log`**) to see why Cassandra was unavailable.

Log Entry: `Forwarding of request to NFS master` *`ip`*`:2009 failed with error kTimeout.`

```
W1002 18:50:59.248074 26086 base_op.cc:752] Forwarding of request to NFS master
 172.17.141.32:2009 failed with
```

```
error kTimeout
```

This message indicates that Stargate cannot connect to the NFS leader on the node specified.

Review the Stargate logs on the node specified in the error.

Cassandra Logs

After analyzing `Stargate` logs, if you suspect an issue with Cassandra/Medusa, analyze the `Cassandra` logs. This topic discusses common entries found in `system.log` and what they mean.

Cassandra Database Logs

The `/home/nutanix/data/logs/cassandra` directory stores the Cassandra metadata database logs. The Nutanix process that starts the Cassandra database (`cassandra_monitor`) logs to the `/home/nutanix/data/logs` directory. However, the most useful information relating to the Cassandra is found in the `system.log*` files located in the `/home/nutanix/data/logs/cassandra` directory. When the file reaches a certain size, it rolls over to a sequentially numbered file (example, `system.log.1`, `system.log.2`, and so on).

This directory contains the output for each of these commands, along with the corresponding timestamps.

Log Entry: `batch_mutate 0 writes succeeded and 1 column writes failed for keyspace:medusa_extentgroupidmap`

```
INFO [RequestResponseStage:3] 2013-09-10 11:51:15,780 CassandraServer.java (line 1290)
 batch_mutate 0
writes succeeded and 1 column writes failed for keyspace:medusa_extentgroupidmap
 cf:extentgroupidmap
row:lr280000:1917645 Failure Details: Failure reason:AcceptSucceededForAReplicaReturnedValue :
 1
```

This is a common log entry and can be ignored. It is equivalent to the CAS errors in the `stargate.ERROR` log. It simply means that another Cassandra node updated the keyspace first.

Log Entry: `InetAddress /x.x.x.x is now dead.`

```
INFO [ScheduledTasks:1] 2013-06-01 10:14:29,767 Gossiper.java (line 258) InetAddress /x.x.x.x
 is now dead.
```

This message indicates that the node could not communicate with the Cassandra instance at the specified IP address.

Either the Cassandra process is down (or failing) on that node or there are network connectivity issues. Check the node for connectivity issues and Cassandra process restarts.

Log Entry: `Caught Timeout exception while waiting for paxos read response from leader: x.x.x.x`

```
ERROR [EXPIRING-MAP-TIMER-1] 2013-08-08 07:33:25,407 PaxosReadDoneHandler.java (line 64) Caught
 Timeout
exception while waiting for paxos read reponse from leader: 172.16.73.85. Request Id: 116.
Proto Rpc Id : 2119656292896210944. Row no:1. Request start time: Thu Aug 08 07:33:18 PDT
 2013.
Message sent to leader at: Thu Aug 08 07:33:18 PDT 2013 # commands:1 requestsSent: 1
```

This message indicates that the node encountered a timeout while waiting for the Paxos leader.

Either the Cassandra process is down (or failing) on that node or there are network connectivity issues. Check the node for connectivity issues or for the Cassandra process restarts.

Prism Gateway Log

This section discusses common entries found in `prism_gateway.log` and what they mean. The Prism log is located at `/home/nutanix/data/logs/prism_gateway.log` on the Prism leader. The Prism leader is the node which is running the web server for the Nutanix UI. This is the log to analyze if there are problems with the UI such as long loading times.

To identify the Prism leader, you can run `cluster status | egrep "CVM|Prism"` and determine which node has the most processes. In the output below, x.x.x.242 is the Prism leader.

```
nutanix@cvm$ cluster status | egrep "CVM|Prism"
```

```
2013-09-10 16:06:42 INFO cluster:946 Executing action status on CVMs
 x.x.x.240,x.x.x.241,x.x.x.242 2013-09-10
16:06:45 INFO cluster:987 Success!
       CVM: x.x.x.240 Up
                          Prism    UP[32655, 32682, 32683, 32687]
       CVM: x.x.x.241 Up
                          Prism    UP[11371, 25913, 25925, 25926]
       CVM: x.x.x.242 Up, ZeusLeader
                          Prism    UP[4291, 4303, 4304, 19468, 20072, 20074, 20075, 20078,
 20113]
```

Log Entry: `Error sending request: java.net.NoRouteToHostException: Cannot assign requested address`

The `stats_aggregator` component periodically issues an RPC request for all Nutanix vdisks in the cluster. It is possible that all the ephemeral ports are exhausted.

The `ss -s` command shows you the number of open ports.

```
nutanix@cvm$ ss -s
```

```
Total: 277 (kernel 360)
TCP:   218 (estab 89, closed 82, orphaned 0, synrecv 0, timewait 78/0), ports 207

Transport Total     IP        IPv6
*         360       -         -
RAW       1         1         0
UDP       23        13        10
TCP       136       84        52
INET      160       98        62
FRAG      0         0         0
```

If there are issues with connecting to the Nutanix UI, escalate the case and provide the output of the `ss -s` command as well as the contents of `prism_gateway.log`.

Zookeeper Logs

The Zookeeper logs are located at `/home/nutanix/data/logs/zookeeper.out` and contains the status of the Zookeeper service.

More often than not, there is no need to look at this log. However, if one of the other logs specifies that it is unable to contact Zookeeper and it is affecting cluster operations, you may want to look at this log to find the error Zookeeper is reporting.

Genesis.out

When checking the status of the cluster services, if any of the services are down, or the Controller VM is reporting `Down` with no process listing, review the log at `/home/nutanix/data/logs/genesis.out` to determine why the service did not start, or why Genesis is not properly running.

Check the contents of **`genesis.out`** if a Controller VM reports multiple services as DOWN, or if the entire Controller VM status is DOWN.

Like other component logs, **`genesis.out`** is a symbolic link to the latest **`genesis.out`** instance and has the format **`genesis.out.`***`date-timestamp`*.

An example of steady state output:

```
nutanix@cvm$ tail -F ~/data/logs/genesis.out
```

```
2017-03-23 19:24:00 INFO node_manager.py:2070 Certificate cache in sync

2017-03-23 19:24:00 INFO node_manager.py:4732 Checking if we need to sync the local SVM and
 Hypervisor DNS configuration
with Zookeeper

2017-03-23 19:24:38 ERROR lcm_zeus.py:96 Failed to read zknode /appliance/logical/lcm/operation
 with error: no node

2017-03-23 19:24:39 INFO framework.py:637 No other LCM operation in progress

2017-03-23 19:26:00 INFO node_manager.py:1960 Certificate signing request data is not available
 in Zeus configuration

2017-03-23 19:26:00 INFO node_manager.py:1874 No CA certificates found in the Zeus
 configuration

2017-03-23 19:26:00 INFO node_manager.py:1877 No Svm certificates found in the Zeus
 configuration

2017-03-23 19:26:00 INFO node_manager.py:1880 No Svm certificate maps found in the Zeus
 configuration

2017-03-23 19:26:00 INFO node_manager.py:2070 Certificate cache in sync

2017-03-23 19:26:00 INFO node_manager.py:4732 Checking if we need to sync the local SVM and
 Hypervisor DNS configuration
with Zookeeper

2017-03-23 19:28:00 INFO node_manager.py:1960 Certificate signing request data is not available
 in Zeus configuration

2017-03-23 19:28:00 INFO node_manager.py:1874 No CA certificates found in the Zeus
 configuration

2017-03-23 19:28:00 INFO node_manager.py:1877 No Svm certificates found in the Zeus
 configuration

2017-03-23 19:28:00 INFO node_manager.py:1880 No Svm certificate maps found in the Zeus
 configuration

2017-03-23 19:28:00 INFO node_manager.py:2070 Certificate cache in sync
```

Under normal conditions, the **`genesis.out`** file logs the following messages periodically:

```
Unpublishing service Nutanix Controller
Publishing service Nutanix Controller
Zookeeper is running as [leader|follower]
```

Prior to these occasional messages, you should see **`Starting [n]th service`**. This is an indicator that all services were successfully started.

Tip: You can ignore any INFO messages logged by Genesis by running the command:

```
grep -v -w INFO /home/nutanix/data/logs/genesis.out
```

Possible Errors

```
2017-03-23 19:28:00 WARNING command.py:264 Timeout executing scp -q -o CheckHostIp=no -o
 ConnectTimeout=15 -o
StrictHostKeyChecking=no -o TCPKeepAlive=yes -o UserKnownHostsFile=/dev/null -o
PreferredAuthentications=keyboard-interactive,password -o BindAddress=x.x.x.254
'root@[x.x.x.1]:/etc/resolv.conf' /tmp/resolv.conf.esx: 30 secs elapsed
```

```
2017-03-23 19:28:00 ERROR node_dns_ntp_config.py:287 Unable to download ESX DNS configuration
 file, ret -1,
stdout , stderr
```

```
2017-03-23 19:28:00 WARNING node_manager.py:2038 Could not load the local ESX configuration
```

```
2017-03-23 19:28:00 ERROR node_dns_ntp_config.py:492 Unable to download the ESX NTP
 configuration file,
ret -1, stdout , stderr
```

Any of the above messages means that Genesis was unable to log on to the ESXi host using the configured password.

Diagnosing a Genesis Failure

About this task

Determine the cause of a Genesis failure based on the information available in the log files.

Procedure

1. Examine the contents of the **`genesis.out`** file and locate the stack trace (indicated by the CRITICAL message type).
2. Analyze the ERROR messages immediately preceding the stack trace.

```
  ...
2017-03-23 19:30:00 INFO node_manager.py:4170 No cached Zeus configuration found.
2017-03-23 19:30:00 INFO hyperv.py:142 Using RemoteShell ...
2017-03-23 19:30:00 INFO hyperv.py:282 Updating NutanixUtils path
2017-03-23 19:30:00 ERROR hyperv.py:290 Failed to update the NutanixUtils path: [Errno 104]
 Connection reset by peer
2017-03-23 19:30:00 CRITICAL node_manager.py:3559 File "/home/nutanix/cluster/bin/genesis",
 line 207, in <module>
    main(args)
  File "/home/nutanix/cluster/bin/genesis", line 149, in main
    Genesis().run()
  File "/home/nutanix/jita/main/28102/builds/build-danube-4.1.3-stable-release/python-tree/
bdist.linux-x86_64/egg/util/misc/decorators.py", line 40, in wrapper
  File "/home/nutanix/jita/main/28102/builds/build-danube-4.1.3-stable-release/python-tree/
bdist.linux-x86_64/egg/cluster/genesis/server.py", line 132, in run
  File "/home/nutanix/jita/main/28102/builds/build-danube-4.1.3-stable-release/python-tree/
bdist.linux-x86_64/egg/cluster/genesis/node_manager.py", line 502, in initialize
  File "/home/nutanix/jita/main/28102/builds/build-danube-4.1.3-stable-release/python-tree/
bdist.linux-x86_64/egg/cluster/genesis/node_manager.py", line 3559, in discover
```

```
...
```

In the example above, the certificates in **`AuthorizedCerts.txt`** were not updated, which means that you failed to connect to the NutanixHostAgent service on the host.

> Note: NutanixHostAgent is Hyper-V specific.

ESXi Log Files

These log files are present on ESXi hosts.

Table 10: Location: **`/var/logs`**

Log	Contents
hostd.log	**`hostd`** (daemon to communicate with vmkernel) process output
vmkernel.log	vmkernel activity
vpxa.log	**`vpxa`** (daemon to communicate with vCenter) process output

Table 11: Location: **`/vmfs/volumes/`**

Log	Contents
`datastore`/*`vm_name`*/vmware.log	Virtual machine activity and health

Nutanix Calm Log Files

The following table provides Nutanix Calm logs related information.

Table 12: Nutanix Calm Log Files

Log	Description
`/home/docker/nucalm/logs`	Logs of microservices from Nutanix Calm container.
`/home/docker/epsilon/logs`	Logs of microservices from Epsilon Container.
`/home/nutanix/data/logs/genesis.out`	Logs containing information about enabling container service and starting Nutanix Calm and epsilon containers.
`/home/nutanix/data/logs/epsilon.out`	Logs containing information about starting epsilon service. Note: This log contains information about epsilon container crashes.
`/home/nutanix/data/logs/nucalm.out`	Logs containing information about starting Nutanix Calm service.

Log	Description
`/home/docker/docker-latest/plugins/*/rootfs/nvp.log`	Logs containing the docker volume plug-in, used to create or mount the Nutanix Calm volume group when epsilon container starts. If volumes are not listed by **`docker volume ls`** you can check here to know why the volume group is not mounted.
`/home/log/messages`	Logs containing information regarding the communication with Prism web console cluster for mounting volume group and network communication issues with the Prism web console cluster.

TRAFFIC MARKING FOR QUALITY OF SERVICE

To prioritize outgoing (or egress) traffic as required, you can configure quality of service on the traffic for a cluster.

There are two distinct types of outgoing or egress traffic:

- Management traffic (mgmt)
- Data services (data-svc)

Data services traffic consists of the following protocols:

Table 13: Data Services Protocols

Protocol	Port	Nutanix Services
NFS	Source port (src_port)	-Nutanix Files-
SMB	Source port (src_port)	-Nutanix Files-
Cluster-to-cluster replications (external or inter-site)	Source and destination ports	Stargate and Cerebro
Node-to-node replications (internal or intra-site)	Source and destination ports	Stargate and Cerebro
iSCSI	Data services IP address targets	Nutanix Files and Volumes

Traffic other than data services traffic is management traffic.

When you enable QoS, you can mark both - types of traffic with QoS values. AOS considers the values in hexadecimal even if you provide the values in decimal. When you view or get the QoS configuration enabled on the cluster, nCLI provides the QoS values in hexadecimal format (0xXX where XX is hexadecimal value in the range 00–3f).

> Note: Set any QoS value in the range 0–0x3f. The default QoS values for the traffic are as follows:
>
> - Management traffic (mgmt) = 0x10
> - Data services (data-svc) = 0xa

Configuring QoS

Configure Quality of Service (QoS) for management and data services traffic using nCLI.

About this task

To perform the following operations for QoS on the egress traffic of a cluster, use the nCLI commands in this section:

- Enable QoS on the cluster.
- View or get the QoS configuration enabled on the cluster.

- Set QoS values for all traffic types or specific traffic types.
- Disable QoS on the cluster.

When you run any of the QoS configuration commands and the command succeeds, the console displays the following output indicating the successful command run:

`QoSUpdateStatusDTO(status=true, message=null)`

Where:

- `status=true` indicates that the command succeeded.
- `message=null` indicates that there is no error.

When you run any of the QoS configuration commands and the command fails, the console displays the following sample output indicating the failure:

`QoSUpdateStatusDTO(status=false, message=QoS is already enabled.)`

Where:

- `status=false` indicates that the command failed.
- `message=QoS is already enabled.` indicates why the command failed. This sample error message indicates that the `net enable-qos` command failed because QoS enable command was run again when QoS is already enabled.

Procedure

- Activate QoS on a cluster by running the following commands on all the CVMs in the cluster:

  ```
  nutanix@cvm$ echo --qos_enabled=true >> ~/config/genesis.gflags
  nutanix@cvm$ allssh genesis restart
  ```

- To enable QoS on a cluster, run the following command:

  ```
  ncli> net enable-qos [data-svc="data-svc value"][mgmt="mgmt value"]
  ```

  If you run the command as `net enable-qos` without the options, AOS enables QoS with the default values (`mgmt=0x10` and `data-svc=0xa`).

  > Note: After you run the `net enable-qos` command, if you run it again, the command fails and AOS displays the following output:
  >
  > `QoSUpdateStatusDTO(status=false, message=QoS is already enabled.)`

  > Note: If you need to change the QoS values after you enable it , run the `net edit-qos` command with the option (`data-svc` or `mgmt` or both as necessary).

  > Note: Set any QoS value in the range 0x0-0x3f.

- To view or get the QoS configuration enabled on a cluster, run the following command:

  ```
  ncli> net get-qos
  ```

  > Note: When you get the QoS configuration enabled on the cluster, nCLI provides the QoS values in hexadecimal format (0xXX where XX is hexadecimal value in the range 00-3f).

  A sample output on the console is as follows:

  ```
  QoSDTO(status=true, isEnabled=true, mgmt=0x10, dataSvc=0xa, message=null)
  ```

  > Note:
  >
  > Where:
  >
  > - **`status=true`** indicates that the **`net get-qos`** command passed. **`status=false`** indicates that the **`net get-qos`** command failed. See the **`message=`** value for the failure error message.
  > - **`isEnabled=true`** indicates that QoS is enabled. **`isEnabled=false`** indicates that QoS is not enabled.
  > - **`mgmt=0x10`** indicates that QoS value for Management traffic (***`mgmt`*** option) is set to 0x10 (represented in hexadecimal value as **`0x10`**. If you disabled QoS, then this parameter is displayed as **`mgmt=null`**.
  > - **`dataSvc=0xa`** indicates that QoS value for data services traffic (***`data-svc`*** option) is set to 0xa (represented in hexadecimal value as **`0xa`**. If you disabled QoS, then this parameter is displayed as **`dataSvc=null`**.
  > - **`message=null`** indicates there is no error message. **`message=`** parameter provides the command failure error message if the command fails.

- To set the QoS values for the traffic types on a cluster after you enabled QoS on the cluster, run the following command:

  ```
  ncli> net edit-qos [data-svc="data-svc value"][mgmt="mgmt value"]
  ```

  You can provide QoS values between 0x0-0x3f for one or both the options. The value is hexadecimal representation of a value between decimal 0-63 both inclusive.

- To disable QoS on a cluster, run the following command:

  ```
  ncli> net disable-qos
  ```

  ```
  QoSDTO(status=true, isEnabled=false, mgmt=null, dataSvc=null, message=null)
  ```

BLOCKSTORE SUPPORT WITH SPDK

At the heart of the Nutanix data fabric, Nutanix Blockstore is a block management system implemented in AOS. It is available on all hypervisors. Blockstore moves device interactions into user space eliminating any context switching or kernel driver invocation. It creates an extensible file system and block management layer, fully managed in user space. This removes the need to invoke any file system kernel driver for non-boot disks.

Newer storage media such as NVMe have user space libraries (such as SPDK) to manage device I/O directly. eliminating the need to make any system calls (context switches). Blockstore enables AOS to leverage Intel Storage Performance Development Kit (SPDK) for direct access of NVMe backed disks.

With Blockstore, you replace the Linux-kernel file system with a leaner one running inside the user space. This approach enables direct communication with the underlying storage media using SPDK APIs and avoiding in-kernel system-calls.

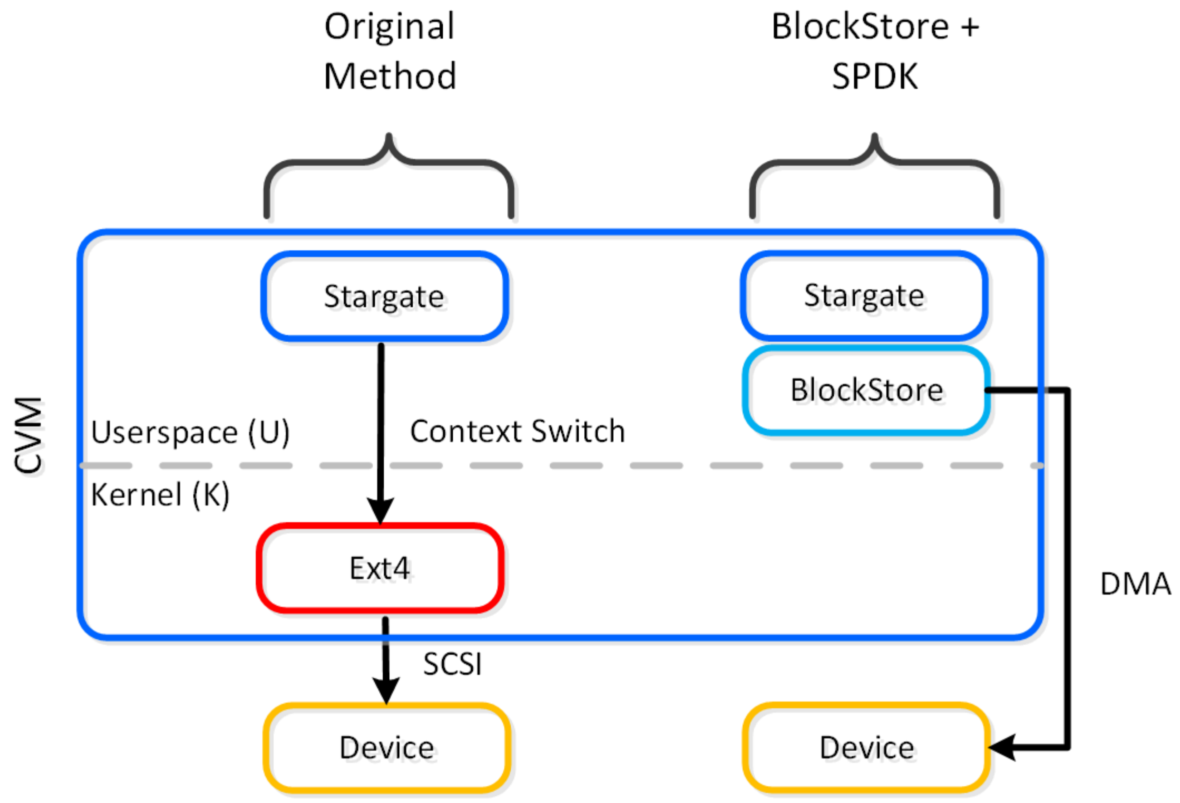


Figure 9: Blockstore + SPDK Approach

The combination of Blockstore and SPDK provides the following benefits:

- Ultra low latency
- Higher performance
- Greater host CPU efficiency.

The following image shows a high-level overview of the updated I/O path with Blockstore + SPDK:

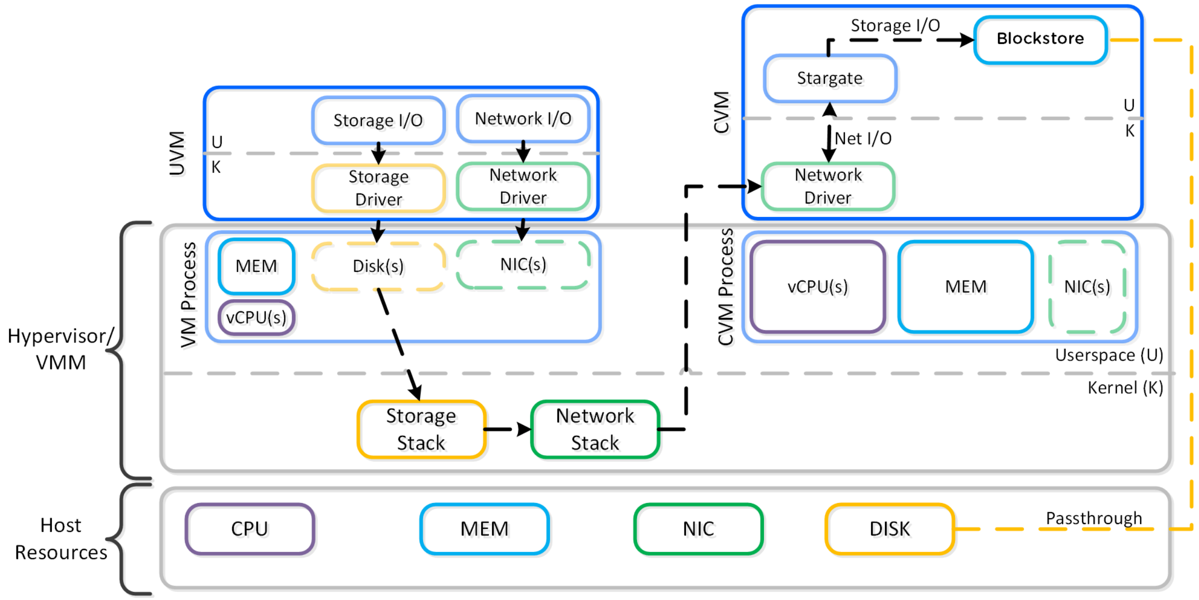


Figure 10: High-level Overview

Considerations for Blockstore with SPDK

When you install a cluster using a minimum AOS version of 6.1 or later, Blockstore is available on storage devices such as SAS/SATA SSD, and NVMe SSD. SPDK is available on NVMe and Optane SSDs subject to conditions.

SPDK is automatically enabled on a cluster when Blockstore is active on the cluster with NVMe (and) or Optane SSDs and other Blockstore requirements.

Requirements

Requirements for Blockstore and SPDK include the following:

- Ensure that the cluster is created on a compatible version of AOS. For information about the AOS versions compatible with Blockstore, see the AOS Family Release Notes.
- Ensure that the all-flash (AF) cluster, created using a minimum AOS version of 6.1 or later, has at least one non-boot flash device for Blockstore (without SPDK) to be enabled on that device.

> Note: Blockstore is enabled only on non-boot flash devices, in other words, the flash device that is not used as part of RAID partition to boot the CVM.
>
> If flash devices need to be configured as part of CVM boot (RAID) partition, then ensure that the all-flash cluster has more than two flash devices on at least one node. For example, in an all-flash cluster in which each node has only two flash devices, the blockstore is not enabled in that cluster. When you add one more flash device on any node in that cluster as a non-boot device, Blockstore is enabled on the newly added non-boot flash device.

Nutanix supports additional storage device configurations in each AOS release. Contact Nutanix Support for information about the storage device configurations for the AOS release you deploy.

- Ensure that the Controller VM memory is set to a minimum of 48 GB memory. This requirement is only applicable to Blockstore with SPDK. If you are not deploying SPDK, ignore this requirement.

  For more information about CVM memory settings, see the *CVM Specifications - Cluster Feature Based* table in the Controller VM (CVM) Field Specifications on page 90 topic.

Verifying that Blockstore with SPDK is Enabled

About this task

To verify that SPDK is enabled, do any of the following.

Procedure

1. Check if Blockstore is enabled.

   ```
   nutanix@cvm$ df | grep fuse
   ```

   If SPDK is not enabled, check is the Controller VM memory set to 48 GB or more.

   Ensure that the requirements for Blockstore with SPDK are fulfilled. Blockstore is not enabled if any of requirements are not met in the cluster. See Considerations for Blockstore with SPDK on page 83.

2. Check the `/dev/spdk/` table to see if any NVMe devices are displayed therein using one of the following methods:

   - On each node separately:

     ```
     nutanix@cvm$ ls /dev/spdk/*
     ```

     Or

   - On all nodes collectively:

     ```
     nutanix@cvm$ allssh ls /dev/spdk/*
     ```

   If you see NVMe devices in the `/dev/spdk/` table, then Blockstore with SPDK is enabled.

3. Alternatively, check the `iostat` command output to see if any NVMe devices are displayed therein.

   ```
   nutanix@cvm$ iostat
   ```

   If you see SPDK NVMe devices in device list, then Blockstore with SPDK is enabled on the NVMe devices.

What to do next

If all the aforesaid checks are met and SPDK is still not enabled, contact Nutanix Support.

INTEL OPTANE PERFORMANCE TIER FOR ALL-NVME CLUSTERS

Intel Optane NVMe drives provide high performance with low latency for I/O operations for frequently read data segments or read intensive workloads.

I/O Tiers

Before describing the Intel Optane performance tier, this section describes storage tiers and the structure of read-write operations.

Distributed Storage Fabric (DSF) I/O Path

Within the CVM, the Stargate process is responsible for handling storage I/O for user VMs (UVMs) and persistence (RF, etc.). The Autonomous Extent Store (AES) is used to handle sequential and sustained random workloads (subject to certain conditions). Oplog that extends to both tiers in a hybrid drive configuration handles purely random read-write data and drains into the Extent Store.

Information Lifecycle Management (ILM) Tiers

Nutanix Information Lifecycle Management (ILM) uses storage tiers to place data based on how different types of storage devices like SATA SSD or NVMe SSDs perform. ILM monitors how data is being accessed and places the most accessed data in high performance, low latency drives to achieve the best possible access speeds. ILM distinguishes read intensive data from write intensive data.

Nutanix ILM defines storage tiers based on categorization of storage drives. Nutanix categorizes drives into three types: DAS-SATA, SSD-SATA, and SSD-PCIe. DAS-SATA includes HDDs. SSD-SATA includes SATA SSDs. The SSD-PCIe category includes NVMe drives like Intel P4510. ILM uses a preference list for data placement using Extent groups. This list, in the descending order of preference, is: SSD-PCIe, SSD-SATA and DAS-SATA.

ILM deploys a multi tier architecture in hybrid drive configurations (such as SATA SSD + HDD) to enable faster processing of high frequency data. In a hybrid drive configuration, tier 0 is always the fast tier with low latency drives such as SSDs. Tier 1 is the slower tier with higher latency devices such as HDDs.

In general, in all-flash drive configurations including all-NVMe configurations, ILM uses single tier architecture.

In a hybrid drive configuration, the Nutanix ILM based I/O path consists of the following components and pathways where Tier 1 is HDD and Tier 0 is low latency SATA SSD. Bursty Rand refers to purely random writes and Sustained Rand refers to sustained random writes.

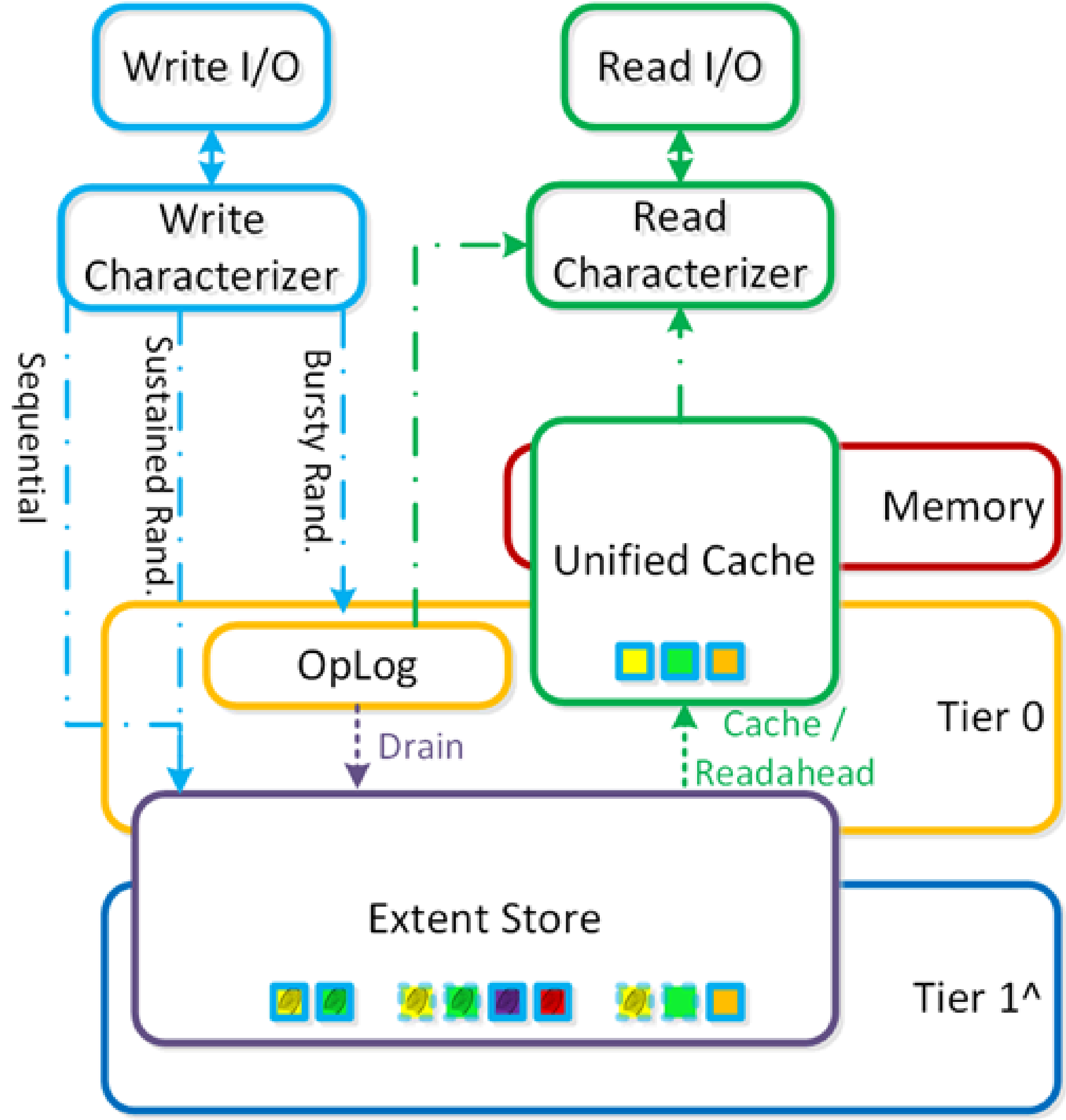


Figure 11: I/O in Hybrid Drive Configuration

Typically, in an all-flash drive configuration, the I/O path using single tier architecture is described by the following figure.

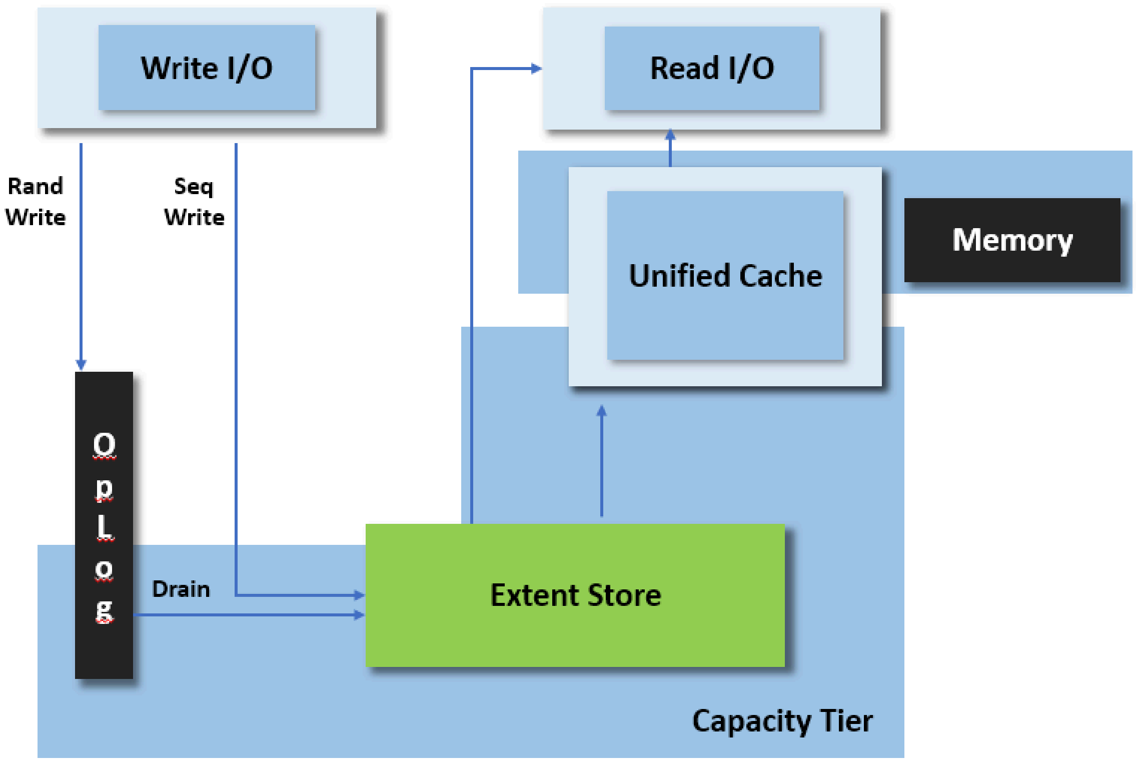


Figure 12: I/O in All-flash Drive Configuration

Intel Optane And NVMe Drive Configurations

In an all-NVMe configuration having Intel Optane drives, in fresh installations of supporting AOS version, ILM enables a multi tier architecture, similar to that in hybrid drive configurations, with an additional Optane tier.

> Tip: A typical all-NVMe configuration with Intel Optane would consist of 2 x Optane SSDs + 6 x NVMe SSDs with size being 750GB and 1.5TB for Optane and NVMe SSDs respectively.

From a storage tier perspective, ILM categorizes Optane as SSD-MEM-NVMe tier and gives it the highest weight to enable migration of frequently accessed data. With this additional Optane tier, the multi tier preference list for random data migration using Extent groups, in the descending order of preference, is: SSD-MEM-NVMe, SSD-PCIe, SSD-SATA and DAS-SATA. Thus, Optane tier is the most preferred tier used primarily for migrating random data that is read with high frequency.

> Note: AOS prefers non-Optane NVMe SSDs as first choice for initial placement to avoid filling up the scarce Optane resources.

Reads in all-NVMe configurations with Intel Optane (Intel Optane + NVMe SSD) are managed with two tiers - the highest performance media being Intel Optane is tier 0 (SSD-MEM-NVMe) and the lower performance media being the other (non-Optane) NVMe SSDs are treated as tier 1 (SSD-PCIe). AOS migrates frequently accessed (frequently read) random data to the Optane tier to leverage the low latency high read speed. Less frequently read random data is retained in Tier 1 that is the normal NVMe SSD tier. Random data when it has higher read frequency versus write frequency is defined as read-hot. AOS (via ILM) migrates Extent groups for read-hot random data (read intensive workloads) to the Optane tier based on the pre-configured migration weights assigned to the tiers.

Important: Ensure that all the nodes in the cluster have only Intel Optane and NVMe SSD drives (all-NVMe with Intel Optane). See Requirements and Limitations for Performance Tier on page 89.

The I/O path is described in the following figure.

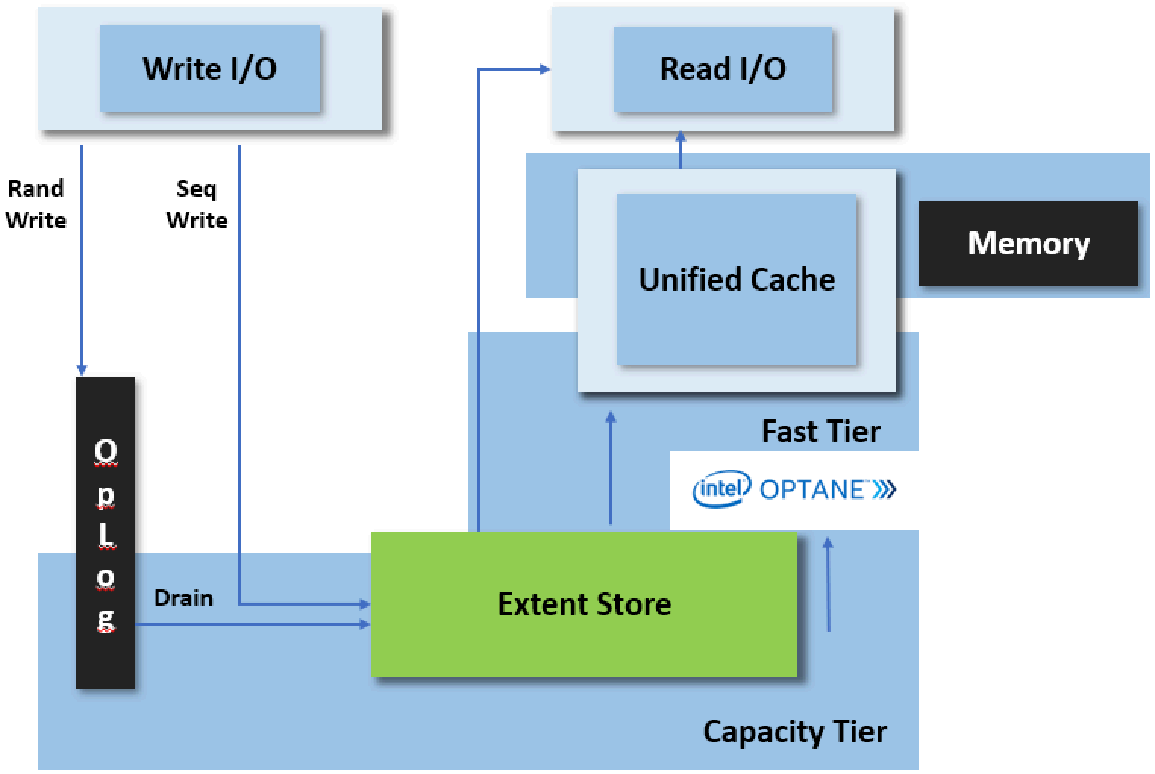


Figure 13: I/O in All-NVMe Drive Configuration with Optane

Summary: AOS Optimizations for Optane

AOS optimizes the use of Optane drives in the NVMe drive configurations in the following ways:

- AOS migrates the random data read with high frequency (read-hot random data) to the Optane performance tier (SSD-MEM-NVMe) since Optane drives provide very low read latency. AOS uses the Read count versus write count and migration weights for a random data segment to decide whether it should be migrated to the Optane tier.

  Note: AOS (via ILM) uses pre-configured read and write weights for tier-based migration or migration decisions. When a data segment is no longer read-hot or random, it is migrated to non-Optane NVMe tier to conserve Optane resources.

- AOS (Cassandra service) uses only the non-Optane NVMe drives for metadata.
- AOS prefers non-Optane NVMe drives as first choice for initial placement to avoid filling up the scarce Optane resources.
- AOS uses both Optane and non-Optane NVMe tiers for Oplog and Extent store.

- AOS reserves capacity for different replication factors, for example, in an Optane drive of 750GB capacity as follows:
    - With RF2 enabled: 190 GB is reserved capacity.
    - With RF3 enabled: 280 GB is reserved capacity.

Requirements and Limitations for Performance Tier

Requirements

The Intel Optane Performance tier is only supported is the following requirements are fulfilled:

- The AOS installation is a fresh installation of an AOS version that support Optane Tier for NVMe (minimum AOS version of 6.1).
- Ensure that the clusters have only Intel Optane and other NVMe SSD drives.
- Ensure that all the containers in the cluster have AES enabled.

The following limitations apply to the Optane performance tier.

- The Optane performance tier is not supported on clusters upgraded to supporting AOS versions.
- Performance tier is auto-disabled in a cluster if any node in the cluster has an HDD or a SATA SSD. If any node in the cluster has a drive configuration that includes a SATA SSD or an HDD, AOS disables the performance tier for the cluster as a whole including on all the other nodes that have only NVMe and Intel Optane drives.

  When the Optane performance tier is disabled, an alert is raised. To clear the alert and auto-enable the performance tier, remove the SATA SSD or HDD that is present on any node in the cluster.

  Introducing (or adding) a SATA SSD or HDD in any node in the cluster also disables the performance tier and raises the alert.
- The Optane performance tier is auto-disabled if any node in the cluster that does not have any Optane drives. In other words, all the nodes in the cluster must have only NVMe and Optane drives.

MINIMUM FIELD REQUIREMENTS FOR NUTANIX CLOUD INFRASTRUCTURE (NCI)

This section provides the information about the minimum field requirements to set up Nutanix Cloud Infrastructure (NCI) with HCI nodes.

For information about the maximum supported NCI node capacity, see Nutanix Configuration Maximums.

Controller VM (CVM) Field Specifications

This section provides the information about the minimum CVM configurations based on the platform category available at field:

> Note: For the platforms described in this section, if two SSD devices (SAS/SATA or NVMe) are used for Cassandra metadata, the SSD devices must be ext4 format (non-blockstore) except for NX-1065 platform that uses one SSD.

To interpret the acronyms and symbols used in the Table 15: CVM Field Specifications on page 91 table, see the following information:

Table 14: Acronyms and Symbols - Description

Attribute	Specification	Description
Symbols	Column headings placed under “$”	$ Indicates the intended functionality or setup to be achieved or set $
	Column Headings placed under “#”	# Indicates the existing field configuration #
	Column Headings placed under “&“	& Indicates the minimum configurations to be met at the field for NCI &
Minimum	Minimum Requirements	Indicates the mandatory minimum requirements for CVM and HCI nodes to operate. If you proceed with the mandatory minimum requirements, the performance of the system gets impacted based on the workloads and any expansion case that arises due to field customization scenarios.

Attribute	Specification	Description
Recommended	Nutanix-Recommended Requirements	Indicates the Nutanix-recommended minimum requirements for the CVM and HCI nodes that is based on the future expansion considerations. If you proceed with the recommended minimum requirements, the system accommodates the additional field customizations and provides less latency.
-	Configurations marked with "-"	Indicates the configuration is either not applicable or not restricted to be set to any specific value.

Table 15: CVM Field Specifications

$ Platform Category to be deployed at the field $	# Platform Details/Specifications #	& Minimum CVM Field Requirements &	
		vCPU	vRAM (in GiB)
Virtual Desktop Infrastructure (VDI) General x86 server virtualization -		8	20

AMD - Naples, Rome		Naples = 8
		Rome =12
			vRAM depends on the node capacity (All-Flash or Hybrid HCI Node used).
			For more information about vRAM requirements based on node capacity, see Snapshot Frequency Field Requirements on page 104
			.

IBM Power PC - 1U, 2U (EIA rack unit) Power8		6	32
Dense Storage* (Minimum 2 Numa nodes and HDD >= 32 TB or SSD >= 48 TB)		*Minimum*: (8 to 14)*** *Recommended:* 14	40

$ Platform Category to be deployed at the field $	# Platform Details/Specifications #	& Minimum CVM Field Requirements &	
		vCPU	vRAM (in GiB)
High Performance** (Numa nodes >= 2, NVMe Drives >= 2, Cores >= 8) RDMA (2 or 4 Socket) is enabled. iSER is enabled 8 or more physical core present in each NUMA node. .		(8 to 12)***	*Minimum:* • 32 = For 2- socket RDMA • 48 = For 4-socket RDMA • 64 = For iSER
High Performance** (Numa nodes >= 2, NVMe Drives >= 2, Cores >= 8) RDMA (2 or 4 Socket) and Hyperthreading is enabled. or 12 or more physical core present in each NUMA node.		(12 to 16)***	
High Performance^ (Total NVMe capacity > 96 TB)		*Minimum*: (12 to 16)*** *Recommended:* 14	64
Generic Number of physical cores in each NUMA node: • 8 or more • 6 or more and Hyperthreading is enabled		(8 to 12)***	20
Remote or Branch Office (ROBO) Specific Deployments and EDGE Virtualization Specific Deployments		6	20
Storage-Only Node (Minimal Compute Node)^^		Number of CPUs on physical host minus 2^^.	Available RAM minus 16 GB, limited to a maximum of 256 GiB^^.

$ Platform Category to be deployed at the field $	# Platform Details/Specifications #	& Minimum CVM Field Requirements &	
		vCPU	vRAM (in GiB)

**Dense Storage specifications include:*

- Number of NUMA nodes = 2 or more.
- Either of the following storage specification:
  - HDD = 32 TB or more.
  - SSD = 48 TB or more (Any combination: SAS/SATA only, NVMe, only, or a mix of SAS/SATA and NVMe).
  - HDD + SSD = 120 TB or more

***High performance specifications include:*

- Number of NUMA nodes = 2 or more.
- Number of NVMe Drives = 2 or more.
- Number of physical cores in each NUMA node = 8 or more / 12 or more.

**** The vCPU value is based on the number of physical cores per socket in each Numa node. For example:*

- If the number of physical cores is 12, the minimum vCPU should be:
  - (12 -2) = 10 when the supported range specified is 8 to 12 or 8 to 14.
  - 12 when the supported range specified is 12 to 16.
- If the number of physical cores is 16, the minimum vCPU should be:
  - (16 -2) = 14 when the supported range specified is 12 to 16 or 8 to 14.
  - 12 when the supported range specified is 8 to 12.
- If the number of physical cores is 24, the minimum vCPU should be:
  - 16 when the supported range is 12 to 16.
  - 12 when the supported range specified is 8 to 12.
  - 14 when the supported range specified is 8 to 14.

*^ Total NVMe capacity = nvme_slots * max_nvme_drive_size. For example, if 20 slots are available for NVMe and one of the slot consists of an NVMe drive of 7 TB, then the total NVMe capacity = 20 * 7 TB = 140 TB.*

^^ Storage-Only node (Minimal Compute Node):

- CVM vCPU = Till Foundation version 5.3.x, the vCPU is capped to a maximum of 22 vCPUs. From Foundation version 5.4 onwards, the capping of maximum 22 vCPUs is not applicable.
- CVM vRAM = The memory capping to a maximum of 256 GiB is applicable from Foundation version 5.3 and above.For example, if the available RAM is 512 GiB, the system allocates a maximum of 256 GiB and does not considers the 512-16 = 496 GiB value. However, if you change the system allocated vRAM, the vRAM gets overridden with the supplied value.

  In the earlier Foundation versions, the memory allocation happens without capping to 256 GiB.

CVM Specifications - Cluster Feature Based

The following table provides the CVM Specifications based on the features enabled in the cluster:

Table 16: CVM Specifications - Cluster Feature Based

$ Cluster Feature $	& CVM Field Requirements - vCPU and vRAM (in GiB): Minimum, Recommended &	
	Minimum	Recommended
Fault Tolerance (FT) 2 with Replication Factor (RF) 3 enabled	vRAM: • Default platform memory + 8 GiB or 40 GiB whichever is lower. • For dense nodes, minimum is 40 GiB	
De-duplication	vRAM: • Performance Tier de-duplication: Platform vRAM + 8 GiB • Capacity and Performance Tier de-duplication: Platform vRAM + 12 GiB	32 GiB
Performance tier de-duplication	vRAM: Platform vRAM + 8 GiB	32 GiB
Storage Performance Development Kit (SPDK) is enabled with at least 1 NVMe	vRAM = 48 GiB vCPU = Equal to the number of physical cores in the CVM NUMA node.	

HCI Node Capacity Guidelines

This section provides the information about the generic storage and capacity guidelines that must be followed for the following deployment scenarios:

- New HCI node. See New HCI Node on page 97.
- Modification in deployed HCI Node. See Modification in Deployed HCI Node on page 97.
- Heterogeneous Clusters Management. See Heterogeneous Cluster Management on page 97.

> Note: Ensure that the maximum capacity of the node is inline with qualified HCI and NUS limits.

> Important: NVMe is considered as one SSD type. All the global rules, guidelines, or requirements specified for SSDs are also applicable to NVMe; however, any distinct requirements that apply only to NVMe are specified exclusively.

Global Storage Guidelines for Hybrid HCI Node

This section defines the global storage rules that apply to Hybrid HCI nodes.

Rules for Minimum SSD capacity and a minimum number of SSD Drives

The following rules apply for minimum SSD capacity and minimum number of SSD drives in Hybrid HCI node:

- Every node should have a minimum of 2 SSD drives.
- A minimum ratio of 1:2 ratio should be maintained for SSD:HDD drives.
- Minimum SSD capacity is maintained as a percentage of overall node size.

The following table provides the guidelines for SSD tier capacity based on node type and overall node capacity:

Table 17: Guidelines for SSD Tier Capacity

Node Type	Overall Node Capacity	Guidelines for SSD Tier Capacity
Nodes with snapshots/ DR (NearSync, 1 hr. to less than 6 hrs., 6 hrs. to less than 24 hrs.)	< = 60 TB	SSD tier must be a minimum 4% of the total node capacity or active Working Set Size (WSS), whichever is greater.
	> 60 TB	SSD tier must be a minimum 10% of the total node capacity or 4% of the total node capacity plus active Working Set Size (WSS), whichever is greater.
Nodes without snapshot/DR or 24 hours RPO	< = 136 TB	SSD tier must be a minimum 4% of the total node capacity or active Working Set Size (WSS), whichever is greater.
	> 136 TB	SSD tier must be a minimum 10% of the total node capacity or 4% of the total node capacity plus active Working Set Size (WSS), whichever is greater.

Note:

- If nodes are configured with 2 NVMe + 2 HDDs configuration, the SSD capacity should be a minimum of 30% or more of the overall node capacity.
- For more information about requirements for Hybrid HCI nodes with snapshots/DR, see Synchronous (0 seconds RPO), Asynchronous (NearSync, 1 hour to less than 6 Hours), Asynchronous (6 hours to less than 24 hours) and Asynchronous (24 Hours or more) sections in *Nutanix Disaster Recovery Guide*.
- For more information about requirements for Hybrid HCI nodes without snapshots/ DR, see Hybrid HCI Node on page 99.

- Drive sizes within a tier (HDD tier or SSD (SAS/SATA or NVMe) tier) should be less than a skew of 20% difference. For more information, see New HCI Node on page 97 and Modification in Deployed HCI Node on page 97.
- A combination of different SSDs (SAS/SATA + NVMe) and HDD is not allowed in the same HCI node. The SSDs can be either SAS/SATA or NVMe, and mixing of NVMe and HDD, or SSD (SAS/SATA) and HDD is allowed.

Rule for Expansion

The latest generation servers can be used to expand clusters with older generation server nodes. For information about hardware mixing restrictions, see Product Mixing Restrictions in *NX Series Hardware Administration Guide*.

All the SSDs must be added upfront to ensure that the SSD tier is big enough to accommodate the meta-data for the maximum capacity when a node is fully populated. The SSDs added later are not used for Cassandra. HDDs can be added if the above global rules; Rules for Minimum SSD capacity and a minimum number of SSD Drives on page 95, are met at your site.

Rule for Partial Population - Hybrid Node

Partial Population on Hybrid HCI nodes is allowed if the following global rules are met at your site:

- Rules for Minimum SSD capacity and a minimum number of SSD Drives on page 95 as specified above.
- Rule for Expansion on page 96 as specified above.

Global Storage Guidelines for All-Flash HCI Node

This section defines the global storage rules that apply to All-Flash HCI nodes.

The following generic storage guidelines are applicable for All-Flash HCI nodes:

- Drive sizes within a node tier should be less than a skew of 20% difference. For more information, see New HCI Node on page 97 and Modification in Deployed HCI Node on page 97.

Rules for Partial Population - All Flash Node

Partial population is allowed on any individual SSDs or any combination of SSDs.

Global Storage Guidelines for Heterogeneous Clusters

This section defines the global storage rules that apply to heterogeneous clusters.

Mixing of Hybrid and All-Flash Nodes in a Cluster

The following rules apply when you mix Hybrid and All-Flash nodes in a cluster:

- Adding All-Flash HCI nodes in a hybrid cluster is supported.
- Adding Hybrid HCI nodes in an All-Flash cluster is not supported.

DR implications of mixing of nodes of different sizes and types in a Cluster

When you mix nodes of different sizes and types within a cluster, the highest RPO level applicable to any one node in the cluster is applicable to the entire cluster. For example:

- If you have a 4-node All-Flash cluster with 2 Nodes of 90TB which can support 1HR RPO and 2 nodes of 135 TB which can support only 6HR RPO, the highest RPO value of *6 hours to less than 24 hours* is applicable for the entire cluster.
- If you have 2 AF nodes of 135 TB which can support 6HR RPO and 2 Hybrid nodes of 135 TB which can support 1HR RPO, the RPO supported for the cluster is the highest of the two, which is *6 hours to less than 24 hours*.

New HCI Node

When you deploy a new HCI node (Hybrid or All-Flash) to a new or existing cluster, ensure that the capacity of the drives in the HCI node is similar within a tier (HDD tier or SSD tier (SAS / SATA SSD or NVMe)). A greater capacity difference leads to performance inconsistency as the larger drive sizes are targeted for initial writes and, in the back-end, the system constantly tries to balance the data on the drives. This leads to an increased back-end workload leading to fluctuation in performance.

For example:

- For All-Flash HCI, all nodes must use identical capacity SSDs (for example, 1.92 TB), regardless of the SSD type (SATA/SAS/NVMe) or mixed SSDs (SATA/SAS + NVMe)
- The Hybrid HCI node must have the same capacity SSDs (for example, 3.86 TB SATA SSD) and the same capacity HDD (for example, 8 TB HDD)

Cluster Expansion

The latest generation servers can be used to expand clusters with older generation server nodes. For information about hardware mixing restrictions, see Product Mixing Restrictions in *NX Series Hardware Administration Guide*.

Modification in Deployed HCI Node

This section describes the guidelines for the addition of new drives or replacement of existing drives in the deployed HCI node of a cluster.

When you add new drives to an existing HCI node, in addition to adhering to the maximum node capacity limits qualified for HCI and NUS, ensure that the drive sizes in a node are similar within a tier (HDD tier or SSD (SAS/SATA or NVMe) tier) with a difference of no more than 20%.

For example:

- *Supported combination* – A mix of 16 TB and 18 TB HDD (Less than 20% skew in drive sizes within the HDD tier in a node).
- *Unsupported combination* – A mix of 14 TB and 18 TB HDD (More than 20% skew in drive sizes within the HDD tier in a node).

Greater than 20% skew in drive sizes might lead to performance inconsistency as the larger drive sizes are targeted for initial writes, and in the back-end, the system constantly tries to balance the data on the drives, which leads to an increased workload leading to fluctuation in performance.

> Note: Starting with AOS 6.0, during drive replacements in the flash tier, in case of non-availability of drive sizes that conform to the above rules, bigger drives can be configured. However, the bigger drives are downsized by AOS to match the capacity of the rest of the drives in the flash tier on the node to avoid performance issues.

Heterogeneous Cluster Management

If the HCI nodes in the cluster are of different sizes, ensure that you adhere to the following guidelines:

- The capacity that is allocated to the largest HCI node is available with the remaining HCI nodes of the cluster. This management technique enables you to handle any failure that occurs on the largest HCI node. For example, in a 4-Node cluster, if the largest HCI node capacity is 75 TB, ensure that the total capacity of the remaining three HCI nodes is equal to or greater than 75 TB.

- When you add new nodes to the cluster, you must follow the Redundancy Factor (RF) of the cluster. If the cluster is an RF-2 cluster, you must utilize the limit and deploy 2 HCI nodes of the same type (either Hybrid or All-Flash). In case you add a smaller number of nodes than the RF of the cluster, then you can perform any capacity increase action only based on the available capacity of the remaining nodes in the cluster.

  For example, if you have set a cluster with three nodes of 10 TB capacity each and the fourth node of 80 TB capacity, the total capacity of the cluster becomes 110 TB. In this case, the usable raw capacity is only 60 TB (with 30 TB logical capacity for RF2). The total capacity of 80 TB of the fourth node cannot be used due to the unavailability of required space on the rest of the nodes for replica placement.

If you mix nodes of different sizes and types in a heterogeneous cluster, ensure that you observe the DR implications specified in Global Storage Guidelines for Heterogeneous Clusters on page 96.

HCI Node - Recommended Drive Addition / Replacement Instructions

This section describes the Nutanix-recommended instructions to be followed when you add or replace a drive in an existing HCI node.

New Drive Addition

Nutanix recommends you add all the new physical drives to one cluster node at the same time. This minimizes the risk of data unavailability when you perform the drive addition for multiple cluster nodes.

> Note: If you proceed with one physical drive addition, at a time to one cluster node or at the same time to multiple cluster nodes, the following issues might occur in the system:
>
> - Number of Stargate restarts increases, and the local Stargate becomes unavailable for a short term.
> - Stargate across multiple cluster nodes can restart at the same time.

The new physical drives that are added to the deployed HCI node of a cluster must be equally distributed between the nodes in the cluster in a round-robin fashion and RF of the cluster must be maintained.

Existing Drive Replacement

The drive replacement should be done one at a time in the cluster nodes. You must wait for the rebuild to complete before replacing the next drive in the HCI node of a cluster.

HCI Node Field Requirements

This section provides information about the minimum SSD configurations required for the following types of HCI nodes:

- *Hybrid HCI Node* - Involves both SSDs and HDDs. In the case of Hybrid HCI nodes, the SSDs can be either SAS/SATA or NVMe. A combination of different SSDs (SAS/SATA + NVMe) and HDD is not allowed in the same HCI node.
- *All-Flash HCI Node* - Involves only SSDs. In the case of All-Flash HCI nodes, the SSDs can be either SAS/SATA, Optane, NVMe, or a combination of any of these SSDs.

  > Note: The Optane SSD cannot be used as a standalone SSD in an All-Flash HCI Node. It can be used only in combination with SAS/SATA or NVMe.

Important: NVMe is considered as one SSD type. All the global rules, guidelines, or requirements specified for SSDs are also applicable to NVMe; however, any distinct requirements that apply only to NVMe are specified exclusively.

Hybrid HCI Node

The following table provides the information about the minimum and recommended SSD requirements for Hybrid HCI node:

Note:

- Ensure that you adhere to the capacity guidelines described in HCI Node Capacity Guidelines on page 94.
- Ensure that you follow the Nutanix-recommended new drive addition and replacement instructions described in HCI Node - Recommended Drive Addition / Replacement Instructions on page 98.

To interpret the acronyms and symbols used in the Table 19: Hybrid HCI Node - SSD Requirements on page 100 table, see the following information:

Table 18: Acronyms and Symbols - Description

Attribute	Specification	Description
Symbols	Column headings placed under "$"	$ Indicates the intended functionality or setup to be achieved or set $
	Column Headings placed under "#"	# Indicates the existing field configuration #
	Column Headings placed under "&"	& Indicates the minimum configurations to be met at the field for NCI &
M	Minimum Requirements	Indicates the mandatory minimum requirements for CVM and HCI nodes to operate. If you proceed with the mandatory minimum requirements, the performance of the system gets impacted based on the workloads and any expansion case that arises due to field customization scenarios.

Attribute	Specification	Description
R	Nutanix-Recommended Requirements	Indicates the Nutanix-recommended minimum requirements for the CVM and HCI nodes that is based on the future expansion considerations. If you proceed with the recommended minimum requirements, the system accommodates the additional field customizations and provides less latency.
-	Configurations marked with "-"	Indicates the configuration is either not applicable or not restricted to be set to any specific value.

Table 19: Hybrid HCI Node - SSD Requirements

$ Hybrid HCI Node $						
# Node Specification #			& SSD Requirement for HCI Node &			
SSD Type	Number of Drives	Number of HDDs Available	Number of SSDs; Minimum, Recommended		SSD Capacity (in TB): Minimum, Recommended	
			M	R	M	R
SAS/SATA	<= 14	Maximum up to 12 LFF HDDs	2		*If the overall node capacity is up to 136 TB, the minimum SSD capacity should be 4% of the overall capacity or active Working Set Size (WSS)*, whichever is greater.* *If the overall node capacity is greater than 136 TB, the minimum SSD capacity should be 10% of overall capacity or 4 % of the overall capacity plus active Working Set Size (WSS)*, whichever is greater.* Note: If nodes are configured with 2 NVMe + 2 HDDs configuration, the SSD capacity should be a minimum 30% or more of the overall node capacity.	
	> 14	Maximum up to 20 LFF HDDs	4			
	> = 24	Maximum up to 24 LFF HDDs	8			
NVMe**	< = 12	Maximum up to 10 LFF HDDs	2			

$ Hybrid HCI Node $						
# Node Specification #			& SSD Requirement for HCI Node &			
SSD Type	Number of Drives	Number of HDDs Available	Number of SSDs; Minimum, Recommended		SSD Capacity (in TB): Minimum, Recommended	
			M	R	M	R

**The active Working Set Size (WSS) is the amount of data that the application reads/writes frequently. Ensure that the application has enough SSD storage to accommodate its active WSS; otherwise, the application experiences performance degradation. Nutanix sizing tool considers both while sizing for the cluster.*

***The following conditions apply for NVMe drives:*

- No support available for Blockstore + SPDK in NVMe + HDD configuration.
- Hot-Swap for NVMe drives is supported.

> Important:
>
> - Ensure that you refer to the section Snapshot Frequency Field Requirements on page 104 before you use the snapshots at your site.
> - Partial population is supported for Hybrid nodes. For more information, see Global Storage Guidelines for Hybrid HCI Node on page 94.
> - A minimum of 2:1 ratio for HDD:SSD is required to provide sufficient bandwidth on the slower tier to absorb ILM down migrations from the faster tier. In case it is difficult to maintain a 2:1 ratio for HDD:SSD, an All-Flash node is recommended.
>
>   The following server platforms are some exceptions where configuration mechanism of 2:1 ratio for HDD:SSD is not followed:
>
>   - 4 drive slot Dell XC, Lenovo HX, Fujitsu XF platforms (2 SSD + 2 HDD)
>   - 10 drive slot Dell XC, Lenovo HX, Fujitsu XF platforms (4 SSD + 6 HDD)
>   - NX-1065 (1 SSD + 2 HDD), NX-1175S (2 SSD + 2 HDD)
>   - HPE DX360 4 LFF Gen10 and Gen10 Plus, DX320 4 LFF Gen11

All-Flash HCI Node

The following table provides the information about the minimum and recommended SSD requirements for an All-Flash HCI node:

> Note:
>
> - Ensure that you adhere to the capacity guidelines described in HCI Node Capacity Guidelines on page 94.
> - Ensure that you follow the nutanix-recommended new drive addition and replacement instructions described in HCI Node - Recommended Drive Addition / Replacement Instructions on page 98.

To interpret the acronyms and symbols used in the Table 21: All-Flash HCI Node - SSD Requirements on page 103 table, see the following information:

Table 20: Acronyms and Symbols - Description

Attribute	Specification	Description
Symbols	Column headings placed under "$"	$ Indicates the intended functionality or setup to be achieved or set $
	Column Headings placed under "#"	# Indicates the existing field configuration #
	Column Headings placed under "&"	& Indicates the minimum configurations to be met at the field for NCI &
M	Minimum Requirements	Indicates the mandatory minimum requirements for CVM and HCI nodes to operate. If you proceed with the mandatory minimum requirements, the performance of the system gets impacted based on the workloads and any expansion case that arises due to field customization scenarios.
R	Nutanix-Recommended Requirements	Indicates the Nutanix-recommended minimum requirements for the CVM and HCI nodes that is based on the future expansion considerations. If you proceed with the recommended minimum requirements, the system accommodates the additional field customizations and provides less latency.
-	Configurations marked with "-"	Indicates the configuration is either not applicable or not restricted to be set to any specific value.

Table 21: All-Flash HCI Node - SSD Requirements

$ All-Flash HCI Node $					
# Node Specification #		& SSD Requirements for HCI Node &			
SSD Type	Number of Drives	Number of SSDs; Minimum, Recommended		SSD Capacity (in TB): Minimum, Recommended	
		M	R	M	R
SAS / SATA Only	-	2*	4 or more	• ** For metadata and boot drives* • *** SPDK is enabled if a node contains at least 1 NVMe with a minimum 48 GiB vRAM for CVM .* • *For nodes with a capacity greater than 136 TB, a minimum of 64 GiB vRAM is required for CVM.* • *For nodes with all NVMe and capacity greater than 96 TB, a minimum of 64 Gib vRAM is required for CVM.* • **** All 24 NVMe drives are supported with AOS 6.7 and above release.*	
All NVMe***	-	2*	4 or more		
NVMe + SAS / SATA	<=10	2* SAS / SATA 1** NVMe	-		
	> 10	4* SAS/SATA 1** NVMe			
Optane + NVMe	-	2 Optane 2 NVMe	2 Optane 4 NVMe	*For Optane:* 1.5 TB (2 x 750 GB) *For NVMe:* 3.84 TB (2 x 1.92 TB)	*For Optane:* 3 TB (2 x 1.5 TB) *For NVMe:* 7.68 TB (4 x 1.92 TB)
Optane + SAS / SATA	-	2 Optane		*For Optane:* 1.5 / 1.6 TB (2 x 750 / 800 GB)	*For Optane:* 3 / 3.2 TB (2 x 1.5/1.6 TB)

$ All-Flash HCI Node $					
# Node Specification #		& SSD Requirements for HCI Node &			
SSD Type	Number of Drives	Number of SSDs; Minimum, Recommended		SSD Capacity (in TB): Minimum, Recommended	
		M	R	M	R

Important:

- Ensure that you refer to section Snapshot Frequency Field Requirements on page 104 before you use the snapshots at your site.
- The following conditions apply to all the above node specifications:
    - Partial population is allowed on any individual SSDs or any combination of SSDs.
    - Oplog can be on any NVMe device for better performance.
- The following conditions apply to all the above node specifications with Optane SSD only:
    - CVM boot drive must be a non-block store NVMe drive.
    - SPDK is enabled if a node contains at least 1 NVMe SSD.
    - The server platform examples for Optane +NVMe are:
        - NX-8170
        - HPE DX-360
        - Intel DCS LCY2224NX3
    - The server platform examples for Optane + SAS / SATA are:
        - Lenovo HX: Cascade lakeHX3320 (12 SFF) / HX7520 (24 SFF)
        - Intel DCS: LCY2216NX2, LCY2224NX2, LCY2224NX3
- The following conditions apply to all the above node specifications with NVMe only:
    - A maximum of 24 NVMe drives are supported in a node.

Snapshot Frequency Field Requirements

For information about the minimum SSD and CVM requirements based on Snapshot frequency needed at field to define the Recovery Point Objective (RPO), refer any of the following documents:

- *Data Protection and Recovery with Prism Element Guide* - See Resource Requirements Supporting Snapshot Frequency (Asynchronous, NearSync and Metro) information.
- *Nutanix Disaster Recovery Guide* - See On-Prem Hardware Resource Requirements information.

COPYRIGHT

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