

v4.1 | July 2023 | BP-2071

BEST PRACTICES

Nutanix AHV Networking

Copyright

Copyright 2023 Nutanix, Inc.

Nutanix, Inc.

1740 Technology Drive, Suite 150

San Jose, CA 95110

All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. Nutanix and the Nutanix logo are registered trademarks of Nutanix, Inc. in the United States and/or other jurisdictions. All other brand and product names mentioned herein are for identification purposes only and may be trademarks of their respective holders.

Contents

1. Executive Summary.....	5
2. Introduction.....	6
Audience.....	6
Purpose.....	6
Document Version History.....	6
3. AHV Networking Overview.....	9
Open vSwitch.....	9
Bridges.....	9
Virtual Switch.....	9
Ports.....	10
Bonds.....	10
Bridge Chaining.....	12
Virtual Local Area Networks.....	14
Nutanix AHV IP Address Management.....	15
Nutanix AHV MAC Address Management.....	17
4. AHV Network Management.....	18
View Network Status.....	18
Prism Uplink Configuration and Virtual Switches.....	22
Production Network Changes.....	22
OVS Command Line Configuration.....	24
5. AHV Networking Best Practices.....	26
Open vSwitch Bridge and Bond Recommendations.....	26
Load Balancing in Bond Interfaces.....	35
VLANs for AHV Hosts and CVMs.....	43
VLAN for Guest VMs.....	46
IP Address Management Best Practices.....	48
MAC Address Management Best Practices.....	49
CVM Network Segmentation.....	50
Jumbo Frames.....	51

6. Conclusion.....	52
7. Appendix.....	53
AHV Networking Terminology.....	53
AHV Networking Best Practices Checklist.....	53
AHV Command Line Tutorial.....	58
AHV Networking Command Examples.....	61
References.....	63
About the Authors.....	63
About Nutanix.....	64
List of Figures.....	65

1. Executive Summary

The default networking we describe in the [Nutanix AHV best practice guide](#) covers a wide range of scenarios Nutanix administrators encounter. Use this networking guide for situations with unique VM and host networking requirements not covered elsewhere.

The default Nutanix AHV networking configuration provides a highly available network for guest VMs and the Nutanix Controller VM (CVM). This default configuration includes IP address management and simple VM traffic control and segmentation using VLANs. Network visualization for AHV available in Prism also provides a view of the guest and host network configuration for troubleshooting and verification.

This guide is useful when the defaults don't match customer requirements. Configuration options include host networking high availability and load balancing mechanisms beyond the default active-backup, tagged VLAN segmentation for host and CVM traffic, and detailed command-line configuration techniques for situations where a GUI may not be sufficient. The tools we present here enable you to configure AHV to meet the most demanding network requirements.

2. Introduction

Audience

This best practice guide is part of the Nutanix Solutions Library. We wrote it for Nutanix AHV administrators configuring host and VM networking in more complex designs. Readers of this document should already be familiar with the [Nutanix AHV best practice guide](#), which covers basic networking.

Purpose

In this document, we cover the following topics:

- Open vSwitch in Nutanix AHV
 - VLANs for hosts, CVMs, and guest VMs
 - IP address management (IPAM)
 - Network adapter teaming within bonds
 - Network adapter load balancing
 - Command line overview and tips
-

Document Version History

Version Number	Published	Notes
1.0	February 2017	Original publication.
1.1	February 2018	Added jumbo frame configuration, bond name recommendation, and considerations for staging installation with flat switch.

Version Number	Published	Notes
1.2	October 2018	Updated product naming and recommendations regarding the balance-slb bond mode.
1.3	December 2018	Updated the Open vSwitch Bridge and Bond Recommendations section.
2.0	March 2020	Added production workflow instructions, bridge chaining, and VM networking enhancements.
2.1	June 2020	Updated the Nutanix overview, jumbo frame recommendations, and terminology.
3.0	February 2021	Updated links, clarified CVM maintenance mode recommendations, and added AOS 5.19 Virtual Switch recommendations.
3.1	March 2021	Updated the VLANs for AHV Hosts and CVMs section.
3.2	April 2021	Updated the LACP and Link Aggregation section.
3.3	September 2021	Updated the View Network Status, Open vSwitch Bridge and Bond Recommendations, Load Balancing in Bond Interfaces, and AHV Networking Command Examples sections.

Version Number	Published	Notes
3.4	January 2022	Updated Post-Imaging Network State, Network Connections for 2 × 10 GbE NICs, and Network Connections for 2 × 10 GbE and 2 × 1 GbE NICs images.
3.5	August 2022	Updated links.
3.6	September 2022	Removed note about managed network limit from the AHV Networking Overview section.
3.7	January 2023	Replaced reference to KB 2852 with reference to the AHV Administration Guide.
4.0	June 2023	Removed older CLI and GUI commands and clarified LACP speed recommendations. Added duplicate MAC best practices and active adapter configuration details.
4.1	July 2023	Added technical clarification to the Open vSwitch Bridge and Bond Recommendations section.

3. AHV Networking Overview

Nutanix AHV uses Open vSwitch (OVS) to connect the CVM, the hypervisor, and guest VMs to each other and to the physical network on each node. The CVM manages the OVS inside the AHV host. Don't allow any external tool to modify the OVS.

Open vSwitch

OVS is an open source software switch designed to work in a multiserver virtualization environment. In Nutanix AHV VLAN-backed networks, the OVS behaves like a layer 2 learning switch that maintains a MAC address table. The hypervisor host and VMs connect to virtual ports on the switch.

Nutanix AHV exposes many popular OVS features through the Prism GUI, such as VLAN tagging, load balancing, and link aggregation control protocol (LACP).

Bridges

A bridge switches traffic between physical and virtual network interfaces. The default Nutanix AHV configuration includes an OVS bridge called br0 and a native Linux bridge called virbr0. The virbr0 Linux bridge carries management and storage communication between the CVM and AHV host. All other storage, host, and VM network traffic flows through the br0 OVS bridge by default. The AHV host, VMs, and physical interfaces use ports for connectivity to the bridge.

Virtual Switch

In Nutanix AOS versions 5.19 and later, you can use a virtual switch to manage multiple bridges and uplinks in Prism. For more information, view the [Host Network Management](#) section in the AHV Administration Guide.

We designed the virtual switch configuration to provide flexibility in configuring virtual bridge connections. A virtual switch defines a collection of AHV nodes

and the uplink ports on each node. It aggregates similar OVS bridges on all the AHV nodes in a cluster. For example, `vs0` is the default virtual switch and aggregates the `br0` bridges and `br0-up` uplinks from all the nodes.

Ports

Ports are logical constructs created in a bridge that represent connectivity to the virtual switch. Nutanix uses several port types:

- An internal port—with the same name as the default bridge (`br0`)—acts as the AHV host management interface.
 - Tap ports connect VM virtual NICs (vNICs) to the bridge.
 - VXLAN ports are only used for the IP address management (IPAM) functionality provided by AHV.
 - Bonded ports provide NIC teaming for the physical interfaces of the AHV host.
-

Bonds

Bonded ports aggregate the physical interfaces on the AHV host for fault tolerance and load balancing. By default, the system creates a bond named `br0-up` in bridge `br0` containing all physical interfaces. Changes to the default bond (`br0-up`) using `manage_ovs` commands can rename it to `bond0` when using older examples, so your system might be named differently than the following diagram. Nutanix recommends using the name `br0-up` to quickly identify this interface as the bridge `br0` uplink. Using this naming scheme, you can also easily distinguish uplinks for additional bridges from each other.

OVS bonds allow for several load-balancing modes to distribute traffic, including `active-backup`, `balance-slb`, and `balance-tcp`. Administrators can also activate LACP for a bond to negotiate link aggregation with a physical switch. During installation, the `bond_mode` defaults to `active-backup`, which is the configuration we recommend for ease of use.

The following diagram illustrates the networking configuration of a single host immediately after imaging. The best practice is to use only the 10 GbE or more

NICs and to disconnect the 1 GbE NICs if you don't need them. For additional information on bonds, refer to the AHV Networking Best Practices section.

Note: Only use NICs of the same speed in the same bond.

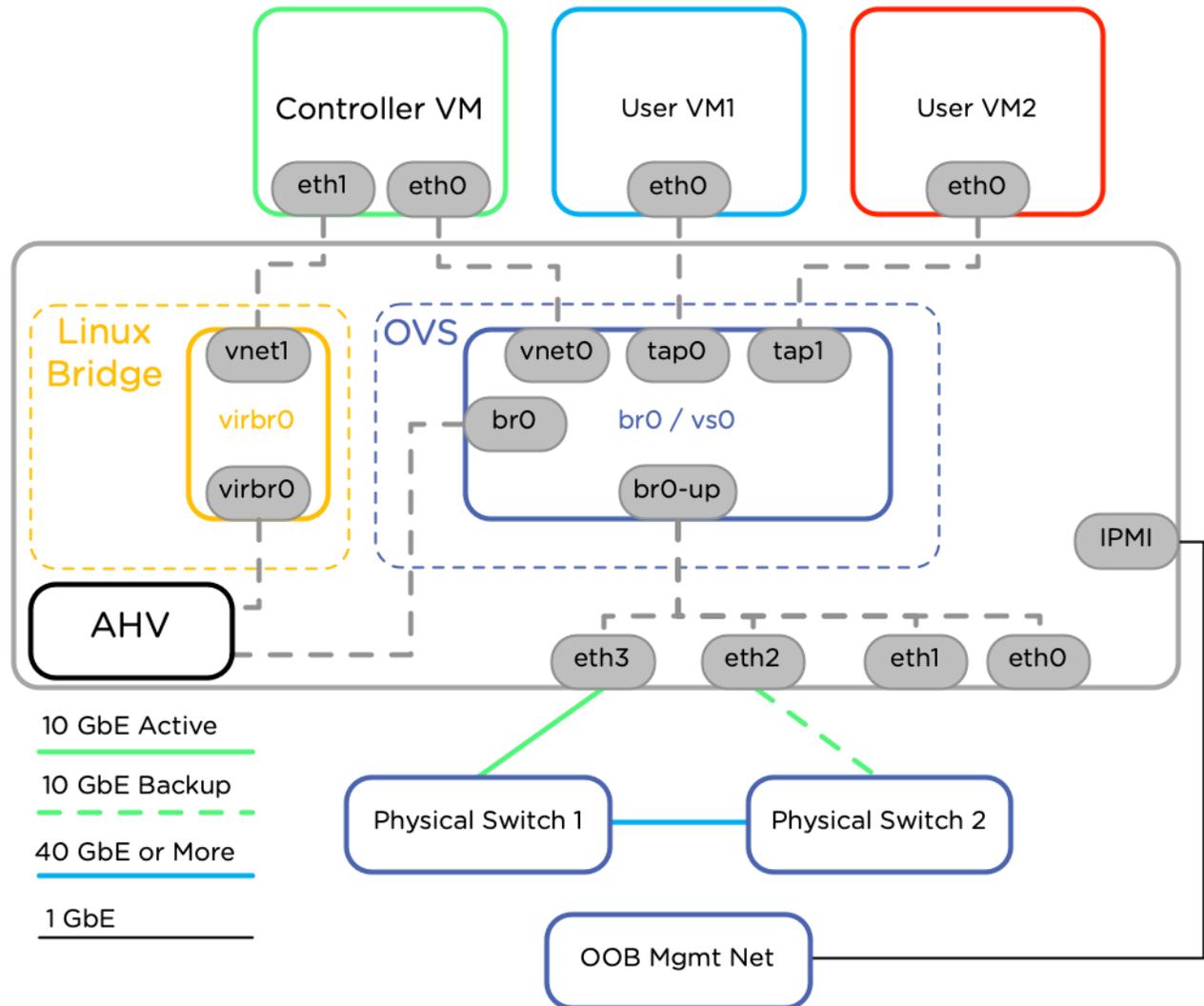


Figure 1: Post-Imaging Network State

Connections from the server to the physical switch use 10 GbE or more networking. You can establish connections between the switches with 40 GbE or more direct links or through a leaf-spine network topology (not shown). The IPMI management interface of the Nutanix node also connects to the out-of-

band management network, which may connect to the production network. Each node always has a single connection to the management network, but we omitted this element from further images in this document for clarity and simplicity.

For more information on the physical network recommendations for a Nutanix cluster, refer to the [Physical Networking best practice guide](#).

Bridge Chaining

From Nutanix AOS 5.5 onward, all AHV hosts use a bridge chain (multiple OVS bridges connected in a line) as the backend for features like microsegmentation. Each bridge in the chain performs a specific set of functions. Physical interfaces connect to bridge brX, and VMs connect to bridge brX.local. Between these two bridges are br.microseg for microsegmentation and br.nf for directing traffic to network function VMs. The br.mx and br.dmx bridges allow multiple uplink bonds in a single AHV host (such as br0-up and br1-up) to use these advanced networking features.

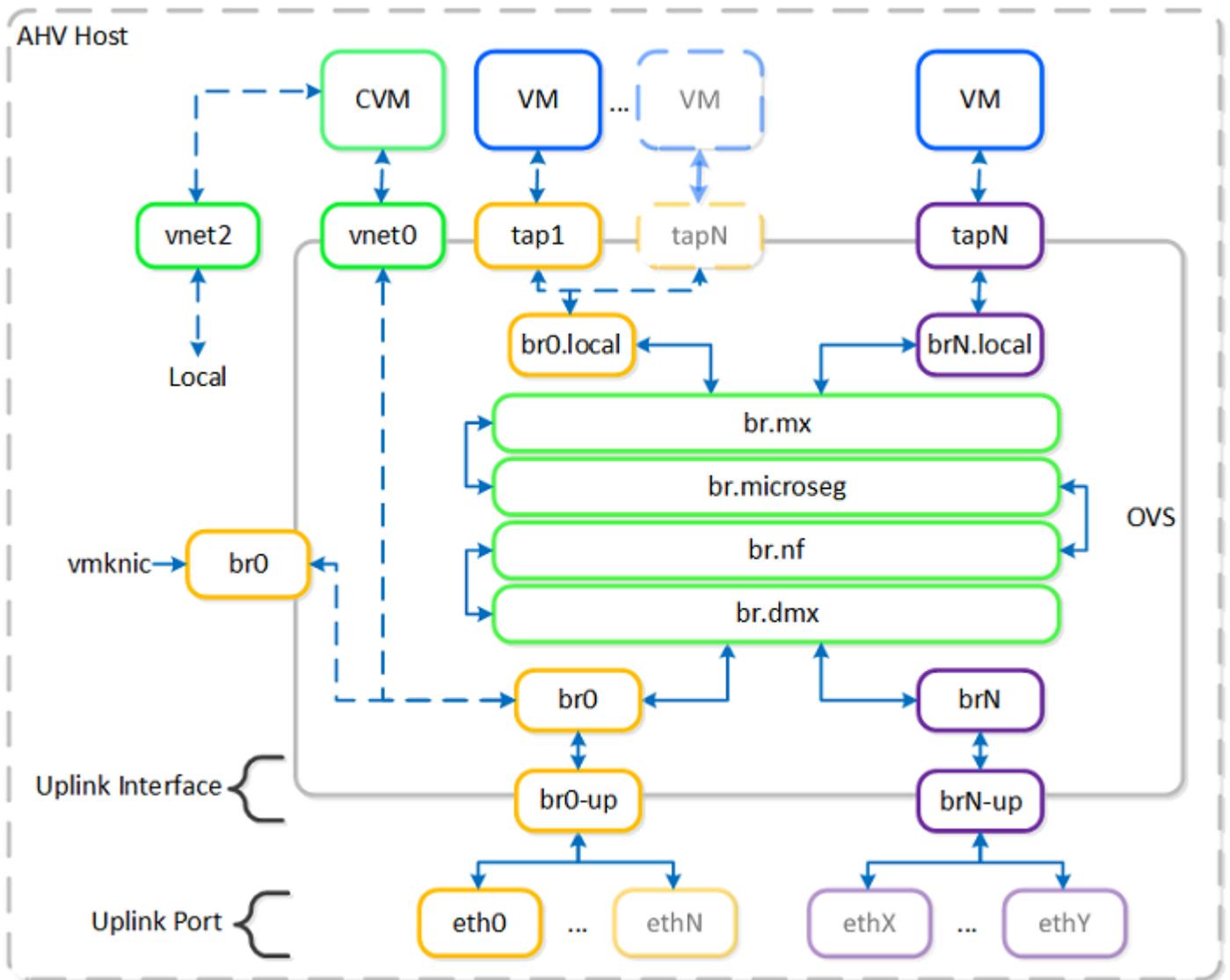


Figure 2: AHV Bridge Chain

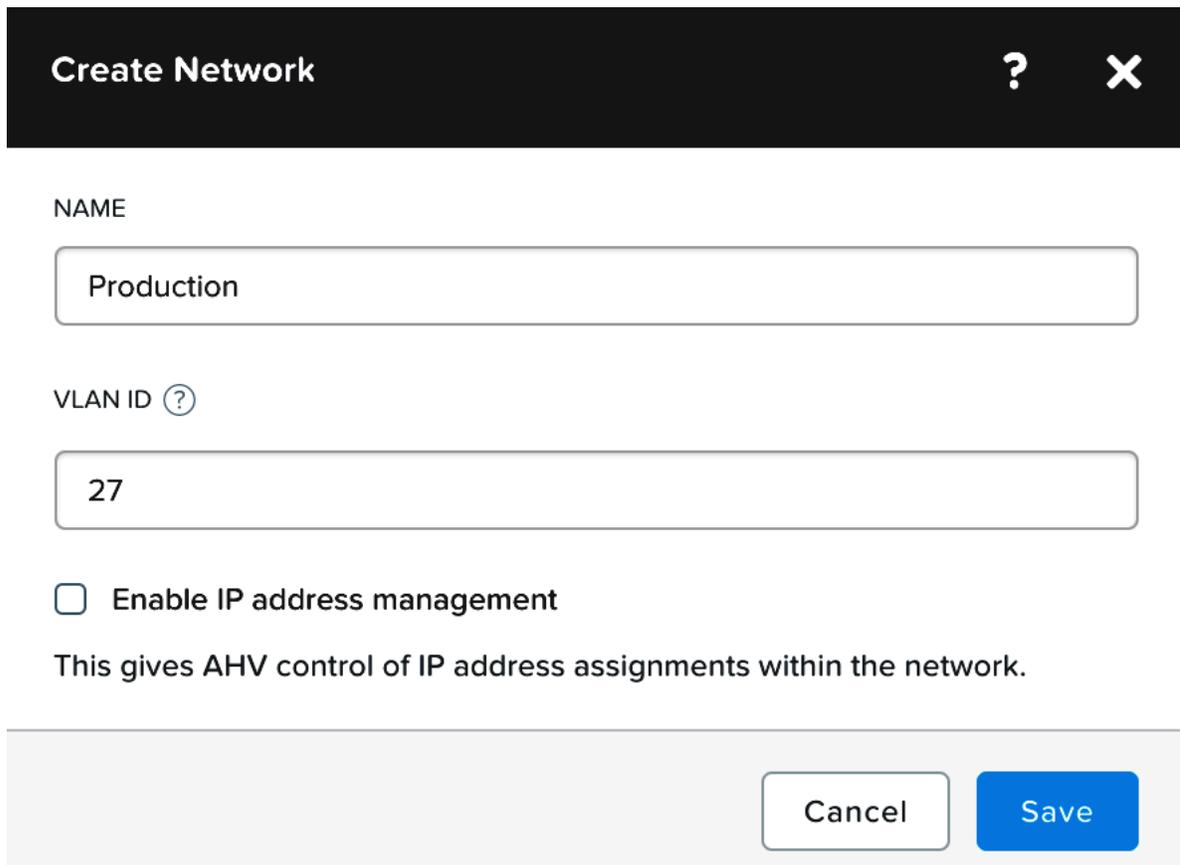
Traffic from VMs enters the bridge chain at brX.local and flows through the chain to brX, which makes local switching decisions. The brX bridge either forwards the traffic to the physical network or sends it back through the chain to reach another VM. Traffic from the physical network takes the opposite path, from brX all the way to brX.local. All VM traffic must flow through the bridge chain, which applies microsegmentation and network functions.

The management of the bridge chain is automated, and no user configuration of the chain is required or supported. Because there are no configurable components, we don't include this bridge chain, which exists between the physical interfaces and the guest VMs, in other diagrams.

Virtual Local Area Networks

Nutanix AHV supports the use of VLANs for the CVM, AHV host, and guest VMs. We discuss the steps for assigning VLANs to the AHV host and CVM in the Best Practices section. You can easily create and manage a vNIC's networks for VMs using the Prism GUI, the Acropolis CLI (aCLI), or REST without any additional AHV host configuration.

Each virtual network in AHV maps to a single VLAN and bridge. You must create each VLAN and virtual network created in AHV on the physical top-of-rack switches as well, but integration between AHV and the physical switch can automate this provisioning. In the following figure, we're using Prism to assign the network the name Production and the VLAN ID 27 for a network on the default bridge, br0. This process adds the VLAN tag 27 to all AHV hosts in the cluster on bridge br0.



Create Network ? X

NAME

Production

VLAN ID ?

27

Enable IP address management

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

Cancel Save

Figure 3: Prism UI Network Creation

By default, all VM vNICs are created in access mode on br0, which permits only one VLAN per virtual network. However, you can choose to configure a vNIC in trunked mode using the aCLI instead, allowing multiple VLANs on a single VM NIC for network-aware VMs. For more information on vNIC modes or multiple bridges, refer to the Best Practices section.

Nutanix AHV IP Address Management

With IP Address Management (IPAM), Nutanix AHV can assign IP addresses automatically to VMs using DHCP. Administrators can configure each virtual network with a specific IP subnet, associated domain settings, and IP address pools available for assignment to VMs.

Create Network ? ✕

Enable IP address management

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

NETWORK IP ADDRESS / PREFIX LENGTH

GATEWAY IP ADDRESS

Configure Domain Settings

IP ADDRESS POOLS ?

+ Create Pool

START ADDRESS	END ADDRESS	
192.168.101.20	192.168.101.200	✎ ✕

Cancel **Save**

Figure 4: IPAM

Administrators can use Nutanix AHV with IPAM to deliver a complete virtualization deployment, including network address management, from the Prism interface.

Nutanix AHV assigns an IP address from the address pool when it creates a managed VM NIC; the address releases back to the pool when the VM NIC or VM is deleted. With a managed network, AHV intercepts DHCP requests from guest VMs and bypasses traditional network-based DHCP servers. AHV uses the last network IP address in the assigned network for the managed network DHCP server unless you select Override DHCP server when you create the network.

Nutanix AHV MAC Address Management

Nutanix AHV clusters use the MAC address prefix OUI **50:6B:8D** by default to create guest VM NICs. AHV assigns a random MAC address in this range that is guaranteed to be unique in the cluster when creating a new NIC. See the MAC Address Management Best Practices section later in this document for information about multicluster environments.

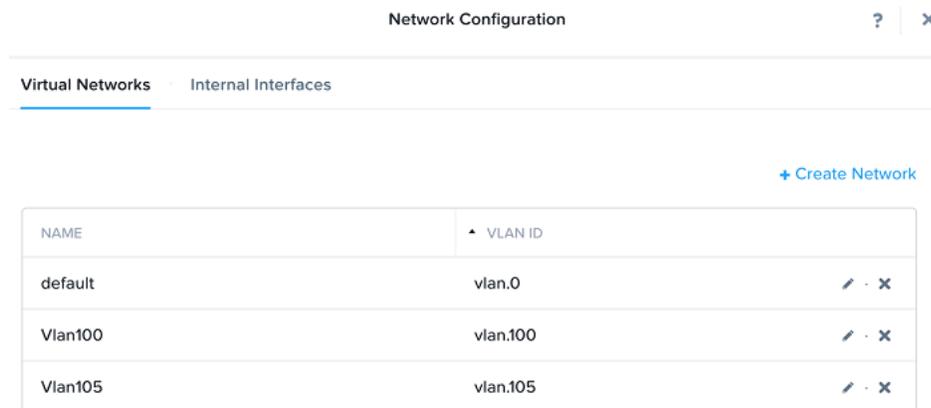
4. AHV Network Management

The following sections illustrate the most common methods used to manage network configuration for VMs and AHV hosts. Some information is visible in both Nutanix Prism and the CLI, and we show both outputs when available. Refer to the AHV Command Line section in the appendix for more information on CLI usage.

View Network Status

View Network Configuration for VMs in Prism

Select Network Configuration, then Virtual Networks to view VM virtual networks from the VM page, as shown in the following figure.



The screenshot shows the 'Network Configuration' page in the Nutanix Prism UI. The page has a title bar with a question mark and a close button. Below the title bar, there are two tabs: 'Virtual Networks' (which is selected and underlined) and 'Internal Interfaces'. A '+ Create Network' button is located in the top right corner of the main content area. Below the tabs is a table with three columns: 'NAME', 'VLAN ID', and a column for actions (edit and delete icons). The table contains three rows of data.

NAME	VLAN ID	
default	vlan.0	 · 
Vlan100	vlan.100	 · 
Vlan105	vlan.105	 · 

Figure 5: Prism UI Network List

You can see individual VM network details under the Table view on the VM page by selecting the desired VM and choosing Update.

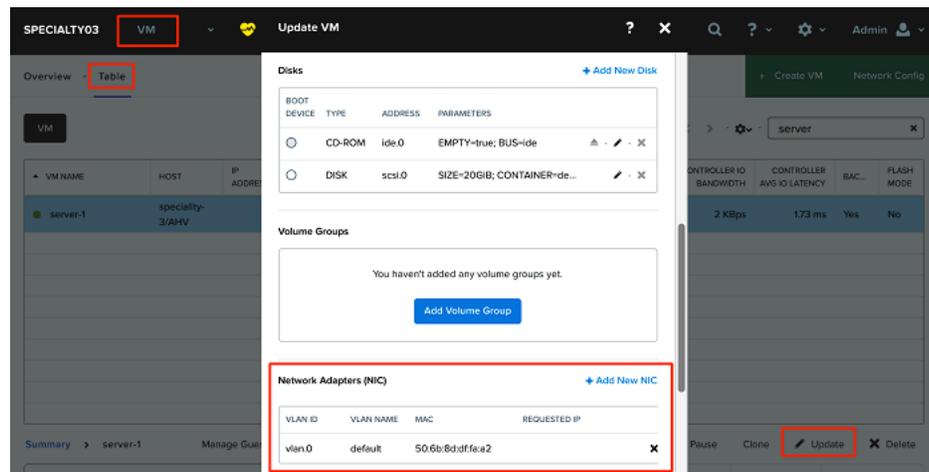


Figure 6: Prism UI VM Network Details

View AHV Host Network Configuration in Prism

Select the Network page to view VM- and host-specific networking details. When you select a specific AHV host, Nutanix Prism displays the network configuration, as shown in the following figure. For more information on the new Network Visualization feature, refer to the [network visualization section of the Prism Web Console Guide](#).

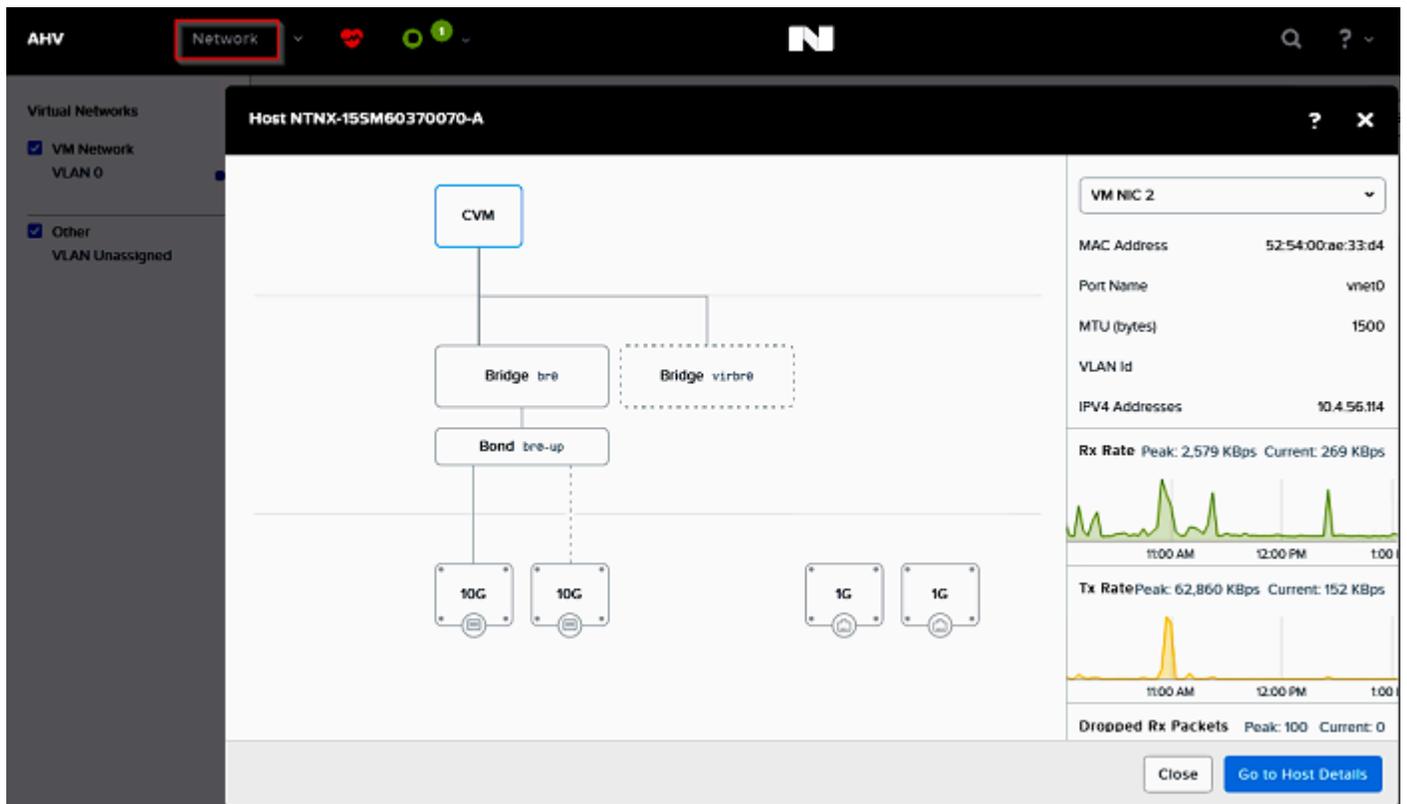


Figure 7: AHV Host Network Visualization

View AHV Host Network Configuration in the CLI

You can view Nutanix AHV network configuration in detail using the aCLI, AHV bash, and OVS commands shown in the appendix. The following sections outline basic administration tasks and the commands needed to review and validate a configuration.

Administrators can perform all management operations through the Prism web interface and APIs or through SSH access to the CVM.

Note: For better security and a single point of management, avoid connecting directly to the AHV hosts. All AHV host operations can be performed from the CVM by connecting to 192.168.5.1, the internal management address of the AHV host.

View Physical NIC Status from the CVM

To verify the names, speed, and connectivity status of all AHV host interfaces, use the `manage_ovs show_uplinks` command.

```
nutanix@CVM$ manage_ovs --bridge_name br0 show_uplinks
Uplink ports: br0-up
Uplink ifaces: eth3 eth2
nutanix@CVM$ manage_ovs show_interfaces
name mode link speed
eth0 1000 True 1000
eth1 1000 True 1000
eth2 10000 True 10000
eth3 10000 True 10000
```

View OVS Bridge and Bond Status from the CVM

Verify bridge and bond details in the AHV host using the `show_uplinks` command.

```
nutanix@CVM$ manage_ovs show_uplinks
Bridge: br0
Bond: br0-up
  bond_mode: active-backup
  interfaces: eth3 eth2
  lacp: off
  lacp-fallback: True
  lacp_speed: off
Bridge: br1
Bond: br1-up
  bond_mode: active-backup
  interfaces: eth1 eth0
  lacp: off
  lacp-fallback: True
  lacp_speed: off
```

View VM Network Configuration from a CVM Using the aCLI

Connect to any CVM in the Nutanix cluster to launch the aCLI and view cluster-wide VM network details.

```
nutanix@CVM$ acli
<acropolis> net.list
Network name      Network UUID      Type  Identifier
Production        ea8468ec-c1ca-4220-bc51-714483c6a266  VLAN  27
vlan.0            a1850d8a-a4e0-4dc9-b247-1849ec97b1ba  VLAN  0
<acropolis> net.list_vms vlan.0
VM UUID          VM name          MAC address
7956152a-ce08-468f-89a7-e377040d5310  VM1             52:54:00:db:2d:11
47c3a7a2-a7be-43e4-8ebf-c52c3b26c738  VM2             52:54:00:be:ad:bc
501188a6-faa7-4be0-9735-0e38a419a115  VM3             52:54:00:0c:15:35
```

Prism Uplink Configuration and Virtual Switches

From Nutanix AOS 5.19 on, you can manage interfaces and load balancing for bridges and bonds from the Nutanix Prism web interface using virtual switches. This method automatically changes all hosts in the cluster and performs the appropriate maintenance mode and VM migration. You don't need a manual maintenance mode entrance or exit, which saves configuration time. Refer to the [AHV Host Network Management documentation](#) for complete instructions and more information on virtual switches.

Nutanix recommends using the Prism web interface exclusively for all network management from 5.19 on. For AOS versions from 5.11 through 5.18, use the Prism uplink configuration to configure networking for systems with a single bridge and bond. For versions prior to 5.11 or for versions from 5.11 through 5.18 in systems with multiple bridges and bonds, use the CLI to modify the network configuration instead. Refer to the appendix for the relevant CLI commands.

Don't use any `manage_ovs` commands to make changes once you've used the Prism virtual switch or uplink configuration; the virtual switch automatically reverts changes made using `manage_ovs`.

Production Network Changes

Note: Exercise caution when you make changes that impact the network connectivity of Nutanix nodes.

When you use the CLI, we strongly recommend that you perform changes on one node (AHV host and CVM) at a time after you ensure that the [cluster can tolerate a single-node outage](#). To prevent network and storage disruption, place the AHV host and CVM of each node in maintenance mode before you make CLI network changes. While in maintenance mode, the system migrates VMs off the AHV host and directs storage services to another CVM.

Follow the steps in this section on one node at a time to make CLI network changes to a Nutanix cluster that is connected to a production network.

- Use SSH to connect to the first CVM you want to update. Check its name and IP address to make sure you're connected to the correct CVM. Verify failure

tolerance and don't proceed if the cluster can't tolerate at least one node failure.

- Verify that the target AHV host can enter maintenance mode.

```
nutanix@CVM$ acli host.enter_maintenance_mode_check <host ip>
```

- Put the AHV host in maintenance mode.

```
nutanix@CVM$ acli host.enter_maintenance_mode <host ip>
```

- Find the <host ID> in the output of the command `ncli host list`.

```
nutanix@CVM$ ncli host list
  Id                               : 00058977-c18c-af17-0000-000000006f89::2872 #-
  "2872" is the host ID
  Uuid                              : ddc9d93b-68e0-4220-85f9-63b73d08f0ff
  ...
```

- Enable maintenance mode for the CVM on the target AHV host. You can skip this step if the CVM services aren't running or the cluster state is stopped.

```
nutanix@CVM$ ncli host edit id=<host ID> enable-maintenance-mode=true
```

- Because network changes can disrupt host connectivity, use IPMI to connect to the host console and perform the desired network configuration changes. Once you have changed the configuration, ping the default gateway and another Nutanix node to verify connectivity.

- After all tests have completed successfully, remove the CVM and AHV host from maintenance mode.

- From a different CVM, run the following command to take the affected CVM out of maintenance mode:

```
nutanix@cvm$ ncli host edit id=<host ID> enable-maintenance-mode=false
```

- Exit host maintenance mode to restore VM locality, migrating VMs back to their original AHV host.

```
nutanix@cvm$ acli host.exit_maintenance_mode <host ip>
```

Move to the next node in the Nutanix cluster and repeat these steps to enter maintenance mode, make the desired changes, and exit maintenance mode. Repeat this process until you have made the changes on all hosts in the cluster.

OVS Command Line Configuration

To view the OVS configuration from the CVM command line, use the AHV-specific `manage_ovs` command. To run a single view command on every Nutanix CVM in a cluster, use the `allssh` shortcut described in the AHV Command Line appendix.

Note: In a production environment, we only recommend using the `allssh` shortcut to view information. Don't use the `allssh` shortcut to make changes in a production environment. When you make network changes, only use the `allssh` shortcut in a nonproduction or staging environment.

Note: The order in which flags and actions pass to `manage_ovs` is critical. The flag must come before the action. Any flag passed after an action isn't parsed.

```
nutanix@CVM$ manage_ovs --helpshort
USAGE: manage_ovs [flags] <action>
```

To list all physical interfaces on all nodes, use the `show_interfaces` command. The `show_uplinks` command returns the details of a bonded adapter for a single bridge.

```
nutanix@CVM$ allssh "manage_ovs show_interfaces"
nutanix@CVM$ allssh "manage_ovs --bridge_name <bridge> show_uplinks"
```

The `update_uplinks` command configures a comma-separated list of interfaces into a single uplink bond in the specified bridge. If the bond doesn't have at least one interface with a physical connection, the `manage_ovs` command issues a warning and exits without configuring the bond. To avoid this error and provision members of the bond even if they are not connected, use the `require_link=false` flag.

Note: If you do not enter a `bridge_name`, the command runs on the default bridge, `br0`.

```
nutanix@CVM$ manage_ovs --bridge_name <bridge> --interfaces <interfaces> --
bond_mode <bond-mode> update_uplinks
nutanix@CVM$ manage_ovs --bridge_name <bridge> --interfaces <interfaces> --
bond_mode <bond-mode> --require_link=false update_uplinks
```

The `manage_ovs update_uplinks` command deletes an existing bond and creates it with the new parameters when you need to change the bond members or load balancing algorithm. Unless you specify the correct bond mode parameter, using `manage_ovs` to update uplinks deletes the bond, then recreates it with the default load balancing configuration. If you use active-

backup load balancing, `update_uplinks` can cause a short network interruption. If you use `balance-slb` or `balance-tcp` (LACP) load balancing and don't specify the correct bond mode parameter, `update_uplinks` resets the configuration to `active-backup`. At this point, the host stops responding to keepalives and network links that rely on LACP go down.

Note: Don't use the command `allssh manage_ovs update_uplinks` in a production environment. If you haven't set the correct flags and parameters, this command can cause a cluster outage.

Only use the `manage_ovs` command with guidance from Nutanix Support.

With AHV clusters that run versions 5.10.x prior to 5.10.4, using `manage_ovs` to make configuration changes on an OVS bridge configured with a single uplink might cause a network loop. If you configured the AHV node with a single interface in the bridge, upgrade AOS to 5.10.4 or later before you make any changes. If you have a single interface in the bridge, engage [Nutanix Support](#) if you can't upgrade AOS and must change the bond configuration.

Compute-Only Node Network Configuration

The `manage_ovs` command runs from a CVM and makes network changes to the local AHV host where it runs. Because Nutanix AHV compute-only nodes don't run a CVM, you use a slightly different process to configure their networks.

- Follow the steps in the Production Network Changes section to put the target AHV host in maintenance mode.
- Run the `manage_ovs` command from any other CVM in the cluster with the `--host` flag.

```
nutanix@CVM$ manage_ovs --host <compute-only-node-IP> --bridge_name <br-name> --bond_name <bond-name> --bond_mode <bond-mode> --interfaces 10g update_uplinks
```

- Exit maintenance mode after you complete the network changes, then repeat these steps on the next compute-only node.

5. AHV Networking Best Practices

The main best practice for Nutanix AHV networking is to keep things simple. The recommended networking configuration—two 10 Gbps or faster adapters using active-backup—provides a highly available environment that performs well with minimal host and switch configuration. Nutanix CVMs and AHV hosts communicate in the untagged VLAN, and tagged VLANs serve guest VM traffic. Use this basic configuration unless there is a compelling business requirement for advanced configuration.

Note: When a GUI configuration exists to achieve the desired configuration, use it. Don't use the CLI unless you can't accomplish your goal in the GUI.

Open vSwitch Bridge and Bond Recommendations

This section addresses advanced bridge and bond configuration scenarios that you can manage in the GUI. For information on CLI commands for end-of-life versions, see the AHV Command Line Tutorial section in the appendix.

Identify the scenario that best matches your desired use case and follow the instructions in the corresponding subsection. With slight modifications, you can use the commands from the second scenario to create as many bridges and bonds as you need. The final two scenarios are variations of the second scenario. In the following sections, we use 10 GbE to refer to the fastest interfaces on the system. If your system uses 25, 40, or 100 GbE interfaces, use the actual speed of the interfaces instead of 10 GbE.

Note: Nutanix recommends that each bond always have at least two interfaces for high availability.

Table: Bridge and Bond Use Cases

Bridge and Bond Scenario	Use Case
2 × 10 GbE (no 1 GbE)	Recommended configuration for easy setup. Use when you can send all CVM and guest VM traffic over the same pair of 10 GbE adapters. Compatible with any load balancing algorithm.
2 × 10 GbE and 2 × 1 GbE separated	Use when you need an additional, separate pair of physical adapters for VM traffic that must be isolated to another adapter or physical switch. In this example we use 10 GbE and 1 GbE, but you can also use this configuration when all adapters are 10 GbE. Keep the CVM traffic on the fastest network. You can place guest VM traffic on either the 10 GbE network or the 1 GbE network. Compatible with any load balancing algorithm.
4 × 10 GbE (2 + 2) and 2 × 1 GbE separated	Use to physically separate CVM traffic such as storage and Nutanix Volumes from guest VM traffic while still providing 10 GbE connectivity for both traffic types. The four 10 GbE adapters are divided into two separate pairs. Compatible with any load balancing algorithm. This case is not illustrated in the following diagrams.
4 × 10 GbE combined and 2 × 1 GbE separated	Use to provide additional bandwidth and failover capacity to the CVM and guest VMs sharing four 10 GbE adapters in the same bond. We recommend using LACP with balance-tcp to take advantage of all adapters. This case is not illustrated in the following diagrams.

Keep the following recommendations in mind for all bond scenarios to prevent undesired behavior and maintain NIC compatibility:

- Don't mix NIC models from different vendors in the same bond.
- Don't mix NICs of different speeds in the same bond.

- Don't mix NICs with different drivers in the same bond.

Note: To verify which NIC driver you're using, run the following command on the AHV host, once for each NIC, replacing `<nic-name>` with the name:

```
ethtool -i <nic-name>
```

Scenario 1: 2 × 10 GbE

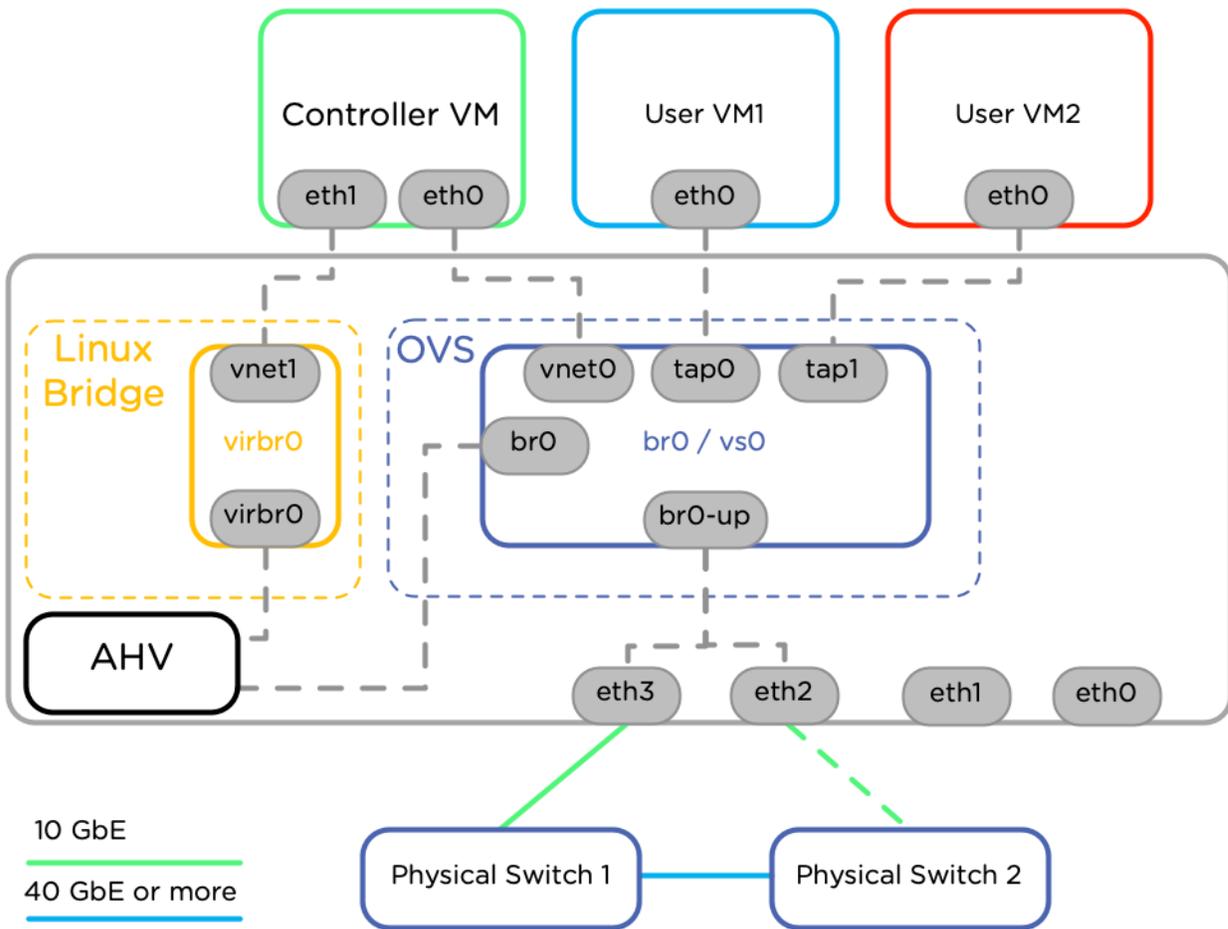


Figure 8: Network Connections for 2 × 10 GbE NICs

The most common network configuration is to use the 10 Gbps or faster interfaces in the default bond for all networking traffic. The CVM and all guest VMs use the 10 GbE interfaces. In this configuration, we don't use the 1 GbE interfaces. Note that this setup differs from the factory configuration because

we removed the 1 GbE interfaces from the OVS bond. For simplicity, we have not included the IPMI connection in these diagrams.

This scenario uses two physical upstream switches, and each 10 GbE interface in the bond plugs into a separate physical switch for high availability. In the bond, only one physical interface is active when using the default active-backup load balancing mode. Nutanix recommends using active-backup because it's easy to configure, works immediately after installation, and requires no upstream switch configuration. See the [Load Balancing in Bond Interfaces](#) section for more information and alternate configurations.

For all clusters with Nutanix AOS versions 5.19 and later, even with multiple bridges, use the [Prism Virtual Switch UI](#) instead of the CLI to create the configuration in the following diagram.

1 General 2 Uplink Configuration

Bond Type

Active-Backup

One adapter in the bond is active. Additional adapters act as backup until the active adapter fails. **Link Aggregation must not be configured on the connected switch.**

Select Hosts

All Hosts x

Select Uplink Ports

Show

- Connected and Unconnected Uplink Ports
- Only Connected Uplink Ports

On Switches (with LLDP)

All switches

Uplink Port Speeds

All Speeds

Select All | Clear All

Hosts	1G NICs	10G NICs
DRMPOC04-2	2 ports	2 ports
	<input type="checkbox"/> eth0	<input checked="" type="checkbox"/> eth2 ?
	<input type="checkbox"/> eth1	<input checked="" type="checkbox"/> eth3 ?
DRMPOC04-4	2 ports	2 ports
DRMPOC04-1	2 ports	2 ports

Figure 9: Virtual Switch Configuration for 2 x 10 GbE Active-Backup

Remove unused NICs from the default virtual switch, especially when they have different speeds. To do so, clear the check box beside each unused NIC in the virtual switch configuration. The virtual switch UI applies the desired configuration to all hosts in the cluster, automatically placing one host at a time in maintenance mode.

Note: Previous versions of this guide used the bond name `bond0` instead of `br0-up`. We recommend using `br0-up` because it identifies the associated bridge and the uplink function of this bond. The virtual switch UI creates bonds in the format `brX-up`.

To remove NICs from the default virtual switch in versions prior to 5.19, refer to the appendix.

Scenario 2: 2 × 10 GbE and 2 × 1 GbE Separated

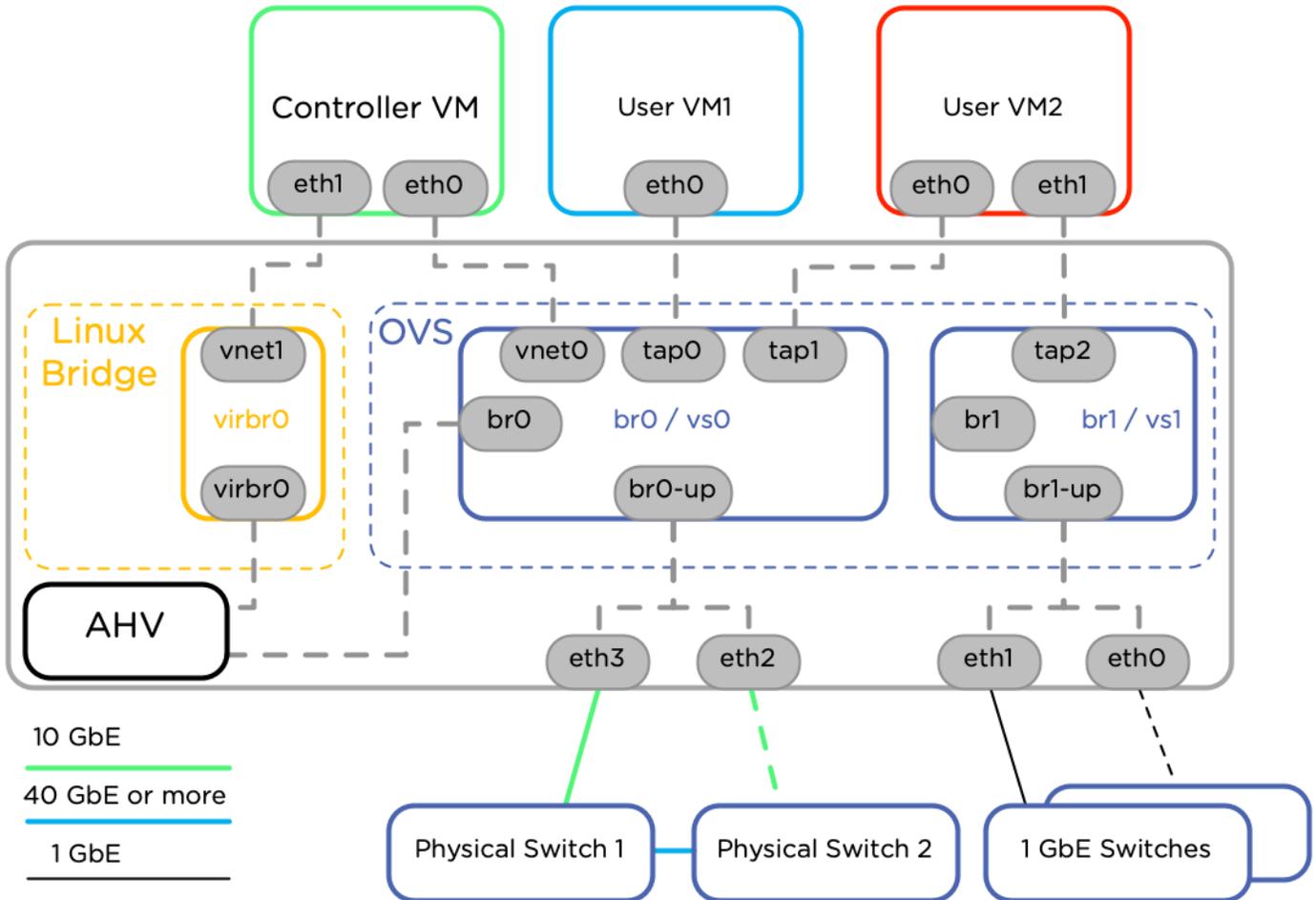


Figure 10: Network Connections for 2 × 10 GbE and 2 × 1 GbE NICs

If you want to use the 1 GbE physical interfaces, separate the 10 GbE and 1 GbE interfaces into different bridges and bonds to ensure that CVM traffic always traverses the fastest possible link.

Here, we've grouped the 10 GbE interfaces (eth2 and eth3) into br0-up and dedicated them to the CVM and VM1. We've grouped the 1 GbE interfaces into br1-up; only a second link on VM2 uses br1. Bonds br0-up and br1-up are added into br0 and br1, respectively.

In this configuration, the CVM and guest VMs use the 10 GbE interfaces on bridge br0. Bridge br1 is available for VMs that require physical network

separation from the CVM and VMs on br0. Devices eth0 and eth1 could alternatively plug into a different pair of upstream switches for further physical traffic separation, as shown. The second pair of physical adapters could be 10 GbE or faster instead of 1 GbE.

For all clusters with AOS versions 5.19 and later, even with multiple bridges, use the [Prism Virtual Switch UI](#) instead of the CLI to create the configuration in the following diagram and new virtual switches.

- 1 General
- 2 Uplink Configuration

Bond Type

Active-Backup

One adapter in the bond is active. Additional adapters act as backup until the active adapter fails. **Link Aggregation must not be configured on the connected switch.**

Select Hosts

All Hosts x

Select Uplink Ports

Show

- Connected and Unconnected Uplink Ports
- Only Connected Uplink Ports

On Switches (with LLDP)

All switches

Uplink Port Speeds

All Speeds

Select All | Clear All

Hosts	1G NICs	10G NICs
DRMPOC04-2	2 ports	2 ports
	<input checked="" type="checkbox"/> eth0	<input type="checkbox"/> eth2 ?
	<input checked="" type="checkbox"/> eth1	<input type="checkbox"/> eth3 ?
DRMPOC04-4	2 ports	2 ports
	<input checked="" type="checkbox"/> eth0	<input type="checkbox"/> eth2 ?

Figure 11: Create New Virtual Switch with 1 GbE Adapters

The virtual switch UI applies the desired configuration to all hosts in the cluster one host at a time, automatically initiating maintenance mode.

To create additional virtual switches in versions prior to AOS 5.19, refer to the appendix.

Creating Networks

In Nutanix AOS versions 5.19 and later, you can create networks on any virtual switch and bridge using a dropdown menu in the Prism UI.

Create Network

Network Name

NewNetwork

Virtual Switch

vs0

✓ vs1

VLAN ID ?

100

Enable IP address management

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

Figure 12: Create Network on Additional Virtual Switch

To create networks on additional virtual switches in AOS versions prior to 5.19, refer to the appendix.

Load Balancing in Bond Interfaces

Nutanix AHV hosts use a bond containing multiple physical interfaces that each connect to a physical switch. To build a fault-tolerant network connection between the AHV host and the rest of the network, connect each physical interface in a bond to a separate physical switch.

A bond distributes traffic between multiple physical interfaces according to the bond mode. In the following table, the throughput rates assume 2 × 10 GbE adapters and simplex speed.

Table: Load Balancing Use Cases

Bond Mode	Use Case	Maximum VM NIC Throughput	Maximum Host Throughput
active-backup	Recommended. Default configuration, transmits all traffic over a single active adapter.	10 Gbps	10 Gbps
balance-slb	Has caveats for multicast traffic. Increases host bandwidth utilization beyond a single 10 GbE adapter. Places each VM NIC on a single adapter at a time. Don't use with link aggregation such as LACP.	10 Gbps	20 Gbps
LACP and balance-tcp	LACP and link aggregation required. Increases host and VM bandwidth utilization beyond a single 10 GbE adapter by balancing VM NIC TCP and UDP sessions among adapters. Also used when network switches require LACP negotiation.	20 Gbps	20 Gbps

Active-Backup

The recommended and default bond mode is active-backup, where one interface in the bond is randomly selected to carry traffic when the AHV host starts. The system only uses other interfaces in the bond when the active link fails. Active-backup is the simplest bond mode, easily allowing connections to multiple upstream switches without additional switch configuration. The limitation is that traffic from all VMs simultaneously uses only the single active link in the bond. All backup links remain unused until the active link fails. In a system with dual 10 GbE adapters, the maximum throughput of all VMs running on a Nutanix node is limited to 10 Gbps, or the speed of a single link.

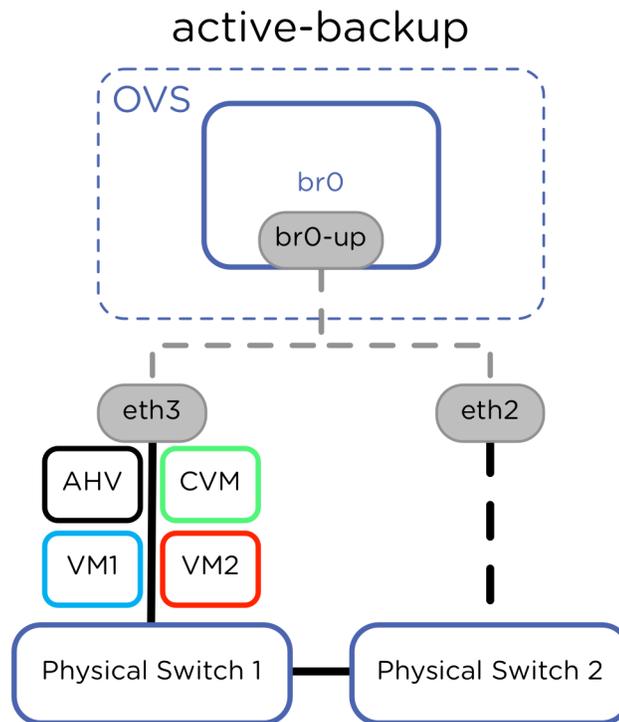


Figure 13: Active-Backup Fault Tolerance

For all clusters with versions 5.19 and later, even with multiple bridges, use the [Prism Virtual Switch UI](#) instead of the CLI to select active-backup mode (configuration shown in the following figure).

Bond Type

Active-Backup



One adapter in the bond is active. Additional adapters act as backup until the active adapter fails. **Link Aggregation must not be configured on the connected switch.**

Figure 14: Virtual Switch Configuration for Active-Backup

For information on configuring the active-backup bond mode in AOS versions prior to 5.19, refer to the appendix.

Balance-SLB

Nutanix doesn't recommend balance-slb because of the multicast traffic caveats noted in this section. To combine the bandwidth of multiple links, consider using link aggregation with LACP and balance-tcp instead of balance-slb. Don't use balance-slb unless you verify that the multicast limitations described here aren't present in your network.

Don't use IGMP snooping on physical switches connected to Nutanix servers that use balance-slb. With balance-slb, the virtual switch forwards inbound multicast traffic on only one active adapter and discards multicast traffic from other adapters. Physical switches with IGMP snooping may discard traffic to the active adapter and only send it to the backup adapters. This mismatch leads to unpredictable multicast traffic behavior. Disable IGMP snooping or configure static IGMP groups for all switch ports connected to Nutanix servers using balance-slb.

Note: IGMP snooping is often enabled by default on physical switches.

The balance-slb bond mode in OVS takes advantage of all links in a bond and uses measured traffic load to rebalance VM traffic from highly used to less-used interfaces. When the configurable bond-rebalance interval expires, OVS uses the measured load for each interface and the load for each source MAC hash to spread traffic evenly among links in the bond. Traffic from some source MAC hashes may move to a less active link to more evenly balance bond member utilization. Perfectly even balancing may not always be possible, depending on the number of source MAC hashes and their stream sizes. With balance-slb,

the virtual switch attempts to maintain less than 10 percent difference in link utilization among members in a bond.

Each VM NIC uses only one bond member interface at a time, but a hashing algorithm distributes multiple VM NICs (multiple source MAC addresses) across bond member interfaces. As a result, it's possible for a Nutanix AHV node with two 10 GbE interfaces to use up to 20 Gbps of network throughput, while an individual VM may have a maximum throughput of 10 Gbps, the speed of a single physical interface.

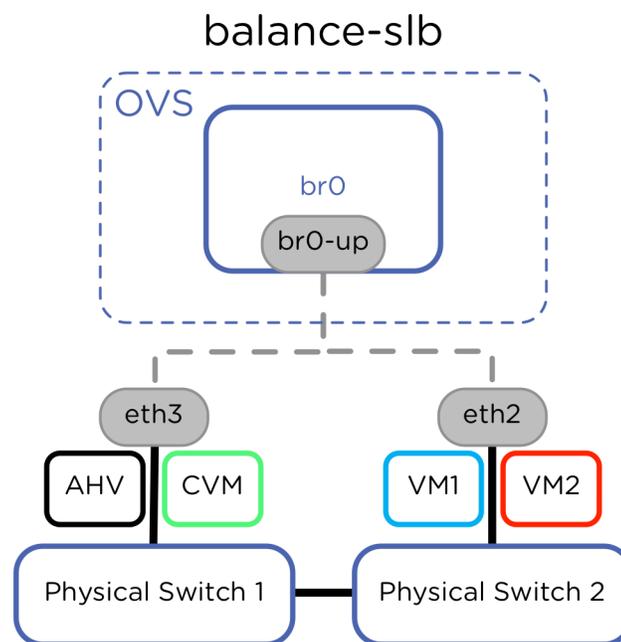


Figure 15: Balance-SLB Load Balancing

For all clusters with versions 5.19 and later, even with multiple bridges, use the [Prism Virtual Switch UI](#) instead of the CLI to configure balance-slb mode by selecting Active-Active with MAC pinning (shown in the following figure).

Bond Type

Active-Active with MAC pinning

Active-Active with MAC pinning is not recommended due to its operational challenges in practice. [Read more here.](#)

Figure 16: Virtual Switch Configuration for Balance-SLB

Note: Don't use link aggregation technologies such as LACP with balance-slb. The balance-slb algorithm assumes that upstream switch links are independent layer 2 interfaces and handles broadcast, unknown, and multicast (BUM) traffic accordingly, selectively listening for this traffic on only a single active adapter in the bond.

To configure the balance-slb bond mode in AOS clusters prior to AOS 5.19, refer to the appendix.

LACP and Link Aggregation

Link aggregation is required to take full advantage of the bandwidth provided by multiple links. In OVS it is accomplished through dynamic link aggregation with LACP and load balancing using balance-tcp.

Nutanix and OVS require dynamic link aggregation with LACP instead of static link aggregation on the physical switch. Do not use static link aggregation such as EtherChannel with AHV.

Nutanix recommends that you enable LACP on the AHV host with fallback to active-backup, then configure the connected upstream switches. Different switch vendors may refer to link aggregation as port channel or LAG. Using multiple upstream switches may require additional configuration, such as a multichassis link aggregation group (MLAG) or virtual PortChannel (vPC). Configure switches to fall back to active-backup mode in case LACP negotiation fails (sometimes called fallback or no suspend-individual). This switch setting assists with node imaging and initial configuration where LACP may not yet be available on the host.

With link aggregation negotiated by LACP, multiple links to separate physical switches appear as a single layer 2 link. A traffic-hashing algorithm such as balance-tcp can split traffic between multiple links in an active-active fashion. Because the uplinks appear as a single layer 2 link, the algorithm can balance traffic among bond members without any regard for switch MAC address tables. Nutanix recommends using balance-tcp when you have LACP and link aggregation configured, because each TCP or UDP stream from a single VM can potentially use a different uplink in this configuration. The balance-tcp algorithm hashes traffic streams by source IP, destination IP, source port, and destination port. With link aggregation, LACP, and balance-tcp, a single VM with multiple TCP or UDP streams could use up to 20 Gbps of bandwidth in an AHV node with two 10 GbE adapters.

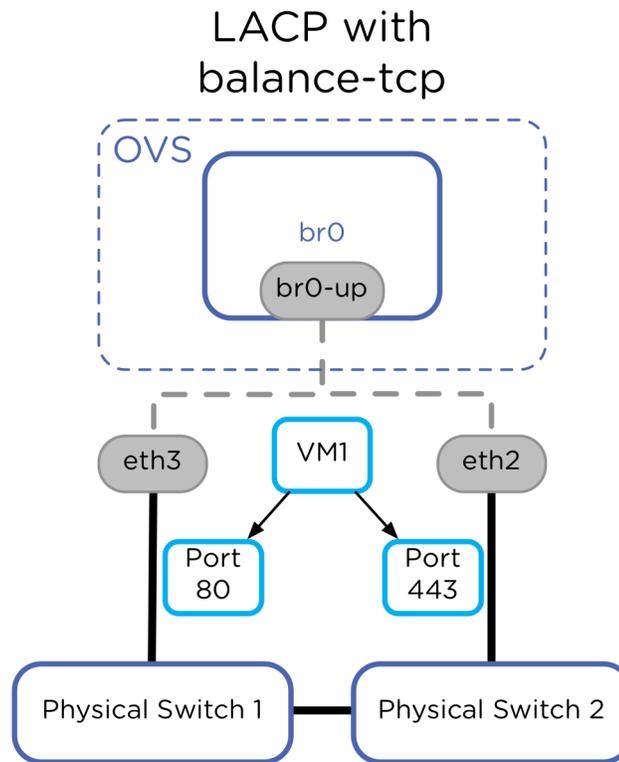


Figure 17: LACP and Balance-TCP Load Balancing

For all clusters with versions 5.19 and later, even with multiple bridges, use the [Prism Virtual Switch UI](#) instead of the CLI to configure balance-tcp with LACP mode by selecting Active-Active (shown in the following figure).

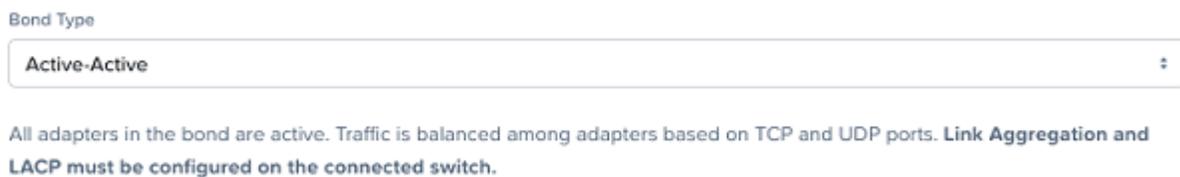


Figure 18: Virtual Switch Configuration for Balance-TCP with LACP

The Prism GUI Active-Active mode configures all AHV hosts with the **fast** setting for LACP speed, causing the AHV host to request LACP control packets at the rate of one per second from the physical switch. In addition, the Prism GUI configuration sets LACP fallback to active-backup on all AHV hosts. You

can't modify these default settings in AHV after you've configured them from the GUI, even by using the CLI.

Note: Upstream physical switch LACP settings such as timers should match the AHV host setting of fast, or 1 second, for configuration consistency and rapid failure detection.

On most physical switches, the default LACP speed configuration is slow (sometimes called normal), or 30 seconds. This value on the physical switch determines how frequently the physical switch wants the AHV host to send LACPDUs. The fast setting (1 second) in AHV requests LACPDUs be sent by the connected physical switch every second, which helps you detect interface failures more quickly. Failure to receive three LACPDUs—in other words, after 3 seconds with the fast setting—shuts down the link in the bond. Nutanix recommends setting `lacp-time` to fast on the physical switch to decrease link failure detection time from 90 seconds to 3 seconds.

For information on LACP configuration in versions prior to AOS 5.19, refer to the appendix.

Note: If you've upgraded from an AOS version prior to 5.19 and were previously using LACP, your AHV hosts may still have the LACP slow setting. Refer to the LACP CLI verification in the appendix to confirm that all AHV hosts are set to the fast LACP setting.

Disable LACP on the Host

To safely turn off the LACP configuration so you can use another load balancing algorithm, configure the connected physical switches to pass traffic even when LACP negotiation fails. Consult the switch vendor's documentation for LACP fallback or suspend behavior. When the physical switches are ready, change the virtual switch setting to automatically initiate the configuration across all hosts in the cluster. The virtual switch configuration places each host into maintenance mode, performs the configuration, then resumes regular activity.

To disable LACP in AOS versions prior to 5.19, refer to the appendix.

Storage Traffic Between CVMs

Using the default active-backup or any other OVS load balancing method, you can't use Prism to select the active adapter for the CVM such that it persists between host reboots. When multiple uplinks from the AHV host connect to multiple switches, ensure that there's enough bandwidth to support Nutanix

CVM replication traffic between nodes. Nutanix recommends redundant 40 Gbps or faster connections between switches. A leaf-spine configuration or direct inter-switch link can satisfy this recommendation. Review the [Physical Networking best practice guide](#) for more information.

In AOS version 5.19 and later, AHV hosts accept a manual OVS configuration to set the preferred active adapter. Run the following command on each AHV host to persist the active adapter between reboots. Replace ethX with the desired active adapter. This configuration takes effect immediately.

Note: In a production environment, Nutanix strongly recommends that you make changes on one node at a time, after you verify that the cluster can tolerate a node failure. Follow the steps in the Production Network Changes section when you make changes.

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up other_config:bond-primary=ethX"
```

Run the following command to confirm the configuration is successful.

```
nutanix@CVM$ ssh root@192.168.5.1 ovs-appctl bond/show
```

The active adapter is configured if you see the following text: **active-backup primary: eth2**

Exit maintenance mode and repeat these steps on every node to configure the entire cluster.

VLANs for AHV Hosts and CVMs

The recommended VLAN configuration is to place the CVM and AHV host in the untagged VLAN (sometimes called the native VLAN), as shown in the following figure. Neither the CVM nor the AHV host requires special configuration with this option. Configure the switch to allow tagged VLANs for guest VM networks to the AHV host using standard 802.1Q VLAN tags. Also, configure the switch to send and receive traffic for the CVM and AHV host's VLAN as untagged. Choose any VLAN on the switch other than 1 as the native untagged VLAN on ports facing AHV hosts.

Note: All CVMs and hypervisor hosts must be on the same subnet and broadcast domain. No systems other than the CVMs and hypervisor hosts should be on this network, which should be isolated and protected.

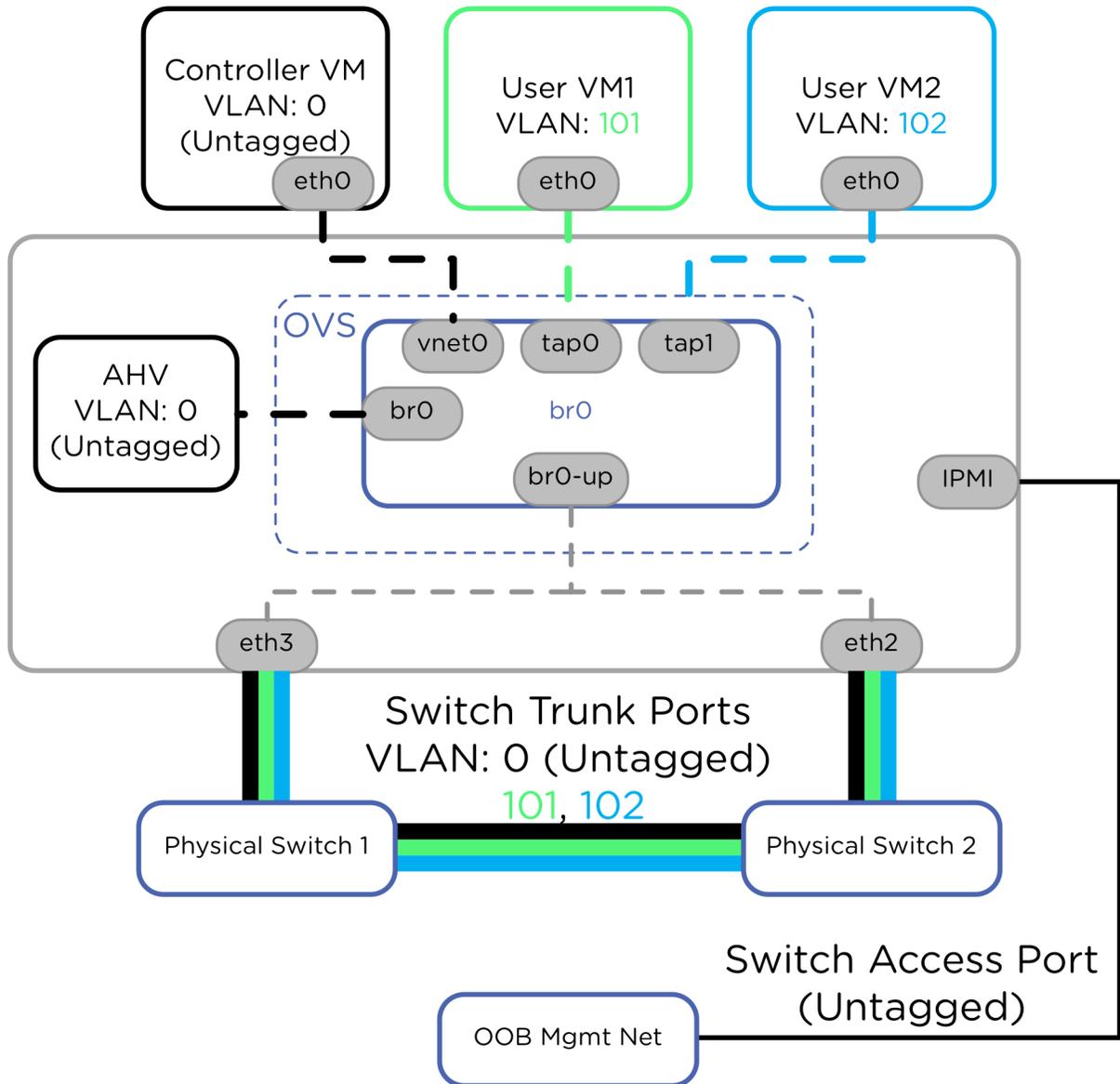


Figure 19: Default Untagged VLAN for CVM and AHV Host

The configuration depicted in the previous figure works well for situations where the switch administrator can set the CVM and AHV VLAN to untagged. However, if you don't want to send untagged traffic to the AHV host and CVM or if your security policy doesn't allow this configuration, you can add a VLAN tag to the host and the CVM with the procedure shown in the next image.

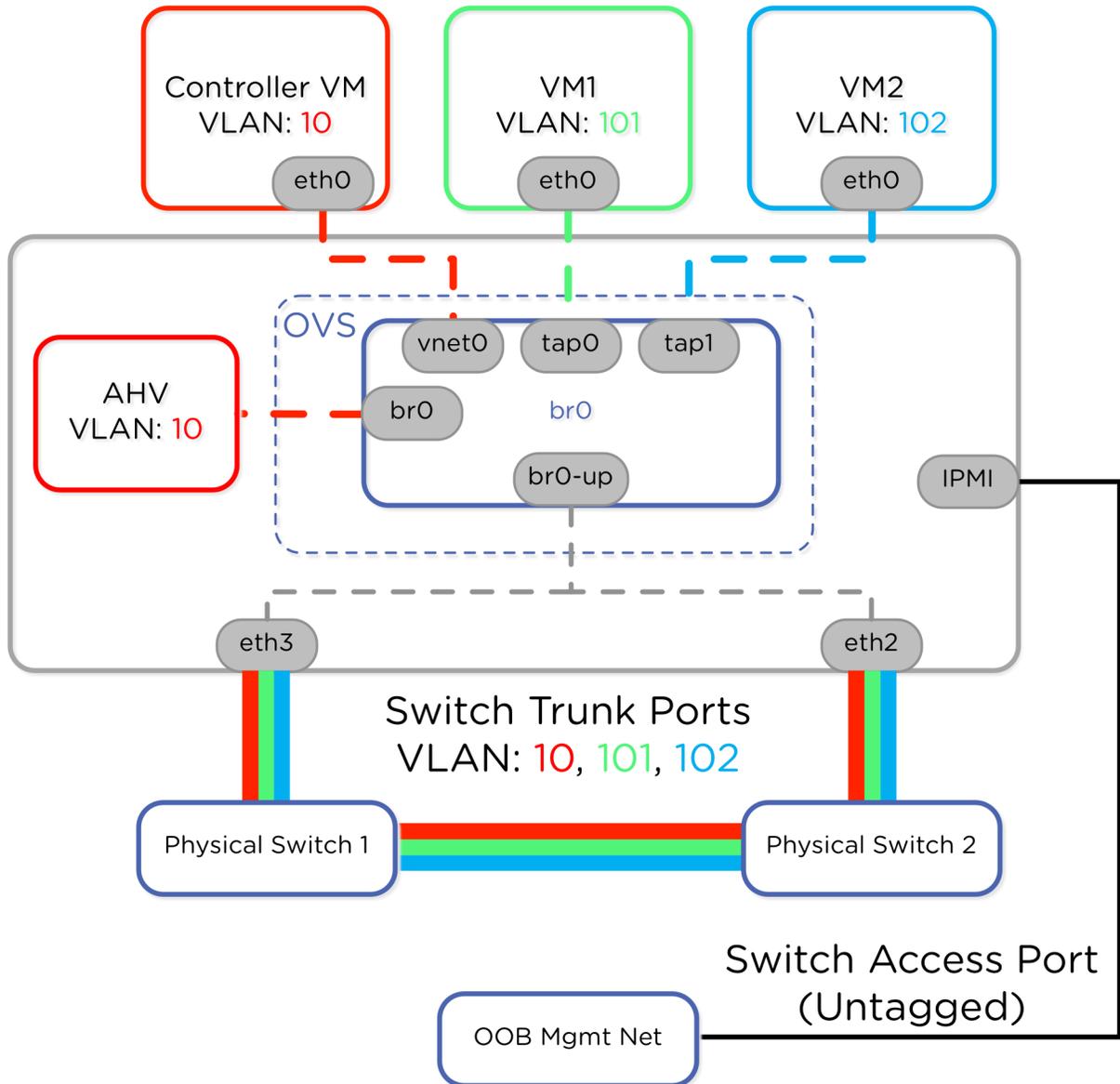


Figure 20: Tagged VLAN for CVM and AHV Host

Note: Ensure that IPMI console access is available for recovery before you start this configuration.

Note: In a production environment, Nutanix strongly recommends that you make changes on one node at a time, after you verify that the cluster can tolerate a node failure. Follow the steps in the Production Network Changes section when you make changes.

- After you enter maintenance mode on the target AHV host and CVM, configure VLAN tags on the AHV host:

Note: Use port br0 in the following command. Port br0 is the internal port assigned to AHV on bridge br0. Do not use any value other than br0.

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0 tag=10"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl list port br0"
```

- Configure VLAN tags for the CVM:

```
nutanix@CVM$ change_cvm_vlan 10
```

Exit maintenance mode and repeat these steps on every node to configure the entire cluster.

Removing VLAN Configuration

Use the following steps to remove VLAN tags and revert to the default untagged VLAN configuration.

- After you enter maintenance mode on the target AHV host and CVM, run the following command for the CVM:

```
nutanix@CVM$ change_cvm_vlan 0
```

- Run the following command for the AHV host:

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0 tag=0"
```

Exit maintenance mode and repeat the previous steps on every node to configure the entire cluster.

VLAN for Guest VMs

The VLAN for a VM is assigned using the VM network. In Nutanix Prism, br0 and vs0 are the default bridge and virtual switch for new networks. Create networks for VMs in bridge br0 in the Prism UI or the aCLI. In Nutanix AOS versions 5.19 and later, use the Prism UI to create networks in any bridge and virtual switch from a dropdown menu. You can also change the network of a VM from the Prism UI.

In AOS versions prior to 5.19, if you need to create networks in additional bridges or modify the NIC of a VM, refer to the appendix.

VM NIC VLAN Modes

VM NICs on AHV can operate in three modes:

1. Access
2. Trunked
3. Direct

Access mode is the default for VM NICs, where a single VLAN travels to and from the VM as untagged but is encapsulated with the appropriate VLAN tag on the physical NIC. VM NICs in trunked mode allow multiple tagged VLANs and a single untagged VLAN on a single NIC for VLAN-aware VMs. You can only add a NIC in trunked mode using the aCLI; you can't distinguish between access and trunked NIC modes in the Prism UI. Direct-mode NICs connect to brX and bypass the bridge chain; don't use them unless advised by Nutanix Support to do so.

Run the following command on any CVM in the cluster to add a new trunked NIC:

```
nutanix@CVM~$ acli vm.nic_create <vm name> network=<network name>
trunked_networks=<comma separated list of allowed VLAN IDs> vlan_mode=kTrunked
```

The native VLAN for the trunked NIC is the VLAN assigned to the network specified in the network parameter. Additional tagged VLANs are designated by the `trunked_networks` parameter.

Run the following command on any CVM in the cluster to verify the VM NIC mode:

```
nutanix@CVM~$ acli vm.get <vm name>
```

Sample output:

```
nutanix@CVM~$ acli vm.get testvm
testvm {
  config {
    ...
    nic_list {
      ip_address: "X.X.X.X"
      mac_addr: "50:6b:8d:8a:46:f7"
      network_name: "network"
      network_type: "kNativeNetwork"
      network_uuid: "6d8f54bb-4b96-4f3c-a844-63ea477c27e1"
      trunked_networks: 3 <--- list of allowed VLANs
      trunked_networks: 4
      trunked_networks: 5
      type: "kNormalNic"
      uuid: "9158d7da-8a8a-44c8-a23a-fe88aa5f33b0"
```

```

        vlan_mode: "kTrunked" <--- mode
    }
    ..:
}
}

```

To change the VM NIC's mode from Access to Trunked, use the command `acli vm.get <vm name>` to find its MAC address. Using this MAC address, run the following command on any CVM in the cluster:

```

nutanix@CVM~$ acli vm.nic_update <vm name> <vm nic mac address>
trunked_networks=<comma separated list of allowed VLAN IDs>
update_vlan_trunk_info=true

```

The `update_vlan_trunk_info=true` parameter is mandatory. If you don't specify this parameter, the command appears to run successfully but the `trunked_networks` setting doesn't change.

To change the mode of a VM NIC from trunked to access, find its MAC address in the output from the `acli vm.get <vm name>` command and run the following command on any CVM in the cluster:

```

nutanix@CVM~$ acli vm.nic_update <vm name> <vm nic mac address> vlan_mode=kAccess
update_vlan_trunk_info=true

```

IP Address Management Best Practices

To avoid duplicate IP addresses in VLAN networks, work with your network team to reserve a range of IP addresses for guest VMs before you enable the IPAM feature.

In multicluster deployments, Nutanix recommends reserving unique IP address ranges in each cluster to simplify provisioning and prevent IP address conflicts between clusters. In scenarios where IP address network ranges overlap in multiple Nutanix AHV clusters, configure each cluster with a unique IP pool in that network range.

In a hypothetical scenario where cluster A and cluster B share the network 172.30.30.0/24, Cluster A uses IP pool 172.30.30.2-125 and Cluster B uses 172.30.30.130-253. If you want the option to break this fictional /24 network into two /25 networks in the future, you can reserve several addresses: 172.30.30.0/25 and 172.30.30.128/25. You can reserve the first IP address in the network ranges, .1 and .129, for the default gateway. Also reserve the IP

addresses .126 and .254 for the Nutanix AHV IPAM process, which takes the last address by default.

In scenarios such as multitenancy or service provider networks where you need to have IP address overlap in the same AHV cluster, consider using Flow Virtual Networking and virtual private clouds instead of AHV VLAN networks.

MAC Address Management Best Practices

To avoid duplicate MAC addresses in VLAN networks, Nutanix recommends that you assign each AHV cluster a set of unique VLANs for guest VMs and that these VLANs don't overlap with other AHV clusters. Nutanix AHV doesn't guarantee unique MAC address assignment by default between Nutanix clusters. Assigning unique VLAN ranges for each cluster reduces the risk of MAC address conflict and follows the general best practice of maintaining small layer 2 broadcast domains with limited numbers of endpoints.

In designs where multiple AHV clusters must share the same VLANs, or when VM MAC addresses must be globally unique even among multiple AHV clusters, Nutanix recommends setting a unique MAC address prefix per cluster.

Consider the following example design scenario with multiple sites and clusters using locally administered MAC addresses.

Define a unique prefix for Site 1, Cluster 1 such as: **02:01:01**

- 02 - Designates a locally administered unicast address. This value could be **x2**, **x6**, **xA**, or **xE**, where x is any valid hexadecimal value.
- 01 - Identifies the site number if desired. Can also be any valid hexadecimal value.
- 01 - Identifies the AHV cluster within the site if desired. Can also be any valid hexadecimal value.

In an example deployment with three sites and five clusters per site, use the following prefixes. The value **xx:xx:xx** represents the usable hexadecimal range for endpoints in each AHV cluster.

- VMs at Site 1 Cluster 1 are assigned MAC addresses **02:01:01:xx:xx:xx**.

- VMs at Site 3 Cluster 5 are assigned MAC addresses 02:03:05:xx:xx:xx.

Contact your Nutanix account and support teams for access to unique MAC prefixes for AOS version 6.6 or earlier if you need them in your environment.

CVM Network Segmentation

The optional backplane LAN creates a dedicated interface in a separate VLAN on all CVMs and AHV hosts in the cluster for exchanging storage replication traffic. The backplane network shares the same physical adapters on bridge br0 by default but uses a different nonroutable VLAN. From Nutanix AOS 5.11.1 on, you can create the backplane network in a new bridge (such as br1). If you place the backplane network on a new bridge, ensure that this bridge has redundant network adapters with at least 10 Gbps throughput and use a fault-tolerant load balancing algorithm.

Use the backplane network only if you need to separate CVM management traffic (such as Prism) from storage replication traffic. The section [Securing Traffic Through Network Segmentation](#) in the Nutanix Security Guide includes diagrams and configuration instructions.

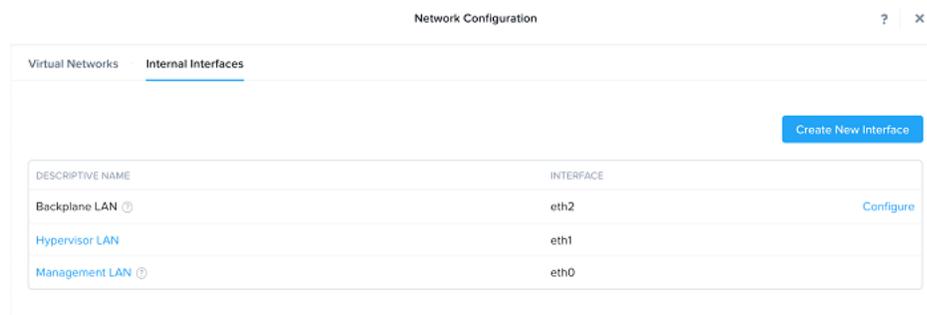


Figure 21: Prism UI CVM Network Interfaces

You can also separate iSCSI and disaster recovery traffic onto dedicated virtual network interfaces on the CVMs using the Create New Interface dialog. The new virtual network interface can use a shared or dedicated bridge. Ensure that the selected bridge uses multiple redundant uplinks.

Jumbo Frames

The Nutanix CVM uses the standard Ethernet MTU (maximum transmission unit) of 1,500 bytes for all the network interfaces by default. The standard 1,500-byte MTU delivers excellent performance and stability. Nutanix doesn't support configuring the MTU on a CVM's network interfaces to higher values.

You can enable jumbo frames (MTU of 9,000 bytes) on the physical network interfaces of AHV, ESXi, or Hyper-V hosts and guest VMs if the applications on your VMs require them. If you choose to use jumbo frames on hypervisor hosts, enable them end to end in the desired network and consider both the physical and virtual network infrastructure affected by the change.

When using Flow Virtual Networking, refer to the MTU recommendations in the [Flow Virtual Networking guide](#).

6. Conclusion

Nutanix recommends using the default AHV networking settings, configured through the Prism GUI, for most Nutanix deployments. However, if your requirements demand specific configuration outside the defaults, this networking guide provides detailed configuration examples that can help.

Administrators can use the virtual switch UI to configure networking features on all hosts. VLAN trunking for guest VMs allows a single VM NIC to pass traffic on multiple VLANs for network-intensive applications. You can apply VLAN tags to the AHV host and CVM in situations that require all traffic to be tagged. Grouping host adapters in different ways can provide physical traffic isolation or allow load balancing and link aggregation to provide maximum throughput and redundancy for VMs and hosts.

With these networking techniques, administrators can configure a Nutanix system with AHV to meet the requirements of any VM or application, no matter how demanding.

For feedback or questions, contact us using the [Nutanix NEXT Community forums](#).

7. Appendix

AHV Networking Terminology

Table: Networking Terminology Matrix

AHV Term	VMware Term	Microsoft Hyper-V or SCVMM Term
Bridge, virtual switch	vSwitch, Distributed Virtual Switch	Virtual switch, logical switch
Bond	NIC team	Team or uplink port profile
Port or tap	Port	N/A
Network	Port group	VLAN tag or logical network
Uplink	pNIC or vmnic	Physical NIC or pNIC
VM NIC	vNIC	VM NIC
Internal port	VMkernel port	Virtual NIC
Active-backup	Active-standby	Active-standby
Balance-slb	Route based on source MAC hash combined with route based on physical NIC load	Switch independent / dynamic
LACP with balance-tcp	LACP and route based on IP hash	Switch dependent (LACP) / address hash

AHV Networking Best Practices Checklist

Command Line

- Use the Nutanix Prism network visualization feature before you use the command line to view the network.

- For Nutanix AOS versions from 5.11 through 5.18, use Prism uplink configuration for clusters with a single bridge and bond.
- From Nutanix AOS 5.19 on, use Prism virtual switch configuration for all network management and don't use the CLI.
- Use the CLI configuration for clusters with Nutanix AOS versions prior to 5.11 or versions from 5.11 through 5.18 with multiple bridges and bonds.
- Don't use `manage_ovs` to make network changes once you have used Prism uplink configuration or virtual switch configuration.
- Use the `allssh` and `hostssh` shortcuts only with `view` and `show` commands. Use extreme caution with commands that make configuration changes, as these shortcuts run them on every CVM or AHV host. Running a disruptive command on all hosts risks disconnecting all hosts. When you make network changes, only use the `allssh` or `hostssh` shortcuts in a staging environment.
- Ensure that IPMI console connectivity is available and place the host and CVM in maintenance mode before you make any CLI host networking changes.
- Connect to a CVM instead of to the AHV hosts when you use SSH. Use the `hostssh` or `192.168.5.1` shortcut for any AHV host operation.
- For high availability, connect to the cluster virtual IP for cluster-wide commands entered in the aCLI.
- Use the `manage_ovs --host` shortcut to configure networking for compute-only nodes.

Open vSwitch

- Don't modify the OpenFlow tables associated with any OVS bridge.
- Although it's possible to set QoS policies and other network configuration on the VM tap interfaces manually (using the `ovs-vsctl` command), we don't recommend or support it. Policies don't persist across VM power cycles or migrations between hosts.
- Don't delete, rename, or modify the OVS bridge `br0` or the bridge chain.
- Don't modify the native Linux bridge `virbr0`.

OVS Bonds

- Include at least two physical interfaces in every bond.
- Aggregate the 10 Gbps or faster interfaces on the physical host to an OVS bond named br0-up on the default OVS bridge br0 and trunk VLANs to these interfaces on the physical switch.
- Use active-backup load balancing unless you have a specific need for LACP with balance-tcp such as increased throughput.
- Create a separate bond and bridge for the connected 1 GbE interfaces or remove them from the primary bond br0-up.
- Don't mix NIC models from different vendors in the same bond.
- Don't mix NICs of different speeds in the same bond.
- Use LACP with balance-tcp only if VMs require link aggregation for higher speed or better fault tolerance. Ensure that you have completed LACP configuration on the physical switches after enabling LACP on AHV.
- Set the LACP speed setting to fast on the physical switches.
- Verify that the LACP speed setting is fast on AHV hosts that you've upgraded using the ovs-appctl commands later in the appendix.
- Don't use the balance-tcp algorithm without LACP upstream switch link aggregation.
- Don't use the balance-slb algorithm if the physical switches use IGMP snooping and pruning.
- Don't use the balance-slb algorithm with link aggregation such as LACP.
- Don't use static link aggregation such as EtherChannel with AHV.

Physical Network Layout

- Refer the [Physical Networking best practice guide](#) for more information.
- Use redundant top-of-rack switches in a leaf-spine architecture. This simple, flat network design is well suited for a highly distributed, shared-nothing compute and storage architecture.

- Connect all the nodes that belong to a given cluster to the same layer 2 network segment.
- If you need more east-west traffic capacity, add spine switches or uplinks between the leaf and spine.
- Use redundant 40 Gbps (or faster) connections to the spine to ensure adequate bandwidth between upstream switches.

Upstream Physical Switch Specifications

- Connect the 10 Gbps or faster uplink ports on the AHV node to switch ports that are nonblocking datacenter-class switches that provide line-rate traffic throughput.
- Use an Ethernet switch that has a low-latency design and provides predictable, consistent traffic latency regardless of packet size, traffic pattern, or the features enabled on the 10 Gbps or faster interfaces. Port-to-port latency should be no higher than two microseconds.
- Use fast-convergence technologies (such as Cisco PortFast) on switch ports connected to the AHV host (sometimes called edge ports).
- To prevent packet loss from oversubscription, avoid switches that use a highly oversubscribed port-buffer architecture, where many ports share the same small buffer.

Switch and host VLANs

- Keep the CVM and AHV host in the same VLAN. By default, the CVM and the hypervisor are placed on the native untagged VLAN configured on the upstream physical switch.
- Configure switch ports connected to AHV as VLAN trunk ports.
- Configure a dedicated native untagged VLAN other than 1 on switch ports facing AHV hosts to carry CVM and AHV host traffic.

Guest VM VLANs

- Use the Prism GUI to configure VM network VLANs on br0.

- Use VLANs other than the dedicated CVM and AHV VLAN.
- In Nutanix AOS versions prior to 5.19, use the aCLI to add VM network VLANs for additional bridges. Include the bridge name in the network name for easy bridge identification. In Nutanix AOS 5.19 and later, use the Prism UI.
- Use VM NIC VLAN trunking only in cases where VMs require multiple VLANs on the same NIC. In all other cases, add a new VM NIC with a single VLAN in access mode to bring new VLANs to VMs.
- Don't use direct-mode NICs unless Nutanix Support tells you to.

CVM Network Configuration

- Don't remove the CVM from the OVS bridge br0 or the native Linux bridge virbr0.
- If required for security, add a dedicated CVM backplane VLAN with a nonroutable subnet to separate CVM storage backplane traffic from CVM management traffic.
- Don't use backplane segmentation or additional service segmentation unless separation of backplane or storage traffic is a mandatory security requirement.
- If the network for the backplane or additional services is connected to a bridge other than br0, ensure that this bridge has redundant uplinks with fault-tolerant load balancing.

Jumbo Frames

- Nutanix doesn't support configuring the MTU on a CVM's network interfaces to higher values.
- If you choose to use jumbo frames on hypervisor hosts, enable them end to end in the desired network and consider both the physical and virtual network infrastructure impacted by the change.
- Nutanix recommends using jumbo frames on the AHV host for Flow Virtual Networking, but CVMs can still use a 1,500 byte MTU.

Address Management

- Coordinate the configuration of AHV-managed IP address pools to avoid address overlap and conflict with existing network DHCP pools.
- Confirm IP address availability with the network administrator before you configure an IPAM address pool in AHV.
- Configure AHV clusters with unique VLANs to avoid VM MAC address conflict.
- If your AHV clusters require overlapping VLANs, contact your account team and Nutanix Support to assign MAC prefixes for each cluster to avoid VM MAC conflicts.

IPMI Ports

- Don't allow multiple VLANs on switch ports that connect to the IPMI interface. For management simplicity, only configure the IPMI switch ports as access ports in a single VLAN.

AHV Command Line Tutorial

Nutanix systems have a number of command-line utilities that make it easy to inspect the status of network parameters and adjust advanced attributes that may not be available in the Prism GUI. In this section, we address the three primary locations where you can enter CLI commands.

The first location is the CVM BASH shell. A command entered here takes effect locally on a single CVM. Administrators can also enter CLI commands in the CVM aCLI shell. Commands entered in the aCLI operate on the level of an entire Nutanix cluster, even though you're accessing the CLI from one CVM. Finally, administrators can enter CLI commands in an AHV host's BASH shell. Commands entered here take effect only on that AHV host. The following diagram illustrates the basic CLI locations.

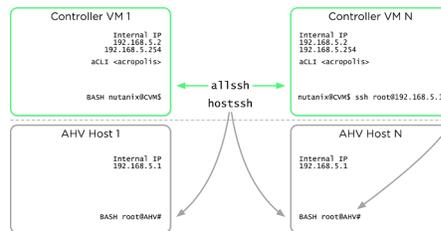


Figure 22: Command Line Operation Overview

CLI shortcuts exist to make cluster management easier. Often, you need to run a command on all CVMs or on all AHV hosts, rather than on a single host. It's tedious to log on to every system and enter the same command on each of them, especially in a large cluster. That's where the `allssh` and `hostssh` shortcuts come in. `allssh` takes a given command entered on the CVM BASH CLI and runs that command on every CVM in the cluster. `hostssh` works similarly, taking a command entered on the CVM BASH CLI and running it on every AHV host in the cluster, as shown in the previous figure.

To streamline the management of CVMs and AHV hosts, the SSH shortcut connects a single CVM directly to the local AHV host. From any single CVM, you can use SSH to connect to the AHV host's local address at IP address 192.168.5.1. Similarly, any AHV host can SSH to the local CVM using the static IP address 192.168.5.254. Because the address 192.168.5.2 on a CVM is used for dynamic high availability purposes in the AHV host, it may not always direct to the local CVM. This SSH connection uses the internal Linux bridge `virbr0`.

Let's take a look at a few examples to demonstrate the usefulness of these commands.

Example 1: `allssh`

Imagine that you need to determine which network interfaces are plugged in on all nodes in the cluster and the link speed of each interface. You could use `manage_ovs show_interfaces` at each CVM, but instead you use the `allssh` shortcut. First, SSH into any CVM in the cluster as the `nutanix` user, then run the command `allssh "manage_ovs show_interfaces"` at the CVM BASH shell:

```
nutanix@NTNX-A-CVM:~$ allssh "manage_ovs show_interfaces"
```

In the sample output, we've truncated the results after the second node to save space.

```

Executing manage_ovs show_interfaces on the cluster
===== a.b.c.d =====
name mode link speed
eth0 1000 True 1000
eth1 1000 True 1000
eth2 10000 True 10000
eth3 10000 True 10000
Connection to a.b.c.d closed.
===== e.f.g.h =====
name mode link speed
eth0 1000 True 1000
eth1 1000 True 1000
eth2 10000 True 10000
eth3 10000 True 10000
Connection to e.f.g.h closed.

```

Example 2: hostssh

If you wanted to view the MAC address of the eth0 interface on every AHV host, you could connect to each AHV host individually and use `ifconfig eth0`. To make things faster, use the `hostssh` shortcut instead. In this example, we still use SSH to connect to the CVM BASH shell, then prefix our desired command with `hostssh`:

```

nutanix@NTNX-A-CVM~$ hostssh "ifconfig eth0 | grep HWaddr"
===== a.b.c.d =====
eth0      Link encap:Ethernet  HWaddr 0C:C4:7A:46:B1:FE
===== e.f.g.h =====
eth0      Link encap:Ethernet  HWaddr 0C:C4:7A:46:B2:4E

```

Example 3: aCLI

Administrators can use the aCLI shell to view Nutanix cluster information that might not be easily available in the Prism GUI. For example, let's list all of the VMs in a given network. First, connect to any CVM using SSH, then enter the aCLI.

```

nutanix@NTNX-A-CVM~$ acli
<acropolis> net.list_vms 1GBNet
VM UUID                               VM name  MAC address
0d6afd4a-954d-4fe9-a184-4a9a51c9e2c1  VM2     50:6b:8d:cb:1b:f9

```

Example 4: SSH Root

The shortcut between the CVM and AHV host can be helpful when you're connected directly to a CVM but need to view some information or run a command against the local AHV host instead. In this example, we verify the localhost line of the `/etc/hosts` file on the AHV host while we're already connected to the CVM.

```
nutanix@NTNX-14SM36510031-A-CVM~$ ssh root@192.168.5.1 "cat /etc/hosts | grep
127"
127.0.0.1    localhost localhost.localdomain localhost4 localhost4.localdomain4
```

With these command-line utilities, you can manage a large number of Nutanix nodes at once. Centralized management helps administrators apply configuration consistently and verify configuration across a number of servers.

AHV Networking Command Examples

- Network view commands:

```
nutanix@CVM$ manage_ovs --bridge_name br0 show_uplinks
nutanix@CVM$ ssh root@192.168.5.1 "ovs-appctl bond/show br0-up"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl show"
nutanix@CVM$ acli
<acropolis> net.list
<acropolis> net.list_vms vlan.0
nutanix@CVM$ manage_ovs --help
nutanix@CVM$ manage_ovs show_interfaces
nutanix@CVM$ allssh "manage_ovs --bridge_name <bridge> show_uplinks"
## Use the Prism Virtual Switch UI instead in Nutanix AOS 5.19 and later.
nutanix@CVM$ manage_ovs --bridge_name <bridge> --interfaces <interfaces> --
bond_mode <bond-mode> update_uplinks
nutanix@CVM$ manage_ovs --bridge_name <bridge> --interfaces <interfaces> --
bond_mode <bond-mode> --require_link=false update_uplinks
```

- Bond configuration for 2 × 10 GbE:
 - › Use the Prism Virtual Switch UI instead in Nutanix AOS 5.19 and later.

```
nutanix@CVM$ manage_ovs --bridge_name br1 create_single_bridge
nutanix@CVM$ manage_ovs --bridge_name br0 --bond_name br0-up --interfaces 10g --
bond_mode <bond-mode> update_uplinks
nutanix@CVM$ manage_ovs --bridge_name br1 --bond_name br1-up --interfaces 1g --
bond_mode <bond-mode> update_uplinks
nutanix@cvm$ acli net.create br1_vlan99 vswitch_name=br1 vlan=99
```

- Bond configuration for 4 × 10 GbE:
 - › Use the Prism Virtual Switch UI instead in Nutanix AOS 5.19 and later.

```
nutanix@CVM$ manage_ovs --bridge_name br1 create_single_bridge
nutanix@CVM$ manage_ovs --bridge_name br2 create_single_bridge
nutanix@CVM$ manage_ovs --bridge_name br0 --bond_name br0-up --interfaces
eth4,eth5 --bond_mode <bond-mode> update_uplinks
nutanix@CVM$ manage_ovs --bridge_name br1 --bond_name br1-up --interfaces
eth2,eth3 --bond_mode <bond-mode> update_uplinks
nutanix@CVM$ manage_ovs --bridge_name br2 --bond_name br2-up --interfaces
eth0,eth1 --bond_mode <bond-mode> update_uplinks
nutanix@cvm$ acli net.create br1_vlan99 vswitch_name=br1 vlan=99
nutanix@cvm$ acli net.create br2_vlan100 vswitch_name=br2 vlan=100
```

- Load balance view command:

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-appctl bond/show"
```

- Load balance active-backup configuration:

- › Use the Prism Virtual Switch UI instead in Nutanix AOS 5.19 and later.

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up bond_mode=active-backup"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up other_config:bond-primary=ethX"
```

- Load balance balance-slb configuration:

- › Use the Prism Virtual Switch UI instead in Nutanix AOS 5.19 and later.

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up bond_mode=balance-slb"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up other_config:bond-rebalance-interval=30000"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-appctl bond/show br0-up"
```

- Load balance balance-tcp and LACP configuration:

- › Use the Prism Virtual Switch UI instead in Nutanix AOS 5.19 and later.

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up other_config:lacp-fallback-ab=True"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up other_config:lacp-time=fast"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up lacp=active"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up bond_mode=balance-tcp"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-appctl bond/show br0-up"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-appctl lacp/show br0-up"
```

- Disable LACP:

- › Use the Prism Virtual Switch UI instead in Nutanix AOS 5.19 and later.

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up bond_mode=active-backup"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up lacp=off"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0-up other_config:lacp-fallback-ab=True"
```

- CVM and AHV host tagged VLAN configuration:

```
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0 tag=10"
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl list port br0"
nutanix@CVM$ change_cvm_vlan 10
nutanix@CVM$ change_cvm_vlan 0
nutanix@CVM$ ssh root@192.168.5.1 "ovs-vsctl set port br0 tag=0"
```

- VM VLAN configuration:

```
nutanix@cvm$ acli vm.nic_update <vm_name> <nic mac address> network=<network name>
```

```
nutanix@CVM~$ acli vm.nic_update <vm name> <vm nic mac address>  
trunked_networks=<comma separated list of allowed VLAN IDs>  
update_vlan_trunk_info=true  
nutanix@CVM~$ acli vm.nic_update <vm name> <vm nic mac address> vlan_mode=kAccess  
update_vlan_trunk_info=true
```

References

1. [AHV Best Practice Guide](#)
 2. [AHV Administration Guide: Host Network Management](#)
 3. [Nutanix Security Guide: Securing Traffic Through Network Segmentation](#)
 4. [Open vSwitch Documentation](#)
 5. [Physical Networking Best Practice Guide](#)
 6. [Prism Web Console Guide: Network Visualization](#)
 7. [Flow Virtual Networking Guide](#)
-

About the Authors

Jason Burns is a Technical Marketing Engineer at Nutanix.

Lakshana Rajendran is a Technical Marketing Engineer at Nutanix.

About Nutanix

Nutanix offers a single platform to run all your apps and data across multiple clouds while simplifying operations and reducing complexity. Trusted by companies worldwide, Nutanix powers hybrid multicloud environments efficiently and cost effectively. This enables companies to focus on successful business outcomes and new innovations. Learn more at [Nutanix.com](https://www.nutanix.com).

List of Figures

Figure 1: Post-Imaging Network State.....	11
Figure 2: AHV Bridge Chain.....	13
Figure 3: Prism UI Network Creation.....	15
Figure 4: IPAM.....	16
Figure 5: Prism UI Network List.....	18
Figure 6: Prism UI VM Network Details.....	19
Figure 7: AHV Host Network Visualization.....	20
Figure 8: Network Connections for 2 x 10 GbE NICs.....	28
Figure 9: Virtual Switch Configuration for 2 x 10 GbE Active-Backup.....	30
Figure 10: Network Connections for 2 x 10 GbE and 2 x 1 GbE NICs.....	32
Figure 11: Create New Virtual Switch with 1 GbE Adapters.....	34
Figure 12: Create Network on Additional Virtual Switch.....	35
Figure 13: Active-Backup Fault Tolerance.....	37
Figure 14: Virtual Switch Configuration for Active-Backup.....	38
Figure 15: Balance-SLB Load Balancing.....	39
Figure 16: Virtual Switch Configuration for Balance-SLB.....	39
Figure 17: LACP and Balance-TCP Load Balancing.....	41
Figure 18: Virtual Switch Configuration for Balance-TCP with LACP.....	41
Figure 19: Default Untagged VLAN for CVM and AHV Host.....	44
Figure 20: Tagged VLAN for CVM and AHV Host.....	45
Figure 21: Prism UI CVM Network Interfaces.....	50
Figure 22: Command Line Operation Overview.....	59