

TECH NOTE

Nutanix Cloud Clusters on Microsoft Azure

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
Document Version History.....	6
2. Nutanix Cloud Clusters Portal.....	7
Azure Bare Metal as a Service (BMaaS).....	8
3. Flow Virtual Networking.....	10
Flow Virtual Networking Gateway.....	11
Open vSwitch.....	12
Creating a VPC.....	14
Creating a Subnet.....	16
4. Migration.....	18
5. Storage Availability in Azure.....	19
Respond to Failures.....	21
Prevent Network Partition Errors.....	22
Proactively Resolve Bad Disk Resources.....	22
Maintain Availability: Disk Failure.....	23
Maintain Availability: Availability Zone Failure.....	23
6. Outbound Communication.....	24
7. Capacity Optimization.....	26
Compression.....	26
Deduplication.....	26
8. Encryption.....	28
9. Virtual Machine High Availability.....	29
VMHA Recommendations and Requirements.....	31

10. Acropolis Dynamic Scheduler.....	32
Affinity Policies.....	32
11. Conclusion.....	33
About Nutanix.....	34
List of Figures.....	35

1. Executive Summary

Nutanix designed its software to give customers running workloads in a hybrid cloud environment the same experience they expect from on-premises Nutanix clusters. Because Nutanix in a hybrid multicloud runs AOS and AHV with the same CLI, UI, and APIs, existing IT processes and third-party integrations continue to work regardless of where they run.

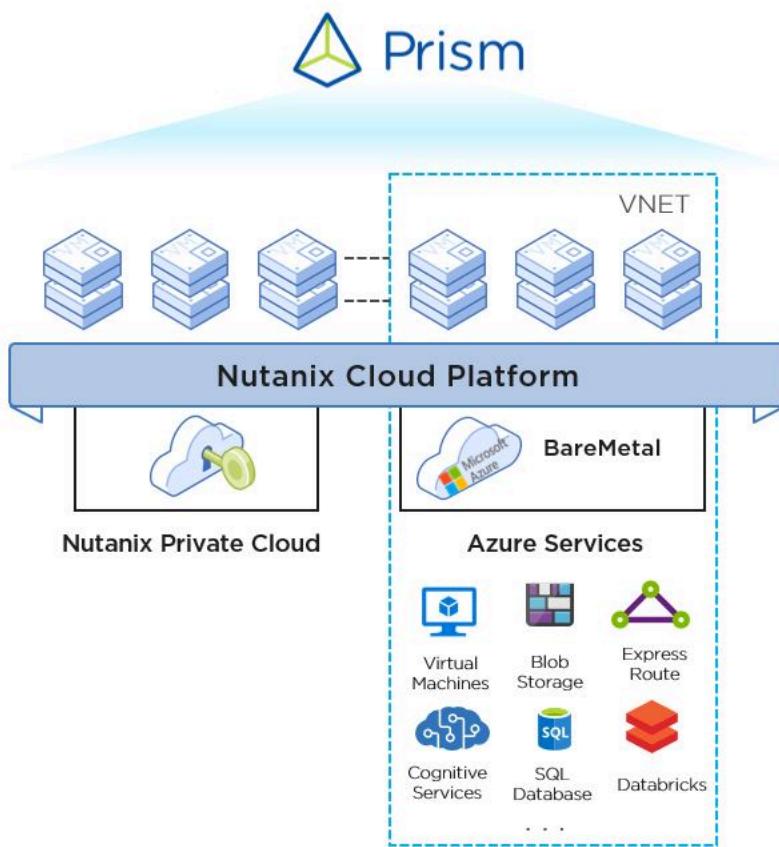


Figure 1: Overview of the Nutanix Cloud Platform with Azure

NC2 (Nutanix Cloud Clusters) on Microsoft Azure is a solution jointly engineered by Nutanix and Azure teams. It offers bare metal as a service to customers from

a hardware consumption perspective, and NC2 provides a consistent experience for provisioning and managing clusters deployed in Azure.

NC2 situates the complete Nutanix hyperconverged infrastructure (HCI) stack directly on a BareMetal instance. This bare-metal instance runs a Controller VM (CVM) and Nutanix AHV as the hypervisor just like any on-premises Nutanix deployment, using the Azure Virtual Network (VNet) to connect to the network.

AHV runs an efficient embedded distributed network controller that integrates user VM networking with Prism Element. Prism Central then works with AHV and NC2 to use Flow Virtual Networking to create an overlay that provides granular control. Flow Virtual Networking enables connectivity to all Azure services and enables workloads running on the clusters to send and receive north- and south-bound traffic.

AOS can withstand hardware failures and software glitches and ensures that application availability and performance are never compromised. Combining features like native rack awareness with the Azure bare-metal service allows Nutanix to operate freely in a dynamic cloud environment.

NC2 on Azure gives on-premises workloads a home in the cloud, offering native access to available cloud services without requiring you to reconfigure your software.

Document Version History

Version Number	Published	Notes
1.0	September 2022	Original publication.
1.1	March 2023	Updated console.nutanix.com to cloud.nutanix.com.
1.1.1	March 2023	Added outbound management firewall requirements.

2. Nutanix Cloud Clusters Portal

Customers access the NC2 Portal through their existing accounts at my.nutanix.com. You can use the portal to deploy Azure clusters and to manage tasks like health remediation and expanding and condensing your clusters. Prism Central is automatically deployed with Flow Virtual Networking to provide overlay networking. Prism Central can manage your deployed NC2 clusters alongside your on-premises clusters. For easy day-two operations, Prism Central can also manage AOS upgrades for on-premises, remote or branch office, and cloud-based Nutanix clusters.

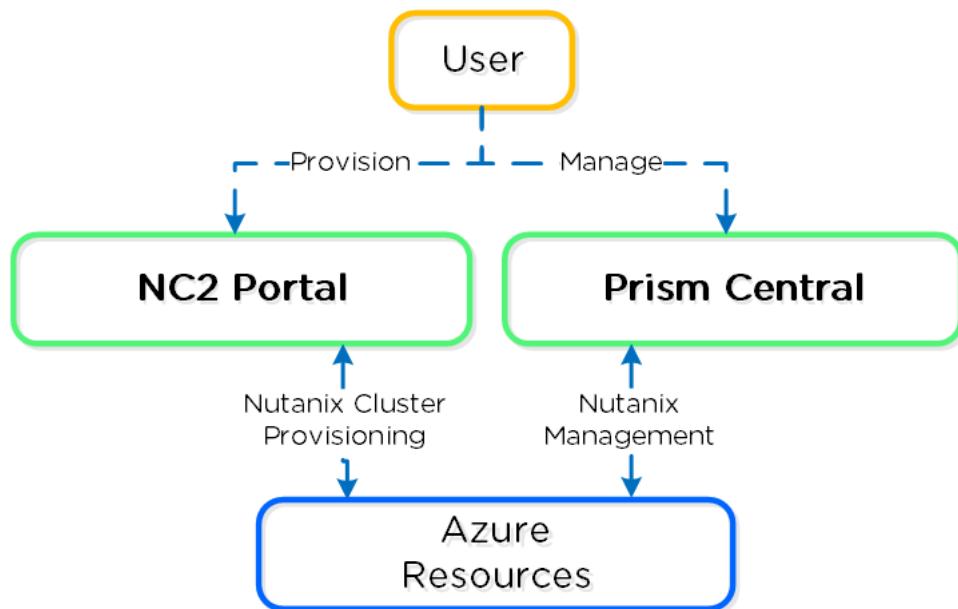


Figure 2: NC2 Management

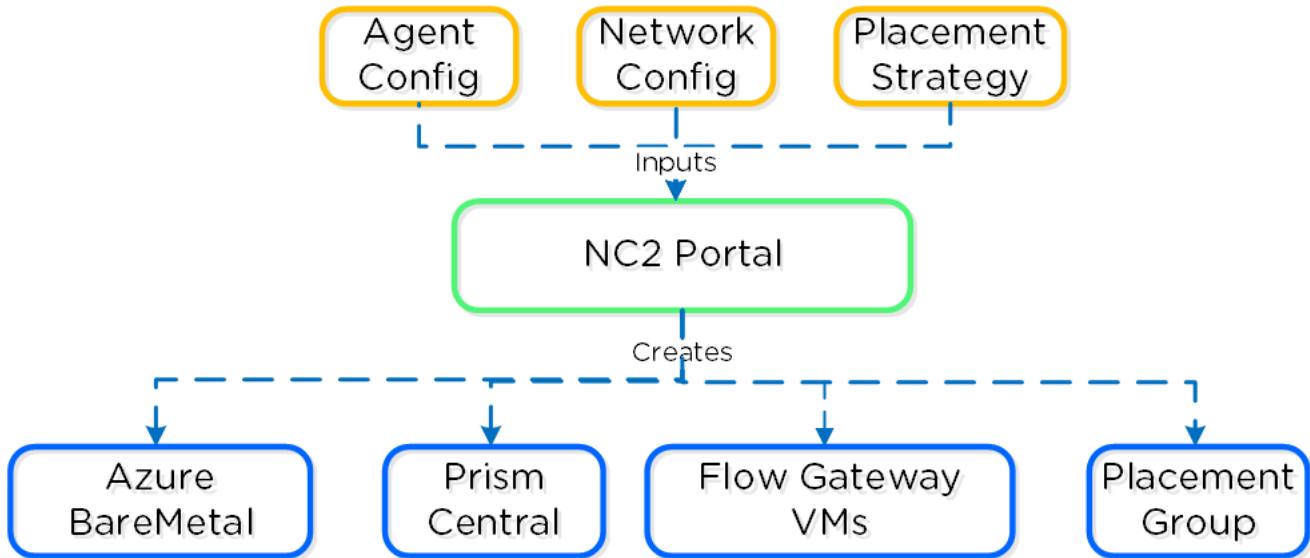


Figure 3: NC2 Portal

The NC2 Portal provides the following services:

- Obtaining and managing bare-metal resources.
- Deploying and setting up Prism Central and Flow Virtual Networking.
- Managing node placement strategy and removing or adding nodes based on the health of the cluster.

For a successful deployment, NC2 needs outbound access to the NC2 portal, using either a Network Address Translation (NAT) gateway or an on-premises VPN with outbound access. Your Nutanix cluster can sit in a private subnet that can be accessed only from your VPN, limiting exposure to your environment.

Azure Bare Metal as a Service (BMaaS)

The NC2 Portal communicates with Azure BMaaS (bare metal as a service) to deploy, remediate, and install Nutanix software on bare-metal nodes.

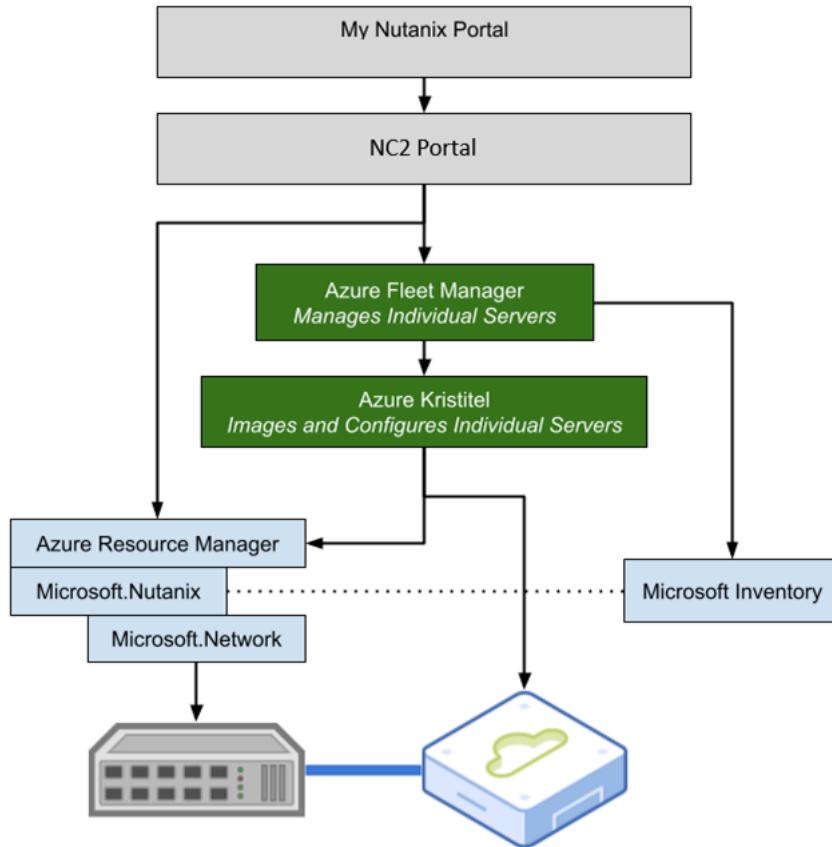


Figure 4: NC2 and Azure Bare Metal as a Service

Azure BMaaS consists of two main services: Azure Fleet Manager and Azure Kristitel. Azure Fleet Manager tracks server allocations and monitors fleet capacity. Fleet Manager also allocates servers to ensure rack awareness inside Azure using placement groups. This service tracks health and outages and works with the NC2 portal to provision new nodes.

Azure Kristitel moves servers into and out of staging. Kristitel installs Nutanix software on the nodes and configures the nodes under the customer's account.

When you register your Azure account with the NC2 portal, you see `microsoft.network`, `microsoft.nutanix`, and `microsoft.BareMetal` listed as the resource providers in your Azure account.

3. Flow Virtual Networking

NC2 uses Flow Virtual Networking in Azure to create an overlay network that simplifies administration and reduces networking constraints across cloud vendors. Flow Virtual Networking masks or reduces cloud constraints by providing an abstraction layer and allows the network substrate (and its associated features and functionalities) to be consistent with the customer's on-premises Nutanix deployments. You can create new virtual networks (called virtual private clouds or VPCs) in Nutanix with subnets in any address range, including those from the RFC1918 (private) address space, and define DHCP, Network Address Translation (NAT), routing, and security policies from the familiar Prism Central interface.

The simplicity provided by Flow Virtual Networking can be seen in the way it allows you to handle subnets. Subnet delegation enables you to designate a specific subnet for an Azure platform as a service (PaaS) that you need to inject into your virtual network, but Azure only allows one delegated subnet per VNet. NC2 needs a management subnet delegated to the Microsoft.BareMetal/AzureHostedService in order to deploy Nutanix clusters, and every subnet used for user-native VM networking also needs to be delegated to the same service. Because a VNet can have only one delegated subnet, networking configuration can quickly get out of hand with VNets peered among each other to allow communication.

Flow Virtual Networking drastically reduces the number of VNets needed to enable the workloads running on NC2 and Azure to communicate by allowing you to create over 500 subnets while only consuming one Azure VNet.

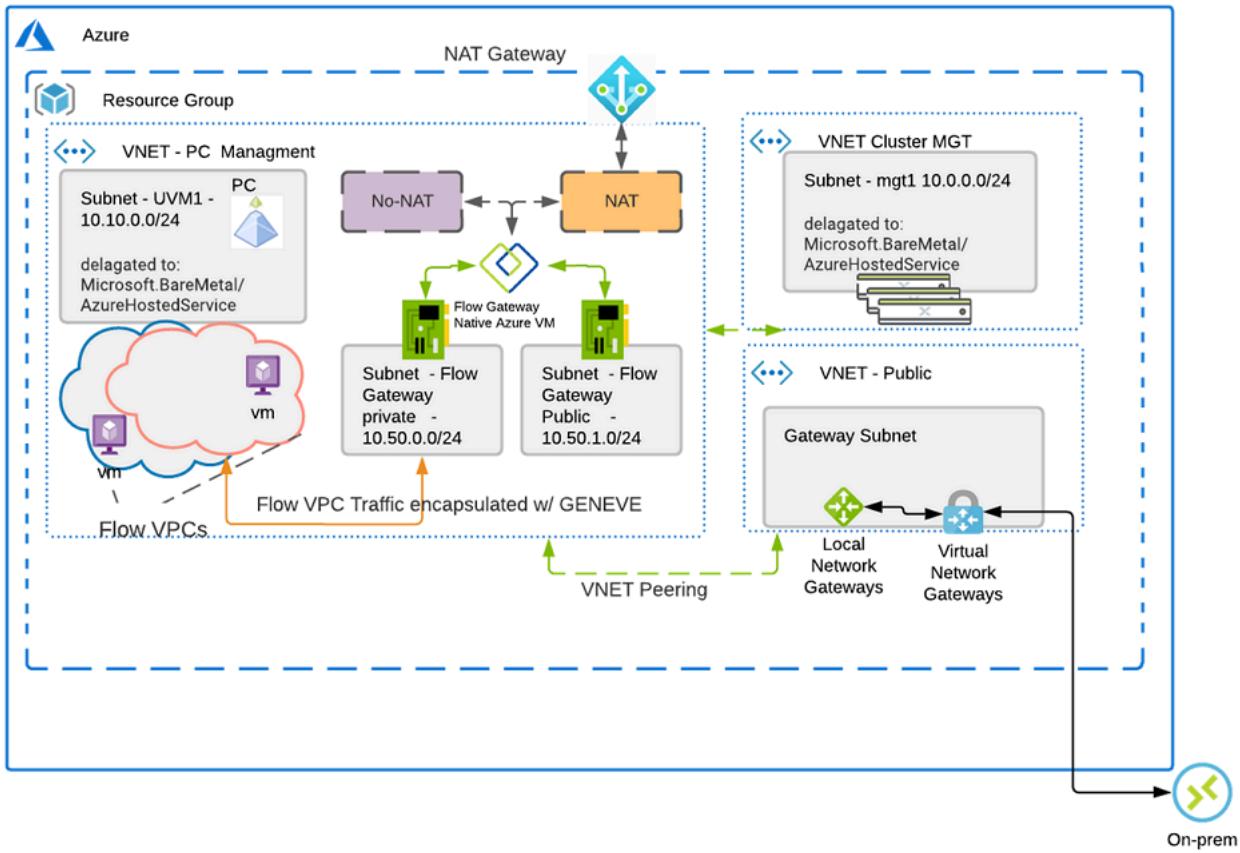


Figure 5: Azure Network Design

Flow Virtual Networking Gateway

Prism Central provides the control plane for Flow Virtual Networking. The subnet for Prism Central is delegated to Microsoft.BareMetal/AzureHostedService so that you can use native Azure networking to distribute IP addresses for Prism Central.

Once you deploy Prism Central, the Flow Virtual Networking gateway deploys into the same subnet Prism Central is using. The Flow Virtual Networking gateway allows the user VMs using the VPCs to communicate with native Azure services and have parity with native Azure VMs for elements such as:

- User-defined routes: You can create custom or user-defined (static) routes in Azure to override Azure's default system routes or to add additional routes to a subnet's route table. In Azure, you create a route table, then associate the route table to zero or more virtual network subnets.
- Load balancer deployment: You can balance services offered by user VMs with the Azure-native load balancer.
- Network security groups: You can write stateful firewall policies.

The Flow Virtual Networking gateway VM is responsible for all VM traffic going north and south from the cluster. During deployment you can pick different sizes for the Flow Virtual Networking gateway VM based on how much bandwidth you need.

Note: CVM replication between other CVMs and on-premises clusters doesn't go through the Flow Virtual Networking gateway VM so you don't have to size for that traffic.

User VMs that want to communicate with AHV, CVM, Prism Central, and Azure resources go through the external network card on the Flow Virtual Networking gateway VM, and NAT uses a native Azure address to ensure routing to all resources. You can also use user-defined routes in Azure to communicate directly with Azure resources if you don't want to use NAT. This method allows fresh installs to communicate with Azure right away and gives customers options for more advanced configurations.

Open vSwitch

AHV uses Open vSwitch (OVS) for all VM networking. You can configure VM networking through Prism or the aCLI, and each vNIC connects to a tap interface. The following figure shows a conceptual diagram of the OVS architecture.

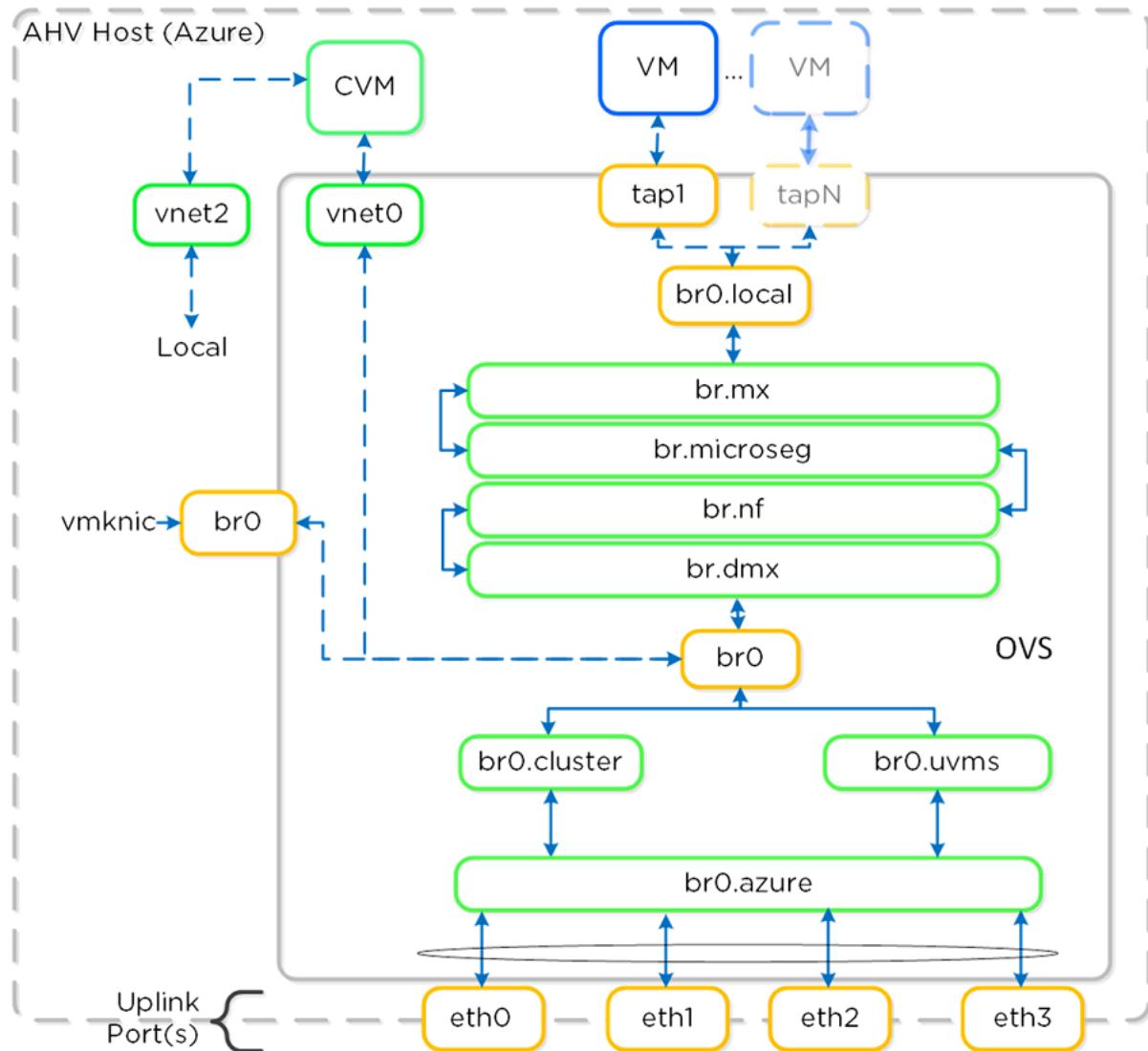


Figure 6: OVS Conceptual Architecture

The Nutanix OVS implementation is similar to the on-premises implementation. The br0 bridge splits traffic between br0.cluster (AHV and CVM IP addresses) and br0.uvms (user VM IP addresses). The AHV and CVM traffic from br0.cluster passes through to the br0.azure bridge with no modification to the data packets. The top-of-rack switches provide security for br0.cluster traffic. The user VM traffic from br0.uvms is passed through to br0.azure by the OVS rules installed for VLAN ID translation.

br0.azure has the OVS bond br0.azure-up, which forms a bonded interface with the bare-metal attached physical NICs. br0.azure hides the bonded interface from br0.uvms and br0.cluster.

Creating a VPC

Successfully deploying the first cluster in Azure automatically creates a transit-vpc that contains an external subnet called overlay-external-subnet-nat. This external subnet is required if the VPC needs to send traffic to an external destination. The external subnet is tied to the Flow Virtual Networking gateway VM, which has 50 IP addresses assigned to its external NIC. You can use these addresses to assign source NAT IP addresses to reach user-created VPCs for outbound communication and for Nutanix user VMs that need direct inbound communication. NAT gateways perform the IP address translations required for external routing, but you can also have external connectivity without NAT.

To create a VPC, perform the following steps:

- Sign in to the Prism Central web console.
- Click the entities menu in the main menu, expand Network & Security, then select Virtual Private Clouds. The Virtual Private Clouds List page appears.

The screenshot shows the Nutanix Prism Central interface. The left sidebar has a dark theme with white text. It includes sections for Dashboard, Compute & Storage, Network & Security (with Subnets, Virtual Private Clouds, Floating IPs, Connectivity, and Security Policies), Data Protection, Hardware, Activity, Operations, Administration, Services, and Prism Central Settings. The Network & Security section is expanded, and the Virtual Private Clouds item is selected, highlighted with a blue border. The main content area is titled "Virtual Private Clouds" and "List". It features a "Create VPC" button and an "Actions" dropdown. A search bar says "Type text to filter by". Below it, a section titled "Viewing all 3 Virtual Private Clouds" lists three entries: "CloudReady", "SecureNet", and "transit-vpc". The "transit-vpc" entry is selected, indicated by a blue border around its name and a "Transit VPC" label next to it.

Figure 7: Virtual Private Clouds List

- Click Create VPC. The Create VPC dialog opens.

The screenshot shows the 'Create VPC' dialog box. At the top, it says 'Create VPC' and has a close button ('X'). Below that is a 'Name' field containing 'ACME-AZURE'. Under 'External Connectivity', there's a note about VLAN Subnets required for external connectivity. It lists an 'External Subnet' named 'overlay-external-subnet-nat' with a note that maximum of 1 NAT and 1 No-NAT Subnet is allowed. There's also a section for 'Externally Routable IP Addresses' which is marked as optional and non-overlapping with other VPCs. A 'Domain Name Servers (DNS)' field is also present. At the bottom right are 'Cancel' and 'Create' buttons.

Figure 8: Create VPC

In the previous example, you only need an externally routable IP address if you plan on creating subnets that talk directly to resources outside the cluster without using NAT.

Creating a Subnet

Once you've created a VPC, you can use Flow Virtual Networking to create subnets for your user VMs.

Create Subnet

Name: server-web

Type: Overlay VPC: NewOnprem

IP Address Management

Network IP Prefix: 10.19.160.0/24 Gateway IP Address: 10.19.160.5

DHCP Settings

IP Address Pools

+ Create Pool

Start Address	End Address	Actions
10.19.160.100	10.19.160.240	Edit Remove

Cancel Save

Figure 9: Creating a Subnet with Flow Virtual Networking

Subnets you create have their own built-in IPAM. If outside applications need to talk directly to your user VM inside the subnet, you can also assign floating IP addresses from an Azure pool that comes from the external network of the Flow Virtual Networking gateway.

4. Migration

There are many reasons to move your applications to Azure, including consolidation, bursting, or wanting them on a cloud-based service. Once you configure networking from Azure to on-premises, you can choose any proven method for moving applications to an AHV-based cluster, which saves time and money. The following are most common ways to migrate data to NC2 in Azure:

Nutanix Disaster Recovery (disaster recovery orchestration)

If you want to take advantage of protection policies and recovery plans to protect applications across multiple Nutanix clusters, set up Nutanix Disaster Recovery (previously Leap) from Prism Central by selecting the checkbox. Whether you're doing disaster recovery or migrations, Nutanix Disaster Recovery stages your applications to be restored in the right order. You can also use the protection policies to quickly revert to on-premises if desired.

Nutanix Move

Nutanix Move is a cross-hypervisor migration solution that migrates VMs with minimal downtime. Nutanix Move supports three migration types: VMs running on ESXi managed by vCenter, EC2 instances backed by Elastic Block Storage (EBS) running on AWS, and VMs running on Hyper-V. Nutanix Move also supports migrating Azure EC2 VMs to AHV on the Nutanix cluster, though this use case is minimal.

AHV-based backups

You can use any third-party backup product to restore applications to NC2, which is important when you need to migrate or do testing and development work.

5. Storage Availability in Azure

An Azure Region is a separate geographic area. Each Region has multiple, isolated locations known as Availability Zones. An Availability Zone is a logical datacenter available for any Azure customer in that Region to use. Each zone in a Region has redundant and separate power, networking, and connectivity to reduce the likelihood of two zones failing simultaneously. Nutanix uses a partition placement strategy when deploying nodes inside an Azure Availability Zone. One Nutanix cluster can't span different Availability Zones in the same Region, but you can have multiple Nutanix clusters replicating between each other in different zones or Regions.

Nutanix places the Azure bare-metal nodes in different Azure racks and stripes new hosts across the partitions.

A partition placement strategy with n partitions or failure domains per node type is used.

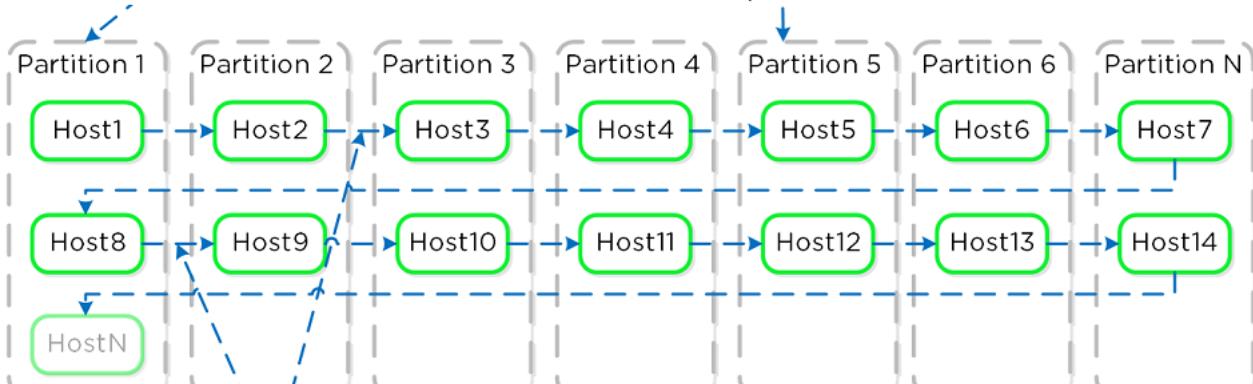


Figure 10: Partition Placement

When you have formed the Nutanix cluster, the partition groups map to the Nutanix rack-awareness feature. AOS storage writes data replicas to other racks in the cluster to ensure that the data remains available for both replication factor 2 and replication factor 3 scenarios in the case of a rack failure or planned downtime.

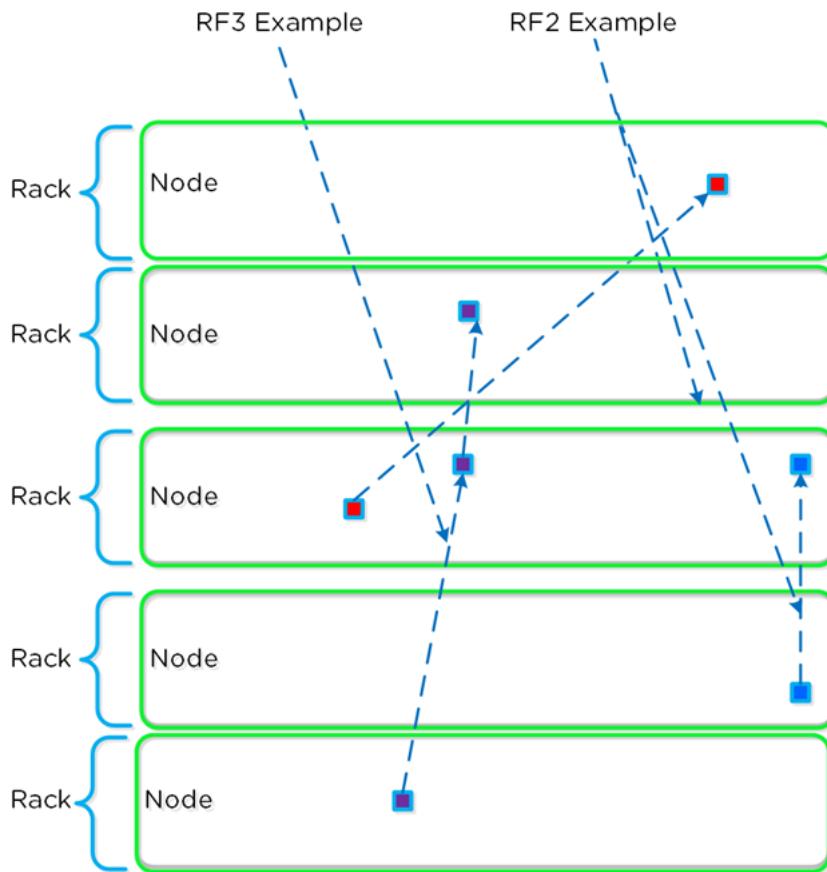


Figure 11: Replication Factor Data Placement Across Racks

The following table highlights the minimum number of racks required in your cluster to withstand a given number of rack failures. Nutanix Erasure Coding (EC-X) is one of the storage reduction technologies available in AOS. EC-X takes one or two data copies and creates a parity you can use to recreate the data if required.

Table: Desired Fault Tolerance and Required Nodes

Desired Awareness Type	Fault Tolerance Level	EC-X Enabled	Minimum Units in the Cluster	Simultaneous Failure Tolerance
Rack	1	No	3 racks	1 rack
Rack	1	Yes	4 racks	1 rack
Rack	2	No	5 racks	2 racks
Rack	2	Yes	6 racks	2 racks

Administrators can use replication factor 3 when high availability requirements exceed the data protection level that replication factor 2 provides. We also recommend using replication factor 3 in larger clusters (32 or more nodes). Your environment's specific availability requirements should dictate what replication factor you use.

Respond to Failures

AOS storage withstands a variety of hardware failures and builds strong redundancy into the software stack. Nutanix software processes that encounter a serious error are designed to fail fast. This design principle quickly restarts normal operations instead of waiting for a potentially faulty process to complete. Because Nutanix storage continuously monitors components, it can stop and restart them when an error occurs to recover as quickly as possible, rather than letting them linger in an unresponsive state. Each host relies on its local CVM to service all storage requests. AOS storage continuously monitors the health of all CVMs in the cluster. If an unrecoverable error occurs on a particular CVM, Nutanix autopathing automatically reroutes requests from the host to a healthy CVM on another node, providing data path redundancy.

This redirection continues until the local CVM failure issue is resolved. Because the cluster has a global namespace and access to replicas for all the data on that node, it can service requests immediately. This ability provides a high degree of fault tolerance and failover for all VMs in a Nutanix cluster. If the node's CVM continues to be unavailable for a prolonged period, data automatically replicates to maintain the necessary replication factor.

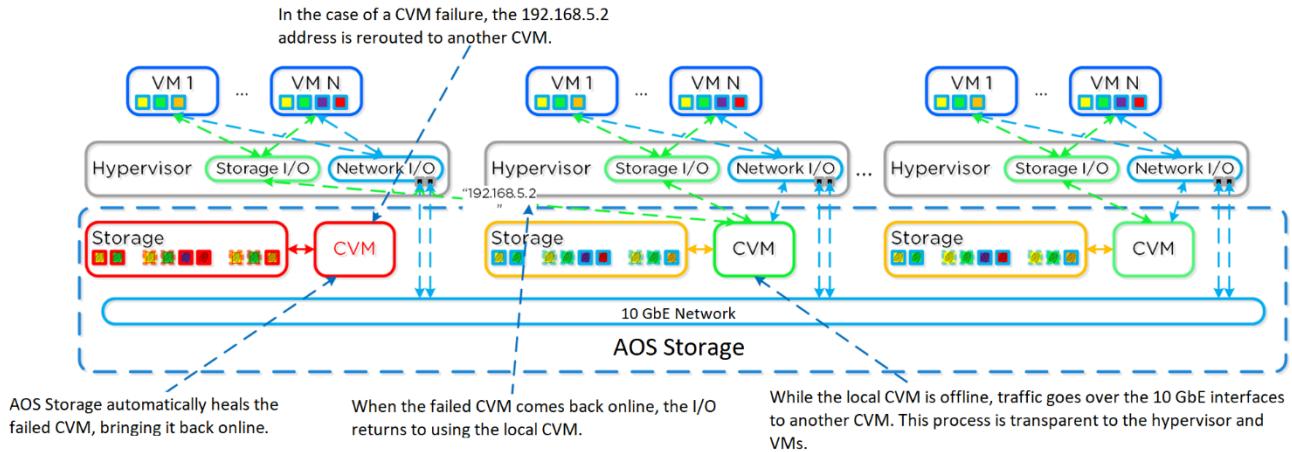


Figure 12: Data Path Redundancy

Prevent Network Partition Errors

Nutanix uses the Paxos algorithm to avoid split-brain scenarios. Paxos is a proven protocol for reaching consensus or quorum among several participants in a distributed system. Before any file system metadata is written to Cassandra, Paxos ensures that all nodes in the system agree on the value. If the nodes don't reach a quorum, the operation fails in order to prevent any potential corruption or data inconsistency. This design protects against events like network partitioning, where communication between nodes may fail or packets may become corrupt, leading to a scenario where nodes disagree on values. AOS storage also uses time stamps to ensure that updates are applied in the proper order.

Proactively Resolve Bad Disk Resources

Nutanix storage incorporates a Curator process that performs background housekeeping tasks to keep the entire cluster running smoothly. Among Curator's multiple responsibilities is ensuring file system metadata consistency and combing the extent store for corrupt and underreplicated data.

Additionally, Curator scans extents in successive passes, computes each extent's checksum, and compares it with the metadata checksum to validate consistency. If the checksums don't match, the corrupted extent is replaced

with a valid extent from another node. This proactive data analysis protects against data loss and identifies bad sectors you can use to detect disks that are about to fail.

Maintain Availability: Disk Failure

The Nutanix unified component Stargate receives and processes data. All read and write requests for a node are sent to the Stargate process on that node. The Hades service simplifies the break-fix procedures for disks and automates several tasks that previously required manual user actions. Hades helps fix failing devices before they become unrecoverable.

Once Stargate sees delays in responses to I/O requests to a disk, it marks the disk offline. Hades then automatically removes the disk from the data path and runs smartctl checks against it. If the checks pass, Hades marks the disk online and returns it to service. If the checks fail or if Stargate marks a disk offline three times in one hour (regardless of the smartctl check results), Hades automatically starts the BMaaS removal process. Removing the Azure node triggers an API call to the cluster portal, which notifies the NC2 portal. The NC2 portal allocates a new instance from BMaaS, adds it to the cluster, and marks the Azure node with the unresponsive disk for removal. The cluster software automatically replicates the data on the bad Azure node to other instances, then finishes the removal process for the Azure node.

Maintain Availability: Availability Zone Failure

Availability Zones go offline for a variety of reasons—issues with power, cooling, or networking as well as scheduled system maintenance. We need to ensure that your NC2 on Azure instance meets your availability needs. To avoid downtime in Azure, protect your workloads with Nutanix Disaster Recovery. The destination for Nutanix Disaster Recovery could be another on-premises cluster or another NC2 on Azure instance in a different Availability Zone.

6. Outbound Communication

There are a few general outbound requirements for deploying a Nutanix cluster in Azure in addition to the existing requirements that on-premises clusters use for support services. The following tables show the endpoints the Nutanix cluster needs to communicate with for a successful deployment. Your Azure cluster needs DNS access from the VNet that you deployed it to.

Note: Many of the destinations listed here use DNS failover and load balancing. For this reason, the IP address returned when resolving a specific domain may change rapidly. We can't provide specific IP addresses in place of domain names.

Table: Cluster Outbound to the Cluster Portal

Source	Destination	Protocol	Purpose
Management subnet	https://download.nutanix.com/*	TCP 443 (HTTPS)	Life Cycle Manager (LCM) required to upgrade NCI and NC2 components
Management subnet	https://insights.nutanix.com/*	TCP 443 (HTTPS)	Pulse telemetry provides diagnostic system data to Nutanix Support
Management subnet	gateway-external-api.cloud.nutanix.com	TCP 443 (HTTPS)	NC2 portal orchestration
Management subnet	https://azure-support.nutanix.com/	TCP 443 (HTTPS)	Remote tunnel support
Management subnet	138.236.128.112	TCP 443 (HTTPS)	Azure NTP server

Table: Cluster Outbound to Azure

Source	Destination	Protocol	Purpose
Management subnet	management.azure.com	TCP 443 (HTTPS)	Make API calls from NC2 to manage Azure resources

Source	Destination	Protocol	Purpose
Management subnet	downloads.cloud.nutanix.com	TCP 443 (HTTPS)	Download NC2 RPM Package Manager (RPM) packages

The Ports and Protocols guide lists general firewall support requirements, with additional focus in the [Disaster Recovery \(formerly Leap\)](#) section.

7. Capacity Optimization

Nutanix Cloud Infrastructure (NCI) software offers capacity optimization features that improve storage utilization and performance. The two key features are compression and deduplication.

Compression

Nutanix systems currently offer two types of compression policies.

Inline

The system compresses data synchronously as it's written to optimize capacity and maintain high performance for sequential I/O operations. Inline compression only compresses sequential I/O to avoid degrading performance for random write I/O.

Post-process

For random workloads, data writes to the SSD tier uncompressed for high performance. Compression occurs after cold data migrates to lower-performance storage tiers. Post-process compression acts only when data and compute resources are available, so it doesn't affect normal I/O operations.

Nutanix recommends that all customers carefully consider the advantages and disadvantages of compression for their specific applications. For further information on compression, refer to the [Nutanix Data Efficiency tech note](#).

Deduplication

The software-driven Elastic Deduplication Engine increases the effective capacity in the disk tier and the utilization of the performance tiers (RAM and flash) by eliminating duplicate data. By providing larger effective cache sizes in the performance tier, this feature substantially increases performance for certain workloads.

Deduplication savings vary greatly depending on workload and data types, but, in general, deduplication provides the largest benefit for common data sets, such as full-clone VDI workloads. Nutanix doesn't recommend deduplication for general-purpose server workloads, including business-critical applications.

Note: For containers hosting business-critical applications, VDI, general server workloads, and big data, we recommend disabling deduplication for all except full-clone VDI VMs. Increase CVM memory to at least 24 GB.

Nutanix recommends that all customers carefully consider the advantages and disadvantages of deduplication for their specific applications. For further information on deduplication, refer to the Nutanix [Data Efficiency tech note](#).

8. Encryption

To help reduce cost and complexity, Nutanix supports a native local key manager (LKM) for all clusters with three or more nodes. The LKM runs as a service distributed among all the nodes. You can activate it easily from Prism Element to enable encryption without adding another silo to manage.

Organizations often purchase external key managers (EKMs) separately for both software and hardware. However, because the Nutanix LKM runs natively in the CVM, it's highly available and there's no variable add-on pricing based on the number of nodes. Every time you add a node, you know the final cost. You also gain peace of mind because when you upgrade your cluster, the key management services are also upgraded. When you upgrade the infrastructure and management services in lockstep, you're ensuring your security posture and availability by staying in line with the support matrix.

Nutanix software encryption provides native AES-256 data-at-rest encryption, which can interact with any KMIP- or TCG-compliant external key management service (KMS) server (such as Vormetric and SafeNet) and the native Nutanix KMS, introduced in AOS 5.8. The system uses Intel AES-NI acceleration for encryption and decryption processes to minimize any potential performance impacts.

We recommend using the native Nutanix KMS to provide additional security for your workloads in the cloud.

Note: The first copy of the data (written locally) is encrypted. The copy sent over the wire is also encrypted and stored on a remote node.

9. Virtual Machine High Availability

VM high availability (VMHA) ensures that VMs restart on another AHV host in the cluster if a host fails. VMHA considers RAM when calculating available resources throughout the cluster for starting VMs.

VMHA respects affinity and anti-affinity rules. For example, with VM-host affinity rules, VMHA doesn't start a VM pinned to AHV host 1 and host 2 on another host when those two are down unless the affinity rule specifies an alternate host.

There are two VM high availability modes:

Default

This mode requires no configuration and is included by default when you deploy an AHV-based Nutanix cluster. When an AHV host becomes unavailable, the VMs that were running on the failed AHV host restart on the remaining hosts, depending on the available resources. If the remaining hosts don't have sufficient resources, some of the failed VMs may not restart.

Guarantee

This nondefault configuration reserves space throughout the AHV hosts in the cluster to guarantee that all VMs can restart on other hosts in the AHV cluster during a host failure. To enable Guarantee mode, select the Enable HA Reservation checkbox shown in the following figure. A message then displays the amount of memory reserved and how many AHV host failures the system can tolerate.

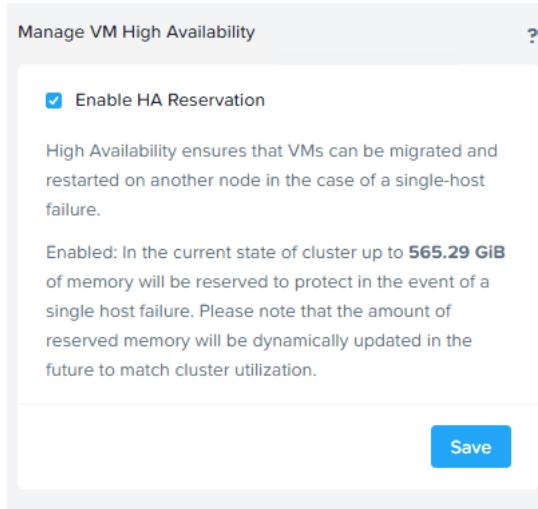


Figure 13: Enable VM High Availability Reservation

The VMHA configuration reserves resources to protect against:

- One AHV host failure, if all Nutanix containers are configured with replication factor 2.
- Two AHV host failures, if any Nutanix container is configured with replication factor 3.

Administrators can use the aCLI to manage protection against two AHV host failures when using replication factor 3. Use the following command to designate the maximum number of tolerable AHV host failures:

```
nutanix@cvm\$ acli ha.update num_host_failures_to_tolerate=x
```

When an unavailable AHV host comes back online after a VMHA event, VMs previously running on that host migrate back to maintain data locality.

To disable VMHA per VM, set a negative value (-1) when creating or updating the VM. This configuration removes the VM from the VMHA resource calculation.

```
nutanix@cvm\$ acli vm.update \<VM Name\> ha_priority=-1
nutanix@cvm\$ acli vm.create \<VM Name\> ha_priority=-1
```

In this configuration, the VM doesn't start on a new AHV host when its host fails; it only starts again when the failed host comes back online.

VMHA Recommendations and Requirements

- Use the nondefault VMHA Guarantee mode when you need to ensure that all VMs can restart if an AHV host fails.
- When using Guarantee mode, keep the default reservation type of kAcropolisHAReserveSegments; don't alter this setting.

Note: The VMHA reservation type kAcropolisHAReserveHosts is deprecated. Never change the VMHA reservation type to kAcropolisHAReserveHosts.

- Consider storage availability requirements when using VMHA Guarantee mode. Ensure that the parameter num_host_failures_to_tolerate is less than the configured storage availability. If there are only two copies of the VM data, the VM data could be unavailable if two hosts are down at the same time even though there are CPU and RAM resources to run the VMs.
- You must disable VMHA before you can use the Acropolis Dynamic Scheduler (ADS) VM-host affinity feature to pin a VM to one AHV host. However, we don't recommend pinning VMs to a particular AHV host, as we discuss in the following section.

10. Acropolis Dynamic Scheduler

ADS ensures that compute (CPU and RAM) and storage resources are available for VMs and volume groups in the Nutanix cluster. You can also use ADS to define affinity policies.

Affinity Policies

You define affinity policies one of two ways: manually (if you're a Nutanix administrator) or with a VM-provisioning workflow. There are two affinity policies:

VM-host affinity

This configuration keeps a VM on a specific set of AHV hosts. It is useful when you need to limit VMs to a subset of available AHV hosts because of application licensing, host resources (such as available CPU cores or CPU gigahertz speed), available RAM or RAM speed, or local SSD capacity. Host affinity is a must rule: AHV always honors the specified rule.

Note: We recommend against using VM-host affinity.

VM-VM antiaffinity

This configuration ensures that two or more VMs don't run on the same AHV host. It's useful when an application provides high availability and an AHV host must not be the application's single point of failure. Antiaffinity is a should rule that's honored only when there are enough resources available to run VMs on separate hosts.

For additional information about ADS and affinity policies, read the [ADS section](#) of the [Nutanix AHV best practice guide](#).

11. Conclusion

Nutanix software running in the cloud provides an easy extension for your on-premises datacenter. If you're already consuming cloud resources, the native Nutanix integration with Azure means that you don't need any additional skills to get your workloads running in the cloud. Management overhead shrinks when you no longer need an additional overlay network to secure and lock down networking between your on-premises environment and the cloud. Once you have NC2 running in the cloud, you can enjoy native networking speeds between migrated workloads and the new cloud services you want to consume. For more information, check out the [NC2 website](#).

About Nutanix

Nutanix is a global leader in cloud software and a pioneer in hyperconverged infrastructure solutions, making clouds invisible and freeing customers to focus on their business outcomes. Organizations around the world use Nutanix software to leverage a single platform to manage any app at any location for their hybrid multicloud environments. Learn more at www.nutanix.com or follow us on Twitter [@nutanix](https://twitter.com/nutanix).

List of Figures

Figure 1: Overview of the Nutanix Cloud Platform with Azure.....	5
Figure 2: NC2 Management.....	7
Figure 3: NC2 Portal.....	8
Figure 4: NC2 and Azure Bare Metal as a Service.....	9
Figure 5: Azure Network Design.....	11
Figure 6: OVS Conceptual Architecture.....	13
Figure 7: Virtual Private Clouds List.....	15
Figure 8: Create VPC.....	16
Figure 9: Creating a Subnet with Flow Virtual Networking.....	17
Figure 10: Partition Placement.....	19
Figure 11: Replication Factor Data Placement Across Racks.....	20
Figure 12: Data Path Redundancy.....	22
Figure 13: Enable VM High Availability Reservation.....	30