v1.6.1 | March 2023 | TN-2028

TECH NOTE

Nutanix Cloud Clusters on AWS

NUTANIX
YOUR ENTERPRISE CLOUD

Copyright

Contents

1. Executive Summary

Nutanix designed its software to give customers running workloads on cloud computing providers like Amazon Web Services (AWS) the same experience they expect from on-premises Nutanix clusters. Because Nutanix Cloud Clusters on AWS (NC2 on AWS) runs Nutanix 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.

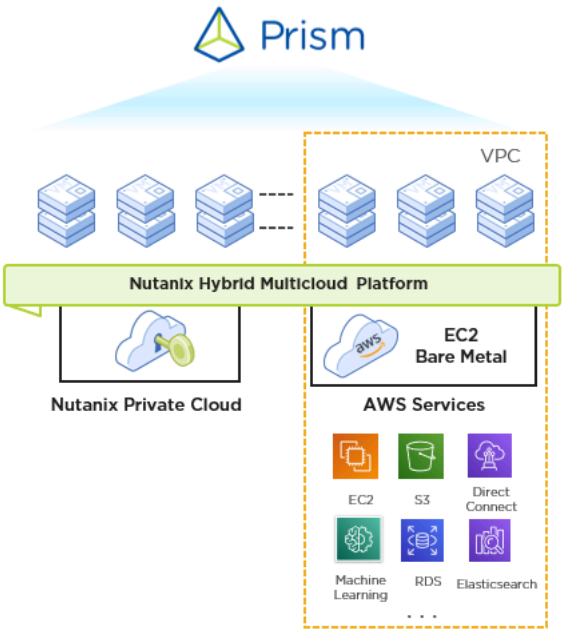

Figure 1: Overview of the Nutanix Cloud Infrastructure

NC2 on AWS situates the complete Nutanix hyperconverged infrastructure (HCI) stack directly on an Amazon Elastic Compute Cloud (EC2) bare-metal 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 AWS elastic network interface (ENI) to connect to the network. AHV user VMs don't require any additional configuration to access AWS services or other EC2 instances.

AHV runs an efficient embedded distributed network controller that integrates user VM networking with AWS networking. AHV assigns all user VM IP addresses to the bare-metal host where the VMs run. Instead of creating an overlay network, the AHV embedded network controller simply provides the networking information of the VMs running on NC2 on AWS, even as a VM moves around the AHV hosts. Because NC2 on AWS integrates IP address management with AWS Virtual Private Cloud (VPC), AWS allocates all user VM IP addresses from the AWS subnets in the existing VPCs.

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 AWS partition placement groups allows Nutanix to operate freely in a dynamic cloud environment. Customers can also save money by using hibernation to shut down clusters that aren't being used. Hibernation saves all cluster data to S3 before releasing the bare-metal instances back to the AWS pool to reduce EC2 costs.

NC2 on AWS quickly 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	August 2020	Original publication.
1.1	September 2020	Updated the Cluster Outbound to the Cluster Portal table.
1.2	September 2021	Added information on the hibernation feature and updated Storage Availability in AWS and Hibernate and Resume sections.
1.3	January 2022	Updated the Cluster Outbound to the Cluster Portal table.

Version Number	Published	Notes
1.4	March 2022	Updated product names to align with Nutanix Cloud Platform packaging.
1.5	October 2022	Added heterogeneous cluster support.
1.6	March 2023	Updated console.nutanix.com to cloud.nutanix.com.
1.6.1	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 AWS clusters and to manage tasks like health remediation and expanding and condensing your clusters. On-premises Prism Central can manage your deployed NC2 cluster 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.

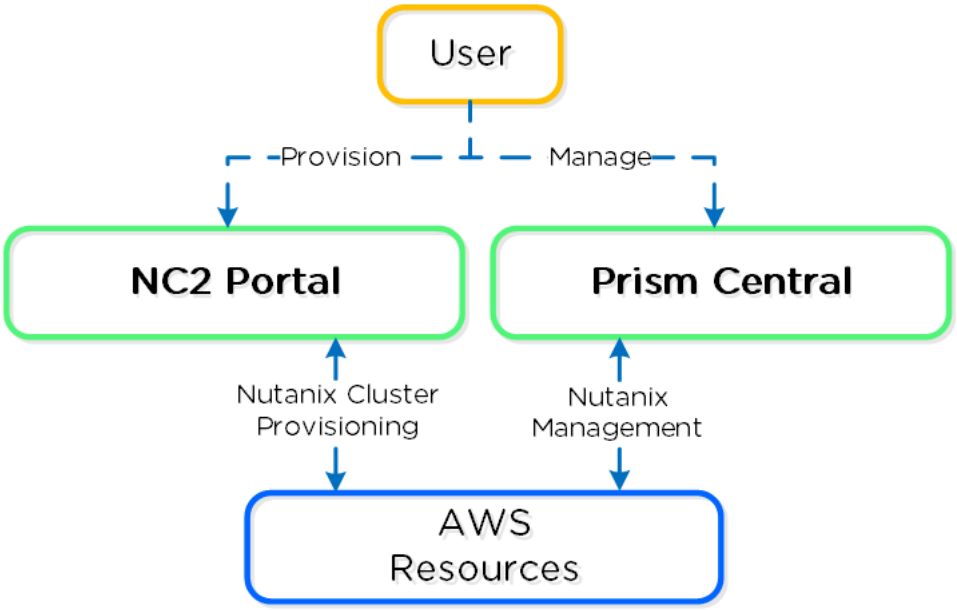

Figure 2: Cloud Clusters Management

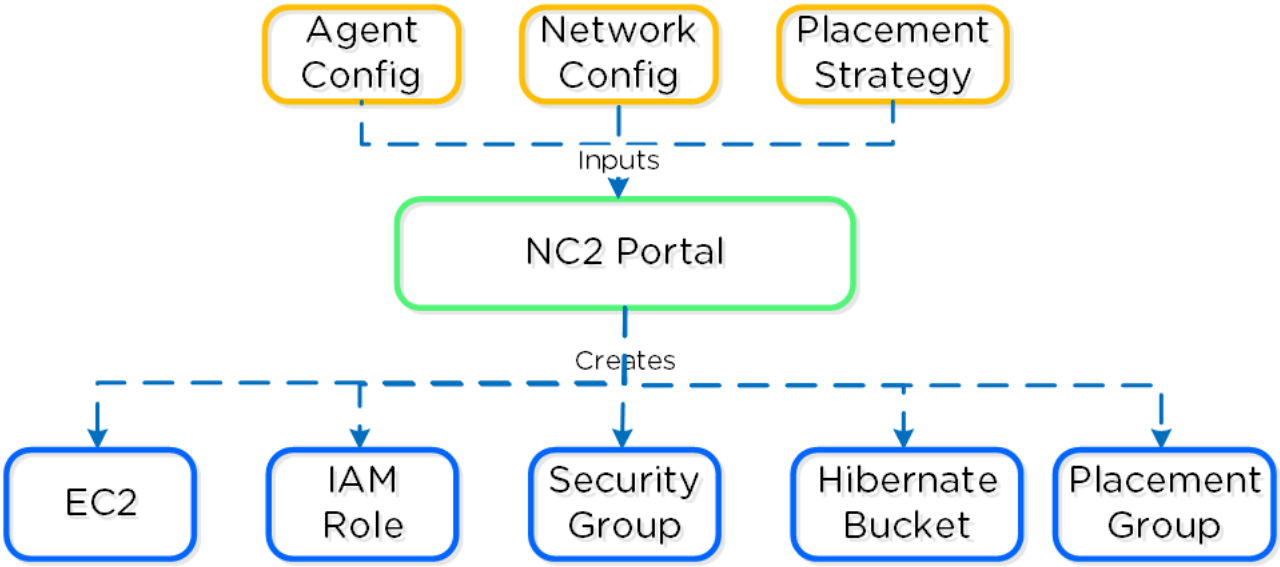

Figure 3: Cloud Clusters Portal

The NC2 portal provides the following services:

- Obtaining and managing bare-metal resources.
- Ensuring the correct IAM roles are created and used for deployment.
- Creating AWS security group rules to help lock down your AWS resources.
- Performing hibernate and resume operations, including S3 bucket creation.
- Managing node placement strategy and removing or adding nodes based on the health of the cluster.

3. Nutanix Cloud Networking

Nutanix can deliver a true hybrid multicloud experience because it has native cloud networking. Nutanix integration with the AWS networking stack means that every VM deployed on NC2 on AWS gets a native AWS IP address, so as soon as you migrate or create an application on NC2 on AWS, it has full access to all AWS resources. Integration also removes the burden of managing and deploying an additional network overlay. Because the Nutanix network capabilities are directly on top of the AWS overlay, network performance remains high and additional network controllers don't consume host resources.

With native network integration, you can deploy NC2 in existing AWS VPCs. As existing AWS environments have gone through change control and security processes already, you don't need to do anything except allow NC2 on AWS to talk to an NC2 portal. With this integration, you can increase security in your cloud environments.

Nutanix uses native AWS API calls to deploy AOS on bare-metal EC2 instances and consume network resources. Each bare-metal EC2 instance has full access to its bandwidth through an elastic network interface (ENI). For example, if you deploy Nutanix to an i3.metal instance, each node has access to 25 Gbps. With AHV, the ENI ensures that you don't need to set up additional networking high availability for redundant network paths to the top-of-rack switch.

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.

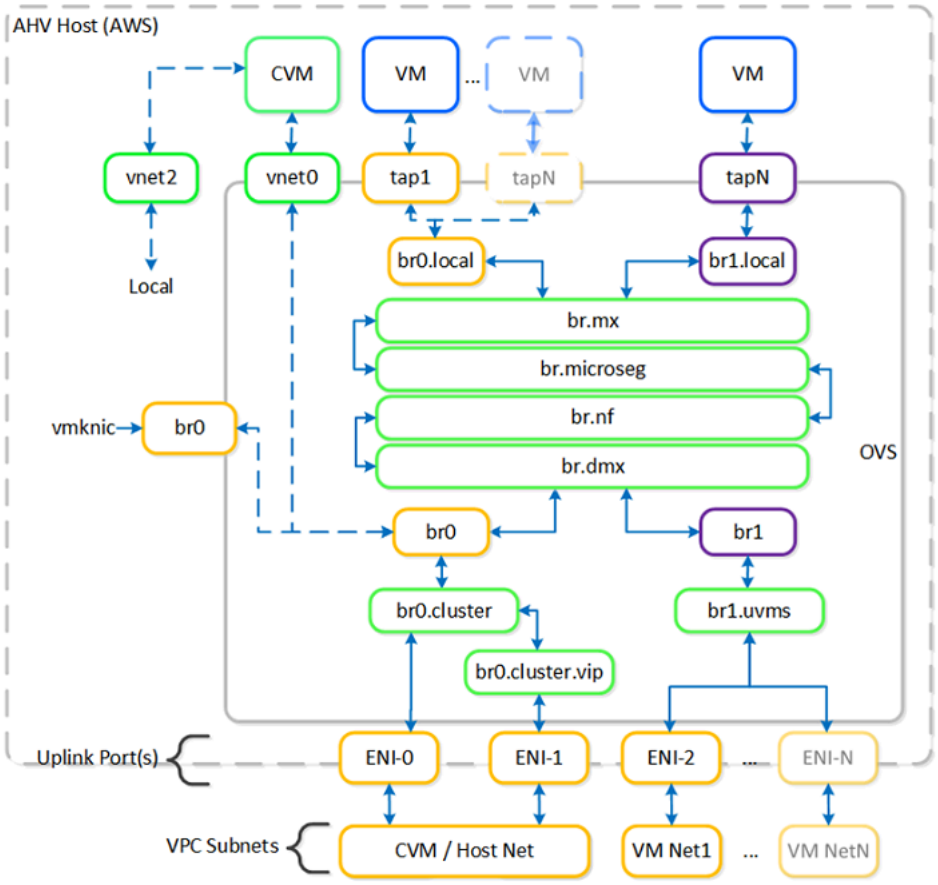

Figure 4: OVS Conceptual Architecture

When AOS runs on AWS, you can rely on the AWS overlay network to provide the best possible throughput automatically, leading to a very consistent and simple network configuration.

The ENI is a logical networking component in a VPC that represents a virtual network card. An ENI can have one primary IP address and up to 50 secondary IP addresses. All deployed user VMs (UVMs) use the secondary IP addresses to get direct AWS network access. AHV hosts deployed in AWS have separate ENIs for management traffic (AHV and CVM) and UVMs, which means customers can have different AWS security groups for management and UVMs. NC2 on AWS creates a single default security group for UVMs running in the Nutanix cluster. Any ENIs created to support UVMs are members of this default security group, which allows all UVMs in a cluster to communicate with each

other. In addition to the security groups, customers can use Nutanix Flow Network Security to provide greater security controls for east-west network traffic.

Creating a Subnet

A customer first creates a subnet in AWS in a VPC, then connects it to AOS in Prism Element. Cloud network, a new service in the CVM, works in conjunction with AOS configuration and assigns a VLAN ID (or VLAN tag) to the AWS subnet and fetches relevant details about the subnet from AWS. The network service keeps customers from using the AHV or CVM subnet for UVMs by not allowing them to create a network with the same subnet.

You can use each ENI to manage 49 secondary IP addresses. A new ENI is also created for each subnet you use. The AHV host, VMs, and physical interfaces use ports to connect to the bridges, and both bridges communicate with the AWS overlay network. Because each host already has the drivers needed for a successful deployment, you don't need to do any additional work to use the AWS overlay network. Just keep the following best practices in mind:

- Don't share AWS UVM subnets between clusters.
- Have separate subnets for management (AHV and CVM) and user VMs.
- If you plan to use VPC peering, use nondefault subnets to ensure uniqueness across AWS Regions.
- Divide your VPC network range evenly across all usable Availability Zones in a Region.
- In each Availability Zone, create one subnet for each group of hosts that has unique routing requirements (for example, public versus private routing).
- Size your VPC CIDR and subnets to support significant growth.

Guest AHV IP Address Management

AHV uses IP address management (IPAM) to integrate with native AWS networking. NC2 on AWS uses the native AHV IPAM to inform the AWS DHCP server of all IP address assignments using API calls. NC2 relies on AWS to send

gratuitous Address Resolution Protocol (ARP) packets for any additions to an ENI's secondary IP addresses. We rely on these packets to ensure that each hypervisor host is notified when an IP address moves or new IP addresses become reachable. For user VMs, you can't share a given AWS subnet between two NC2 on AWS deployments. You can, however, use the same management subnet (AHV and CVMs) for multiple clusters.

We added a new service called the cloud network controller (CNC) to the AHV host to help with ENI creation. CNC runs an OpenFlow controller, which manages the OVS in the AHV hosts and handles mapping, unmapping, and migrating UVM secondary IP addresses between ENIs or hosts. A subcomponent of CNC called cloud port manager provides the interface and manages AWS ENIs.

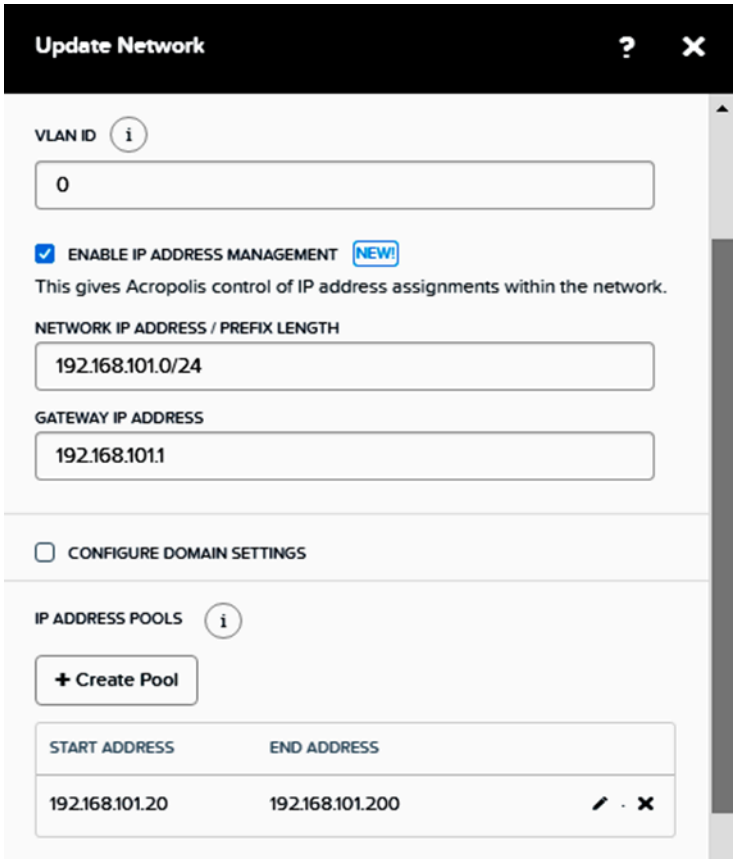

Figure 5: IPAM with AWS

IPAM avoids address overlap by sending AWS API calls to inform AWS which addresses are being used.

The AOS leader assigns an IP address from the address pool when it creates a managed vNIC and releases the address back to the pool when the vNIC or VM is deleted.

Note: You can't use or assign the first four IP addresses or the last IP address in each subnet.

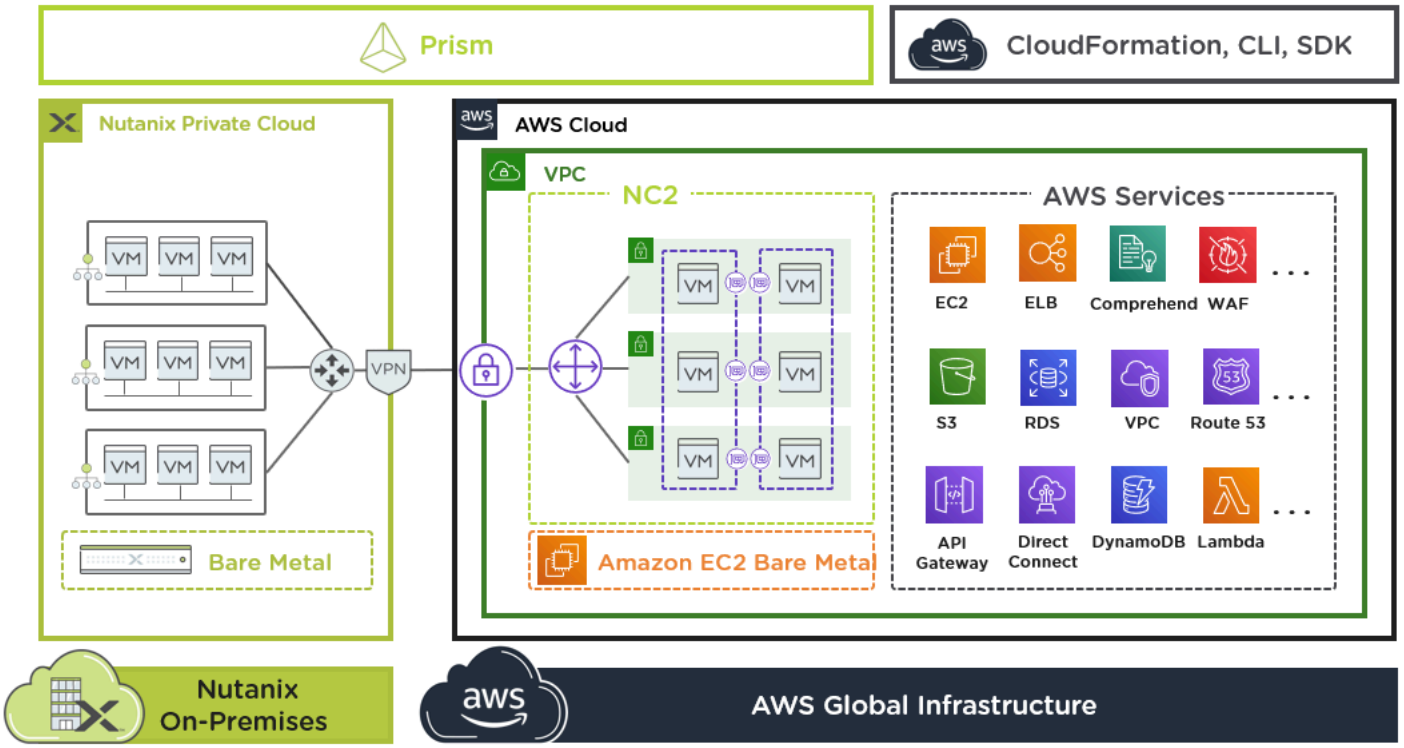

Figure 6: AWS Networking

Using native AWS networking allows you to quickly establish connectivity and eliminate costly performance impacts from third-party networking overlays. Cloud administrators can focus on their tasks instead of managing additional networking technologies. Let's walk through a typical deployment to see the AWS constructs in action.

- Click Create Cluster in the NC2 portal.
- Provide the name and URL name for the cluster.

- Select AWS as the cloud provider.
- Fill in the other information for your specific cluster, then click Validate Network Configuration to test your configuration.
- Click Next.

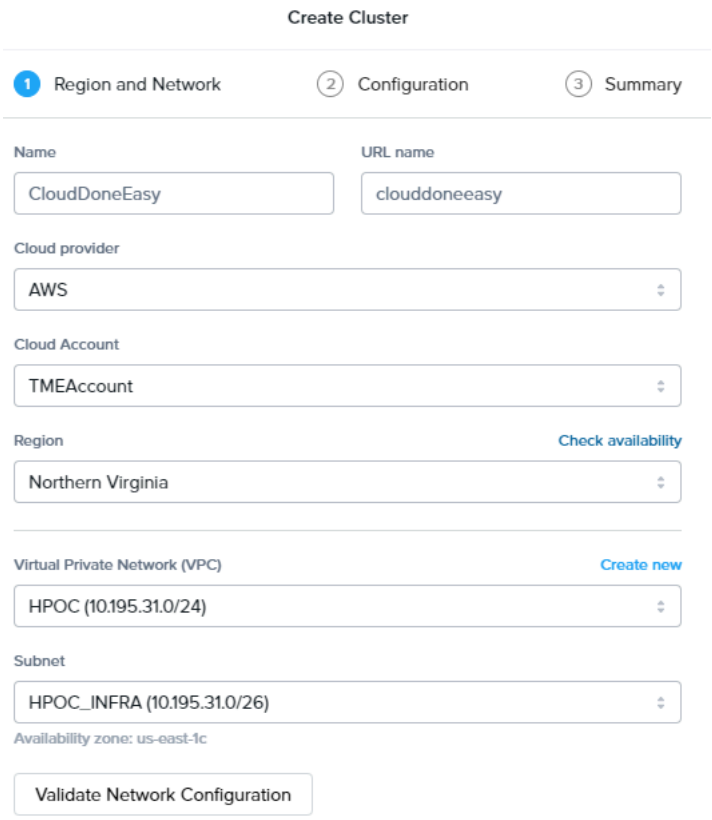

Figure 7: Cluster Deployment Network on AWS

- In the Configuration section, select your AOS version.
- Configure the Cluster Capacity as desired, and make sure you select Disabled for Prism access from the public internet.
- We selected Terminate at a specific time for Scheduled Account Termination because we're using this instance as a test. Use the setting that best suits your cluster needs.

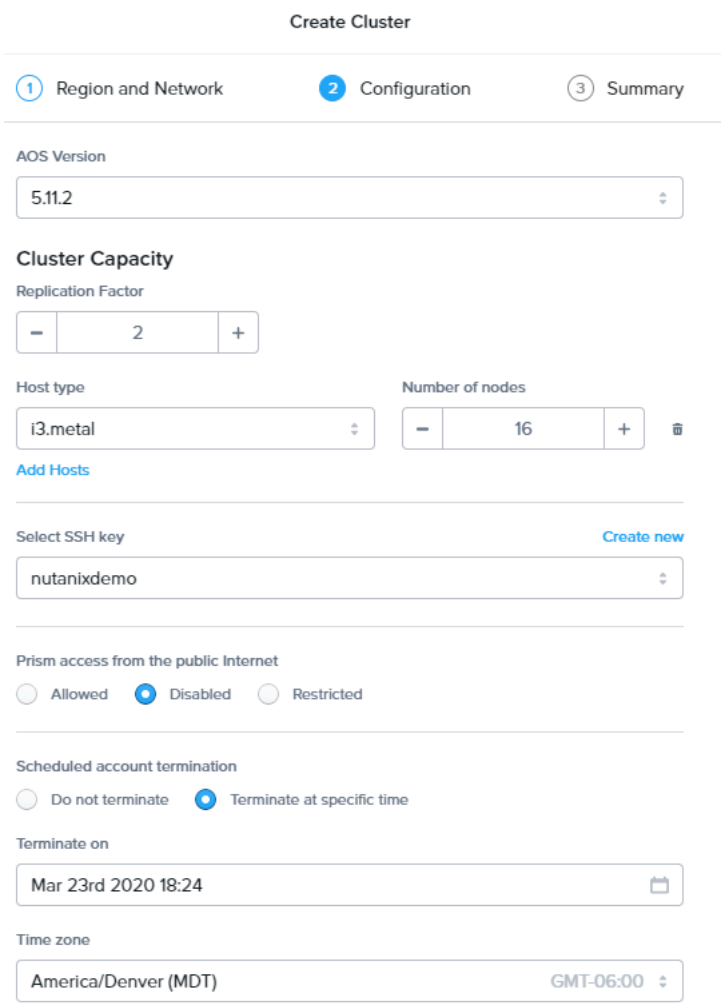

Figure 8: Cluster Deployment Configuration on AWS

After you fill in both parts of the configuration, click Submit. Your provisioned cluster is available 30 to 40 minutes later. Native integration with the AWS network provides flexibility and makes it easy to access other AWS services.

You can also either use an existing AWS subnet that follows your architectural standards or create a new one. The subnet you select or create becomes the management subnet and provides the IP addresses for the hypervisor (AHV) host and the CVM.

Once you deploy the cluster, you can set up a VPN gateway in AWS and create a site-to-site VPN connection. The following figure shows a high-level overview of a VPN connection for a typical NC2 on AWS deployment.

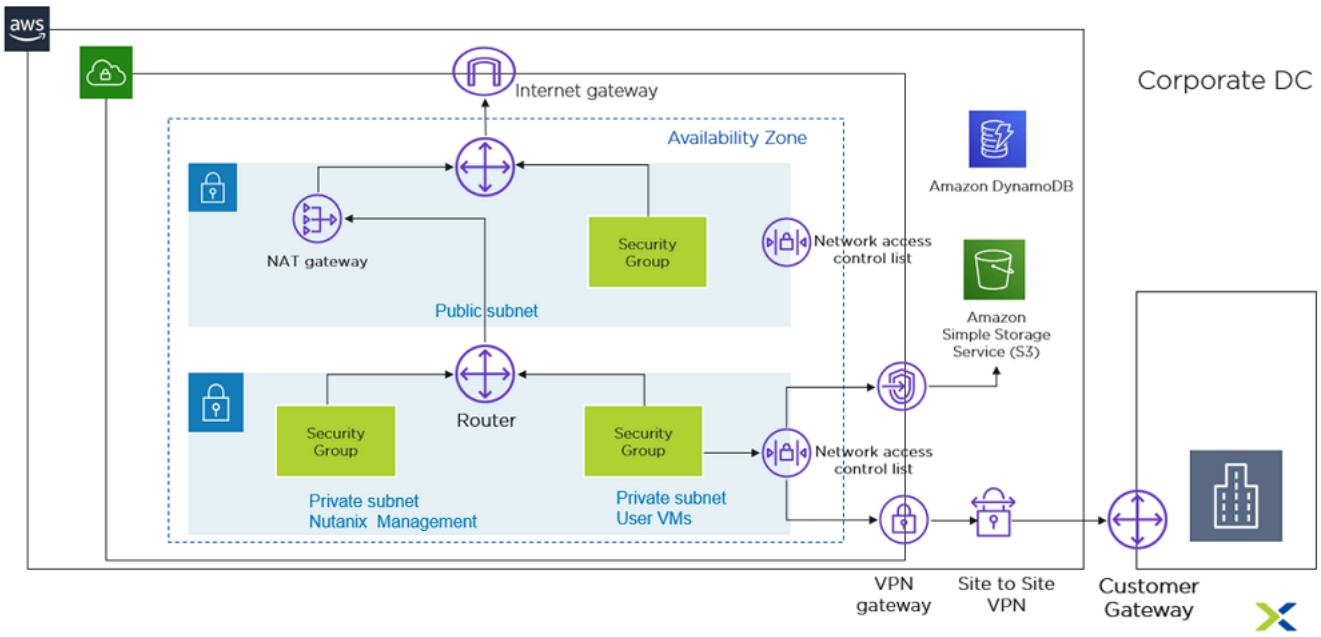

Figure 9: VPN Connection

For a successful deployment, NC2 needs outbound access to the NC2 portal, either using an internet gateway or an on-premises VPN with outbound access. Your Nutanix cluster can sit in a private subnet that can only be accessed from your VPN, limiting exposure to your environment. Ensure that redundant paths are available for outbound internet access, as you use the NC2 portal to add and remove AWS nodes based on the health of the system.

You can use AWS security groups and network access control lists to secure your cluster relative to other AWS or on-premises resources. Nutanix automatically creates three security groups to limit traffic to the cluster:

1. Internal management: Allows all internal traffic between all CVMs and all AHV hosts (EC2 bare-metal hosts). Don't edit this group without approval from Nutanix Support.
2. User management: Allows users to access Prism Element and some other services running on the CVM.

3. UVM: Allows UVMs to talk to each other. By default, all UVMs on all subnets can talk to each other, but you can edit the policy to lock down more traffic. You could alternatively use Flow Network Security to prevent east-west traffic.

You can also use existing groups in your environment. In the security groups settings in the following figure, ports 2009 and 2020 support replication to an on-premises Nutanix cluster.

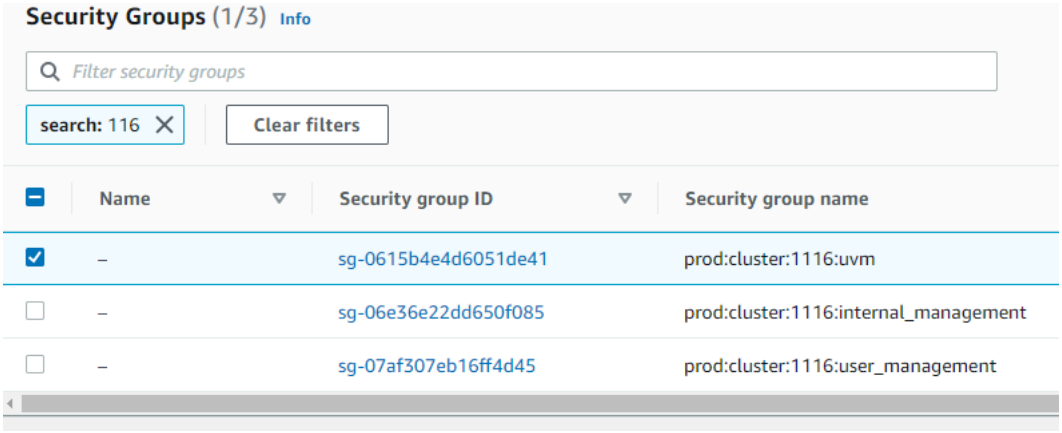

Figure 10: AWS Security Groups

With AWS security groups, you can choose to allow access to the AWS CVMs and AHV host only from your on-premises management network and CVMs. You can limit replication from on-premises to AWS at the granularity of the specific port. Because all the replication software is embedded in the CVM on both sides, you can easily migrate your workloads back and forth.

The simplicity of NC2 can save you money. Because you don't need any additional overlay networks, you save on the cost of additional compute that overlays require. You avoid the costs for management gateways, network controllers, edge devices, and storage incurred from adding appliances. With a simpler system, you also realize significant operational savings on maintenance and troubleshooting.

4. Migration

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

Native data protection

You can use this method for ESXi- and AHV-based clusters. Creating a remote site for your new NC2 on AWS deployment and setting up the native networking integration only takes a few minutes; you simply need to ensure that the ports are open on the management security group you need for replication. All the existing data protection best practices apply because a bare-metal AWS deployment essentially acts as an additional supported original equipment manufacturer (OEM).

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 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 AWS 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 on AWS, which is important when you need to migrate or do testing and development work.

5. Storage Availability in AWS

An AWS Region is a distinct geographic area. Each Region has multiple, isolated locations known as Availability Zones. An Availability Zone is a logical datacenter available for any AWS 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 AWS 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. Using up to seven partitions, Nutanix places the AWS bare-metal nodes in different AWS racks and stripes new hosts across the partitions.

NC2 on AWS supports combining heterogenous node types in a cluster. You can deploy a cluster of one node type and then expand that cluster's capacity by adding heterogenous nodes to it. This feature protects your cluster if its original node type runs out in the Region and provides flexibility when expanding your cluster on demand. If you're looking to right-size your storage solution, support for heterogenous nodes can give you more instance options to choose from.

> Note: Nutanix recommends keeping the minimum number of additional nodes greater than or equal to your cluster's redundancy factor. You should expand the cluster in multiples of your redundancy factor for the additional nodes.

When combining instance types in a cluster, you must always maintain at least three nodes of the original type you deployed the base cluster with. You can expand or shrink the base cluster with any number of heterogenous nodes if at least three nodes of the original type remain and the cluster size stays within the limit of 28 nodes.

The following table and figure both refer to the cluster's original instance type as Type A and its compatible heterogenous type as Type B.

Table: Supported Instance Type Combinations

Type A	Type B
i3.metal	i3en.metal
i3en.metal	i3.metal
z1d.metal	m5d.metal
m5d.metal	z1d.metal

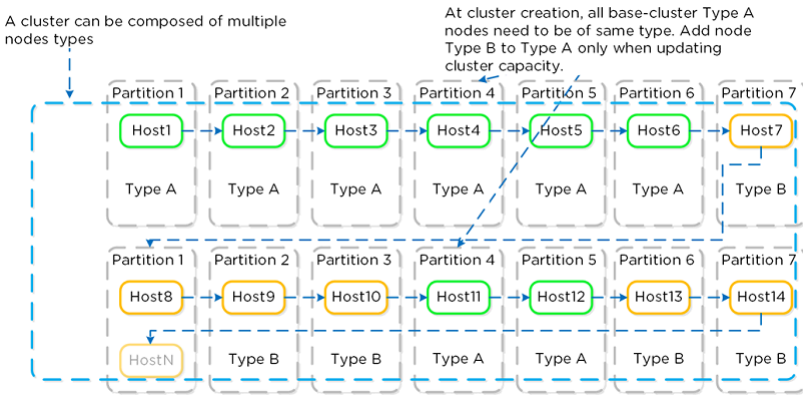

Figure 11: Partition Placement

When you've 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.

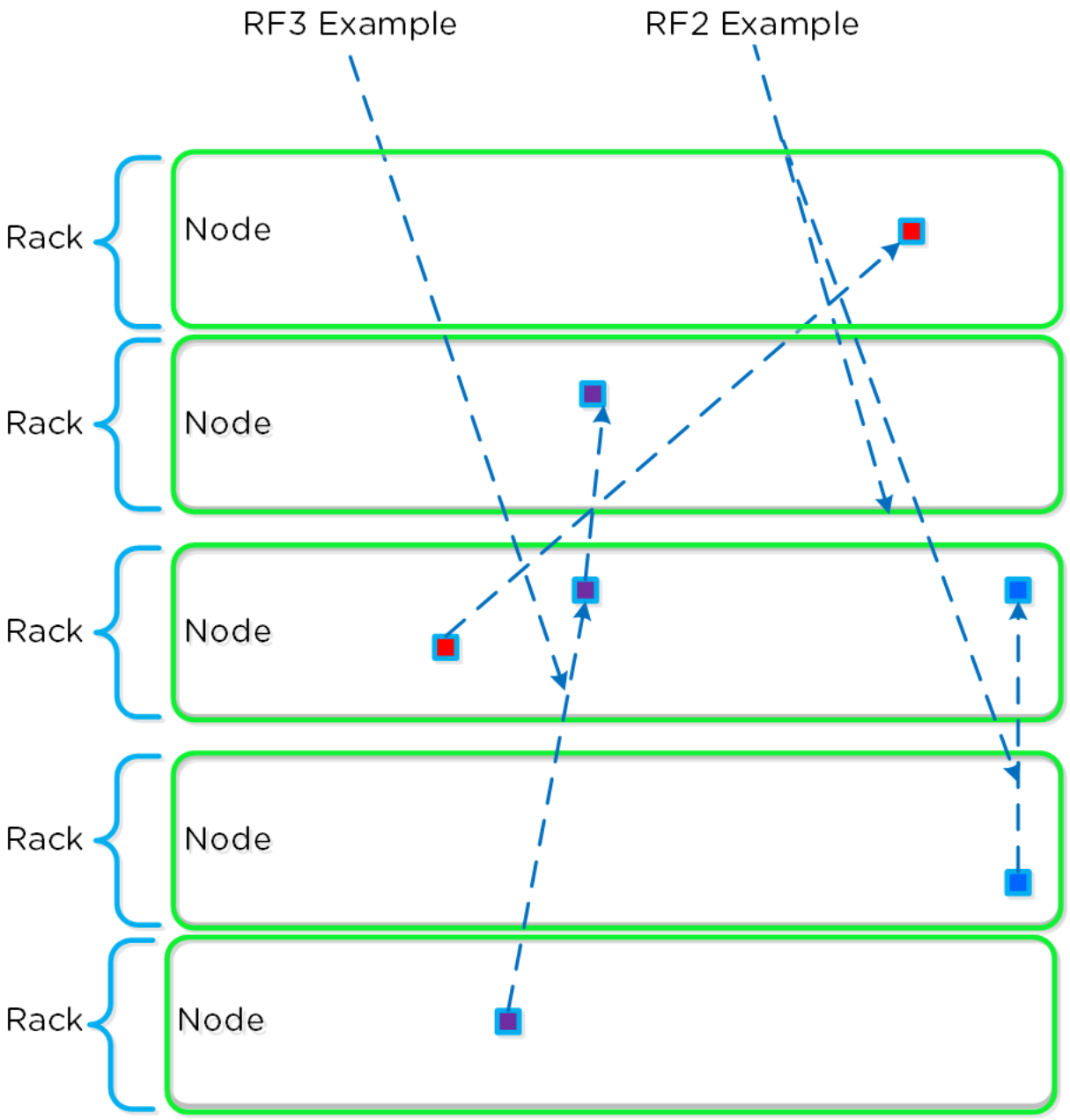

Figure 12: Replication Factor Data Placement Across Racks

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 calculates a parity bit 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

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.

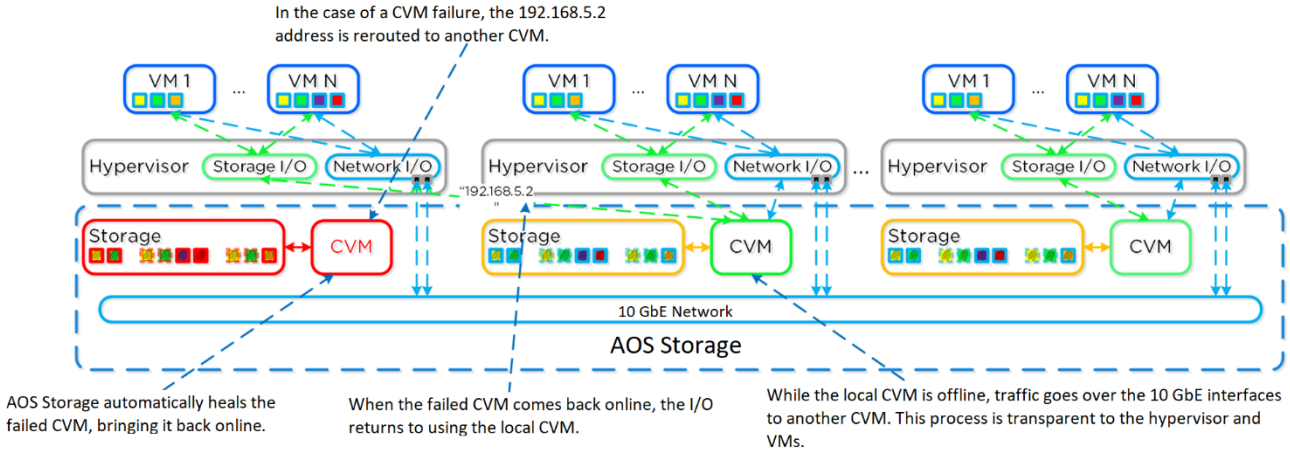

Figure 13: Data Path Redundancy

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 fails or packets become corrupt, leading to a scenario where nodes disagree on values. AOS storage also uses timestamps to ensure that updates are applied in the proper order.

Proactively Resolve Bad Disk Resources

AOS 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.

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 EC2 removal process. Removing the EC2 instance triggers an API call to the cluster portal, which notifies the NC2 portal. The NC2 portal allocates a new instance, adds it to the cluster, and marks the EC2 instance with the unresponsive disk for removal. The cluster software automatically replicates the data on the bad EC2 instance to other instances, then deletes the bad EC2 instance.

Maintain Availability: Availability Zone Failure

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

6. Hibernate and Resume

NC2 on AWS has the unique ability to preserve customer data while bare-metal nodes are shut down. Hibernation can save you money when the cluster isn't in use—for example, in test or development environments that aren't used on the weekends or for disaster recovery when recovery point objectives (RPOs) are a day or more.

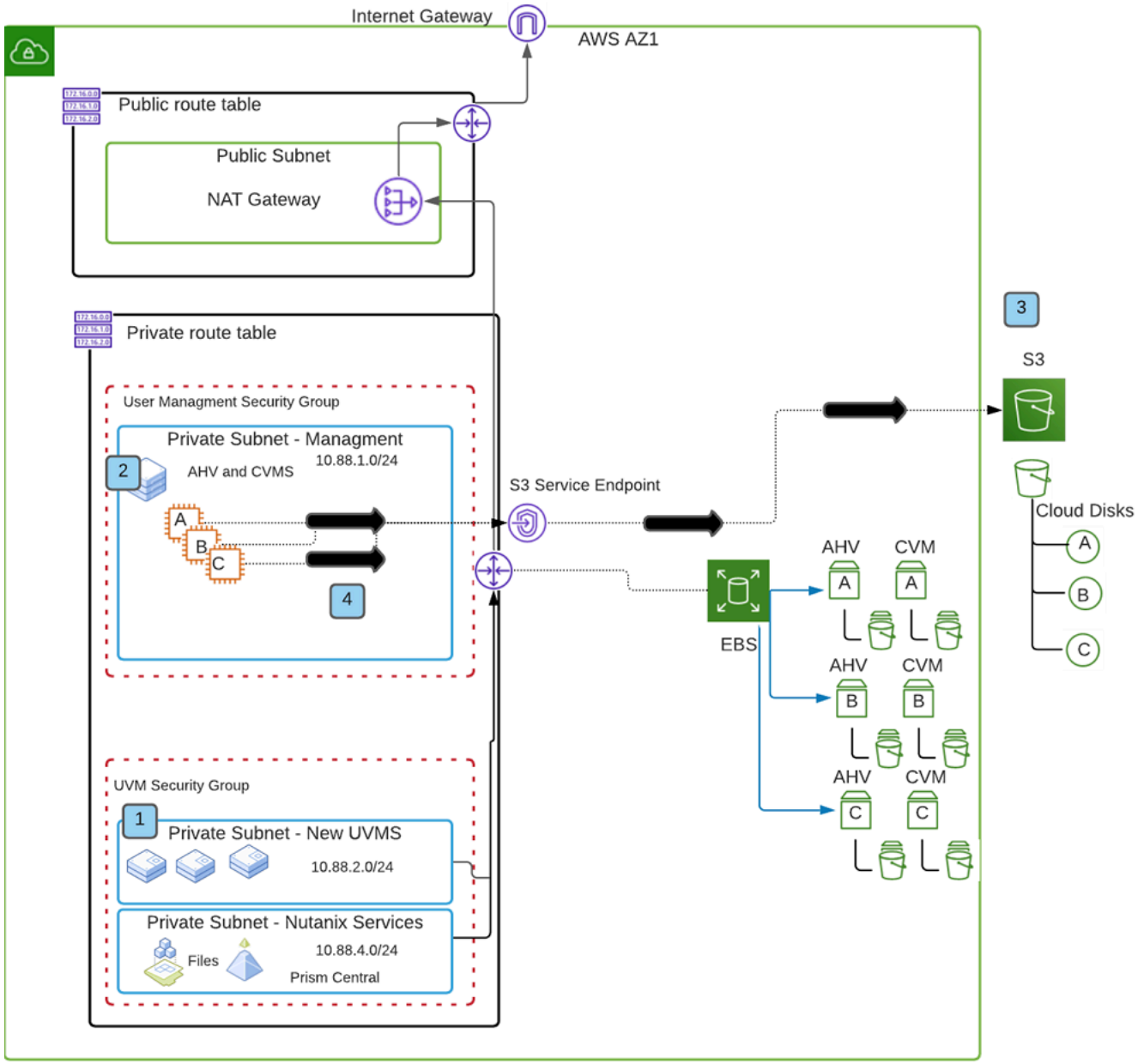

Figure 14: Hibernation

The following steps give an overview of the hibernation process:

Note: Shut down all the VMs on the cluster before starting hibernation.

1. NC2 on AWS runs prechecks against the cluster to verify that no VMs, upgrades, or other workflows (such as cluster expansion) are running.
2. NC2 on AWS creates and adds at least one S3 disk per node or CVM in the cluster. The S3 disks form a cloud storage tier used by the same internal process that tiers data between SSD and HDD for on-premises clusters.
3. NC2 on AWS puts all hosts into maintenance mode, which prevents UVM starts and stops all I/O operations. Curator and Stargate first migrate extent store data to the cloud, maintaining replication factor 1 (a single copy), then migrate Cassandra metadata.
4. NC2 on AWS protects the cluster configuration and state information maintained on CVM boot disks during hibernation by snapshotting the respective EBS volumes.
5. NC2 on AWS releases the bare-metal nodes back to the AWS pool.

If your cluster is hibernating, you can bring the clusters back online with the resume operation. Resume deploys the nodes you need into the original AWS region and migrates your data back to them.

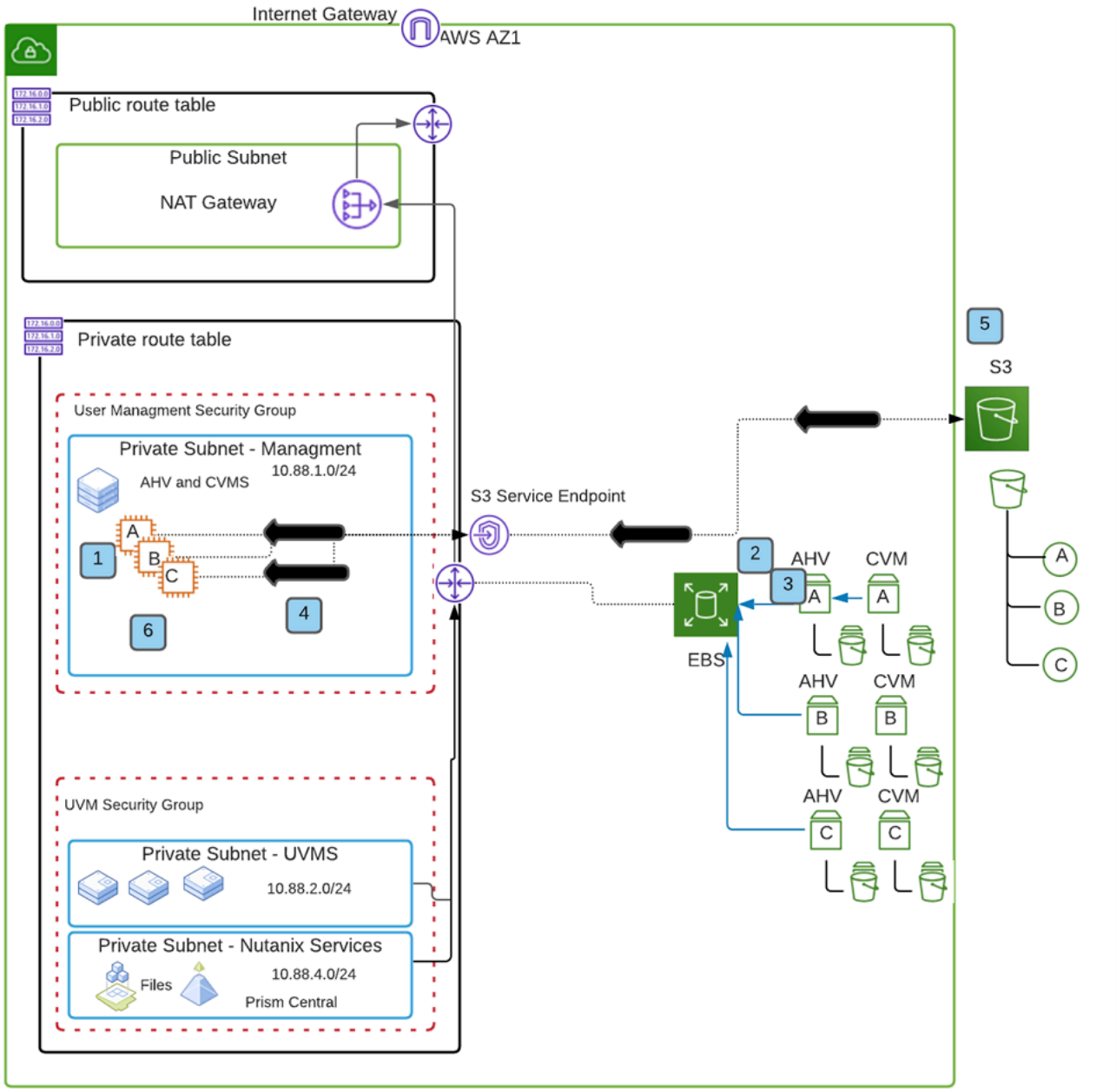

Figure 15: Resume Operation

The following steps give an overview of the resume process:

1. When you start the resume process, NC2 on AWS deploys new EC2 instances and attaches cloud disks to them to form the AWS Nutanix cluster.
2. NC2 on AWS restores the cluster boot disk's Zeus configuration from the EBS volume snapshot.
3. NC2 on AWS restores the cluster configuration from the EBS volume snapshot.
4. Genesis starts the necessary cluster services, and the Cassandra dynamic ring changer restores metadata.
5. NC2 on AWS creates and adds at least one S3 disk per node or CVM in the cluster. The S3 disks form a cloud storage tier.
6. Curator restarts and registers the kSelectiveClusterHibernate scan for the restore operation. Curator schedules the selective hibernate scan to migrate extent store data from S3 to NVMe.
7. Genesis transitions to kNormal mode once all the data is restored. NC2 on AWS marks the hibernated disks as `to-remove` and removes them from the cluster configuration.

Data Throughput

Moving data between the cluster in AWS and S3 occurs in four main stages: metadata migration, data migration, AOS processing, and NC2 portal processing. We tested throughput on a three-node i3.metal cluster with VMs consuming 18 virtual disks (vDisks) with a total usable capacity of 9.67 TB and 4.74 GB of metadata. The following list provides the test details and shows the timing results for all phases:

- Test details:
  - Three-node cluster i3.metal
  - 18 vDisks with 250 GB each
  - Total capacity consumed by the VMs = 9.74 TB
  - Total metadata size on cluster = 4.74 GB
- Metadata migration phase: 34 minutes
- Data migration phase: 59 minutes

- Total AOS processing time: 1 hour, 38 minutes
- Total NC2 portal processing time: 1 hour, 50 minutes

7. Deployment Models

As customer environments can vary greatly, we provide several example deployment models. Regardless of the model you use, there are a few general outbound requirements for deploying a Nutanix cluster in AWS on top of 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.

> 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://portal.nutanix.com/*	TCP 443 (HTTPS)	Nutanix service portal
Management subnet	https://gateway-external-api.cloud.nutanix.com	TCP 443 (HTTPS)	NC2 portal orchestration
Management subnet	https://downloads.cloud.nutanix.com/clusters/*	TCP 443 (HTTPS)	Download NC2 RPM Package Manager (RPM) packages
Management subnet	https://downloads.cloud.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	169.254.169.123	TCP 443 (HTTPS)	NTP server

Source	Destination	Protocol	Purpose
Management subnet	169.254.169.254/*	TCP 443 (HTTPS)	Access instance metadata related to the AWS service role

Table: Cluster Outbound to EC2

Source	Destination	Protocol	Purpose
Management subnet	ec2.<region>.amazonaws.com (for example, a cluster in us-west-2 requires ec2.us-west-2.amazonaws.com)	TCP 443 (HTTPS)	Access AWS metadata
Management subnet	aws.amazon.com	TCP 443 (HTTPS)	Access AWS metadata

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

Multicluster Deployment

To protect your NC2 on AWS cluster in the event of an Availability Zone failure, use your existing on-premises NC2 instance as a disaster recovery target. There are many options when it comes to Nutanix disaster recovery, but here we're focusing on the native data protection included with every base Nutanix license.

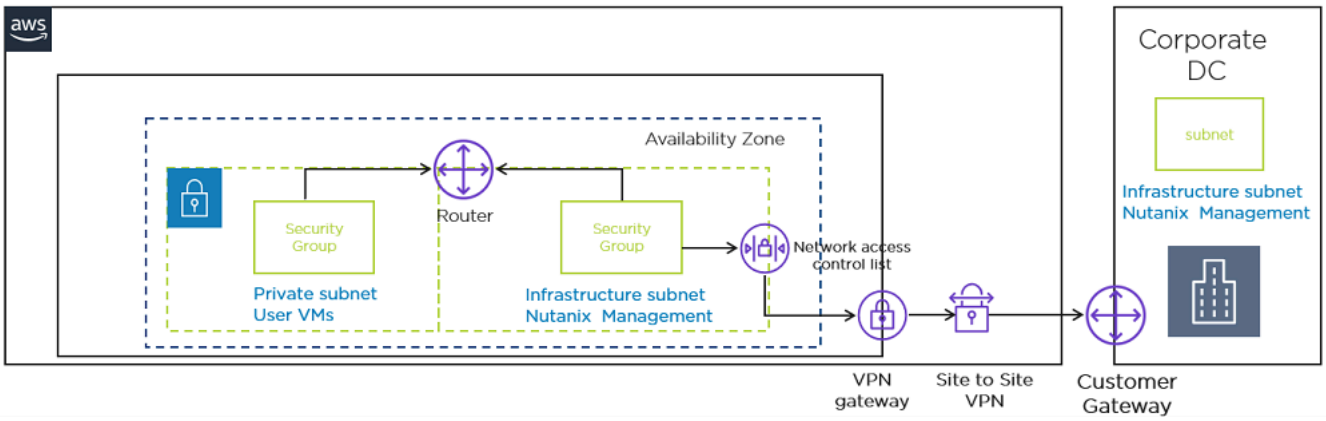

Figure 16: Multicluster Deployment

The following table details the inbound ports you need to establish replication between an on-premises cluster and a Nutanix cluster running in AWS. You can create these ports on the infrastructure subnet security group that was automatically created when you deployed NC2 on AWS. The ports need to open in both directions.

Table: Inbound Security Group Rules for AWS

Type	Protocol	Port Range	Source	Description
SSH	TCP	22	On-premises CVM subnet	SSH into the AHV node
Custom TCP rule	TCP	2222	On-premises CVM subnet	SSH access to the CVM
Custom TCP rule	TCP	9440	On-premises CVM subnet	UI access
Custom TCP rule	TCP	2020	On-premises CVM subnet	Replication
Custom TCP rule	TCP	2009	On-premises CVM subnet	Replication

Make sure you set up the cluster virtual IP address for your on-premises and AWS clusters. This IP address is the destination address for the remote site.

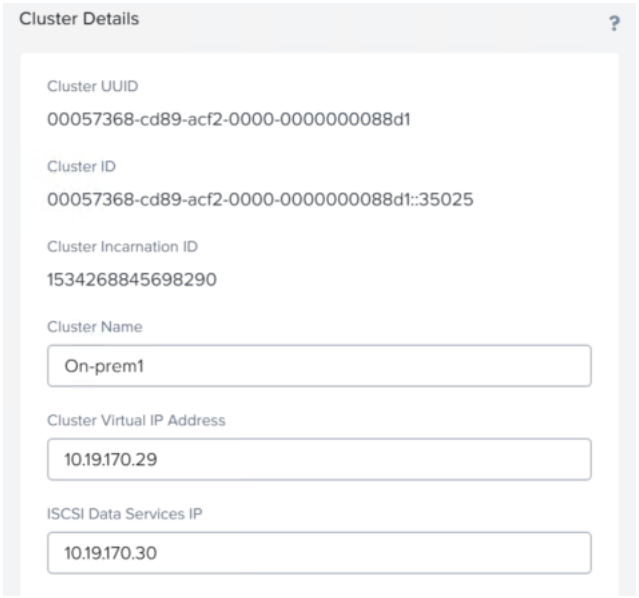

Figure 17: Cluster Details

Near-synchronous (NearSync) and asynchronous replication are both valid options for customers running AHV on-premises. You can set your RPO to be as little as one minute with NearSync and one hour with asynchronous replication. If you want to use NearSync, your on-premises cluster needs to meet the requirements listed in the Requirements of Data Protection with NearSync Replication section of the Data Protection and Recovery with Prism Element guide.

Multiple Availability Zone Deployment

If you don't have an on-premises cluster available for data protection or you want to use the low-latency links between Availability Zones, you can create a second NC2 on AWS deployment in a different Availability Zone. With this method, there is no data transfer charge between Amazon EC2 and other Amazon Web Services in the same AWS Region (for example, between Amazon EC2 US West and Amazon S3 US West), which can be a significant benefit.

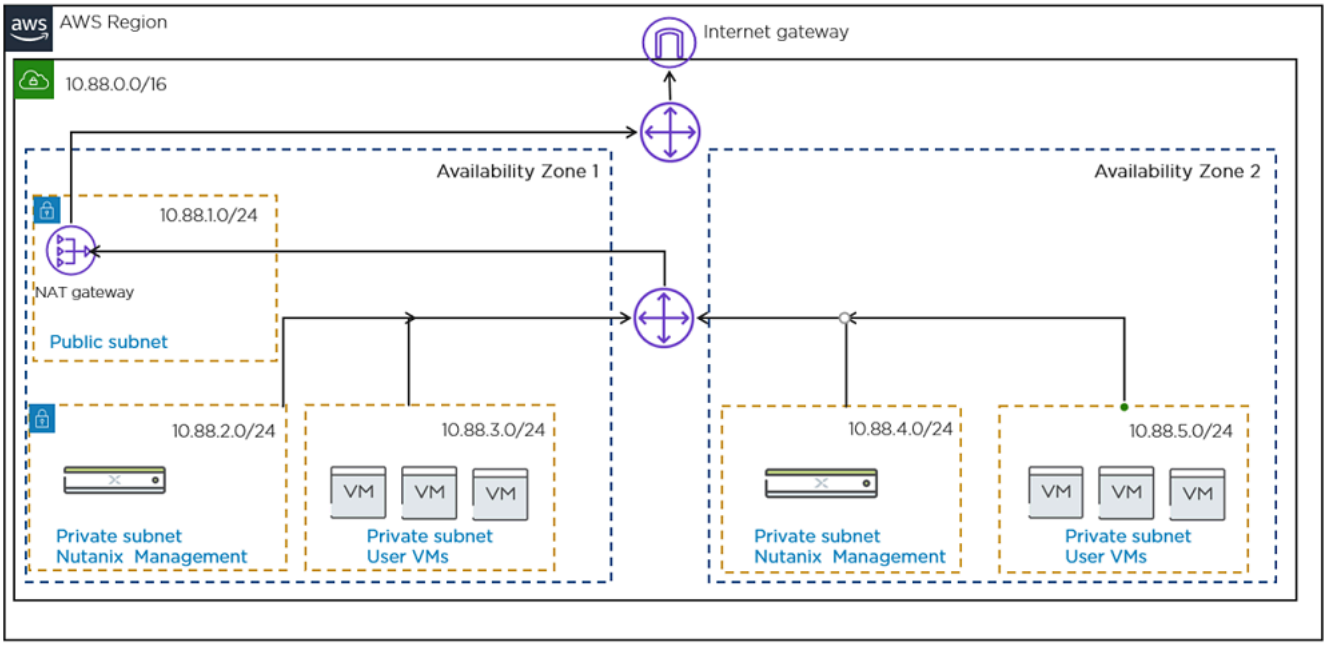

Figure 18: Multiple Availability Zone Deployment

Because you're still in your own custom VPC, your network design is very simple. You can isolate your private subnets for UVMs between clusters and use the private Nutanix management subnets to allow replication traffic between them. All private subnets can share the same routing table. Edit the inbound access in each Availability Zone's security group as shown in the following tables to allow replication traffic.

Table: Availability Zone 1 NC2 on AWS Security Group Settings

Type	Protocol	Port Range	Source	Description
Custom TCP rule	TCP	9440	10.88.4.0/24	UI access
Custom TCP rule	TCP	2020	10.88.4.0/24	Replication
Custom TCP rule	TCP	2009	10.88.4.0/24	Replication

Table: Availability Zone 2 NC2 on AWS Security Group Settings

Type	Protocol	Port Range	Source	Description
Custom TCP rule	TCP	9440	10.88.2.0/24	UI access
Custom TCP rule	TCP	2020	10.88.2.0/24	Replication
Custom TCP rule	TCP	2009	10.88.2.0/24	Replication

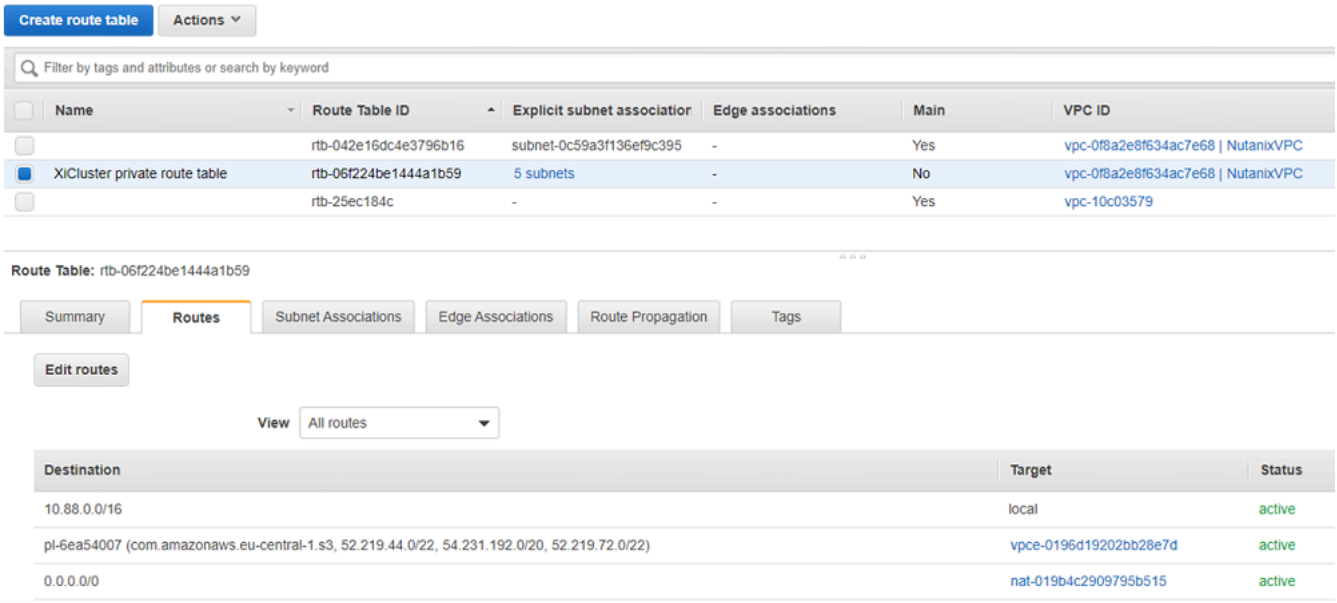

Figure 19: Private Route Table

If Availability Zone 1 goes down, you can activate protected VMs on the cluster in Availability Zone 2. Once Availability Zone 1 comes back online, you can redeploy a Nutanix cluster in Availability Zone 1 and reestablish data protection. New clusters require full replication.

Single Availability Zone Deployment

Deploying a single cluster in AWS is great for more ephemeral workloads where you want to take advantage of performance improvements and use the same automation pipelines you use on-premises. If you decide not to use native Nutanix data protection to back up your workloads to your on-premises

cluster or a second NC2 on AWS instance, you must use another method; if your Availability Zone fails, you aren't guaranteed the same EC2 instances when it resumes operation. You can use AHV-compatible backup products to target S3 as the backup destination, and, depending on the failure mode you want to recover from, you can also replicate that S3 bucket to a different Region. HYCU and Veeam are two backup methods proven to work with AHV.

HYCU

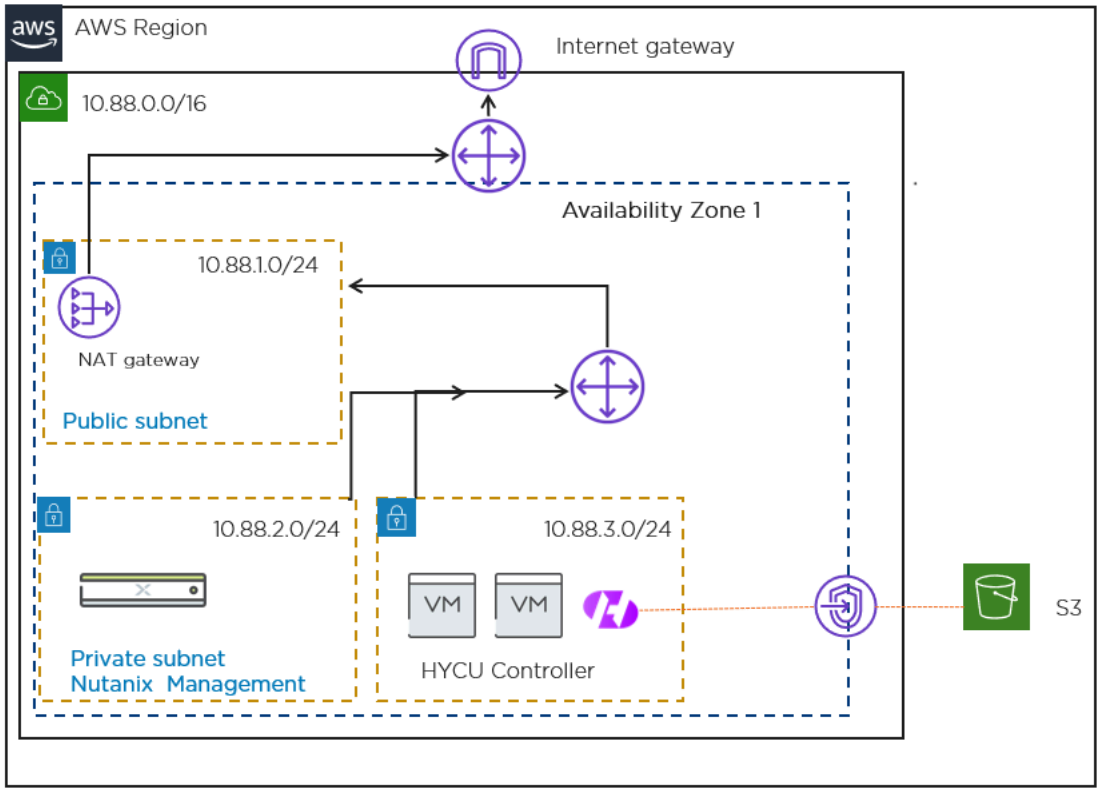

Figure 20: HYCU Backup Architecture

HYCU has the following limitations:

- HYCU doesn't support AWS S3 targets that use the Glacier storage class.
- HYCU currently only supports AWS S3 Signature Version 4.

The HYCU backup controller runs on the deployed Nutanix cluster as a VM. It needs to communicate with the private subnet to register with Prism Element.

Once the HYCU controller is running on the Nutanix cluster, you can configure the S3 endpoint.

> Note: Deploy a VPC endpoint for S3 so your backup traffic doesn't go over the internet.

The following procedure walks you through creating an S3 endpoint:

- Open the Amazon VPC console. In the navigation pane, choose Endpoints.
- The opened page asks you to create your first endpoint. Click Create Endpoint.
- Choose your VPC and specify a policy that controls access to the AWS service. Allow full access; if you don't, only IAM users can access the service.
  - If you want to create a locked user profile, you must specify at least these AWS S3 permissions: s3:GetObject, s3:DeleteObject, s3:PutObject, s3:ListBucket, s3:GetBucketAcl, s3:ListBucketMultipartUploads, and s3:GetBucketLocation.
- Associate the endpoint with your private subnet.

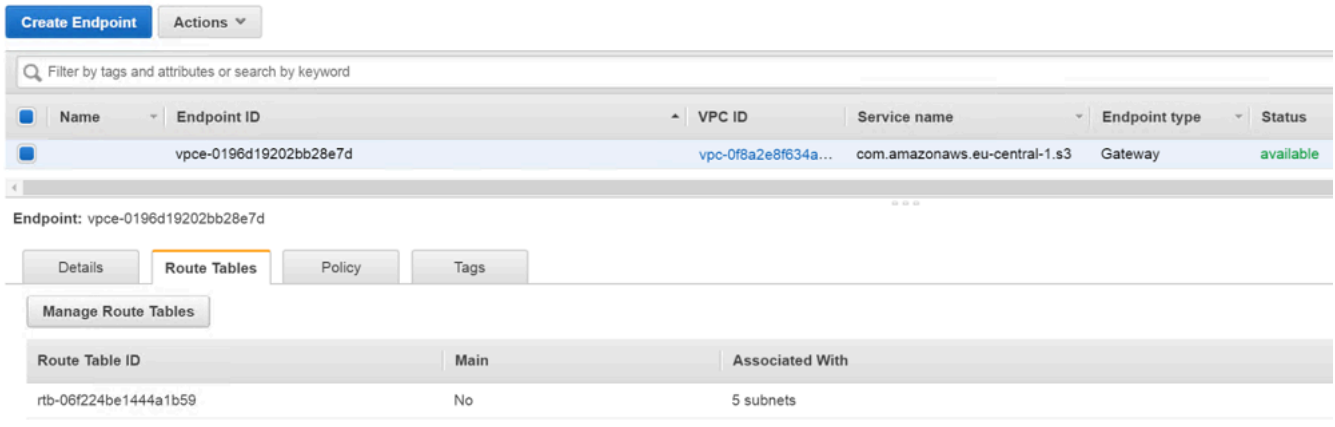

Figure 21: S3 VPC Endpoint

You can use the VPC endpoint to block all public access to your public S3 bucket.

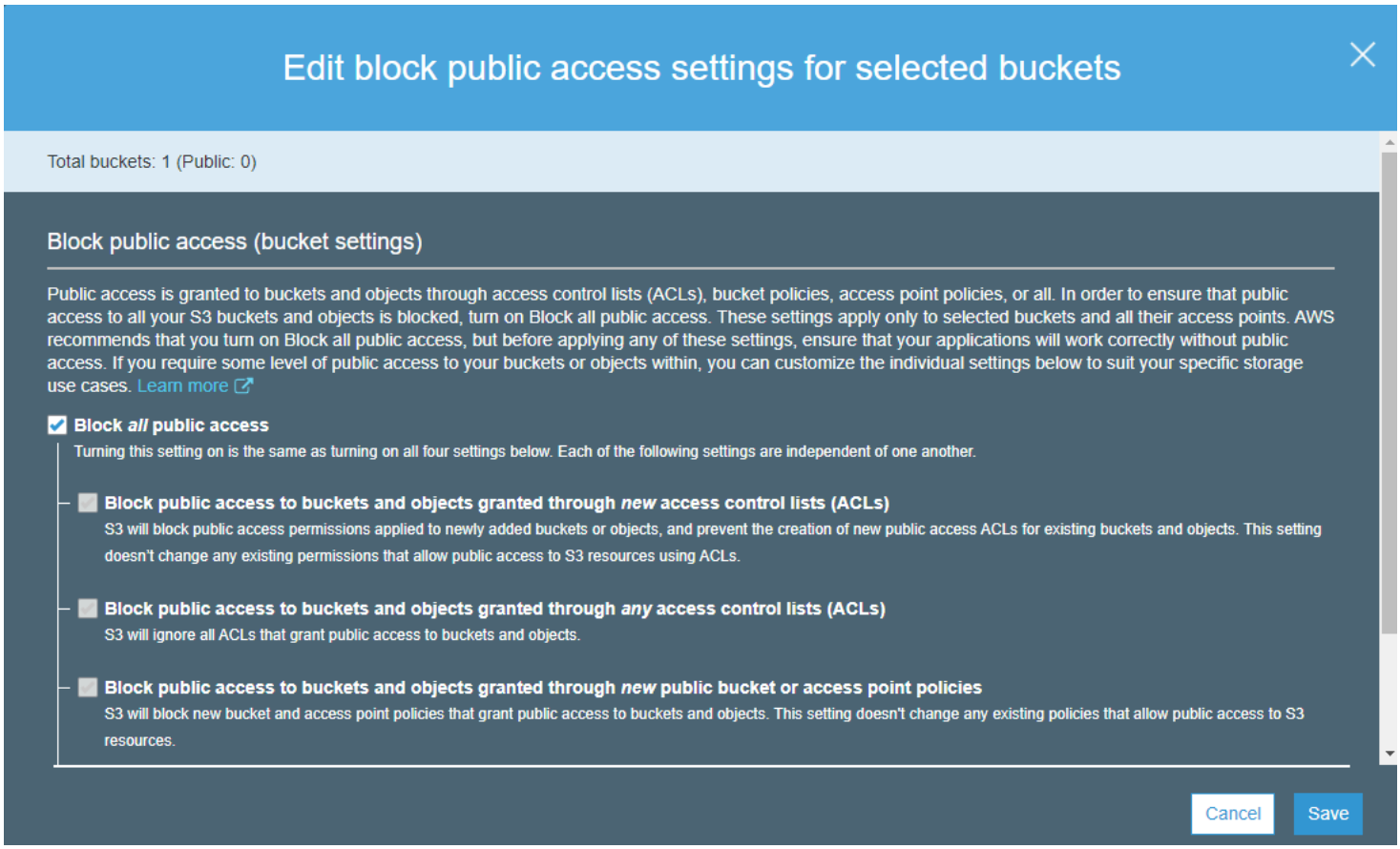

Figure 22: Block Public Access to Backup S3 Bucket

To ensure that you can restore your VMs to a new NC2 on AWS instance in the case of an unrecoverable NC2 on AWS failure:

- Back up the VMs and the HYCU controller to an S3 target.
- Once you deploy a new NC2 on AWS instance, create a temporary HYCU controller on that cluster.
- Before you import existing targets, suspend all other activities on the controller.
- For faster restores, put the HYCU controller backup in its own S3 bucket.

Follow these steps to restore your VMs to the new NC2 on AWS instance:

- Deploy a temporary HYCU backup controller.

- Import the targets that store the backup of the original HYCU backup controller.
- Add the new NC2 on AWS instance as the target location for restoring your HYCU backup controller. If you plan to restore VMs and applications, add locations for those as well.

You can find more detailed instructions for restoring VMs in the HYCU support documentation.

Replicate your S3 bucket to a different Region. AWS provides cross-Region replication (CRR) to copy objects in Amazon S3 buckets across Regions. With this tool, you can deploy NC2 on AWS in the new Region and replicate your S3 bucket to the original site once the source Region comes back. To use CRR, you need to enable versioning on your source and destination S3 buckets, as shown in the following image.

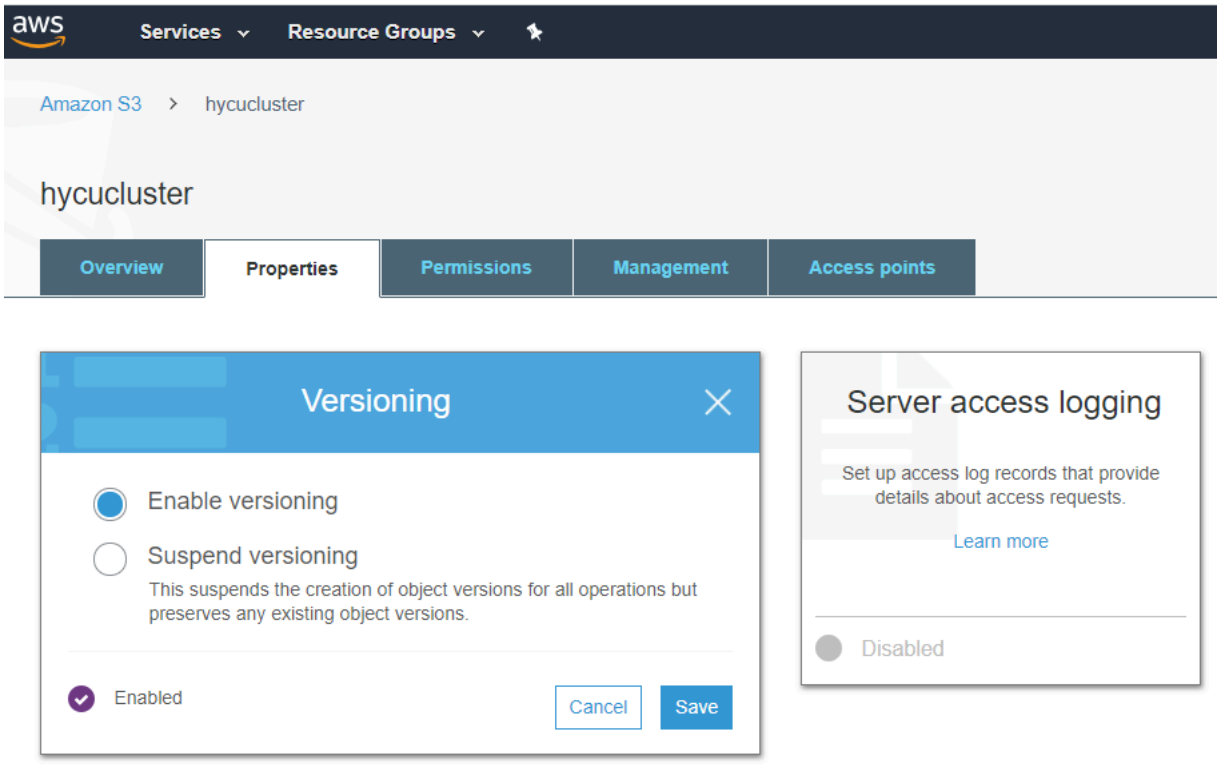

Figure 23: Enable Versioning to Use CRR

Veeam Backup & Replication

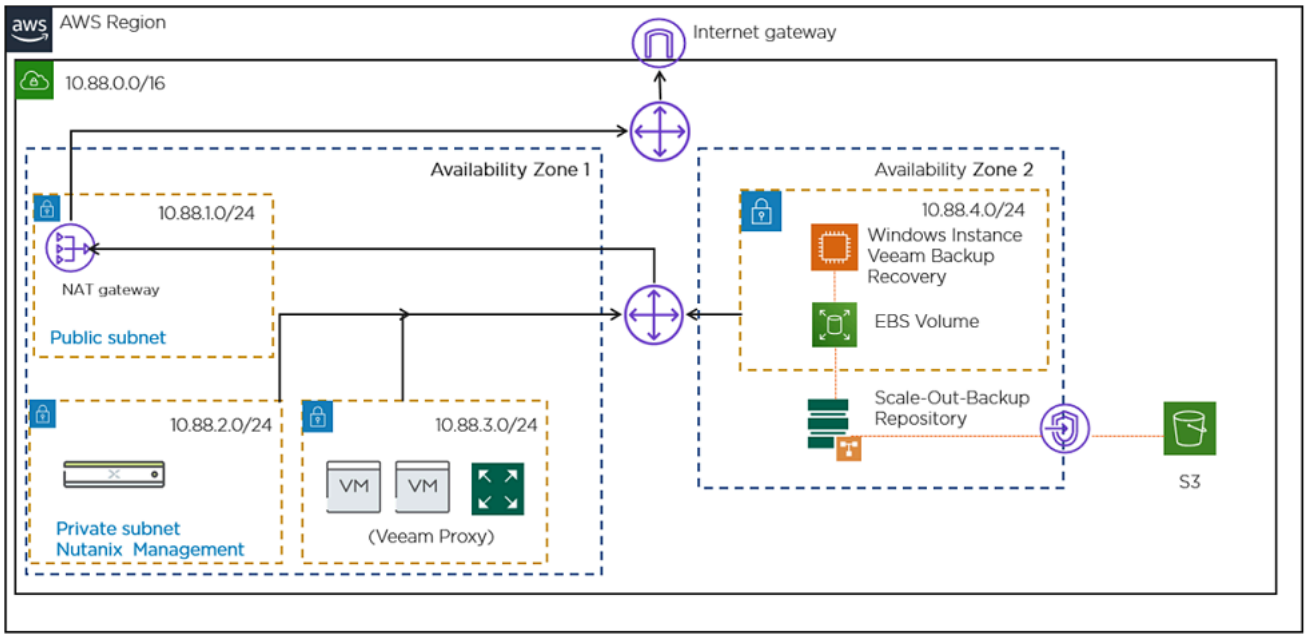

Figure 24: Veeam Backup Architecture

You can deploy Veeam Backup & Replication for NC2 on AWS in several ways. Regardless of the overall deployment strategy, you must run a Windows EC2 instance because Veeam requires all block storage in the scale-out backup repository to use S3. For this reason, you need to have block storage and S3 available to create the scale-out repository. You must also deploy an AHV proxy connected to Veeam Backup & Replication and Prism Element on NC2 on AWS. You could also run Veeam Backup & Replication directly as a VM on NC2 on AWS and use a Linux EC2 appliance as the backup repository.

Table: Veeam Ports between the Nutanix Cluster and Veeam Backup & Replication

From	To	Port	Protocol	Description
Browser	Veeam Availability for Nutanix AHV server	8100	HTTPS	Web UI of the proxy appliance

From	To	Port	Protocol	Description
Veeam Availability for Nutanix AHV (proxy appliance)	Nutanix REST API	9440	HTTPS	Port used to connect with Nutanix REST API
Veeam Availability for Nutanix AHV (proxy appliance)	Veeam Backup & Replication server	10006	TCP	Port used to connect to Veeam Backup & Replication
Veeam Availability for Nutanix AHV (proxy appliance)	Nutanix AHV server	3260	TCP/iSCSI	For connecting to disks on Nutanix AHV
Veeam Availability for Nutanix AHV (proxy appliance)	Veeam Agent server	2500–5000	TCP	Default range of ports used as transmission channels for jobs and restore sessions; for every TCP connection a job uses, one port from this range is assigned

To see all the ports used for Veeam Backup & Replication, backup proxy, and backup repositories, see the Used Ports section of the Veeam Backup & Replication User Guide.

The AWS S3 setup for your bucket is the same as the configuration when using HYCU, with the following exceptions:

- AWS S3 only supports the Standard, Standard-Infrequent Access, and One Zone-IA storage classes. Use S3 Standard to store frequently accessed data and use S3 Standard-IA and S3 One Zone-IA to store long-lived but

less frequently accessed data. Choose your storage class based on your requirements.

- You can't use any mechanism other than Veeam Backup & Replication to manage data and data retention in an object storage bucket or container.
- Veeam Backup & Replication doesn't support enabling life cycle rules; doing so may result in backup and restore failures.

Use the VPC endpoint to keep your S3 public bucket blocked from any public access.

If you run Veeam Backup & Replication in a different Availability Zone than your NC2 on AWS instance, it's easy to recover if the Availability Zone where NC2 on AWS runs fails. To use Veeam Backup & Replication to recover from an Availability Zone failure, deploy a new Veeam AHV proxy and register it with Veeam Backup & Replication. We recommend keeping a configuration backup of the Veeam AHV proxy. If you decide to keep a configuration backup of the proxy, use the settings in the following figure.

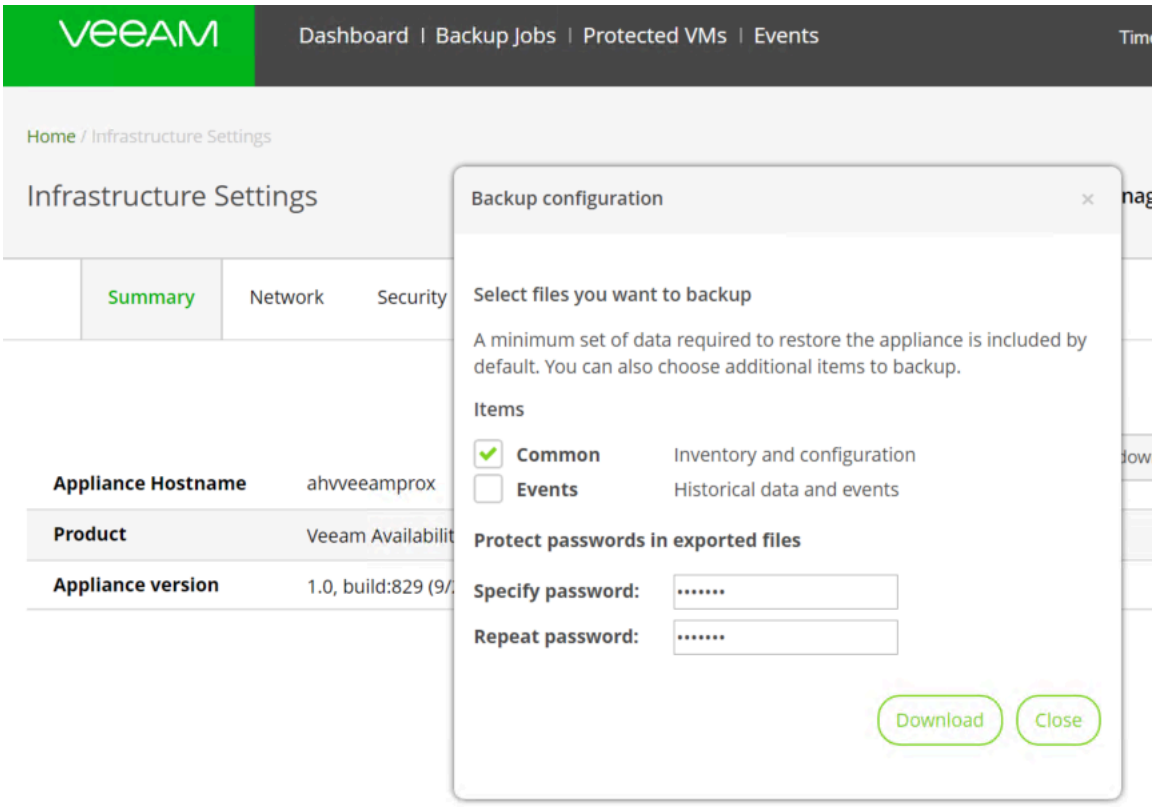

Figure 25: Backup Configuration of AHV Proxy

The following high-level instructions walk you through restoring your NC2 on AWS instance after an Availability Zone failure:

- Deploy a new Veeam proxy to the new NC2 on AWS cluster.
- Register the new Veeam proxy with the NC2 on AWS cluster.
- In the Veeam Backup & Replication console, scan the repository.
- In the Veeam proxy appliance console, click the gear icon, click Manage Veeam Servers, then select the required Veeam backup server from the list.
- Click Import Backups, then click Proceed to confirm the action. The proxy appliance scans the backup repositories and imports all compatible backups of AHV VMs.
- Restore the imported backups.

To recover from a Region failure, you must have an additional scale-out backup repository set up in a different AWS Region using a Linux-based EC2 VM. Because the scale-out backup uses EBS and S3, retention rules in the backup policy don't guarantee that all data has been moved to S3. Run a backup copy job after your backup to ensure that you get the latest copies:

- In the Veeam Backup & Replication console, click New Backup Copy Job.
- In the window that opens, click Target. Select the correct backup repository from the dropdown menu and enter 7 in Restore points to keep. Click Next.
- Define your backup window and finish the configuration.

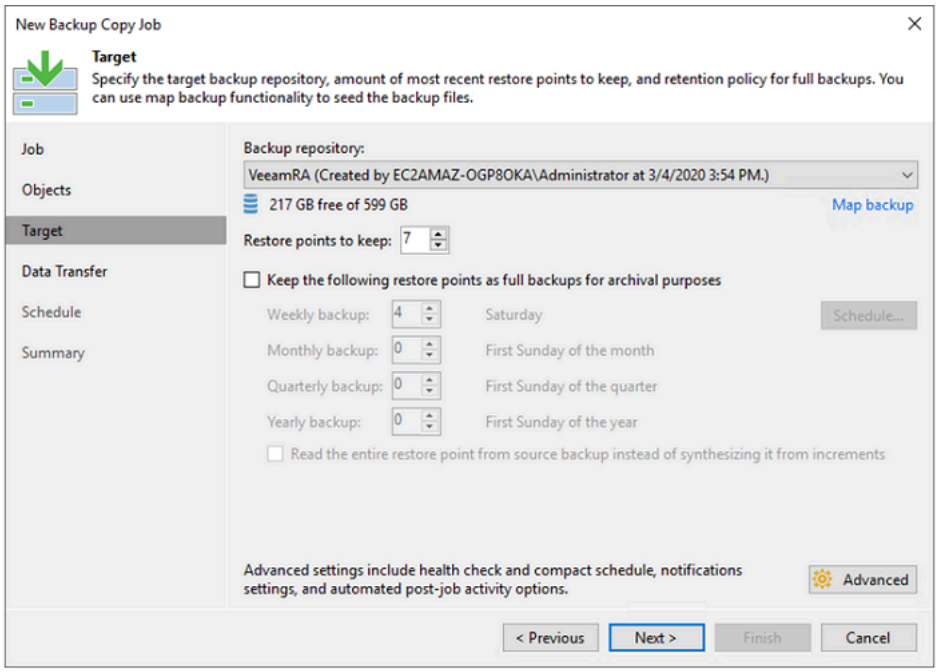

Figure 26: Backup Copy Job

8. 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 is written to optimize capacity and to 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, as well as the utilization of the performance tiers (RAM and flash), by eliminating duplicate data. By providing for 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.

9. 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. Customers looking to simplify their infrastructure operations can now have one-click infrastructure for their key manager as well.

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 is 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 upgrading 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 or 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.

10. 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 antiaffinity 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 VMHA 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, as shown in the following figure. A message then displays the amount of memory reserved and how many AHV host failures the system can tolerate.

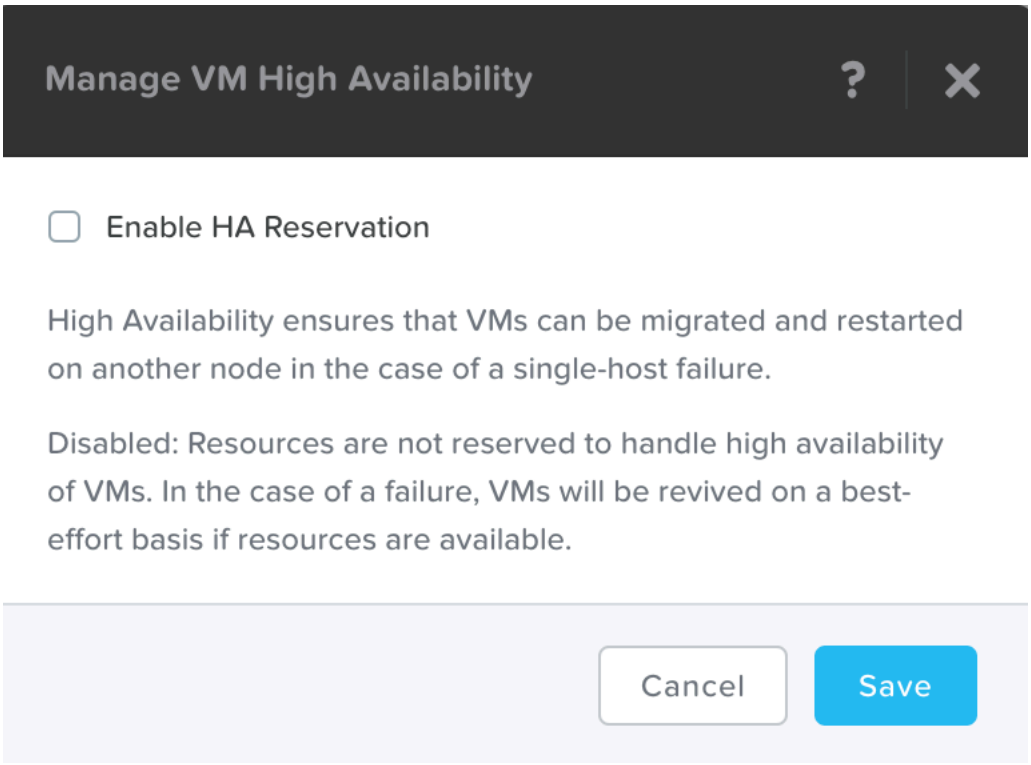

Figure 27: VM High Availability Reservation

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.

11. Acropolis Dynamic Scheduler

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's 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 do not 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.

12. 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 AWS 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 Nutanix Cloud Clusters running, you can enjoy native networking speeds between migrated workloads and the new cloud services you want to consume. For more information, check out the Nutanix Cloud Clusters 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.

List of Figures