AWS guest-connected storage disaster recovery

NetApp Solutions

NetApp
October 04, 2022

Table of Contents

AWS guest-connected storage disaster recovery

Overview - AWS guest-connected storage disaster recovery

Previous: Technology.

This section provides instructions to help users verify, configure, and validate their on-premises and cloud environments for use with NetApp and VMware. Specifically, this solution is focused on the VMware guest-connected use case with ONTAP AFF on-premises and VMware Cloud and AWS FSx ONTAP for the cloud. This solution is demonstrated with two applications: Oracle and MS SQL in a disaster recovery scenario.

Next: Requirements.

Requirements

Previous: Overview - AWS guest-connected storage disaster recovery.

This section details the requirements to access and configure on-premises resources, VMware Cloud, and Amazon FSx ONTAP.

Skills and knowledge

The following skills and information are required to access Cloud Volumes Service for AWS:

- Access to and knowledge of your VMware and ONTAP on-premises environment.
- Access to and knowledge of VMware Cloud and AWS.
- Access to and knowledge of AWS and Amazon FSx ONTAP.
- Knowledge of your SDDC and AWS resources.
- Knowledge of the network connectivity between your on-premises and cloud resources.
- Working knowledge of disaster recovery scenarios.
- Working knowledge of applications deployed on VMware.

Administrative

Whether interacting with resources on-premises or in the cloud, users and administrators must have the ability and entitlements to provision those resources where they need them when they need according to their entitlements. The interaction of your roles and permissions for your on-premises systems, including ONTAP and VMware, and your cloud resources, including VMware Cloud and AWS, is paramount for a successful hybrid cloud deployment.

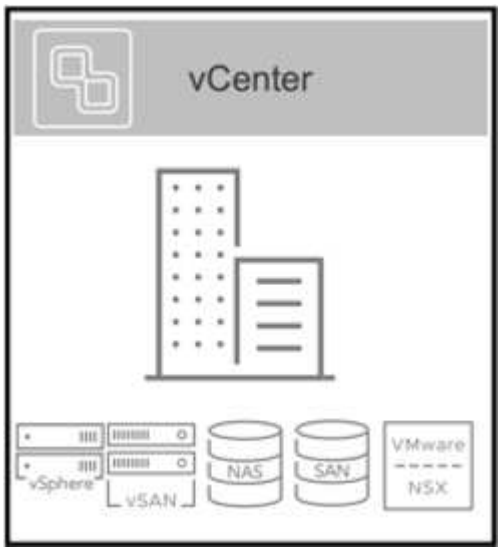
The following administrative tasks must be in place to construct a DR solution with VMware and ONTAP on-premises and VMware Cloud on AWS and FSx ONTAP.

- Roles and accounts enabling provisioning of the following:
  - ONTAP storage resources
  - VMware VMs, datastores, and so on
  - AWS VPC and security groups

- Provisioning of on-premises VMware environment and ONTAP
- VMware Cloud environment
- An Amazon for FSx for ONTAP file system
- Connectivity between your on-premises environment and AWS
- Connectivity for your AWS VPC

On-premises

The VMware virtual environment includes licensing of ESXi hosts, VMware vCenter Server, NSX networking, and other components, as can be seen in the following figure. All are licensed differently, and it is important to understand how the underlying components consume the available licensed capacity.



ESXi hosts

Compute hosts in a VMware environment are deployed with ESXi. When licensed with vSphere at various capacity tiers, virtual machines can take advantage of the physical CPUs on each host and applicable entitled features.

VMware vCenter

Managing ESXi hosts and storage is one of the many capabilities made available to the VMware administrator with vCenter Server. As of VMware vCenter 7.0, there are three editions of VMware vCenter available, depending on the license:

- vCenter Server Essentials
- vCenter Server Foundation
- vCenter Server Standard

VMware NSX

VMware NSX provides administrators with the flexibility required to enable advanced features. Features are enabled depending upon the version of NSX-T Edition that is licensed:

- Professional

- Advanced
- Enterprise Plus
- Remote Office/Branch Office

NetApp ONTAP

Licensing with NetApp ONTAP refers to how administrators gain access to various capabilities and features within NetApp storage. A license is a record of one or more software entitlements. Installing license keys, also known as license codes, enables you to use certain features or services on your storage system. For instance, ONTAP supports all major industry-standard client protocols (NFS, SMB, FC, FCoE, iSCSI, and NVMe/FC) through licensing.

Data ONTAP feature licenses are issued as packages, each of which contains multiple features or a single feature. A package requires a license key, and installing the key enables you to access all features in the package.

License types are as follows:

- **Node-locked license.** Installing a node-locked license entitles a node to the licensed functionality. For the cluster to use the licensed functionality, at least one node must be licensed for the functionality.
- **Master/site license.** A master or site license is not tied to a specific system serial number. When you install a site license, all the nodes in the cluster are entitled to the licensed functionality.
- **Demo/temporary license.** A demo or temporary license expires after a certain time. This license enables you to try certain software functionality without purchasing an entitlement.
- **Capacity license (ONTAP Select and FabricPool only).** An ONTAP Select instance is licensed according to the amount of data that the user wants to manage. Starting with ONTAP 9.4, FabricPool requires a capacity license to be used with a third-party storage tier (for example, AWS).

NetApp SnapCenter

SnapCenter requires several licenses to enable data protection operations. The type of SnapCenter licenses you install depends on your storage environment and the features that you want to use. The SnapCenter Standard license protects applications, databases, files systems, and virtual machines. Before you add a storage system to SnapCenter, you must install one or more SnapCenter licenses.

To enable the protection of applications, databases, file systems, and virtual machines, you must have either a Standard controller-based license installed on your FAS or AFF storage system or a Standard capacity-based license installed on your ONTAP Select and Cloud Volumes ONTAP platforms.

See the following SnapCenter Backup prerequisites for this solution:

- A volume and SMB share created on the on-premises ONTAP system to locate the backed- up database and configuration files.
- A SnapMirror relationship between the on-premises ONTAP system and FSx or CVO in the AWS account. Used for transporting the snapshot containing the backed up SnapCenter database and configuration files.
- Windows Server installed in the cloud account, either on an EC2 instance or on a VM in the VMware Cloud SDDC.
- SnapCenter installed on the Windows EC2 instance or VM in VMware Cloud.

MS SQL

As part of this solution validation, we use MS SQL to demonstrate disaster recovery.

For more information regarding best practices with MS SQL and NetApp ONTAP, follow this link.

Oracle

As part of this solution validation, we use ORACLE to demonstrate disaster recovery. For more information regarding best practices with ORACLE and NetApp ONTAP, follow this link.

Veeam

As part of this solution validation, we use Veeam to demonstrate disaster recovery. For more information regarding best practices with Veeam and NetApp ONTAP, follow this link.

Cloud

AWS

You must be able to perform the following tasks:

- Deploy and configure domain services.
- Deploy FSx ONTAP per application requirements in a given VPC.
- Configure VMware Cloud on the AWS Compute gateway to allow for traffic from FSx ONTAP.
- Configure an AWS security group to allow communication between the VMware Cloud on AWS subnets to the AWS VPC subnets where FSx ONTAP service is deployed.

VMware Cloud

You must be able to perform the following tasks:

- Configure the VMware Cloud on AWS SDDC.

Cloud Manager account verification

You must be able to deploy resources with NetApp Cloud Manager. To verify that you can, complete the following tasks:

- Sign up for Cloud Central if you haven't already.
- Log into Cloud Manager.
- Set up Workspaces and Users.
- Create a connector.

Amazon FSx for NetApp ONTAP

You must be able to perform the following task after you have an AWS account:

- Create an IAM administrative user capable of provisioning Amazon FSx for the NetApp ONTAP file system.

Configuration prerequisites

Given the varying topologies that customers have, this section focuses on the ports necessary to enable communication from on-premises to cloud resources.

Required ports and firewall considerations

The following tables describe the ports that must be enabled throughout your infrastructure.

For a more comprehensive list of required ports for Veeam Backup & Replication software, follow this link.

For a more comprehensive list of port requirements for SnapCenter, follow this link.

The following table lists the Veeam port requirements for Microsoft Windows Server.

From	To	Protocol	Port	Notes
Backup server	Microsoft Windows server	TCP	445	Port required for deploying Veeam Backup & Replication components.
Backup proxy		TCP	6160	Default port used by the Veeam Installer Service.
Backup repository		TCP	2500 to 3500	Default range of ports used as data transmission channels and for collecting log files.
Mount server		TCP	6162	Default port used by the Veeam Data Mover.

For every TCP connection that a job uses, one port from this range is assigned.

The following table lists the Veeam port requirements for Linux Server.

From	To	Protocol	Port	Notes
Backup server	Linux server	TCP	22	Port used as a control channel from the console to the target Linux host.
		TCP	6162	Default port used by the Veeam Data Mover.

From	To	Protocol	Port	Notes
		TCP	2500 to 3500	Default range of ports used as data transmission channels and for collecting log files.

For every TCP connection that a job uses, one port from this range is assigned.

The following table lists the Veeam Backup Server port requirements.

From	To	Protocol	Port	Notes
Backup server	vCenter Server	HTTPS, TCP	443	Default port used for connections to vCenter Server. Port used as a control channel from the console to the target Linux host.
	Microsoft SQL Server hosting the Veeam Backup & Replication configuration database	TCP	1443	Port used for communication with Microsoft SQL Server on which the Veeam Backup & Replication configuration database is deployed (if you use a Microsoft SQL Server default instance).
	DNS Server with name resolution of all backup servers	TCP	3389	Port used for communication with the DNS Server

If you use vCloud Director, make sure to open port 443 on underlying vCenter Servers.

The following table lists Veeam Backup Proxy port requirements.

From	To	Protocol	Port	Notes
Backup server	Backup proxy	TCP	6210	Default port used by the Veeam Backup VSS Integration Service for taking a VSS snapshot during the SMB file share backup.

From	To	Protocol	Port	Notes
Backup proxy	vCenter Server	TCP	1443	Default VMware web service port that can be customized in vCenter settings.

The following table lists SnapCenter port requirements.

Port Type	Protocol	Port	Notes
SnapCenter management port	HTTPS	8146	This port is used for communication between the SnapCenter client (the SnapCenter user) and the SnapCenter Server. Also used for communication from the plug-in hosts to the SnapCenter Server.
SnapCenter SMCore communication port	HTTPS	8043	This port is used for communication between the SnapCenter Server and the hosts where the SnapCenter plug-ins are installed.
Windows plug-in hosts, installation	TCP	135, 445	These ports are used for communication between the SnapCenter Server and the host where the plug-in is being installed. The ports can be closed after installation. In addition, Windows Instrumentation Services searches ports 49152 through 65535, which must be open.
Linux plug-in hosts, installation	SSH	22	These ports are used for communication between the SnapCenter Server and the host where the plug-in is being installed. The ports are used by SnapCenter to copy plug-in package binaries to Linux plug-in hosts.
SnapCenter Plug-ins Package for Windows / Linux	HTTPS	8145	This port is used for communication between SMCore and hosts where the SnapCenter plug-ins are installed.

Port Type	Protocol	Port	Notes
VMware vSphere vCenter Server port	HTTPS	443	This port is used for communication between the SnapCenter Plug-in for Vmware vSphere and vCenter server.
SnapCenter Plug-in for Vmware vSphere port	HTTPS	8144	This port is used for communication from the vCenter vSphere web client and from the SnapCenter Server.

Next: Networking.

Networking

Previous: Requirements.

This solution requires successful communication from the on-premises ONTAP cluster to AWS FSx for NetApp ONTAP interconnect cluster network addresses to perform NetApp SyncMirror operations. Also, a Veeam backup server must have access to an AWS S3 bucket. Instead of using Internet transport, an existing VPN or Direct Connect link can be used as a private link to an S3 bucket.

On premises

ONTAP supports all major storage protocols used for virtualization, including iSCSI, Fibre Channel (FC), Fibre Channel over Ethernet (FCoE), or Non-Volatile Memory Express over Fibre Channel (NVMe/FC) for SAN environments. ONTAP also supports NFS (v3 and v4.1) and SMB or S3 for guest connections. You are free to pick what works best for your environment, and you can combine protocols as needed on a single system. For example, you can augment general use of NFS datastores with a few iSCSI LUNs or guest shares.

This solution leverages NFS datastores for on-premises datastores for guest VMDKs and both iSCSI and NFS for guest application data.

Client networks

VMkernel network ports and software-defined networking provide connectivity to ESXi hosts allowing them to communicate with elements outside the VMware environment. Connectivity depends on the type of VMkernel interfaces used.

For this solution, the following VMkernel interfaces were configured:

- Management
- vMotion
- NFS
- iSCSI

Storage networks provisioned

A LIF (logical interface) represents a network access point to a node in the cluster. This allows communication with the storage virtual machines that house the data accessed by clients. You can configure LIFs on ports

over which the cluster sends and receives communications over the network.

For this solution, LIFs are configured for the following storage protocols:

- NFS
- iSCSI

Cloud connectivity options

Customers have a lot of options when connecting their on-premises environment to cloud resources, including deploying VPN or Direct Connect topologies.

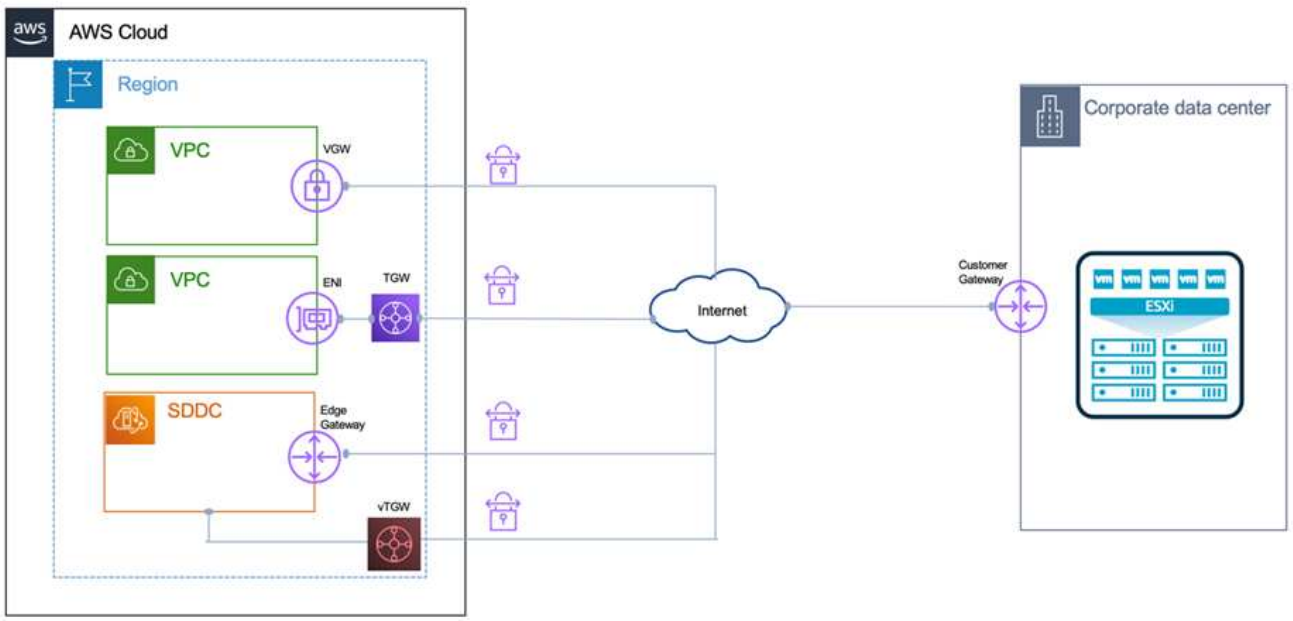
Virtual Private Network (VPN)

VPNs (Virtual Private Networks) are often used to create a secure IPSec tunnel with internet-based or private MPLS networks. A VPN is easy to set up, but it lacks reliability (if internet-based) and speed. The end point can be terminated at the AWS VPC or at the VMware Cloud SDDC. For this disaster recovery solution, we created connectivity to AWS FSx for NetApp ONTAP from the on-premises network. So, it can be terminated at the AWS VPC (Virtual Private Gateway or Transit Gateway) where FSx for NetApp ONTAP is connected.

VPN setup can be route-based or policy-based. With a route-based setup, the endpoints exchange the routes automatically and setup learns the route to the newly created subnets. With a policy- based setup, you must define the local and remote subnets, and, when new subnets are added and allowed to communicate in the IPSec tunnel, you must update the routes.

If the IPSec VPN tunnel is not created on the default gateway, remote network routes must be defined in route tables via the local VPN tunnel end point.

The following figure depicts typical VPN connection options.

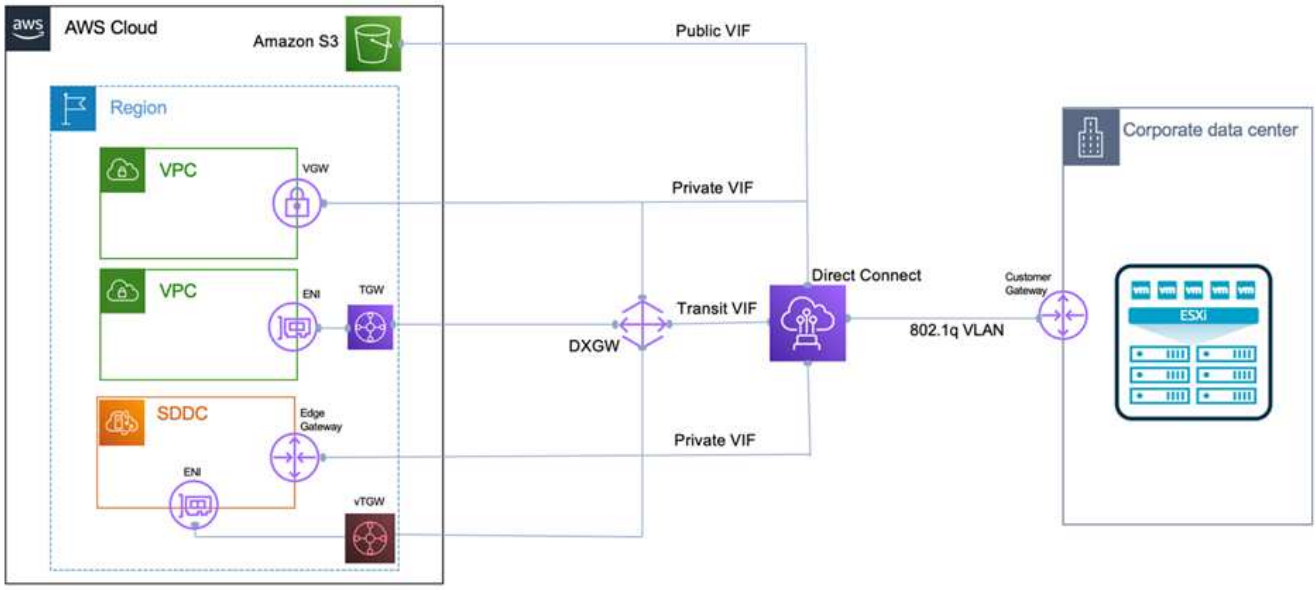


Direct Connect

Direct Connect provides a dedicated link to the AWS network. Dedicated connections create links to AWS using a 1Gbps, 10Gbps, or 100Gbps Ethernet port. AWS Direct Connect partners provide hosted connections using pre-established network links between themselves and AWS and are available from 50Mbps up to 10Gbps. By default, the traffic is unencrypted. However, options are available to secure traffic with MACsec or IPsec. MACsec provides layer-2 encryption while IPsec provides layer-3 encryption. MACsec provides better security by concealing which devices are communicating.

Customers must have their router equipment in an AWS Direct Connect location. To set this up, you can work with AWS Partner Network (APN). A physical connection is made between that router and the AWS router. To enable access to FSx for NetApp ONTAP on VPC, you must have either a private virtual interface or a transit virtual interface from Direct Connect to a VPC. With a private virtual interface, the Direct Connect to VPC connection scalability is limited.

The following figure depicts the Direct Connect interface options.



Transit gateway

The transit gateway is a region-level construct that allows increased scalability of a Direct Connect-to-VPC connection within a region. If a cross- region connection is required, the transit gateways must be peered. For more information, check the AWS Direct Connect documentation.

Cloud network considerations

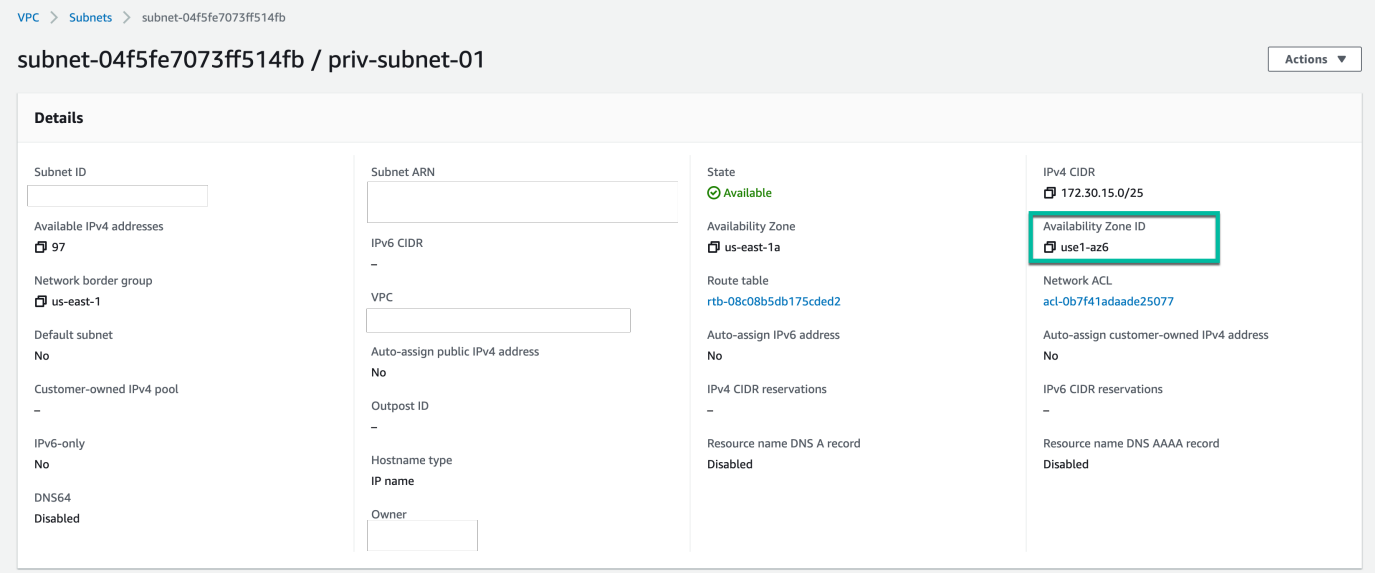
In the cloud, the underlying network infrastructure is managed by the cloud service provider, whereas customers must manage the VPC networks, subnets, route tables, and so on in AWS. They must also manage NSX network segments at the compute edge. SDDC groups routes for the external VPC and Transit Connect.

When FSx for NetApp ONTAP with Multi-AZ availability is deployed on a VPC connected to VMware Cloud, iSCSI traffic receives necessary route table updates to enable communication. By default, there is no route available from VMware Cloud to the FSx ONTAP NFS/SMB subnet on the connected VPC for Multi-AZ deployment. To define that route, we used the VMware Cloud SDDC group, which is a VMware- managed transit gateway, to allow communication between the VMware Cloud SDDCs in the same region as well as to external VPCs and other transit gateways.

There are data transfer costs associated with using a transit gateway. For cost details specific to a region, see this link.

VMware Cloud SDDC can be deployed in a single availability zone, which is like having a single datacenter. A stretch cluster option is also available, which is like a NetApp MetroCluster solution that can provide higher availability and reduced downtime in case of availability-zone failure.

To minimize data-transfer cost, keep the VMware Cloud SDDC and AWS Instances or services in the same availability zone. It is better to match with an availability zone ID rather than with a name because AWS provides the AZ order list specific to the account to spread the load across availability zones. For example, one account (US-East-1a) might point to AZ ID 1 whereas another account (US-East-1c) might point to AZ ID 1. The availability zone ID can be retrieved in several ways. In the following example, we retrieved the AZ ID from the VPC subnet.

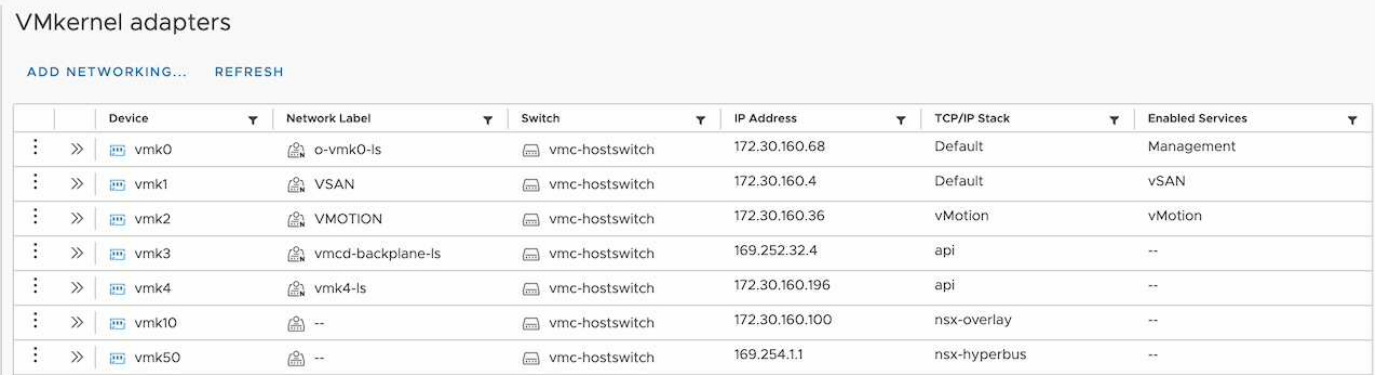


In the VMware Cloud SDDC, networking is managed with NSX, and the edge gateway (Tier-0 router) that handles the north-south traffic uplink port is connected to the AWS VPC. The compute gateway and the management gateways (Tier-1 routers) handle east-west traffic. If the uplink ports of the edge becomes heavily used, you can create traffic groups to associate with specific host IPs or subnets. Creation of a traffic group creates additional edge nodes to separate the traffic. Check the VMware documentation on the minimum number of vSphere hosts required to use a multi-edge setup.

Client networks

When you provision the VMware Cloud SDDC, VMKernel ports are already configured and are ready for consumption. VMware manages those ports and there is no need to make any updates.

The following figure depicts sample Host VMKernel info.



Storage networks provisioned (iSCSI, NFS)

For VM guest storage networks, we typically create port groups. With NSX, we create segments that are consumed on vCenter as port groups. Because storage networks are in a routable subnet, you can access the LUNs or mount the NFS exports using the default NIC even without creating separate network segments. To separate storage traffic, you can create additional segments, define rules, and control the MTU size on those segments. To provide fault tolerance, it is better to have at least two segments dedicated for the storage network. As we mentioned previously, if uplink bandwidth becomes an issue, you can create traffic groups and assign IP prefixes and gateways to perform source-based routing.

We recommend matching the segments in the DR SDDC with the source environment to prevent guessing of mapping network segments during failover.

Security groups

Many security options provide secure communication on the AWS VPC and the VMware Cloud SDDC network. Within the VMware Cloud SDDC network, you can use NSX trace flow to identify the path, including the rules used. Then, you can use a network analyzer on the VPC network to identify the path, including the route tables, security groups, and network access control lists, that is consumed during the flow.

Next: Storage.

Storage

Previous: Networking.

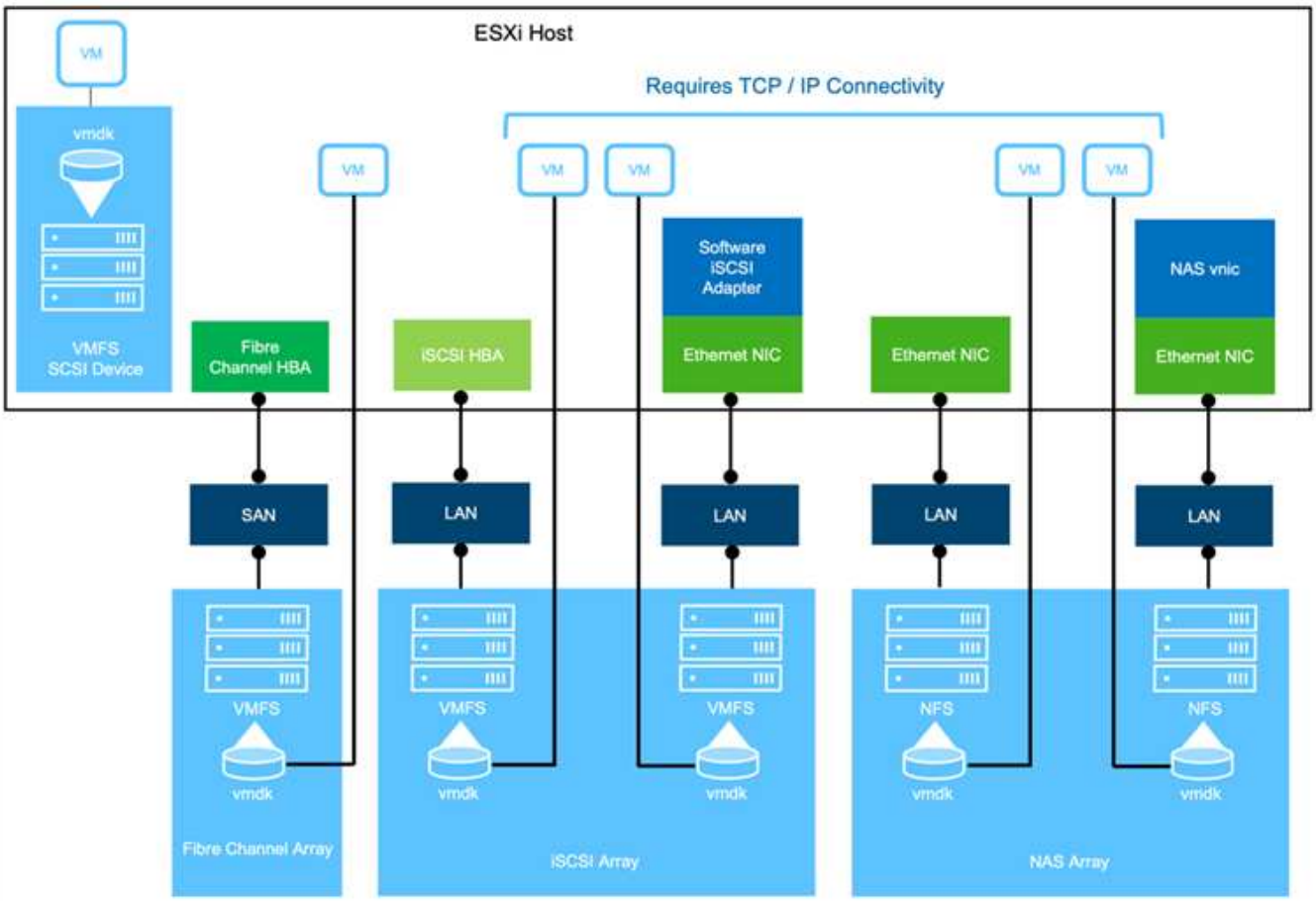
NetApp AFF A-Series systems deliver a high-performance storage infrastructure with flexible data management options that are cloud enabled to meet a wide variety of enterprise scenarios. In this solution, we used an ONTAP AFF A300 as our primary on-premises storage system.

NetApp ONTAP together with ONTAP Tools for VMware and SnapCenter were used in the solution to provide comprehensive management and application backup capabilities that are tightly integrated with VMware vSphere.

On-premises

We used ONTAP storage for the VMware datastores that hosted the virtual machines and their VMDK files. VMware supports multiple storage protocols for connected datastores, and, in this solution, we used NFS volumes for datastores on the ESXi hosts. However, ONTAP storage systems support all protocols supported by VMware.

The following figure depicts VMware storage options.



ONTAP volumes were used for both iSCSI and NFS guest-connected storage for our application VMs. We used the following storage protocols for application data:

- NFS volumes for guest connected Oracle database files.
- iSCSI LUNs for guest connected Microsoft SQL Server databases and transaction logs.

Operating system	Database type	Storage protocol	Volume description
Windows Server 2019	SQL Server 2019	iSCSI	Database files
		iSCSI	Log files
Oracle Linux 8.5	Oracle 19c	NFS	Oracle binary
		NFS	Oracle data
		NFS	Oracle recovery files

We also used ONTAP storage for the primary Veeam backup repository as well as for a backup target for the SnapCenter database backups.

- SMB share for the Veeam backup repository.
- SMB share as a target for the SnapCenter database backups.

Cloud storage

This solution includes VMware Cloud on AWS for hosting virtual machines that are restored as a part of the failover process. As of this writing, VMware supports vSAN storage for the datastores that host the VMs and VMDKs.

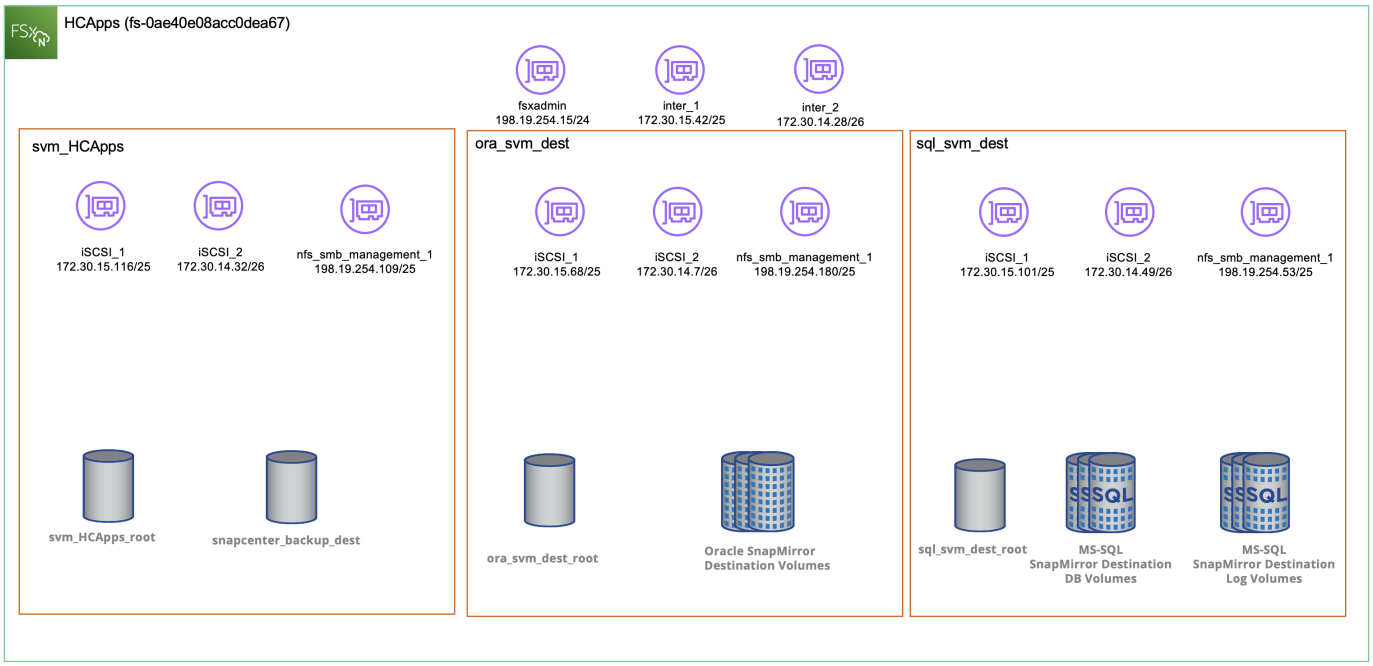
FSX for ONTAP is used as the secondary storage for application data that is mirrored using SnapCenter and SyncMirror. As a part of the failover process, the FSx for ONTAP cluster is converted to primary storage, and the database applications can resume normal function running on the FSx storage cluster.

Amazon FSx for NetApp ONTAP setup

To deploy AWS FSx for NetApp ONTAP using Cloud Manager, follow the instructions at this link.

After FSx ONTAP is deployed, drag and drop the on-premises ONTAP instances into FSx ONTAP to start replication setup of volumes.

The following figure depicts our FSx ONTAP environment.



Network interfaces created

FSx for NetApp ONTAP has network interfaces preconfigured and ready to use for iSCSI, NFS, SMB, and inter-cluster networks.

VM datastore storage

The VMware Cloud SDDC comes with two VSAN datastores named `vsandatastore` and `workloaddatastore`. We used `vsandatastore` to host management VMs with access restricted to cloudadmin credential. For workloads, we used `workloaddatastore`.

Next: Compute.

Compute

Previous: Storage.

VMware vSphere provides virtualized infrastructure in the datacenter and across all the major cloud providers. This ecosystem is ideal for disaster recovery scenarios for which virtualized compute stays consistent regardless of location. This solution uses VMware virtualized compute resources at both the datacenter location and in the VMware Cloud on AWS.

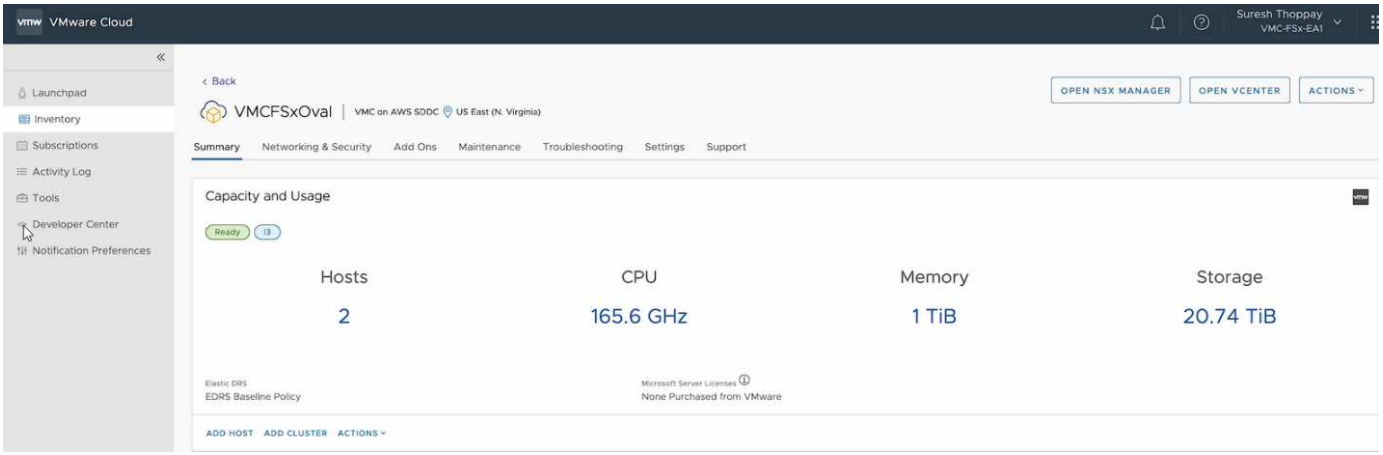
On-premises

This solution uses HPE Proliant DL360 Gen 10 Servers running VMware vSphere v7.0U3. We deployed six compute instances to provide adequate resources for our SQL server and Oracle servers.

We deployed 10 Windows Server 2019 VMs running SQL Server 2019 with varying database sizes and 10 Oracle Linux 8.5 VMs running Oracle 19c, again, with varying database sizes.

Cloud

We deployed an SDDC in VMware Cloud on AWS with two hosts to provide adequate resources to run the virtual machines restored from our primary site.



Next: Cloud Backup Tools.

Cloud Backup Tools

Previous: Compute.

To conduct a failover of our application VMs and database volumes to VMware Cloud Volume services running in AWS, it was necessary to install and configure a running instance of both SnapCenter Server and Veeam Backup and Replication Server. After failover is complete, these tools must also be configured to resume normal backup operations until a failback to the on-premises datacenter is planned and executed.

Deployment of backup tools

SnapCenter server and Veeam Backup & Replication server can be installed in the VMware Cloud SDDC or they can be installed on EC2 instances residing in a VPC with network connectivity to the VMware Cloud environment.

SnapCenter Server

SnapCenter software is available from the NetApp support site and can be installed on Microsoft Windows systems that reside either in a Domain or Workgroup. A detailed planning guide and installation instructions can be found at the NetApp documentation center.

The SnapCenter software can be found at this link.

Veeam Backup & Replication server

You can install the Veeam Backup & Replication server on a Windows server in VMware Cloud on AWS or on an EC2 instance. For detailed implementation guidance, see the Veeam Help Center Technical Documentation.

Backup tools and configuration

After they are installed, SnapCenter and Veeam Backup & Replication must be configured to perform the necessary tasks to restore data to VMware Cloud on AWS.

SnapCenter configuration

To restore application data that has been mirrored to FSx ONTAP, you must first perform a full restore of the on-premises SnapCenter database. After this process is complete, communication with the VMs is reestablished and application backups can now resume using FSx ONTAP as primary storage.

For a list of steps to be completed on the SnapCenter Server residing in AWS, see the section Deploy Secondary Windows SnapCenter Server.

Veeam Backup & Replication configuration

To restore virtual machines that have been backed up to Amazon S3 storage, the Veeam Server must be installed on a Windows server and configured to communicate with VMware Cloud, FSx ONTAP, and the S3 bucket that contains the original backup repository. It must also have a new backup repository configured on FSx ONTAP to conduct new backups of the VMs once they are restored.

For a complete list of steps required to complete failover of the application VMs, see the section Deploy Secondary Veeam Backup & Replication Server.

Next: Overview - Disaster recovery.