Managing the VMware Cloud on AWS Data Center

4 October 2019

VMware Cloud on AWS

You can find the most up-to-date technical documentation on the VMware website at:

https://docs.vmware.com/

If you have comments about this documentation, submit your feedback to

docfeedback@vmware.com

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

Contents

About Managing the VMware Cloud on AWS Data Center

The *Managing the VMware Cloud on AWS Data Center* documentation explains how to set up, examine, and configure the components of your VMware Cloud on AWS data center. This documentation includes information on setting up Hybrid Linked Mode.

You first do initial configuration, such as setting up networking for your VMware Cloud on AWS environment. After that, you can create resource pools and folders, add a vCenter Single Sign-On identity source, and perform other operations that you might already be familiar with from an on-premises environment. You can also use hybrid linked mode to view and manage an on-premises and a VMware Cloud on AWS data center together.

Table 1-1. *Managing the VMware Cloud on AWS Data Center*

Topic	Highlights
vSphere in VMware Cloud on AWS	Introduces components and interfaces and gives pointers to more info.
Hybrid Linked Mode	Explains how you can link your VMware Cloud on AWS instance with an on-premises vCenter Server instance.
vSphere Authentication with vCenter Single Sign-On	■ Explains how vCenter Single Sign-On works ■ Has instructions for making a local AD domain an identity source for vCenter Server
Permissions and User Management Tasks	■ Overview of the vCenter Server permissions model ■ Managing permissions ■ Privileges reference for VMware Cloud on AWS
Clusters and Resource Pools	Explains the clusters and resource pools you see in a VMware Cloud on AWS data center, and the options you have to customize resource pools.
VMware Cloud on AWS Storage	Discusses the vSAN storage available in an VMware Cloud on AWS and how to manage it.
Logical Networking in VMware Cloud on AWS	Gives networking background information and step-by-step instructions for networking setup from the console.
vSphere Tags and Attributes	Explains how to use tags to make things easier to find and manage in your VMware Cloud on AWS data center.

Intended Audience

This information is intended for anyone who wants to use VMware Cloud on AWS to create, configure, and manage an SDDC. The information is written for administrators who have a basic understanding of configuring and managing vSphere in an on-premises environment and are familiar with virtualization concepts. In-depth knowledge of Amazon Web Services is not required.

vSphere in VMware Cloud on AWS

1

VMware vSphere® uses the power of virtualization to transform data centers into simplified cloud computing infrastructures, enabling IT organizations to deliver flexible and reliable IT services.

The two core components of vSphere are VMware ESXi™ and VMware vCenter Server®. ESXi is the hypervisor. Virtual machines run on a cluster of ESXi hosts that are managed by vCenter Server. In VMware Cloud on AWS, VMware takes care of much of the management ESXi hosts and vCenter Server.

In an on-premises vSphere environment, you are responsible for many of the SDDC infrastructure management tasks. In VMware Cloud on AWS, you can focus instead on creating, configuring, and managing virtual machines, VM templates, and OVF templates. The VMware Cloud on AWS provides resource management, authentication and authorization, and other features in the background.

This chapter includes the following topics:

- vSphere Components and Interfaces
- vSphere Managed Inventory Objects
- Finding information about VMware Cloud on AWS

vSphere Components and Interfaces

VMware vSphere is a suite of software components for virtualization. These include ESXi, vCenter Server, and other software components that fulfill a number of different functions in the vSphere environment. VMware manages many parts of your VMware Cloud on AWS SDDC for you, but you can examine all components, and change certain parts of the configuration..

vSphere Components

vSphere includes the following software components:

ESXi

The hypervisor on which you run virtual machines as a set of configuration and disk files that together perform all the functions of a physical machine.

vCenter Server

A service that acts as a central administrator for VMware ESXi hosts.

vCenter Server runs continuously in the background. It performs its monitoring and managing activities even when no clients are connected.

VMware Cloud on AWS includes a single vCenter Server that can be connected to an on-premises vCenter Server using Hybrid Linked Mode.

vCenter Single Sign-On

A service that is part of the vCenter Server management infrastructure. The vCenter Single Sign-On authentication service makes the VMware cloud infrastructure platform more secure by allowing the various vSphere software components to communicate with each other through a secure token exchange mechanism, instead of requiring each component to authenticate a user separately with a directory service like Active Directory.

vSphere Interfaces

The vSphere interface you use depends on the task you want to perform and on the component you want to manage.

vSphere Client

The vSphere Client is an HTML5-based client for managing VMware Cloud on AWS. vSphere Client also performs most configuration tasks for on-premises vSphere SDDCs.

vSphere Command-Line Interfaces

vSphere supports multiple command-line interfaces for configuring virtual machines and other vSphere components.

vSphere SDKs

vSphere supports several SDKs for managing different aspects of your vSphere environment.

Virtual Machine Console

Just as a physical machine, each virtual machine has a console that supports certain management tasks, depending on the operating system.

vCenter Server Features

Many vCenter Server features that required special licensing in earlier versions of the product are available as part of the vSphere Standard license in vSphere 6.x and are also supported for VMware Cloud on AWS.

vCenter Server features include:

vSphere vMotion

Enables you to move running virtual machines from one ESXi host to another ESXi host without service interruption. vSphere HA uses vSphere vMotion to migrate virtual machines if a host becomes unavailable.

Storage vMotion

Allows you to move the disks and configuration file of a running virtual machine from one datastore to another without service interruption.

vSphere High Availability

vSphere High Availability ensures that if a host fails in an SDDC cluster, all virtual machines on the host are restarted on another host in the same cluster. vSphere High Availability settings are preconfigured in VMware Cloud on AWS and cannot be reconfigured by customers.

vSphere DRS

Helps improve resource allocation and power consumption across all hosts and resource pools. vSphere DRS collects resource use information for all hosts and virtual machines in the cluster and or migrates virtual machines in the following situations:

- Initial placement – When you power on a virtual machine in the cluster for the first time, DRS either places the virtual machine or makes a recommendation.
- Load balancing – DRS attempts to improve resource use across the cluster by performing automatic migrations of virtual machines (vMotion) or by providing a recommendation for virtual machine migrations.

See Using Policies and Profilesfor more detail on the storage policies that govern DRS operation in the SDDC.

vSphere Managed Inventory Objects

In vSphere, the inventory is a collection of virtual and physical objects on which you can place permissions, monitor tasks and events, and set alarms. You can group most inventory objects by using folders to more easily manage them.

All inventory objects, with the exception of hosts, can be renamed to represent their purposes. For example, they can be named after company departments or locations or functions.

Note Managed object names cannot exceed 214 bytes (UTF-8 encoded).

vCenter Server monitors and manages the following inventory objects:

Data Centers

Unlike folders, which are used to organize specific object types, a data center is an aggregation of all the different types of objects used to work in virtual infrastructure.

Within each data center, there are four separate hierarchies.

- Virtual machines (and templates)
- Hosts (and clusters)
- Networks
- Datastores

Each VMware Cloud on AWS SDDC has a single data center named SDDC-Datacenter. The data center defines the namespace for networks and datastores. The names for these objects must be unique within a data center. You cannot have two datastores with the same name within a single data center. Virtual machines, templates, and clusters need not be unique within the data center, but must be unique within their folder.

Clusters

A collection of ESXi hosts and associated virtual machines intended to work together as a unit. When you add a host to a cluster, the host's resources become part of the cluster's resources. vCenter Server manages the resources of all hosts in a cluster as one unit.

Datastores

A virtual representation of physical storage resources in the data center. A datastore is the storage location for virtual machine files. In an on-premises SDDC, these physical storage resources can come from the local SCSI disk of the ESXi host, the Fibre Channel SAN disk arrays, the iSCSI SAN disk arrays, or Network Attached Storage (NAS) arrays. For both on-premises and cloud SDDCs, vSAN datastores hide the idiosyncrasies of the underlying physical storage and present a uniform model for the storage resources required by virtual machines.

Folders

Folders allow you to group objects of the same type so you can easily manage them. For example, you can use folders to set permissions across objects, to set alarms across objects, and to organize objects in a meaningful way.

A folder can contain other folders, or a group of objects of the same type: data centers, clusters, datastores, networks, virtual machines, templates, or hosts. For example, one folder can contain hosts and a folder containing hosts, but it cannot contain hosts and a folder containing virtual machines.

Hosts

The physical computer on which ESXi is installed. All virtual machines run on hosts or clusters.

Networks

A set of virtual network interface cards (virtual NICs), distributed switches or vSphere Distributed Switches, and port groups or distributed port groups that connect virtual machines to each other or to the physical network outside of the virtual data center. All virtual machines that connect to the same port group belong to the same network in the virtual environment. You can monitor networks and set permissions and alarms on port groups and distributed port groups.

Resource pools

Resource pools are used to compartmentalize the CPU and memory resources of a host or cluster. Virtual machines run in, and draw their resources from, resource pools. You can create multiple resource pools as direct children of a standalone host or cluster and then delegate control over each resource pool to other individuals or organizations.

vCenter Server DRS provides various options for monitoring the status of the resources and adjusting or suggesting adjustments to the virtual machines using the resources. You can monitor resources and set alarms on them.

Templates

A template is a master copy of a virtual machine that can be used to create and provision new virtual machines. Templates can have a guest operating system and application software installed. They can be customized during deployment to ensure that the new virtual machine has a unique name and network settings.

Virtual machines

A virtualized computer environment in which a guest operating system and associated application software can run. Multiple virtual machines can operate on the same managed host machine concurrently.

vApps

vSphere vApp is a format for packaging and managing applications. A vApp can contain multiple virtual machines.

Finding information about VMware Cloud on AWS

VMware Cloud on AWS includes detailed in-product help for answering many of your questions. In addition, you can search VMware Cloud on AWS documentation, or vSphere documentation.

Note Not all the information in the vSphere documentation set applies to VMware Cloud on AWS. VMware takes care of many of the data center management tasks, such as setting up DRS or HA, for you.

The following documents offer information specific to VMware Cloud on AWS:

- *VMware Cloud on AWS Getting Started* helps you understand how VMware Cloud on AWS works and how to set up networking from the console, and includes other onboarding tasks.
- *Managing Virtual Machines in VMware Cloud on AWS* gives step-by-step instructions for creating and cloning virtual machines and virtual machine templates. It also explains how to install and customize the guest operating system, and how to set up content libraries.
- *Managing the VMware Cloud on AWS Data Center* explains how to set up, examine, and configure the components of your VMware Cloud on AWS data center. This documentation includes information on setting up Hybrid Linked Mode.

To view documentation:

- Go to docs.vmware.com.
- Use search and filters to find the information you need.

You can also go to the VMware Cloud on AWS documentation center directly.

Configuring Hybrid Linked Mode 2

Hybrid Linked Mode allows you to link your VMware Cloud on AWS vCenter Server instance with an on-premises vCenter Single Sign-On domain.

Important Before you can use Hybrid Linked Mode with VMware Cloud on AWS, you must configure your on-premises vCenter to enable single sign-on. See vSphere Authentication with vCenter Single Sign-On for details.

If you link your cloud vCenter Server to a domain that contains multiple vCenter Server instances linked using Enhanced Linked Mode, all of those instances are linked to your cloud SDDC.

Using Hybrid Linked Mode, you can:

- View and manage the inventories of both your on-premises and VMware Cloud on AWS data centers from a single vSphere Client interface, accessed using your on-premises credentials.
- Migrate workloads between your on-premises data center and cloud SDDC.
- Share tags and tag categories from your vCenter Server instance to your cloud SDDC.

Hybrid Linked Mode supports on-premises vCenter Server systems running 6.0 Update 3 patch c and later with either embedded or external Platform Services Controller (both Windows and vCenter Server Appliance). vCenter Server systems with external Platform Services Controller instances linked in Enhanced Linked Mode are also supported, up to the scale limits documented in vSphere 6.0 Configuration Maximums.

You have two options for configuring Hybrid Linked Mode. You can use only one of these options at a time.

- You can install the vCenter Cloud Gateway Appliance and use it to link from your on-premises data center to your cloud SDDC. In this case, SSO users and groups are mapped from your on-premises environment to the SDDC and you do not need to add an identity source to the SDDC LDAP domain.
- You can link your VMware Cloud on AWS SDDC to your on-premises vCenter Server. In this case, you must add an identity source to the SDDC LDAP domain.

This chapter includes the following topics:

- Configuring Hybrid Linked Mode from the Cloud SDDC
- Unlink a Cloud SDDC

Hybrid Linked Mode Prerequisites

Ensure that you have met the following prerequisites before configuring Hybrid Linked Mode.

Common Prerequisites

The following prerequisites are common to both linking from the vCenter Cloud Gateway Appliance and from the cloud SDDC.

- Ensure that your on-premises data center and your cloud SDDC are synchronized to an NTP service or other authoritative time source. When using Hybrid Linked Mode, VMware Cloud on AWS can tolerate a time skew of up to ten minutes between the on-premises data center and the cloud SDDC.
- Configure a management gateway IPsec VPN connection between your on-premises data center and cloud SDDC.
- The maximum latency between your cloud SDDC and on-premises data center must be 100 msec roundtrip.
- Decide which of your on-premises users will have Cloud Administrator permissions. Add these users to a group within your identity source. Ensure that this group has access to your on-premises environment.

Prerequisites for Linking with vCenter Cloud Gateway Appliance

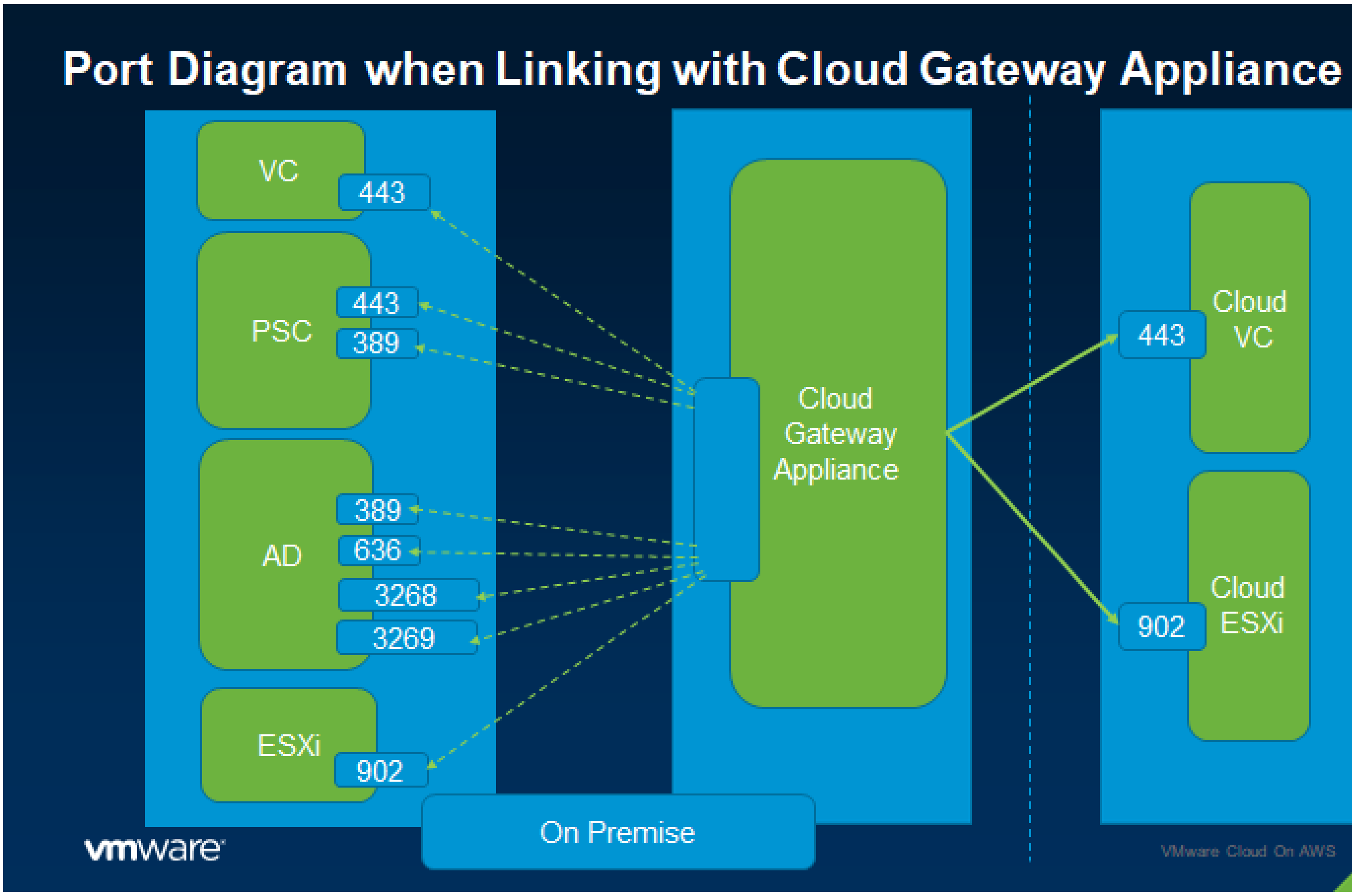
The following prerequisites apply when linking with the vCenter Cloud Gateway Appliance.

- Your on-premises environment is running vSphere 6.5 patch d or later.
- Ensure that the vCenter Cloud Gateway Appliance and your vCenter Server instances can reach each other over your network. Ensure that the following firewall ports are open.

Source	Destination	Port	Purpose
User's web browser	vCenter Cloud Gateway Appliance	5480	Gathering support bundle
vCenter Cloud Gateway Appliance	On-premises vCenter Server	443	Hybrid Linked Mode
vCenter Cloud Gateway Appliance	On-premises Platform Services Controller	443, 389	Hybrid Linked Mode
vCenter Cloud Gateway Appliance	Cloud SDDC vCenter Server	443	Hybrid Linked Mode

Source	Destination	Port	Purpose
vCenter Cloud Gateway Appliance	Cloud ESXi host	902	Virtual Machine Console
vCenter Cloud Gateway Appliance	On-premises Active Directory server (ports dependent on your usage)	389, 636, 3268, 3269	Identity source

The following figure shows the ports required to be open for linking with the vCenter Cloud Gateway Appliance.



Port Diagram when Linking with Cloud Gateway Appliance

- Verify that the host on which you intend to the install the vCenter Cloud Gateway Appliance appliance meets the following hardware requirements

Hardware	Minimum required
CPUs	8
Memory	24 GB
Storage	190 GB

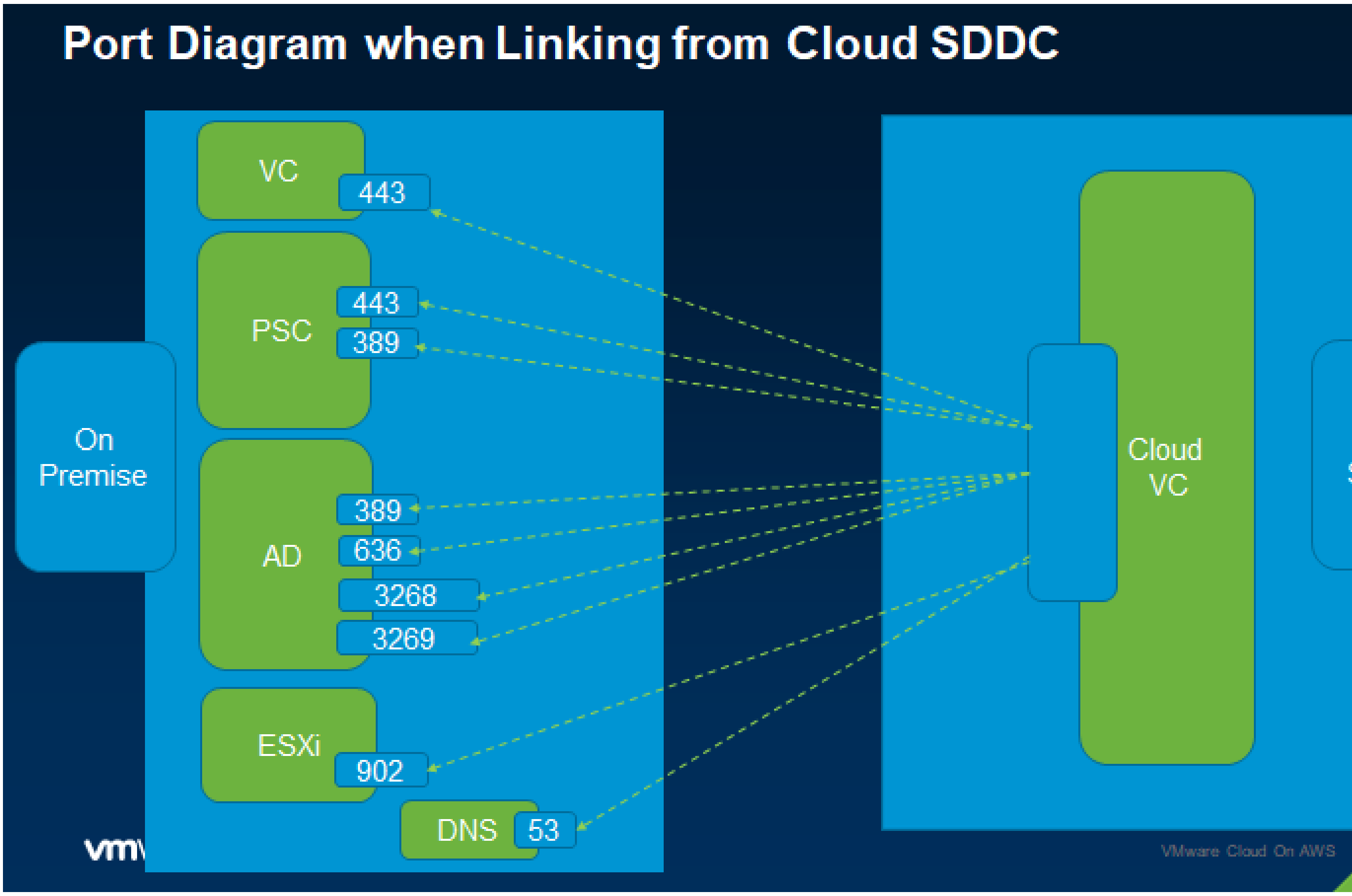
Prerequisites for Linking from the Cloud SDDC

The following prerequisites apply when linking from the cloud SDDC.

- Your on-premises vCenter Server system is running one of the following:
  - vSphere 6.0 Update 3 patch c and later.
  - vSphere 6.5 patch d and later.
- Ensure that you have the login credentials for your on-premises vSphere SSO domain.
- Ensure that you have login credentials for a user who has a minimum of read-only access to the Base DN for users and groups in your on-premises environment. This is used when adding an identity source.
- Ensure that an on-premises DNS server is configured for your management gateway so that it can resolve the FQDN for the identity source and on-premises VMware Cloud on AWS systems.
- Ensure that your on-premises gateway or firewall allows access to the necessary ports from your SDDC for the following services.

Source	Destination	Ports	Purpose
Cloud SDDC	On-premises vCenter Server	443	Hybrid Linked Mode
Cloud SDDC	On-premises Platform Services Controller	389, 443	Hybrid Linked Mode
Cloud SDDC	On-premises Active Directory server (ports dependent on your usage)	389, 636, 3268, 3269	Identity Source
Cloud SDDC	On-premises DNS	53	Resolving FQDN of on-premises vCenter Server and Active Directory Server
Cloud SDDC	On-premises ESXi host	902	Virtual Machine Console

The following figure shows the ports that are required to be open for linking from the cloud SDDC.



- Run the Connectivity Validator tests to check that network connectivity is correctly established for Hybrid Linked Mode. See Validate Network Connectivity for Hybrid Linked Mode.

Configuring Hybrid Linked Mode using the vCenter Cloud Gateway Appliance

Deploy and configure the vCenter Cloud Gateway Appliance to enable Hybrid Linked Mode from your on-premises environment.

In this case, you log into the vCenter Cloud Gateway Appliance to view and manage your on-premises and cloud environments together.

- Install the vCenter Cloud Gateway Appliance from the Graphical Installer

  Download and install the vCenter Cloud Gateway Appliance if you want to link and manage your cloud SDDC from your on-premises data center.

- Link the vCenter Cloud Gateway Appliance to Your Cloud SDDC from the vCenter Cloud Gateway Appliance Installer

  After you have installed the vCenter Cloud Gateway Appliance, you can complete Stage 2 of the installation process to link your appliance to your cloud SDDC.

- Install the vCenter Cloud Gateway Appliance Using the Command-Line Installer

  Use the command-line installer to script or automate your Cloud Gateway Appliance installation.

- Link the vCenter Cloud Gateway Appliance to Your Cloud SDDC from the Gateway Client UI

  Use this procedure to link the vCenter Cloud Gateway Appliance to your cloud SDDC if you did not complete this step as part of the installation process.

- Replace the Certificate for the vSphere Cloud Gateway Appliance

  You can replace the certificate for the vCenter Cloud Gateway Appliance when the certificate expires or when you want to use a certificate from another certificate provider.

- Backing Up the Cloud Gateway Appliance

  Backing up the vCenter Cloud Gateway Appliance is not necessary, because it is stateless and can be redeployed if needed.

Install the vCenter Cloud Gateway Appliance from the Graphical Installer

Download and install the vCenter Cloud Gateway Appliance if you want to link and manage your cloud SDDC from your on-premises data center.

Prerequisites

Ensure that you meet the prerequisites outlined in Hybrid Linked Mode Prerequisites.

Procedure

1. Log in to the VMC Console at https://vmc.vmware.com.
2. Click the **Tools** tab.
3. Click **Download** on the Gateway Appliance card.

   You are directed to My VMware, where you can complete the download of the appliance installer ISO image.
4. In the installer ISO image, browse to the `ui-installer` folder and open the folder for the operating system from which you want to install the appliance.
   - For Windows OS, go the `win32` subdirectory and run the `installer.exe` file.
   - For Linux OS, go to the `lin64` subdirectory, and run the `installer` file.
   - For Mac OS, go to the `mac` subdirectory and run the `Installer.app` file.
5. Click **Get Started**.
6. Under **Deploy Cloud Gateway**, click **Start**.
7. Accept the End User License agreement.

8 Specify the Gateway deployment parameters and click **Next**.

Option	Steps
You can connect to an ESXi host on which to deploy the appliance.	1 Enter the FQDN or IP address of the ESXi host. 2 Enter the HTTPS port of the ESXi host. 3 Enter the user name and password of a user with administrative privileges on the ESXi host, for example, the root user. 4 Click **Next**. 5 Verify that the certificate warning displays the SHA1 thumbprint of the SSL certificate that is installed on the target ESXi host, and click **Yes** to accept the certificate thumbprint.
You can connect to a vCenter Server instance and browse the inventory to select an ESXi host or DRS cluster on which to deploy the appliance.	1 Enter the FQDN or IP address of the vCenter Server instance. 2 Enter the HTTPS port of the vCenter Server instance. 3 Enter the user name and password of user with vCenter Single Sign-On administrative privileges on the vCenter Server instance, for example, the administrator@*your_domain_name* user. 4 Click **Next**. 5 Verify that the certificate warning displays the SHA1 thumbprint of the SSL certificate that is installed on the target vCenter Server instance, and click **Yes** to accept the certificate thumbprint. 6 Select the data center or data center folder that contains the ESXi host or DRS cluster on which you want to deploy the appliance, and click **Next** **Note** You must select a data center or data center folder that contains at least one ESXi host that is not in lockdown or maintenance mode. 7 Select the ESXi host or DRS cluster on which you want to deploy the appliance, and click **Next**.

9 Set up the target appliance VM and click **Next**.

Parameter	Description
VM name	Enter a name for the vCenter Cloud Gateway Appliance VM. The appliance name must not contain a percent sign (%), backslash (\), or forward slash (/) and must be no more than 80 characters in length.
Set root password	Set a root password for the vCenter Cloud Gateway Appliance VM. The password must contain only lower ASCII characters without spaces, at least eight characters, a number, uppercase and lowercase letters, and a special character, for example, an exclamation mark (!), hash key (#), at sign (@), or brackets (()).
Confirm root password	Confirm the password you set above.

10 Select the datastore location for the vCenter Cloud Gateway Appliance and click **Next**.

a Select the datastore where you want to place the vCenter Cloud Gateway Appliance.

b Select **Enable Thin Disk Mode** to conserve disk space by deploying the appliance using a thin disk.

11 Configure the network settings for the appliance and click **Next**.

Parameter	Description
Network	Select the network The networks displayed in the drop-down menu depend on the network settings of the target server. If you are deploying the appliance directly on an ESXi host, non-ephemeral distributed virtual port groups are not supported and are not displayed in the drop-down menu.
IP version	Select the version for the appliance IP address. You can select either IPv4 or IPv6.
IP assignment	Select how to allocate the IP address of the appliance. ■ **static** The wizard prompts you to enter the IP address and network settings. **Note** Avoid using an IP address as a system name. If you use an IP address as a system name, you cannot change the IP address and update the DNS settings after deployment. ■ **DHCP** A DHCP server is used to allocate the IP address. Select this option only if a DHCP server is available in your environment.
FQDN	If you have an enabled DDNS in your environment, you can enter a fully qualified domain name (FQDN) for the appliance. If you enter a FQDN that already exists, the installer warns you that this will cause an error in deployment unless you isolate the network that the appliance is on. For example, you can deploy the appliance on a different port group from the existing FQDN.
IP address	If you selected a static IP address, enter the IP address for the appliance. If you enter an IP address that already exists, the installer warns you that this will cause an error in deployment unless you isolate the network that the appliance is on. For example, you can deploy the appliance on a different port group from the existing IP address.
Subnet mask or prefix length	Enter the subnet mask or prefix length for the IP address.
Default Gateway	Enter the default gateway to be used by the appliance.
DNS Servers	Enter the addresses of the DNS servers used by the appliance.

12 Configure appliance Settings and click **Next.**

- Select **Synchronize Time with NTP servers** and enter the address of one or more NTP servers in the text box to use NTP servers for time synchronization.
- Select **Synchronize Time with ESXi host** to synchronize time to the host you're deploying to.

13 Configure SSO settings.

Option	Description
Platform Services Controller	Enter the IP address or fully qualified domain name of the Platform Services controller in your on-premises environment.
HTTPS Port	Enter the HTTPS port used by the Platform Services Controller.

Option	Description
Single Sign-On Domain	Enter the Single Sign-On domain used by your on-premises Platform Services Controller.
Single Sign-On User Name	Enter the Single Sign-On administrator user name in the form *user@sso-domain*.
Single Sign-On Password	Enter the Single Sign-On administrator password.

14 Select whether to join the vCenter Cloud Gateway Appliance to your Active Directory domain.

Option	Description
Skip	Select this option to skip the step of joining the vCenter Cloud Gateway Appliance to the Active Directory domain. You will need to join the appliance to the domain later before you link the appliance.
Join	Enter the following parameters: a In the **Domain** text box, enter an Active Directory domain name. For example, mydomain.com. b Optionally, in the **Organizational Unit** text box, provide the full OU LDAP FQDN. For example, OU=Engineering,DC=mydomain,DC=com. c In the **Username** text box, enter the user name for the Active Directory administrator in User Principal Name (UPN) format. For example, example@mydomain.com. d In the **Password** field, enter the password for the Active Directory administrator.

15 Click **Finish** to deploy the appliance.

The vCenter Cloud Gateway Appliance is deployed to your on-premises environment. A progress bar shows the progress of deployment.

What to do next

Continue to step 2 of the installer to link the vCenter Cloud Gateway Appliance to your cloud SDDC. See Link the vCenter Cloud Gateway Appliance to Your Cloud SDDC from the vCenter Cloud Gateway Appliance Installer.

Note After you have installed the appliance, consider configuring appliance log collection following the guidelines in VMware Knowledge Base article 67158. Appliance logs are useful when requesting support.

Link the vCenter Cloud Gateway Appliance to Your Cloud SDDC from the vCenter Cloud Gateway Appliance Installer

After you have installed the vCenter Cloud Gateway Appliance, you can complete Stage 2 of the installation process to link your appliance to your cloud SDDC.

Prerequisites

Complete part 1 of the installation process as described in Install the vCenter Cloud Gateway Appliance from the Graphical Installer.

Procedure

1. From the Deployment Stages page of the installer, click **Start** under Stage 2: Configure Hybrid Linked Mode.
2. Connect to the cloud vCenter Server.

Option	Description
vCenter Server	Enter the IP address or FQDN of the vCenter Server instance in your cloud SDDC.
Username	Enter the username for the Cloud Administrator.
Password	Enter the password for the Cloud Administrator.

3. Add the groups you have defined in your on premises environment to serve as cloud administrator groups.

   a. Select the on-premises identity source.

   b. Enter the name of the administrator group in the search box and select the group.

4. Click **Finish**.

   A progress bar shows the progress of the linking operation.

What to do next

When the linking process is complete, you can use the vCenter Cloud Gateway Appliance to view and manage the inventories of your on-premises and cloud SDDCs. Access this interface at http://*cga-address*/ui where *cga-address* is the IP address or FQDN of the vCenter Cloud Gateway Appliance.

Install the vCenter Cloud Gateway Appliance Using the Command-Line Installer

Use the command-line installer to script or automate your Cloud Gateway Appliance installation.

Use the command `vcgw-deploy` to install the vCenter Cloud Gateway Appliance from the command line. In addition to installing the appliance, you can use `vcgw-deploy` to validate your installation templates and run pre-checks on your installation. For a full list of `vcgw-deploy` options, run **`vcgw-deploy install --help`**.

Prerequisites

Ensure that you meet the prerequisites in Hybrid Linked Mode Prerequisites.

Procedure

1. Log in to the VMC Console at https://vmc.vmware.com.
2. Click the **Tools** tab.
3. Click **Download** on the Gateway Appliance card.

   You are directed to My VMware, where you can complete the download of the appliance installer ISO image.

4 Prepare a JSON template for the installation.

   a In the installer ISO image, browse to the `cli-installer/templates` folder.

      This folder contains sample JSON templates for installing the vCenter Cloud Gateway Appliance either directly on an ESXi host or through a vCenter Server system.

   b Copy a template to a working directory and edit it to include the necessary parameters.

      For more information about available template parameters, invoke the installer with the option `--template-help`. For example, on Windows, enter **`vcgw-deploy.exe install --template-help`**.

5 From the command line, change to the `cli-installer` folder and run the installation prechecks.

   - For Windows OS, enter **`vcgw-deploy.exe install`** ***`path-to-template`*** **`--precheck-only`**.
   - For Linux OS, enter **`vcgw-deploy install`** ***`path-to-template`*** **`--precheck-only`**.
   - For Mac OS, enter **`vcgw-deploy install`** ***`path-to-template`*** **`--precheck-only`**.

   The prechecks identify problems with the template and parameters provided, so that you can fix any errors before launching the installation.

6 Launch the installer.

   - For Windows OS, enter **`vcgw-deploy.exe install`** ***`path-to-template`*** **`--accept-eula`**.
   - For Linux OS, enter **`vcgw-deploy install`** ***`path-to-template`*** **`--accept-eula`**.
   - For Mac OS, enter **`vcgw-deploy install`** ***`path-to-template`*** **`--accept-eula`**.

What to do next

Link the vCenter Cloud Gateway Appliance to your cloud SDDC as described in Link the vCenter Cloud Gateway Appliance to Your Cloud SDDC from the Gateway Client UI.

Note After you have installed the appliance, consider configuring appliance log collection following the guidelines in VMware Knowledge Base article 67158. Appliance logs are useful when requesting support.

Link the vCenter Cloud Gateway Appliance to Your Cloud SDDC from the Gateway Client UI

Use this procedure to link the vCenter Cloud Gateway Appliance to your cloud SDDC if you did not complete this step as part of the installation process.

Prerequisites

- You must have Administrator privileges in your on-premises environment in order to perform this task.

Procedure

1 In a web browser, go to http://*cga-address*/ui where *cga-address* is the IP address or FQDN of the vCenter Cloud Gateway Appliance.

2 Log in with your on-premises credentials.

3 Under **Hybrid Cloud**, select **Linked Domains**.

4 Connect to the cloud vCenter Server.

Option	Description
vCenter Server	Enter the IP address or FQDN of the vCenter Server instance in your cloud SDDC.
Username	Enter the username for the Cloud Administrator.
Password	Enter the password for the Cloud Administrator.

5 Add the groups you have defined in your on premises environment to serve as cloud administrator groups.

   a Select the on-premises identity source.

   b Enter the name of the administrator group in the search box and select the group.

6 Click **Finish**.

What to do next

When the linking process is complete, you can use the vCenter Cloud Gateway Appliance to view and manage the inventories of your on-premises and cloud SDDCs. Access this interface at http://*cga-address*/ui where *cga-address* is the IP address or FQDN of the vCenter Cloud Gateway Appliance.

Replace the Certificate for the vSphere Cloud Gateway Appliance

You can replace the certificate for the vCenter Cloud Gateway Appliance when the certificate expires or when you want to use a certificate from another certificate provider.

Prerequisites

Generate certificate signing requests (CSRs) for each certificate you want to replace. Provide the CSR to your Certificate Authority. When the Certificate Authority returns the certificate, place it in a location that you can access from the vCenter Cloud Gateway Appliance.

Procedure

1 In a web browser, go to http://*cga-address*/ui where *cga-address* is the IP address or FQDN of the vCenter Cloud Gateway Appliance.

2 Log in with your on-premises credentials.

3 Navigate to the Certificate Management UI.

   a From the **Home** menu, select **Administration**.

   b Under **Certificates**, click **Certificate Management**.

4 Enter your credentials and click **Login and Manage Certificates**.

5 On the Machine SSL Certificate, select **Actions > Replace**.

6 Click the browse button on the Certificate Chain and provide the path of the certificate chain file.

   This file should contain the machine SSL certificate, the Root CA certificate, and the entire chain of trust.

7 Click the browse button on the private key and provide the private key for the certificate.

8 Click **Replace**.

What to do next

When the certificate is successfully replaced, restart all services on the vCenter Cloud Gateway Appliance. See https://kb.vmware.com/s/article/2109887.

Backing Up the Cloud Gateway Appliance

Backing up the vCenter Cloud Gateway Appliance is not necessary, because it is stateless and can be redeployed if needed.

File-based backup and restore solutions are not supported for the vCenter Cloud Gateway Appliance.

Configuring Hybrid Linked Mode from the Cloud SDDC

As an alternative to using the vCenter Cloud Gateway Appliance, you can configure Hybrid Linked Mode from the cloud SDDC.

In this case, you use your cloud SDDC's vSphere Client to view and manage your complete inventory. When you link from the cloud SDDC, you can link only one on-premises domain.

Validate Network Connectivity for Hybrid Linked Mode

Use the VMC Console Connectivity Validator to check that all required network connectivity is in place for Hybrid Linked Mode.

When you provide the required inputs to the Connectivity Validator, it can verify the network connections required for Hybrid Linked Mode.

Procedure

1 Log in to the VMC Console at https://vmc.vmware.com.

2 Click **View Details** for your SDDC.

3 Click the **Troubleshooting** tab.

4 From the **Use Case** drop down menu, select **Hybrid Linked Mode**.

   The Hybrid Linked Mode connectivity tests are shown. Tests are organized into groups according to the input needed for each group.

5 In the **Input** column, enter the required input for each test you want to run.

6 Run one or more tests.

   - To run all tests, click **Run All**.

- To run a particular test group, click **Run Group** to the right of the group listing.
- To run an individual test, expand the test group and click **Run** next to the individual test.

The status of each test is displayed as it runs. When a test has finished, you can expand the test to see details of the test results.

What to do next

When all tests pass, continue to set up Hybrid Linked Mode. See Add an Identity Source to the SDDC LDAP Domain.

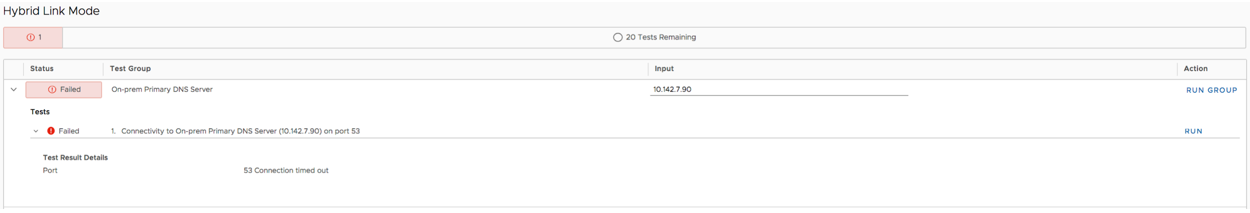
Connectivity Validator: DNS Server Can't Be Reached

On-prem Primary DNS Server or On-prem Secondary DNS Server tests fail in the Connectivity Validator.

Problem

The tests **Connectivity to On-prem Primary DNS Server on Port 53** and/or **Connectivity to On-prem Secondary DNS Server on Port 53** in the Connectivity Validator fail with a message that says `Port 53 Connection timed out`.

Figure 2-1. Image of failed DNS Server connectivity test



Hybrid Link Mode

Status	Test Group	Input	Action
ⓘ Failed	On-prem Primary DNS Server	10.142.7.90	RUN GROUP

Tests

● Failed	1. Connectivity to On-prem Primary DNS Server (10.142.7.90) on port 53	RUN

Test Result Details

Port 53 Connection timed out

Cause

Potential causes of this failure could be:

- The IPsec VPN connection from the cloud SDDC to the on-premises data center might be down.
- The DNS server port 53 is blocked by a firewall rule on the cloud SDDC or on-premises data center.
- You have entered an incorrect IP address for the DNS server.
- The DNS server is down.

Solution

1. Verify that the VPN tunnel from the cloud SDDC to on-premises is up. See View VPN Tunnel Status and Statistics.
2. Inspect the firewall rules in the VMC Consoleto ensure that access to port 53 on the on-premises DNS server is not blocked.
3. Inspect the firewall rules in your on-premises environment to ensure that access to port 53 on the on-premises DNS server is not blocked.

4 Verify that you entered the correct IP address for your on-premises DNS servers. See Specify Management Gateway DNS Servers.

5 Verify that your DNS server is running, and bring it back up if it is down.

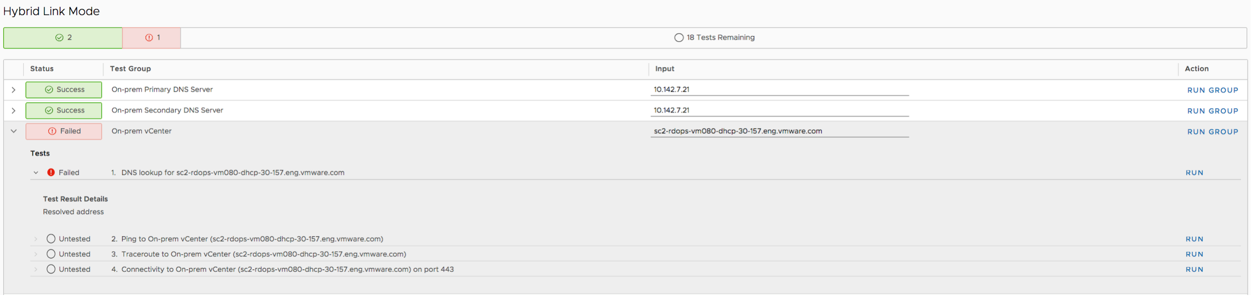
Connectivity Validator: DNS Lookup Failure for a Given FQDN

The DNS lookup test for an on-premises vCenter Server, Platform Services Controller, Active Directory, or ESXi fails.

Problem

One or more DNS lookup tests fails. The **Resolved Address** field in the test results shows no result.

Figure 2-2. Example of DNS lookup test failure



Hybrid Link Mode

✓ 2 | ⓘ 1 | ○ 18 Tests Remaining

Status	Test Group	Input	Action
✓ Success	On-prem Primary DNS Server	10.142.7.21	RUN GROUP
✓ Success	On-prem Secondary DNS Server	10.142.7.21	RUN GROUP
ⓘ Failed	On-prem vCenter	sc2-rdops-vm080-dhcp-30-157.eng.vmware.com	RUN GROUP

Tests

● Failed	1. DNS lookup for sc2-rdops-vm080-dhcp-30-157.eng.vmware.com	RUN

Test Result Details

Resolved address

○ Untested	2. Ping to On-prem vCenter (sc2-rdops-vm080-dhcp-30-157.eng.vmware.com)	RUN
○ Untested	3. Traceroute to On-prem vCenter (sc2-rdops-vm080-dhcp-30-157.eng.vmware.com)	RUN
○ Untested	4. Connectivity to On-prem vCenter (sc2-rdops-vm080-dhcp-30-157.eng.vmware.com) on port 443	RUN

Cause

If the DNS server reachability test succeeded, but the DNS lookup for a given FQDN fails, this could be caused by one of the following:

- The on-premises DNS server does not have an entry for the given FQDN.
- You entered an incorrect FQDN for the test.

Solution

1 Ensure that you entered the correct FQDN.

2 Check that the on-premises DNS server has an entry for the FQDN.

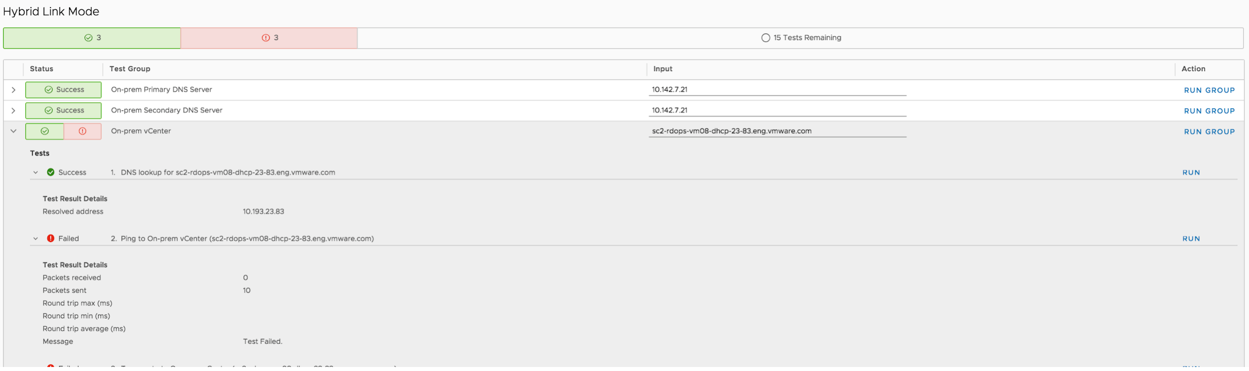
Connectivity Validator: Ping Failure for a Given FQDN

The test that pings an on-premises vCenter Server, Platform Services Controller, Active Directory, or ESXi fails.

Problem

A ping test for a given FQDN fails. Test details show that no responses to ICMP packets were received.

Figure 2-3. Example of a ping test failure



Hybrid Link Mode

✓ 3 | ⓘ 3 | ○ 15 Tests Remaining

Status	Test Group	Input	Action
✓ Success	On-prem Primary DNS Server	10.142.7.21	RUN GROUP
✓ Success	On-prem Secondary DNS Server	10.142.7.21	RUN GROUP
✓ ⓘ	On-prem vCenter	sc2-rdops-vm08-dhcp-23-83.eng.vmware.com	RUN GROUP

Tests

✓ Success 1. DNS lookup for sc2-rdops-vm08-dhcp-23-83.eng.vmware.com RUN

Test Result Details

Resolved address	10.193.23.83

❗ Failed 2. Ping to On-prem vCenter (sc2-rdops-vm08-dhcp-23-83.eng.vmware.com) RUN

Test Result Details

Packets received	0
Packets sent	10
Round trip max (ms)	
Round trip min (ms)	
Round trip average (ms)	
Message	Test Failed.

Cause

Potential causes of this failure could be:

- A firewall rule in the cloud SDDC or the on-premises data center might be blocking ICMP traffic.
- The remote system with the given FQDN is powered-off.

Solution

1. Check your firewall rules set in the VMC Console to ensure that they are not blocking ICMP traffic to the given FQDN.
2. Check your on-premises firewall rules to ensure that they are not blocking ICMP traffic to the given FQDN.
3. Check that the remote system being pinged is powered-on and functioning, and power on or restart if necessary.

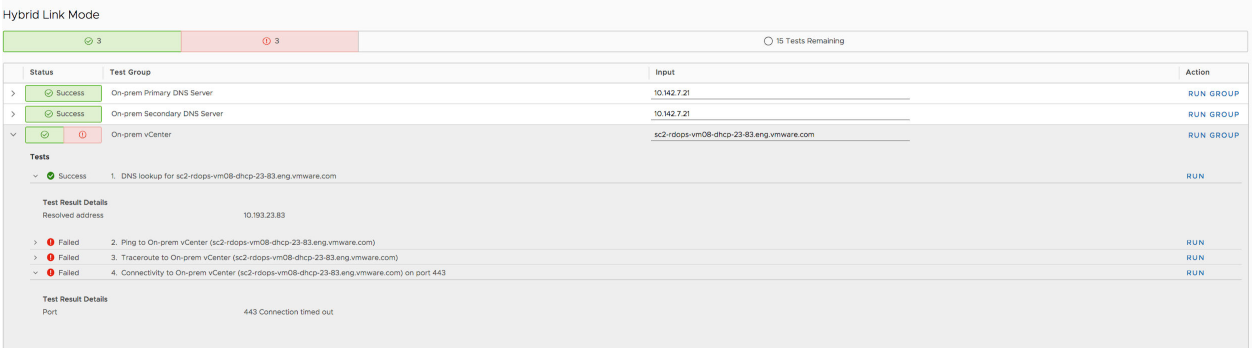
Connectivity Validator: Port Reachability Failure for a Given FQDN

A test to reach a specific port on

Problem

A test for connectivity to a particular port at a given FQDN fails with the message `Port` *`port-number`* `Connection timed out`.

Figure 2-4. Example of a port reachability test failure



Hybrid Link Mode

✓ 3 | ⓘ 3 | ○ 15 Tests Remaining

Status	Test Group	Input	Action
✓ Success	On-prem Primary DNS Server	10.142.7.21	RUN GROUP
✓ Success	On-prem Secondary DNS Server	10.142.7.21	RUN GROUP
✓ ⓘ	On-prem vCenter	sc2-rdops-vm08-dhcp-23-83.eng.vmware.com	RUN GROUP

Tests

✓ Success 1. DNS lookup for sc2-rdops-vm08-dhcp-23-83.eng.vmware.com RUN

Test Result Details

Resolved address	10.193.23.83

❗ Failed 2. Ping to On-prem vCenter (sc2-rdops-vm08-dhcp-23-83.eng.vmware.com) RUN

❗ Failed 3. Traceroute to On-prem vCenter (sc2-rdops-vm08-dhcp-23-83.eng.vmware.com) RUN

❗ Failed 4. Connectivity to On-prem vCenter (sc2-rdops-vm08-dhcp-23-83.eng.vmware.com) on port 443 RUN

Test Result Details

Port	443 Connection timed out

Cause

Potential causes of this failure could be:

- A firewall rule in the cloud SDDC or the on-premises data center might be blocking access to the port.
- The remote system with the given FQDN is powered-off.

Solution

1. Check your firewall rules set in the VMC Console to ensure that they are not blocking access to the specified port.
2. Check your on-premises firewall rules to ensure that they are not blocking access to the specific port.
3. Check that the remote system being pinged is powered-on and functioning, and power on or restart if necessary.

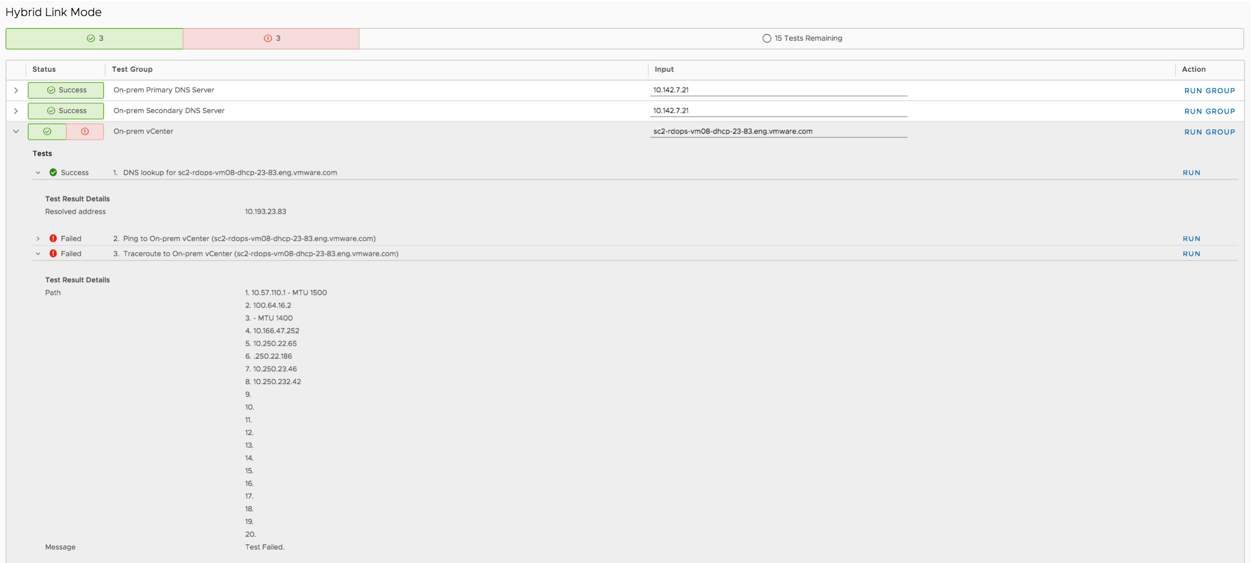
Connectivity Validator: Traceroute Failure for a given FQDN

A traceroute test to a given FQDN fails.

Problem

A traceroute test to an FQDN fails. In the test results, you can see hops to the destination listed without accompanying IP addresses.

Figure 2-5. Example of traceroute test failure



Hybrid Link Mode

✓ 3 | ⓘ 3 | ○ 15 Tests Remaining

	Status	Test Group	Input	Action
>	✓ Success	On-prem Primary DNS Server	10.142.7.21	RUN GROUP
>	✓ Success	On-prem Secondary DNS Server	10.142.7.21	RUN GROUP
˅	✓ ⓘ	On-prem vCenter	sc2-rdops-vm08-dhcp-23-83.eng.vmware.com	RUN GROUP

Tests

˅ ✓ Success 1. DNS lookup for sc2-rdops-vm08-dhcp-23-83.eng.vmware.com RUN

Test Result Details

Resolved address 10.193.23.83

> ! Failed 2. Ping to On-prem vCenter (sc2-rdops-vm08-dhcp-23-83.eng.vmware.com) RUN

˅ ! Failed 3. Traceroute to On-prem vCenter (sc2-rdops-vm08-dhcp-23-83.eng.vmware.com) RUN

Test Result Details

```
Path       1. 10.57.110.1 - MTU 1500
           2. 100.64.16.2
           3. - MTU 1400
           4. 10.166.47.252
           5. 10.250.22.65
           6. .250.22.186
           7. 10.250.23.46
           8. 10.250.232.42
           9.
           10.
           11.
           12.
           13.
           14.
           15.
           16.
           17.
           18.
           19.
           20.
Message    Test Failed.
```

Cause

Potential causes of this failure could be:

- If the ping test to the FQDN itself is successful, a firewall rule in the cloud SDDC or the on-premises data center might be blocking ICMP traffic to one of the hops along the traffic path.
- The remote system with the given FQDN is powered-off.

Solution

1. Check your firewall rules set in the VMC Console to ensure that they are not blocking ICMP traffic to one of the hops along the traffic path.
2. Check your on-premises firewall rules to ensure that they are not blocking ICMP traffic to the one of the hops on the traffic path.
3. Check that the remote system is powered-on and functioning, and power on or restart if necessary.

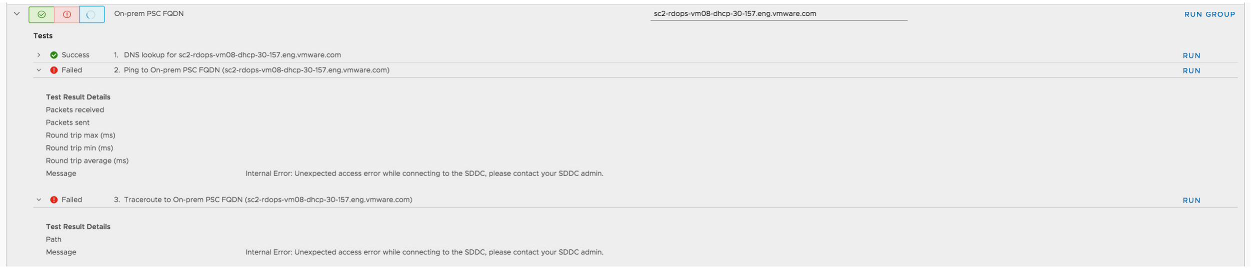
Connectivity Validator: Test Failure Due to Internal Error

A test fails due to an internal error.

Problem

Any of the Connectivity Validator tests might fail with an error message beginning with `Internal Error:`.

Figure 2-6. Example of test failure due to internal error



On-prem PSC FQDN sc2-rdops-vm08-dhcp-30-157.eng.vmware.com

Tests

✓ Success 1. DNS lookup for sc2-rdops-vm08-dhcp-30-157.eng.vmware.com

! Failed 2. Ping to On-prem PSC FQDN (sc2-rdops-vm08-dhcp-30-157.eng.vmware.com)

Test Result Details

Packets received

Packets sent

Round trip max (ms)

Round trip min (ms)

Round trip average (ms)

Message Internal Error: Unexpected access error while connecting to the SDDC, please contact your SDDC admin.

! Failed 3. Traceroute to On-prem PSC FQDN (sc2-rdops-vm08-dhcp-30-157.eng.vmware.com)

Test Result Details

Path

Message Internal Error: Unexpected access error while connecting to the SDDC, please contact your SDDC admin.

Cause

This error most commonly occurs when the Connectivity Validator experiences an internal connectivity problem.

Solution

Most of these failures are intermittent and resolve without you needing to do anything. However, if the error persists, contact VMware customer support.

Add an Identity Source to the SDDC LDAP Domain

The first step toward configuring Hybrid Linked Mode from your SDDC is to add your on-premises LDAP domain as an identity source for the SDDC vCenter Server .

You can configure Hybrid Linked Mode from your SDDC if your on-premises LDAP service is provided by a native Active Directory (Integrated Windows Authentication) domain or an OpenLDAP directory service.

Important If you are using OpenLDAP as the identity source, see the VMware knowledge base article at http://kb.vmware.com/kb/2064977 for additional requirements.

Prerequisites

Ensure that you meet the Common Prerequisites in Hybrid Linked Mode Prerequisites.

Procedure

1 Log in to the vSphere Client for your SDDC.

To add an identity source, you must be logged in as cloudadmin@vmc.local or another member of the Cloud Administrators group.

2 Bring up the **Add Identity Source** dialog.

Use case	Description
Hybrid Linked Mode	a Select **Menu > Administration**. b Under **Hybrid Cloud**, select **Linked Domains**. c Under **Add Cloud Administrator**, select **Add Identity Source** from the **Identity Source** drop-down menu.
All other use cases	a Select **Menu > Administration**. b Under **Single Sign On**, click **Configuration**. c Click **Identity Sources** and click **Add**.

3 Configure the identity source settings.

Option	Description
Identity Source Type	Select **Active Directory as an LDAP Server** for a Windows Active Directory Server or **Open LDAP** for an Open LDAP server.
Name	Enter the name of the identity source.
Base DN for users	Enter the Base Distinguished Name for users.
Base DN for groups	Enter the Base Distinguished Name for groups.
Domain Name	FQDN of the domain. Do not enter an IP address here.
Domain Alias	Enter an alias for the domain. For Active Directory identity sources, the domain's NetBIOS name. Add the NetBIOS name of the Active Directory domain as an alias of the identity source if you are using SSPI authentications.
Username	Enter the ID of a user in the domain who has a minimum of read-only access to Base DN for users and groups. Use UPN format (for example, user@example.com), rather than DN format.
Password	Enter the password of the user who is specified by **Username**.
Connect To	Select which domain controller to connect to. ■ Select **Any domain controller in the domain** to connect to any domain controller. ■ Select **Specific domain controllers** to specify the domain controllers. If you select **Specific domain controllers**, specify the URL for the primary server and the secondary server used for failover. Use the format ldap://hostname:port or ldaps://hostname:port. The port is typically 389 for ldap: connections and 636 for ldaps: connections. For Active Directory multi-domain controller deployments, the port is typically 3268 for ldap: connections and 3269 for ldaps: connections.
SSL Certificates	If you use ldaps:, select **Browse** and select a certificate file to upload to provide security for the ldaps: connection.

When the identity source is added, on-premises users can authenticate to the SDDC, but have the **No access** role. Add permissions for a group of users to give them the Cloud Administrator role.

Link to an On-Premises Data Center

To complete the configuration of Hybrid Linked Mode from the cloud SDDC, link your on-premises data center from your cloud vCenter Server.

Procedure

1 If you haven't already, log in to the vSphere Client for your SDDC and browse to the Linked Domains page.

   a Select **Menu > Administration** to display the Administration page.

   b Under **Hybrid Cloud**, select **Linked Domains**.

2 Connect to the on-premises vCenter Server.

Option	Description
Platform Services Controller	Enter the IP address or FQDN of the Platform Services Controller instance in your on-premise data center.
HTTPS Port	Enter the HTTPS port used by the Platform Services Controller.
Username	Enter the username for the on-premises SSO administrator.
Password	Enter the password for the on-premises SSO administrator.

3 Add the groups you have defined in your on premises environment to serve as cloud administrator groups.

   a Select the on-premises identity source.

   If you haven't already added the on-premises identity source, do so as described in Add an Identity Source to the SDDC LDAP Domain.

   b Enter the name of the administrator group in the search box and select the group.

4 Click **Link**.

Unlink a Cloud SDDC

You can unlink a cloud SDDC from Hybrid Linked Mode when you no longer want linking between your cloud SDDC and a particular on-premises data center..

For example, you might want to link an on-premises data center to your SDDC in order to migrate virtual machines to the SDDC, and then unlink the on-premises data center. If you plan to decomission a linked on-premises data center, unlink it before doing so.

Note Unlinking an on-premises data center from the cloud SDDC does not remove the associated identity source or permissions that you added before linking the domain. Users can still use their on-premises credentials to authenticate to your SDDC, and retain the permissions granted to them. However, they are not able to view the on-premises inventory after unlinking the domain. When you unlink a cloud SDDC from the vCenter Cloud Gateway Appliance, users can't use their on-premises credentials to log into the cloud SDDC any longer.

Unlinking also leaves tags and categories in place, because VMs in your cloud SDDC might still be using those tags.

Prerequisites

Ensure that you have network connectivity between your SDDC management gateway and your SSO Domain.

Procedure

1 Log into the appropriate system.

- If you linked your cloud SDDC and your on-premises data center from the vCenter Cloud Gateway Appliance, log into the vCenter Cloud Gateway Appliance UI.
- If you linked your cloud SDDC and your on-premises data center from the cloud vCenter Server, log in to the vSphere Client for your SDDC.

2 Browse to the Linked Domains page.

   a Select **Menu > Administration** to display the Administration page.

   b Under **Hybrid Cloud**, select **Linked Domains**.

3 Under the name of the linked domain, click **Unlink**.

   A dialog box appears asking you to confirm the unlinking. Note that all currently active sessions are logged out when you unlink a domain.

4 Click **OK**.

   When the unlinking is complete, you are prompted to log out.

5 Click **OK** to log out.

The SSO domain is unlinked. If you want to continue using Hybrid Linked Mode, you can link to another SSO domain or relink to the same domain.

Note After you unlink from the cloud SDDC, new connections to the cloud SDDC vSphere Client cannot view or interact with previously-linked on-premises resources. Currently active sessions in the cloud SDDC vSphere Client continue to be able to view and interact with resources in the previously linked on-premises vCenter Server instances until the users of those sessions log out of the cloud SDDC vSphere Client or the sessions expire. If necessary, log in to each of the previously-linked on-premises vCenter Server instances, and forcibly terminate these sessions.

After you unlink from the vCenter Cloud Gateway Appliance, new connections to the vCenter Cloud Gateway Appliance cannot view or interact with previously-linked cloud resources. Currently active sessions in the vCenter Cloud Gateway Appliance continue to be able to view and interact with resources in the cloud SDDC until the users of those sessions log out or the sessions expire. If necessary, log in to the vCenter Cloud Gateway Appliance and forcibly terminate these sessions.

vCenter Server in VMware Cloud on AWS

3

vCenter Server is the management platform for your virtual machines on-premises as well as in the cloud. VMware performs vCenter Server setup for you by setting up storage, HA and DRS, and so on. You can examine vCenter Server just as you do in an on-premises environment.

In an on-premises vSphere environment, you can change many of the vCenter Server settings, including licensing, statistics collection, logging, and more.

In an VMware Cloud on AWS environment, you can perform the following tasks:

Table 3-1. vCenter Server Tasks in VMware Cloud on AWS

Task	Description
Set a Message of the Day	Set a message and send it to other users of the VMware Cloud on AWS SDDC.
Examine the vCenter Server object hierarchy	Examining the hierarchy is especially helpful if you're migrating from an on-premises environment to the cloud.
Examine events, alarms, and recent tasks	vCenter Server keeps comprehensive logs of events and alarms. An alarms are also displayed directly on the screen where it applies.

This chapter includes the following topics:

- Send a Message to Other Logged In Users
- Examine the vCenter Server General Configuration
- View Events, Alarms, and Recent Tasks

Send a Message to Other Logged In Users

Administrators can send messages to users who are currently logged in to a vCenter Server system. The message might announce maintenance or ask users to log out temporarily.

Procedure

1 In the vSphere Client, navigate to the vCenter Server instance.

2 Click **Configure**.

3 Select **Settings > Message of the Day** and click **Edit**.

4 Enter a message and click **OK**.

The message appears at the top of the vSphere Client in each active user session.

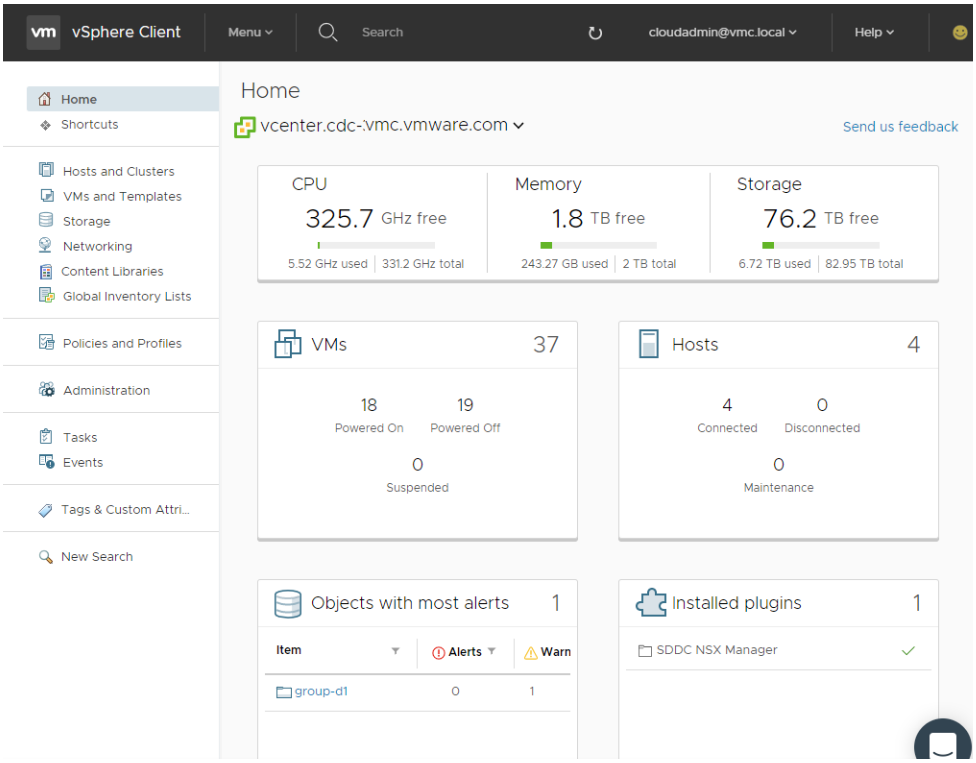
Examine the vCenter Server General Configuration

When you provision a VMware Cloud on AWS SDDC, VMware create a vCenter Server for you. You can examine its current settings.

In an on-premises environment, you might have to fine-tune the vCenter Server configuration. Because VMware managers the vCenter Server instances in VMware Cloud on AWS, you do not need to modify the configuration. You can view it to examine resource allocation, networking, and other attributes.

Procedure

1 In the vSphere Client, click **Home** and examine an overview of your vCenter Server



CPU 325.7 GHz free
5.52 GHz used | 331.2 GHz total

Memory 1.8 TB free
243.27 GB used | 2 TB total

Storage 76.2 TB free
6.72 TB used | 82.95 TB total

VMs 37
18 Powered On
19 Powered Off
0 Suspended

Hosts 4
4 Connected
0 Disconnected
0 Maintenance

Objects with most alerts 1

Item	Alerts	Warn
group-d1	0	1

Installed plugins 1

SDDC NSX Manager ✓

2 To drill down, select **Menu > Hosts and Clusters**, select the vCenter Server, and examine the information in the different tabs.

View Events, Alarms, and Recent Tasks

vCenter Server logs events, alarms and recent tasks You can VMware Cloud on AWS events, alarms, and tasks from the vSphere Client.

- Events are records of user actions or system actions that occur on objects in vCenter Server or on a host.

- Alarms are notifications that are activated in response to an event, a set of conditions, or the state of an inventory object.

See the *vSphere Monitoring and Performance* documentation for details.

Procedure

1 To view the Event Console, select **Menu > Events**.

  You can select events and view details, order the console display, and click **Next** to display earlier events.

2 To view the Task Console, select **Menu > Tasks**.

  You can select tasks and view details and releated events, order the console display, and click **Next** to display earlier tasks.

3 Alarms are displayed in several ways.

  - Any object that has an associated alarm shows an alarm icon in the object hierarchy. Select the object's Summary tab to see details.
  - If you select **Alarms** at the bottom of the vSphere Client, you can display recent alarms and recent tasks. Click the related object to examine it.

    Note If the object with the alarm is a linked vCenter Server system, it is possible that you cannot examine that object if the issue is serious.

vSphere Authentication with vCenter Single Sign-On

4

vCenter Single Sign-On is an authentication broker and security token exchange infrastructure. When a user can authenticate to vCenter Single Sign-On, that user receives a SAML token. Going forward, the user can use the SAML token to authenticate to vCenter services. The user can then perform the actions that user has privileges for.

Because traffic is encrypted for all communications, and because only authenticated users can perform the actions that they have privileges for, your environment is secure.

Starting with vSphere 6.0, vCenter Single Sign-On is part of the Platform Services Controller. The Platform Services Controller contains the shared services that support vCenter Server and vCenter Server components. These services include vCenter Single Sign-On, VMware Certificate Authority, and License Service. See *vCenter Server Installation and Setup* for details on the Platform Services Controller.

For the initial handshake, users authenticate with a user name and password, and solution users authenticate with a certificate. For information about replacing solution user certificates, see vSphere Security Certificates.

The next step is authorizing the users who can authenticate to perform certain tasks. In most cases, you assign vCenter Server privileges, usually by assigning the user to a group that has a role. vSphere includes other permission models such as global permissions. See the *vSphere Security* documentation.

This chapter includes the following topics:

How vCenter Single Sign-On Protects Your Environment

vCenter Single Sign-On allows vSphere components to communicate with each other through a secure token mechanism.

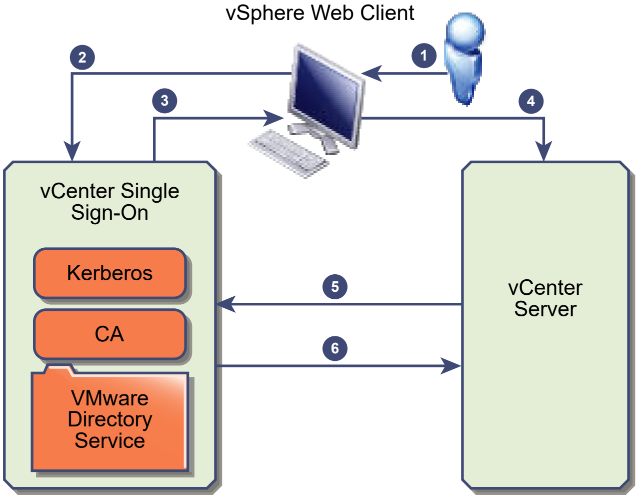
vCenter Single Sign-On uses the following services.

- STS (Security Token Service).
- SSL for secure traffic.
- Authentication of human users through Active Directory or OpenLDAP.

vCenter Single Sign-On Handshake for Human Users

The following illustration shows the handshake for human users.

Figure 4-1. vCenter Single Sign-On Handshake for Human Users



1 A user logs in to the vSphere Client with a user name and password to access the vCenter Server system or another vCenter service.

The user can also log in without a password and check the **Use Windows session authentication** check box.

2 The vSphere Client passes the login information to the vCenter Single Sign-On service, which checks the SAML token of the vSphere Client. If the vSphere Client has a valid token, vCenter Single Sign-On then checks whether the user is in the configured identity source (for example Active Directory).

- If only the user name is used, vCenter Single Sign-On checks in the default domain.
- If a domain name is included with the user name (*DOMAIN*\user1 or user1@*DOMAIN*), vCenter Single Sign-On checks that domain.

3 If the user can authenticate to the identity source, vCenter Single Sign-On returns a token that represents the user to the vSphere Client.

4 The vSphere Client passes the token to the vCenter Server system.

5 vCenter Server checks with the vCenter Single Sign-On server that the token is valid and not expired.

6 The vCenter Single Sign-On server returns the token to the vCenter Server system, using thevCenter Server Authorization Framework to allow user access.

The user can now authenticate, and can view and modify any objects that the user's role has privileges for.

Supported Encryption

AES encryption, which is the highest level of encryption, is supported. The supported encryption affects security when vCenter Single Sign-On uses Active Directory as an identity source.

Using vCenter Single Sign-On with vSphere

When a user logs in to a vSphere component or when a vCenter Server solution user accesses another vCenter Server service, vCenter Single Sign-On performs authentication. Users must be authenticated with vCenter Single Sign-On and have the necessary privileges for interacting with vSphere objects.

vCenter Single Sign-On authenticates both solution users and other users.

- Solution users represent a set of services in your vSphere environment. During installation, VMCA assigns a certificate to each solution user by default. The solution user uses that certificate to authenticate to vCenter Single Sign-On. vCenter Single Sign-On gives the solution user a SAML token, and the solution user can then interact with other services in the environment.
- When other users log in to the environment, for example, from the vSphere Client, vCenter Single Sign-On prompts for a user name and password. If vCenter Single Sign-On finds a user with those credentials in the corresponding identity source, it assigns the user a SAML token. The user can now access other services in the environment without being prompted to authenticate again.

  Which objects the user can view, and what a user can do, is usually determined by vCenter Server permission settings. vCenter Server administrators assign those permissions from the **Permissions** interface in the vSphere Web Client or the vSphere Client, not through vCenter Single Sign-On. See the *vSphere Security* documentation.

vCenter Single Sign-On and vCenter Server Users

Users authenticate to vCenter Single Sign-On by entering their credentials on the login page. After connecting to vCenter Server, authenticated users can view all vCenter Server instances or other vSphere objects for which their role gives them privileges. No further authentication is required.

After installation, the cloudadmin@vmc.local user vCenter Server has administrator access to both vCenter Single Sign-On and vCenter Server. That user can also add identity sources, set the default identity source, and set policies in the vmc.local domain. Certain management operations in the vmc.local domain are restricted to VMware Cloud on AWS operations staff.

Note When you change the password for your SDDC from the vSphere Client, the new password is not synchronized with the password that is displayed on the Default vCenter Credentials page. That page shows only the Default credentials. If you change the credentials, you are responsible for keeping track of the new password. Contact Technical Support and request a password change.

vCenter Single Sign-On Administrator Users

The vCenter Single Sign-On administrative interface is accessible from either the vSphere Client or the vSphere Web Client.

To configure vCenter Single Sign-On and manage vCenter Single Sign-On users and groups, the user administrator@vsphere.local or a user in the vCenter Single Sign-On Administrators group must log in to the vSphere Client . Upon authentication, that user can access the vCenter Single Sign-On administration interface from the vSphere Client and manage identity sources and default domains, specify password policies, and perform other administrative tasks.

Identity Sources for vCenter Server with vCenter Single Sign-On

You can use identity sources to attach one or more domains to vCenter Single Sign-On. A domain is a repository for users and groups that the vCenter Single Sign-On server can use for user authentication.

In VMware Cloud on AWS, the cloudadmin@vmc.local user or other users in the Cloud Administrator group can add identity sources. Having a shared identity source is a prerequisite for setting up Hybrid Linked Mode.

The user and group data is stored in Active Directory or OpenLDAP. Initially, your VMware Cloud on AWS SDDC has the vmc.local identity source.

vCenter Server versions earlier than version 5.1 supported Active Directory and local operating system users as user repositories. As a result, local operating system users were always able to authenticate to the vCenter Server system. vCenter Server version 5.1 and version 5.5 uses vCenter Single Sign-On for authentication. See the vSphere 5.1 documentation for a list of supported identity sources with vCenter Single Sign-On 5.1. vCenter Single Sign-On 5.5 supports the following types of user repositories as identity sources, but supports only one default identity source.

- Active Directory over LDAP. vCenter Single Sign-On supports multiple Active Directory over LDAP identity sources. This identity source type is included for compatibility with the vCenter Single Sign-On service included with vSphere 5.1. Shown as **Active Directory as an LDAP Server** in the vSphere Client.
- OpenLDAP versions 2.4 and later. vCenter Single Sign-On supports multiple OpenLDAP identity sources. Shown as **OpenLDAP** in the vSphere Client.
- Local operating system users. Local operating system users are local to the operating system where the vCenter Single Sign-On server is running. The local operating system identity source exists only in basic vCenter Single Sign-On server deployments and is not available in deployments with multiple vCenter Single Sign-On instances. Only one local operating system identity source is allowed. Shown as **localos** in the vSphere Client.

  Note Do not use local operating system users if the Platform Services Controller is on a different machine than the vCenter Server system. Using local operating system users might make sense in an embedded deployment but is not recommended.

- vCenter Single Sign-On system users. Exactly one system identity source is created when you install vCenter Single Sign-On.

Note At any time, only one default domain exists. If a user from a non-default domain logs in, that user must add the domain name (*DOMAIN\user*) to authenticate successfully.

Add an Identity Source to the SDDC LDAP Domain

The first step toward configuring Hybrid Linked Mode from your SDDC is to add your on-premises LDAP domain as an identity source for the SDDC vCenter Server .

You can configure Hybrid Linked Mode from your SDDC if your on-premises LDAP service is provided by a native Active Directory (Integrated Windows Authentication) domain or an OpenLDAP directory service.

Important If you are using OpenLDAP as the identity source, see the VMware knowledge base article at http://kb.vmware.com/kb/2064977 for additional requirements.

Prerequisites

Ensure that you meet the Common Prerequisites in Hybrid Linked Mode Prerequisites.

Procedure

1 Log in to the vSphere Client for your SDDC.

To add an identity source, you must be logged in as cloudadmin@vmc.local or another member of the Cloud Administrators group.

2 Bring up the **Add Identity Source** dialog.

Use case	Description
Hybrid Linked Mode	a Select **Menu > Administration**. b Under **Hybrid Cloud**, select **Linked Domains**. c Under **Add Cloud Administrator**, select **Add Identity Source** from the **Identity Source** drop-down menu.
All other use cases	a Select **Menu > Administration**. b Under **Single Sign On**, click **Configuration**. c Click **Identity Sources** and click **Add**.

3 Configure the identity source settings.

Option	Description
Identity Source Type	Select **Active Directory as an LDAP Server** for a Windows Active Directory Server or **Open LDAP** for an Open LDAP server.
Name	Enter the name of the identity source.
Base DN for users	Enter the Base Distinguished Name for users.
Base DN for groups	Enter the Base Distinguished Name for groups.

Option	Description
Domain Name	FQDN of the domain. Do not enter an IP address here.
Domain Alias	Enter an alias for the domain. For Active Directory identity sources, the domain's NetBIOS name. Add the NetBIOS name of the Active Directory domain as an alias of the identity source if you are using SSPI authentications.
Username	Enter the ID of a user in the domain who has a minimum of read-only access to Base DN for users and groups. Use UPN format (for example, user@example.com), rather than DN format.
Password	Enter the password of the user who is specified by **Username**.
Connect To	Select which domain controller to connect to. ■ Select **Any domain controller in the domain** to connect to any domain controller. ■ Select **Specific domain controllers** to specify the domain controllers. If you select **Specific domain controllers**, specify the URL for the primary server and the secondary server used for failover. Use the format ldap://hostname:port or ldaps://hostname:port. The port is typically 389 for ldap: connections and 636 for ldaps: connections. For Active Directory multi-domain controller deployments, the port is typically 3268 for ldap: connections and 3269 for ldaps: connections.
SSL Certificates	If you use ldaps:, select **Browse** and select a certificate file to upload to provide security for the ldaps: connection.

When the identity source is added, on-premises users can authenticate to the SDDC, but have the **No access** role. Add permissions for a group of users to give them the Cloud Administrator role.

Managing vCenter Single Sign-On Policies

vCenter Single Sign-On policies enforce the security rules in your environment. You can view and edit the default vCenter Single Sign-On password policy, lockout policy, and token policy.

Edit the vCenter Single Sign-On Password Policy

The vCenter Single Sign-On password policy determines the password format and password expiration. Password policy applies only to users in the vCenter Single Sign-On domain (vmc.local).

Procedure

1 In the vSphere Client select **Menu > Administration**.

2 Under **Single Sign On**, click **Configuration**, and click **Policies.**

3 Select **Password Policies**, click **Edit**, and make changes as needed.

Option	Description
Description	Password policy description.
Maximum lifetime	Maximum number of days that a password is valid before the user must change it.

Option	Description
Restrict reuse	Number of previous passwords that cannot be reused. For example, if you type 6, the user cannot reuse any of the last six passwords.
Maximum length	Maximum number of characters that are allowed in the password.
Minimum length	Minimum number of characters required in the password. The minimum length must be no less than the combined minimum of alphabetic, numeric, and special character requirements.
Character requirements	Minimum number of different character types that are required in the password. You can specify the number of each type of character, as follows: ■ Special: & # % ■ Alphabetic: A b c D ■ Uppercase: A B C ■ Lowercase: a b c ■ Numeric: 1 2 3 The minimum number of alphabetic characters must be no less than the combined uppercase and lowercase characters.
Identical adjacent characters	Maximum number of identical adjacent characters that are allowed in the password. For example, if you enter 1, the following password is not allowed: p@$$word. The number must be greater than 0.

4 Click **OK**.

Edit the vCenter Single Sign-On Lockout Policy

A vCenter Single Sign-On lockout policy specifies when a user's vCenter Single Sign-On account is locked if the user attempts to log in with incorrect credentials. Administrators can edit the lockout policy.

If a user logs in to vmc.local multiple times with the wrong password, the user is locked out. The lockout policy allows administrators to specify the maximum number of failed login attempts, and set the time interval between failures. The policy also specifies how much time must elapse before the account is automatically unlocked.

Procedure

1 In the vSphere Client select **Menu > Administration**.

2 Under **Single Sign On**, click **Configuration**, and click **Policies.**

3 Select **Lockout Policy**, click **Edit**, and make changes as needed.

Option	Description
Description	Optional description of the lockout policy.
Max number of failed login attempts	Maximum number of failed login attempts that are allowed before the account is locked.

Option	Description
Time interval between failures	Time period in which failed login attempts must occur to trigger a lockout.
Unlock time	Amount of time that the account remains locked. If you enter 0, the administrator must unlock the account explicitly.

4 Click **OK**.

Edit the vCenter Single Sign-On Token Policy

The vCenter Single Sign-On token policy specifies token properties such as the clock tolerance and renewal count. You can edit the token policy to ensure that the token specification conforms to security standards in your corporation.

Procedure

1 In the vSphere Client select **Menu > Administration**.

2 Under **Single Sign On**, click **Configuration**, and click **Policies.**

3 Select **Token Policy**, click **Edit**, and make changes as needed.

Option	Description
Clock tolerance	Time difference, in milliseconds, that vCenter Single Sign-On tolerates between a client clock and the domain controller clock. If the time difference is greater than the specified value, vCenter Single Sign-On declares the token invalid.
Maximum token renewal count	Maximum number of times that a token can be renewed. After the maximum number of renewal attempts, a new security token is required.
Maximum token delegation count	Holder-of-key tokens can be delegated to services in the vSphere environment. A service that uses a delegated token performs the service on behalf of the principal that provided the token. A token request specifies a DelegateTo identity. The DelegateTo value can either be a solution token or a reference to a solution token. This value specifies how many times a single holder-of-key token can be delegated.
Maximum bearer token lifetime	Bearer tokens provide authentication based only on possession of the token. Bearer tokens are intended for short-term, single-operation use. A bearer token does not verify the identity of the user or entity that is sending the request. This value specifies the lifetime value of a bearer token before the token has to be reissued.
Maximum holder-of-key token lifetime	Holder-of-key tokens provide authentication based on security artifacts that are embedded in the token. Holder-of-key tokens can be used for delegation. A client can obtain a holder-of-key token and delegate that token to another entity. The token contains the claims to identify the originator and the delegate. In the vSphere environment, a vCenter Server system obtains delegated tokens on a user's behalf and uses those tokens to perform operations. This value determines the lifetime of a holder-of-key token before the token is marked invalid.

4 Click **OK**.

Securing vCenter Server Systems

5

In an on-premises SDDC, you are responsible for ensuring the security of your vCenter Server system. In VMware Cloud on AWS VMware performs most of these tasks for you.

You are responsible for following security best practices, especially for the VMs in your environment, and might want to be aware of some other aspects of vCenter Server and vCenter Single Sign-On such as password and lockout policies.

This chapter includes the following topics:

Security Best Practices and Resources

If you follow best practices, your ESXi and vCenter Server can be as secure as or even more secure than an environment that does not include virtualization.

This manual includes best practices for the different components of your vSphere infrastructure.

Table 5-1. Security Best Practices

vSphere component	Resource
ESXi host	Securing ESXi Hosts
vCenter Server system	vCenter Server Security Best Practices
Virtual machine	Virtual Machine Security Best Practices
vSphere Networking	vSphere Networking Security Best Practices

This manual is only one of the sources you must use to ensure a secure environment.

VMware security resources, including security alerts and downloads, are available on the Web.

Table 5-2. VMware Security Resources on the Web

Topic	Resource
Information on ESXi and vCenter Server security and operations, including secure configuration and hypervisor security.	https://vspherecentral.vmware.com/t/security/
VMware security policy, up-to-date security alerts, security downloads, and focus discussions of security topics.	http://www.vmware.com/go/security
Corporate security response policy	http://www.vmware.com/support/policies/security_response.html VMware is committed to helping you maintain a secure environment. Security issues are corrected in a timely manner. The VMware Security Response Policy states our commitment to resolve possible vulnerabilities in our products.
Third-party software support policy	http://www.vmware.com/support/policies/ VMware supports a variety of storage systems, software agents such as backup agents, system management agents, and so forth. You can find lists of agents, tools, and other software that supports ESXi by searching http://www.vmware.com/vmtn/resources/ for ESXi compatibility guides. The industry offers more products and configurations than VMware can test. If VMware does not list a product or configuration in a compatibility guide, Technical Support attempts to help you with any problems, but cannot guarantee that the product or configuration can be used. Always evaluate security risks for unsupported products or configurations carefully.
Compliance and security standards, and partner solutions and in-depth content about virtualization and compliance	http://www.vmware.com/go/compliance
Information on security certifications and validations such as CCEVS and FIPS for different versions of the components of vSphere.	https://www.vmware.com/support/support-resources/certifications.html
Security configuration guides (formerly known as hardening guides) for different versions of vSphere and other VMware products.	https://www.vmware.com/support/support-resources/hardening-guides.html
Security of the VMware vSphere Hypervisor white paper	http://www.vmware.com/files/pdf/techpaper/vmw-wp-secrty-vsphr-hyprvsr-uslet-101.pdf

Passwords in Your vSphere Environment

Password restrictions, password expiration, and account lockout in your vSphere environment depend on the system that the user targets, who the user is, and how policies are set.

ESXi Passwords

ESXi password restrictions are determined by the Linux PAM module `pam_passwdqc`. See the Linux manpage for `pam_passwdqc` and see #unique_42.

Passwords for vCenter Server and Other vCenter Services

vCenter Single Sign-On manages authentication for all users who log in to vCenter Server and other vCenter services. The password restrictions, password expiration, and account lockout depend on the user's domain and on who the user is.

vCenter Single Sign-On Administrator

The password for administrator@vsphere.local user, or the administrator@*mydomain* user if you selected a different domain during installation, does not expire and is not subject to the lockout policy. In all other regards, the password must follow the restrictions that are set in the vCenter Single Sign-On password policy. See *Platform Services Controller Administration* for details.

If you forget the password for this user, search the VMware Knowledge Base system for information on resetting this password. The reset requires additional privileges such as root access to the vCenter Server system.

Other Users of the vCenter Single Sign-On Domain

Passwords for other vsphere.local users, or users of the domain that you specified during installation, must follow the restrictions that are set by the vCenter Single Sign-On password policy and lockout policy. See *Platform Services Controller Administration* for details. These passwords expire after 90 days by default. Administrators can change the expiration as part of the password policy.

If you forget your vsphere.local password, an administrator user can reset the password using the `dir-cli` command.

Other Users

Password restrictions, password expiration, and account lockout for all other users are determined by the domain (identity source) to which the user can authenticate.

vCenter Single Sign-On supports one default identity source. Users can log in to the corresponding domain with the vSphere Client with just their user names. If users want to log in to a non-default domain, they can include the domain name, that is, specify *user@domain* or *domain\user*. The domain password parameters apply to each domain.

Passwords for vCenter Server Appliance Direct Console User Interface Users

The vCenter Server Appliance is a preconfigured Linux-based virtual machine that is optimized for running vCenter Server and the associated services on Linux.

When you deploy the vCenter Server Appliance, you specify these passwords.

- Password for the root user of the appliance Linux operating system.
- Password for the administrator of the vCenter Single Sign-On domain, administrator@vsphere.local by default.

You can change the root user password and perform other vCenter Server Appliance local user management tasks from the appliance console. See *vCenter Server Appliance Configuration*.

Best Practices for vCenter Server Access Control

Strictly control access to different vCenter Server components to increase security for the system.

The following guidelines help ensure security of your environment.

Use Named Accounts

Make sure that applications use unique service accounts when connecting to a vCenter Server system.

Minimize Access

Do not allow users to log directly in to the vCenter Server host machine. Users who are logged in to the vCenter Server host machine can cause harm, either intentionally or unintentionally, by altering settings and modifying processes. Those users also have potential access to vCenter credentials, such as the SSL certificate. Allow only users who have legitimate tasks to perform to log in to the system and ensure that login events are audited.

Restrict Users From Running Commands in a Virtual Machine

By default, a user with the vCenter Server Administrator role can interact with files and programs within a virtual machine's guest operating system. To reduce the risk of breaching guest confidentiality, availability, or integrity, create a custom nonguest access role without the **Guest Operations** privilege.

Use High RDP Encryption Levels

On each Windows computer in the infrastructure, ensure that Remote Desktop Host Configuration settings are set to ensure the highest level of encryption appropriate for your environment.

Verify vSphere Client Certificates

Instruct users of one of the vSphere Client or other client applications to never ignore certificate verification warnings. Without certificate verification, the user might be subject of a MiTM attack.

vCenter Password Requirements and Lockout Behavior

To manage your vSphere environment, you must be aware of the vCenter Single Sign-On password policy, of vCenter Server passwords, and of lockout behavior.

vCenter Single Sign-On Administrator Password

The password for the administrator of vCenter Single Sign-On, administrator@vsphere.local by default, is specified by the vCenter Single Sign-On password policy. By default, this password must meet the following requirements:

- At least 8 characters

- At least one lowercase character
- At least one numeric character
- At least one special character

The password for this user cannot be more than 20 characters long. Starting with vSphere 6.0, non-ASCII characters are allowed.

vCenter Server Passwords

In vCenter Server, password requirements are dictated by vCenter Single Sign-On or by the configured identity source, which can be Active Directory, OpenLDAP.

vCenter Single Sign-On Lockout Behavior

Users are locked out after a preset number of consecutive failed attempts. By default, users are locked out after five consecutive failed attempts in three minutes and a locked account is unlocked automatically after five minutes.

Starting with vSphere 6.0, the vCenter Single Sign-On domain administrator, administrator@vsphere.local by default, is not affected by the lockout policy. The user is affected by the password policy.

vSphere Permissions and User Management Tasks 6

Authentication and authorization govern access. vCenter Single Sign-On supports authentication, which means it determines whether a user can access vSphere components at all. Each user must also be authorized to view or manipulate vSphere objects.

vSphere supports several different authorization mechanisms, discussed in Understanding Authorization in vSphere. The focus of the information in this section is how the vCenter Server permission model works and how to perform user management tasks.

vCenter Server allows fine-grained control over authorization with permissions and roles. When you assign a permission to an object in the vCenter Server object hierarchy, you specify which user or group has which privileges on that object. To specify the privileges, you use roles, which are sets of privileges.

Initially, only the administrator user for the vCenter Single Sign-On domain, administrator@vsphere.local by default, is authorized to log in to the vCenter Server system. That user can then proceed as follows:

1 Add an identity source in which users and groups are defined to vCenter Single Sign-On. See the *Platform Services Controller Administration* documentation.

2 Give privileges to a user or group by selecting an object such as a virtual machine or a vCenter Server system and assigning a role on that object for the user or group.

Assigning Roles and Permissions Using the vSphere Client (http://link.brightcove.com/services/player/bcpid2296383276001?bctid=ref:video_vsphere67_roles)

This chapter includes the following topics:

Understanding Authorization in vSphere

vSphere supports several models with fine-grained control for determining whether a user is allowed to perform a task. vCenter Single Sign-On uses group membership in a vCenter Single Sign-On group to decide what you are allowed to do. Your role on an object or your global permission determines whether you're allowed to perform other tasks in vSphere.

Authorization Overview

vSphere 6.0 and later allows privileged users to give other users permissions to perform tasks. You can use global permissions, or you can use local vCenter Server permissions to authorize other users for individual vCenter Server instances.

vCenter Server Permissions

The permission model for vCenter Server systems relies on assigning permissions to objects in the object hierarchy. Each permission gives one user or group a set of privileges, that is, a role for a selected object. For example, you can select a virtual machine and select **Add Permission** assign a role to a group of users in a domain that you select. That role gives those users the corresponding privileges on the VM.

Global Permissions

Global permissions are applied to a global root object that spans solutions. For example, if both vCenter Server and vRealize Orchestrator are installed, you can use global permissions. For example, you can give a group of users Read permissions to all objects in both object hierarchies.

Global permissions are not replicated if your environment includes an on-premises vCenter Server and a vCenter Server in the cloud.

Group Membership in vCenter Single Sign-On Groups

For VMware Cloud on AWS, a Cloud Administrator group is predefined in vCenter Single Sign-On. When you use Hybrid Linked mode, you add this group to the linked domain.

Understanding the Object-Level Permission Model

You authorize a user or group to perform tasks on vCenter objects by using permissions on the object. The vSphere permission model relies on assigning permissions to objects in the vSphere object hierarchy. Each permission gives one user or group a set of privileges, that is, a role for the selected object. For example, a group of users might have the ReadOnly role on one VM and the Administrator role on another VM.

The following concepts are important.

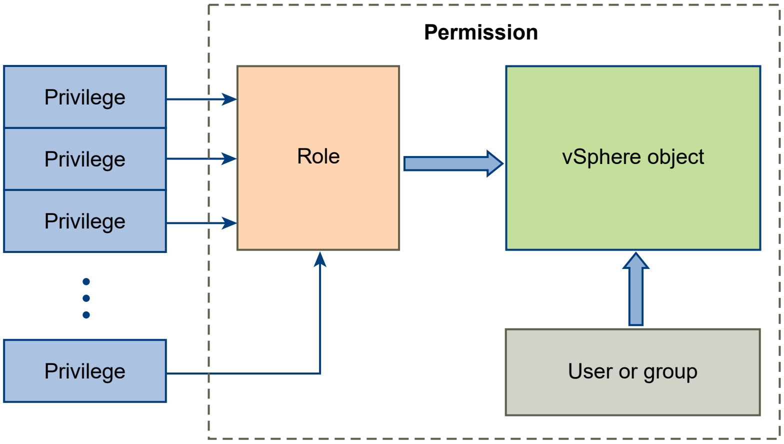
Permissions Each object in the vCenter Server object hierarchy has associated permissions. Each permission specifies for one group or user which privileges that group or user has on the object.

Users and Groups On vCenter Server systems, you can assign privileges only to authenticated users or groups of authenticated users. Users are authenticated through vCenter Single Sign-On. Users and groups must be defined in the identity source that vCenter Single Sign-On uses to authenticate. Define users and groups using the tools in your identity source, for example, Active Directory.

Privileges Privileges are fine-grained access controls. You can group those privileges into roles, which you can then map to users or groups.

Roles Roles are sets of privileges. Roles allow you to assign permissions on an object based on a typical set of tasks that users perform. Default roles, such as Administrator, are predefined on vCenter Server and cannot be changed. Other roles, such as Resource Pool Administrator, are predefined sample roles. You can create custom roles either from scratch or by cloning and modifying sample roles. See Create a Custom Role.

Figure 6-1. vSphere Permissions



To assign permissions to an object, you follow these steps:

1 Select the object to which you want to apply the permission in the vCenter object hierarchy.

In VMware Cloud on AWS you cannot change permissions on objects that VMware manages for you, for example, the vCenter Server instance or ESXi hosts.

2 Select the group or user that should have privileges on the object.

3 Select individual privileges or a role, that is a set of privileges, that the group or user should have on the object.

By default, permissions propagate, that is the group or user has the selected role on the selected object and its child objects.

vCenter Server offers predefined roles, which combine frequently used privilege sets. You can also create custom roles by combining a set of roles.

Permissions must often be defined on both a source object and a destination object. For example, if you move a virtual machine, you need privileges on that virtual machine, but also privileges on the destination data center.

See the following information.

To find out about...	See...
Creating custom roles.	Create a Custom Role
All privileges and the objects to which you can apply the privileges	Defined Privileges
Sets of privileges that are required on different objects for different tasks.	Required Privileges for Common Tasks

vCenter Server User Validation

vCenter Server systems that use a directory service regularly validate users and groups against the user directory domain. Validation occurs at regular intervals specified in the vCenter Server settings. For example, assume that user Smith is assigned a role on several objects. The domain administrator changes the name to Smith2. The host concludes that Smith no longer exists and removes permissions associated with that user from the vSphere objects when the next validation occurs.

Similarly, if user Smith is removed from the domain, all permissions associated with that user are removed when the next validation occurs. If a new user Smith is added to the domain before the next validation occurs, the new user Smith replaces the old user Smith in permissions on any object.

View Permissions and Privileges

For each object in the hierarchy, you can use the vSphere Client view the privileges granted to users in the CloudAdmin role as well as other predefined or custom roles.

Procedure

1. Select an object in the object hierarchy, for example a resource pool or virtual machine, and click **Permissions**.
2. You can then view the privileges associated with each group.
   a. On the vSphere Client Home page, click **Administration**.
   b. Under **Access Control**, click **Roles**.
   c. Click a role name (**CloudAdmin**, for example).
   d. Click the **Privileges** tab on the right.

You can scroll through the list to see the privileges granted to the selected role. See Defined Privileges for a detailed list of all vSphere privileges.

Managing Permissions for vCenter Components

A permission is set on an object in the vCenter object hierarchy. Each permission associates the object with a group or user and the group's or user's access roles. For example, you can select a virtual machine object, add one permission that gives the ReadOnly role to Group 1, and add a second permission that gives the Administrator role to User 2.

By assigning a different role to a group of users on different objects, you control the tasks that those users can perform in your vSphere environment. For example, to allow a group to configure memory for the host, select that host and add a permission that grants a role to that group that includes the **Host.Configuration.Memory Configuration** privilege.

To manage permissions from the vSphere Client, you need to understand the following concepts:

Permissions

Each object in the vCenter Server object hierarchy has associated permissions. Each permission specifies for one group or user which privileges that group or user has on the object.

Users and Groups

On vCenter Server systems, you can assign privileges only to authenticated users or groups of authenticated users. Users are authenticated through vCenter Single Sign-On. Users and groups must be defined in the identity source that vCenter Single Sign-On uses to authenticate. Define users and groups using the tools in your identity source, for example, Active Directory.

Privileges

Privileges are fine-grained access controls. You can group those privileges into roles, which you can then map to users or groups.

Roles

Roles are sets of privileges. Roles allow you to assign permissions on an object based on a typical set of tasks that users perform. Default roles, such as Administrator, are predefined on vCenter Server and cannot be changed. Other roles, such as Resource Pool Administrator, are predefined sample roles. You can create custom roles either from scratch or by cloning and modifying sample roles. See Create a Custom Role.

You can assign permissions to objects at different levels of the hierarchy, for example, you can assign permissions to a host object or to a folder object that includes all host objects. See Hierarchical Inheritance of Permissions. You can also assign permissions to a global root object to apply the permissions to all object in all solutions. See Global Permissions.

Add a Permission to an Inventory Object

After you create users and groups and define roles, you must assign the users and groups and their roles to the relevant inventory objects. You can assign the same permissions to multiple objects simultaneously by moving the objects into a folder and setting the permissions on the folder.

When you assign permissions, user and group names must match Active Directory precisely, including case. If you upgraded from earlier versions of vSphere, check for case inconsistencies if you experience problems with groups.

Prerequisites

On the object whose permissions you want to modify, you must have a role that includes the **Permissions.Modify permission** privilege.

Procedure

1. Browse to the object for which you want to assign permissions in the vSphere Client object navigator.
2. Click the **Permissions** tab.
3. Click the **Add Permission** icon.
4. Select the user or group that will have the privileges defined by the selected role.
   1. From the **User** drop-down menu, select the domain for the user or group.
   2. Type a name in the Search box.

      The system searches user names and group names.
   3. Select the user or group.
5. Select a role from the **Role** drop-down menu.
6. (Optional) To propagate the permissions, select the **Propagate to children** check box.

   The role is applied to the selected object and propagates to the child objects.
7. Click **OK** to add the permission.

Change or Remove Permissions

After a user or group and role pair is set for an inventory object, you can change the role paired with the user or group or change the setting of the **Propagate to children** check box. You can also remove the permission setting.

Procedure

1. Browse to the object in the vSphere Client object navigator.
2. Click the **Permissions** tab.

3 Click a row to select a permission.

Task	Steps
Change permissions	a Click the **Change Role** icon. b Select a role for the user or group from the **Role** drop-down menu. c Toggle the **Propagate to children** check box to change permission inheritance. d Click **OK**.
Remove permissions	Click the **Remove Permission** icon.

Change User Validation Settings

vCenter Server periodically validates its user and group lists against the users and groups in the user directory. It then removes users or groups that no longer exist in the domain. You can disable validation or change the interval between validations. If you have domains with thousands of users or groups, or if searches take a long time to complete, consider adjusting the search settings.

Note This procedure applies only to vCenter Server user lists. You cannot search ESXi user lists in the same way.

Procedure

1 Browse to the vCenter Server system in the vSphere Client object navigator.

2 Select **Configure** and click **Settings > General**.

3 Click **Edit** and select **User directory**.

4 Change the values as needed and click **Save**.

Option	Description
User directory timeout	Timeout interval, in seconds, for connecting to the Active Directory server. This value specifies the maximum amount of time vCenter Server allows a search to run on the selected domain. Searching large domains can take a long time.
Query limit	Toggle on to set a maximum number of users and groups that vCenter Server displays.
Query limit size	Maximum number of users and groups from the selected domain that vCenter Server displays in the **Select Users or Groups** dialog box. If you enter 0 (zero), all users and groups appear.
Validation	Toggle off to disable validation.
Validation Period	Specifies how often vCenter Server validates permissions, in minutes.

Global Permissions

Global permissions are applied to a global root object that spans solutions. In an on-premises SDDC, global permissions might span both vCenter Server and vRealize Orchestrator. But for any vSphere SDDC, global permissions apply to global objects such as tags and content libraries.

You can assign global permissions to users or groups, and decide on the role for each user or group. The role determines the set of privileges that the user or group has for all objects in the hierarchy. You can assign a predefined role or create custom roles. See #unique_53.

It is important to distinguish between vCenter Server permissions and global permissions.

vCenter Server permissions

You usually apply a permission to a vCenter Server inventory object such as a virtual machine. When you do, you specify that a user or group has a role (set of privileges) on the object.

Global permissions

Global permissions give a user or group privileges to view or manage all objects in each of the inventory hierarchies in your deployment. Global permissions also apply to global objects such as tags and content libraries. See Permissions on Tag Objects. Global permissions do not apply to objects such as SDDC hosts and datastores that VMware manages for you.

If you assign a global permission and do not select Propagate, the users or groups associated with this permission do not have access to the objects in the hierarchy. They only have access to some global functionality such as creating roles.

Important Use global permissions with care. Verify that you really want to assign permissions to all objects in all inventory hierarchies.

vCenter Server System Roles

A role is a predefined set of privileges. When you add permissions to an object, you pair a user or group with a role. vCenter Server includes several system roles, which you cannot change.

vCenter Server provides a few default roles. You cannot change the privileges associated with the default roles. The default roles are organized as a hierarchy. Each role inherits the privileges of the previous role. For example, the Administrator role inherits the privileges of the Read Only role.

The vCenter Server role hierarchy also includes several sample roles. You can clone a sample role to create a similar role.

CloudAdmin Role

The CloudAdmin role has the necessary privileges for you to create and manage workloads on your SDDC. However, you cannot access or configuring the certain management components that are supported and managed by VMware, such as hosts, clusters, and management virtual machines.

CloudGlobalAdmin Role

The CoudGlobalAdmin role, which has a subset of the privileges granted to the CloudAdmin role, is deprecated as of SDDC version 1.7.

Administrator Role

Users with the Administrator role for an object are allowed to view and perform all actions on the object. This role also includes all privileges of the

Read Only role. If you have the Administrator role on an object, you can assign privileges to individual users and groups.

If you are acting in the Administrator role in vCenter Server, you can assign privileges to users and groups in the default vCenter Single Sign-On identity source. Supported identity services include Windows Active Directory and OpenLDAP 2.4.

By default, the administrator@vsphere.local user has the Administrator role on both vCenter Single Sign-On and vCenter Server after installation. That user can then associate other users with the Administrator role on vCenter Server.

Read Only Role

Users with the Read Only role for an object are allowed to view the state of the object and details about the object. For example, users with this role can view virtual machine, host, and resource pool attributes, but cannot view the remote console for a host. All actions through the menus and toolbars are disallowed.

No Access Role

Users with the No Access role for an object cannot view or change the object in any way. New users and groups are assigned this role by default. You can change the role on an object-by-object basis.

The administrator of the vCenter Single Sign-On domain, administrator@vsphere.local by default, the root user, and vpxuser are assigned the Administrator role by default. Other users are assigned the No Access role by default.

No Cryptography Administrator Role

Users with the No cryptography administrator role for an object have the same privileges as users with the Administrator role, except for **Cryptographic operations** privileges. This role allows administrators to designate other administrators that cannot encrypt or decrypt virtual machines or access encrypted data, but that can perform all other administrative tasks.

Best Practices for Roles and Permissions

Follow best practices for roles and permissions to maximize the security and manageability of your vCenter Server environment.

VMware recommends the following best practices when configuring roles and permissions in your vCenter Server environment:

- Where possible, assign a role to a group rather than individual users.
- Grant permissions only on the objects where they are needed, and assign privileges only to users or groups that must have them. Use the minimum number of permissions to make it easier to understand and manage your permissions structure.

- If you assign a restrictive role to a group, check that the group does not contain the Administrator user or other users with administrative privileges. Otherwise, you might unintentionally restrict administrators' privileges in the parts of the inventory hierarchy where you have assigned that group the restrictive role.
- Use folders to group objects. For example, to grant modify permission on one set of hosts and view permission on another set of hosts, place each set of hosts in a folder.
- Consider enabling propagation when you assign permissions to an object. Propagation ensures that new objects in the object hierarchy inherit permissions. For example, you can assign a permission to a virtual machine folder and enable propagation to ensure the permission applies to all VMs in the folder.
- Use the No Access role to mask specific areas of the hierarchy. The No Access role restricts access for the users or groups with that role.

CloudAdmin Privileges

Because VMware performs host administration and other tasks for you, a Cloud Administrator requires fewer privileges than an Administrator user on an on-premises data center.

The CloudAdmin role has a set of privileges that is dynamically generated for your SDDC. They include most of the available privileges in all categories. To view the privileges granted to the CloudAdmin role, log into the SDDC vSphere Client, click **Administration > Roles**, select CloudAdmin from the list of roles, then click **PRIVILEGES**.

For more information on the permissions granted by each privilege , see the vSphere Defined Privileges reference.

Clusters and Resource Pools

7

In an on-premises vSphere environment, you configure clusters to group ESXi hosts and to set vSphere HA, vSphere DRS, and other cluster features. You use resource pools group resources. In a VMware Cloud on AWS environment, VMware creates a single cluster with a preset configuration. VMware creates a resource pool for compute VMs and a second resource pool for management VMs. You can view cluster and resource pool settings and create and configure child resource pools.

What you can do in VMware Cloud on AWS depend on what you select.

Table 7-1. Supported Tasks on Clusters and Resource Pools in VMware Cloud on AWS

Object	Supported Tasks
Cluster	A VMware Cloud on AWS environment has one cluster that includes all ESXi hosts that are part of your SDDC. ■ View the cluster configuration, including vSphere DRS and vSphere HA. The cloudadmin user cannot change the cluster configuration. ■ Examine all hosts and all resource pools that are associated with the cluster. You can see the consumed memory and CPU, HA state, and uptime. ■ Examine all VMs, datastores, and networks that are associated with the cluster. ■ Set tags and attributes. See Chapter 10 vSphere Tags and Attributes.
Resource pool	A VMware Cloud on AWS environment has two predefined resource pools. You can perform the following tasks: ■ Create new virtual machines and child resource pools. ■ Change resource allocation settings on child resource pools ■ Rename the resource pools to better match company policy. ■ Monitor the resource pool, its VMs, and its child resource pools, and examine resource pool utilization. ■ Set tags and attributes. See Chapter 10 vSphere Tags and Attributes.

Note Some menu options, such as **Delete**, are available on top-level resource pools but have no effect. As cloudadmin@vmc.local, you do not have permissions to perform those tasks. A warning appears in the Alarms window.

This chapter includes the following topics:

- Examine and Monitor vSphere DRS
- Examine and Monitor vSphere HA
- Examine Cluster Configuration
- Create and Manage Child Resource Pools

Predefined Clusters and Resource Pools

Your VMware Cloud on AWS SDDC includes a cluster and two resource pools. A vSphere cluster organizes and manages all CPU, memory, and storage resources of a set of hosts. Inside each cluster, different resource pools can further carve up the resources. You can create child resource pools of the Compute-ResourcePool.

In a VMware Cloud on AWS SDDC, VMware creates a cluster and manages all the resources in the cluster for you. That means if a VM needs more resources, vSphere DRS migrates it to a different host. If one of the hosts becomes unavailable, vSphere HA immediately brings up all the VMs running on that host on different hosts.

Resource pools allow further division of resource allocation but do not affect the behavior of vSphere DRS and vSphere HA. A VMware Cloud on AWS SDDC has two predefined resource pools. Both resource pools share the same physical hardware but are dedicated to different uses.

Compute-ResourcePool By default, all workload virtual machines are created to the top-level Compute-ResourcePool. You can create child resource pools of this resource pool if you want to prioritize resource allocation to different VMs.

Mgmt-ResourcePool All management virtual machines, for example the NSX Manager and NSX Controller instances, are created in the Mgmt-ResourcePool. Resources in this pool are reserved for management VMs so that they can operate correctly without consuming resources from the Compute-ResourcePool.

Table 7-2. Resources Consumed by Management VMs

Resource Type	Average Resource Consumption
vCPU	32
Memory	104 GB
Storage	3898 GB

Examine VMs and Hosts in the Cluster

In an VMware Cloud on AWS environment, you can examine the VMs and hosts in a cluster.

Procedure

1 In the vSphere Client, click **Menu** and select **Hosts and Clusters**.

2 Select **Cluster-1**.

In VMware Cloud on AWS, a single cluster holds the configuration settings. Because the settings are predefined, a second cluster is not needed.

3 Click **VMs** to examine virtual machines and vApps.

a Check how much CPU and memory that each VM is consuming, and check the allocated and used storage.

b Right-click a down arrow in the title bar to show or hide columns in this display.

c If you want to make changes to a VM or vApp, select it, right-click the VM in the object hierarchy, click **Settings**, and make the change.

See the **Virtual Machine Management** documentation for VMware Cloud on AWS for details.

4 Click **Hosts** to see the resource that the hosts in the cluster are consuming.

You can request additional hosts if current resource useage shows that you'll need them soon.

Examine and Monitor vSphere DRS

vSphere DRS ensures optimal resource allocation across the VMs in your SDDC. In an on-premises environment, you can configure several options, for example, you can use fully automated DRS or decide to receive recommendations. In a VMware Cloud on AWS SDDC, VMware has preconfigured the vSphere DRS options for the cluster of hosts.

VMware has selected settings that ensure optimal resource distribution while minimizing the number of migrations that occur. You don't have to - and you can't - change those selections. You might find it interesting to look at the preconfigured values.

Note For a seamless user experience, the screens for configuring, monitoring, and examining clusters and resource pools are the same in an on-premises SDDC and in a VMware Cloud on AWS SDDC. However, the **Edit** button is grayed out for VMware Cloud on AWS.

Procedure

1 In the vSphere Client, click **Menu** and select **Hosts and Clusters**.

2 Select **Cluster-1**.

In VMware Cloud on AWS, a single cluster holds the configuration settings. Because the settings are predefined, a second cluster is not needed.

3 Examine the DRS settings, which fully automate resource allocation across the cluster.

The UI explains the background information. For detailed information, see the *vSphere Resource Management* documentation.

Examine and Monitor vSphere HA

vSphere HA ensures availability of your virtual machines. If a host fails, vSphere HA restarts the VM on a different host. In a VMware Cloud on AWS SDDC, VMware immediately replaces failed hosts, so you can always count on the full set of resources in your SDDC

Note For a seamless user experience, the screens for configuring, monitoring, and examining clusters and resource pools are the same in an on-premises SDDC and in a VMware Cloud on AWS SDDC. However, the **Edit** button is grayed out for VMware Cloud on AWS.

Procedure

1 In the vSphere Client, click **Menu** and select **Hosts and Clusters**.

2 Select **Cluster-1**.

In VMware Cloud on AWS, a single cluster holds the configuration settings. Because the settings are predefined, a second cluster is not needed.

3 Examine the vSphere HA settings, which are optimized for a cloud environment.

Setting	Discussion
vSphere HA is Turned ON	With vSphere HA turned on, virtual machines are protected against host failure.
Proactive HA is Turned OFF	Because VMware immediately replaces failed hosts, Proactive HA is not needed in the VMware Cloud on AWS environment.
Failure conditions and responses	Because your VMware Cloud on AWS is protected at the hardware level, some vSphere HA features are disabled. It makes no sense to have these features enabled in the cloud.
Admission Control	Shows reserved CPU and memory capacity. If your cluster does not have the specified failover capacity, you cannot start additional VMs.
Advanced Options	Displays the settings for advanced options that are preset in the cloud SDDC.

For detailed discussions of vSphere HA and other features that ensure availability in an on-prem vSphere environment, see the *vSphere Availability* documentation. Many of the features that on-prem administrators have to manage are preset in VMware Cloud on AWS.

4 Click the **vSphere HA Monitoring** link for more information on vSphere HA events.

Examine Cluster Configuration

In your VMware Cloud on AWS SDDC, you can examine all cluster configuration details that you can view and change in an on-premises deployment.

Note For a seamless user experience, the screens for configuring, monitoring, and examining clusters and resource pools are the same in an on-premises SDDC and in a VMware Cloud on AWS SDDC. However, the **Edit** button is grayed out for VMware Cloud on AWS.

Procedure

1 In the vSphere Client, click **Menu** and select **Hosts and Clusters**.

2 Select **Cluster-1**.

In VMware Cloud on AWS, a single cluster holds the configuration settings. Because the settings are predefined, a second cluster is not needed.

3 Examine the following settings.

Configuration Setting	Description
General	Displays the predefined swap file location and VM Compatibility settings.
VMware EVC	VMware Enhanced vMotion Compatibility is disabled. The hosts are uniform, so vMotion Compatibility problems do not occur.
VM/Host Groups	
VM/Host Rules	The administrators of your VMware Cloud on AWS environment create rules to ensure that certain virtual machines never run on the same rules. You can view those rules, but you cannot create rules.
VM Overrides	VM Overrides change the behavior of certain VMs. For example, the vCenter VM has the highest vSphere HA Restart Priority. You can view the overrides that the VMware Cloud on AWS administrator set for some system VMs. You cannot specify overrides for your own VMs.
Host Options	Displays host options, including some information about the host.
Host Profile	All ESXi hosts are managed by VMware and are set up in a uniform fashion. Host profiles are not required.

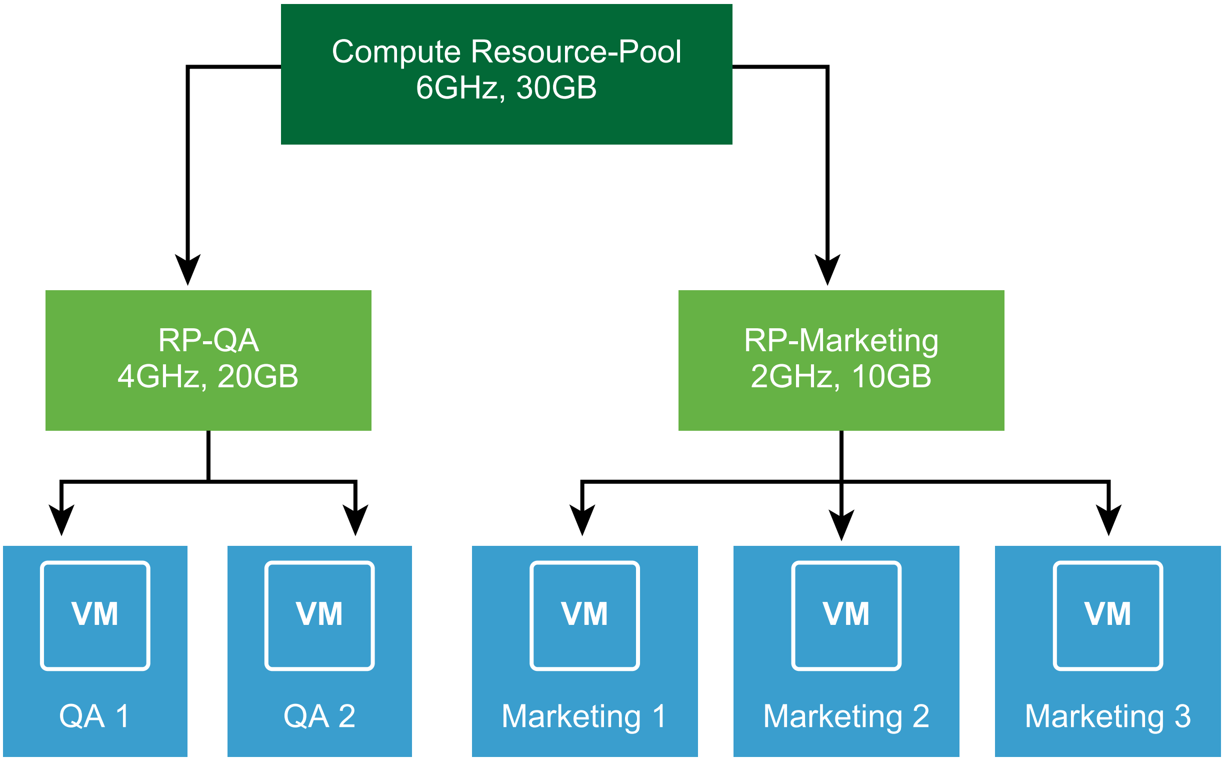
Create and Manage Child Resource Pools

Resource pools allow you perform resource allocation depending on the needs of different groups. You can create a hierarchy of child resource pools for the top-level compute resource pool, called Compute-ResourcePool by default. You can specify resource settings when you create a resource pool, and you can change those settings later.

For example, assume a host has a number of virtual machines. The marketing department uses three of the virtual machines and the QA department uses two virtual machines. Because the QA department needs larger amounts of CPU and memory, the administrator creates one resource pool for each group. The administrator sets **CPU Shares** to **High** for the QA department pool and to **Normal** for the Marketing department pool so that the QA department users can run automated tests. The second resource pool with fewer CPU and memory resources is sufficient for the lighter load of the marketing staff. Whenever the QA department is not fully using its allocation, the marketing department can use the available resources.

The numbers in the following figure show the effective allocations to the resource pools.

Figure 7-1. Parent Resource Pools and Child Resource Pools



Procedure

1 Start the task.

Task	Steps
Create a resource pool	Right-click the parent resource pool and select **New Resource Pool**.
Edit resource pool settings	Right-click a resource pool and select **Edit Resource Settings**. **Note** If you edit the settings of a system-defined resource pools, the changes do not take effect.

2 Specify how to allocate CPU and memory resources.

Option	Description
Name	Name for this resource pool.
Shares	Specify shares for this resource pool with respect to the parent's total resources. Sibling resource pools share resources according to their relative share values bounded by the reservation and limit. ■ Select **Low**, **Normal**, or **High** to specify share values respectively in a 1:2:4 ratio. ■ Select **Custom** to give each virtual machine a specific number of shares, which expresses a proportional weight.
Reservation	Specify a guaranteed CPU or memory allocation for this resource pool. Defaults to 0. A nonzero reservation is subtracted from the unreserved resources of the parent (host or resource pool). The resources are considered reserved, regardless of whether virtual machines are associated with the resource pool.
Expandable Reservation	When the check box is selected (default), expandable reservations are considered during admission control. If you power on a virtual machine in this resource pool, and the combined reservations of the virtual machines are larger than the reservation of the resource pool, the resource pool can use resources from its parent or ancestors.
Limit	Specify the upper limit for this resource pool's CPU or memory allocation. You can usually accept the default (**Unlimited**). To specify a limit, deselect the **Unlimited** check box.

3 Click **OK**.

vSan Storage inVMware Cloud on AWS

8

VMware Cloud on AWS provides two vSAN datastores in each SDDC cluster: WorkloadDatastore, managed by the Cloud Administrator, and vsanDatastore, managed by VMware.

These datastores are logical entities that share a common capacity pool. Each datastore reports the total available free space in the cluster as its **Capacity**. Capacity consumed in either datastore updates the **Free** value for both.

vsanDatastore

This datastore provides storage for the management VMs in your SDDC, such as vCenter Server, NSX controllers, and so on.

The management and troubleshooting of the vSAN storage in your SDDC is handled by VMware. For this reason, you can't edit the vSAN cluster settings or monitor the vSAN cluster. You also do not have permission to browse this datastore, upload files to it, or delete files from it.

WorkloadDatastore

This datastore provides storage for your workload VMs, templates, ISO images, and any other files you choose to upload to your SDDC. You have full permission to browse this datastore, create folders, upload files, delete files, and perform all other operations needed to consume this storage.

The datastores in your SDDC are assigned the default VM storage policy by default. You can define additional storage policies and assign them to either datastore. For more information on vSAN storage policies, see Using vSAN Policies.

This chapter includes the following topics:

Storage Capacity and Data Redundancy

SDDC storage capacity and data redundancy scale with the number of nodes in the SDDC.

Single node SDDCs provide no data redundancy. SDDCs with more nodes support data redundancy through various RAID configurations. All SDDC storage provides data deduplication and compression. Data redundancy in SDDCs that have more than one host is expressed as a number of failures to tolerate (FTT), where a failure can be the loss of a single host in a cluster or a single storage device in an array.

Typical Storage Allocation and Consumption Scenarios

All RAID configurations consume data to support redundancy, as shown in this table. The Usable Storage column shows how much data you could store on a given host under a specific vSAN policy, and accounts for recommended free space by subtracting 30% from that total.

Table 8-1. Usable Storage Per Node

Nodes	FTT and RAID Config	Usable Storage (TB)	Projected Usable Capacity After Dedup/ Compression (TB)
1	none	7	10
3	1 failure, RAID-1 (Mirroring)	3.7	5
4	1 failure, RAID-5 (Erasure Coding)	5.6	8
5	2 failures, RAID-1 (Mirroring)	2.5	3.6
6	2 failures, RAID-6 (Erasure Coding)	4.9	7

RAID overhead can also be expressed in terms of the storage consumed by a VM. This table shows how various RAID and FTT configurations affect the average amount of storage consumed by a 100 GB VM for a given storage allocation.

Table 8-2. Storage Consumed by a 100 GB VM

Nodes	FTT and RAID Config	Actual GB Allocated	Average GB Consumed
1	none	100	69
3	1 failure, RAID-1 (Mirroring)	200	138
4	1 failure, RAID-5 (Erasure Coding)	133	92
5	2 failures, RAID-1 (Mirroring)	300	207
6	2 failures, RAID-6 (Erasure Coding)	150	103

vSAN Deduplication and Compression

vSAN performs block-level deduplication and compression to save storage space. This allows you to make more efficient and cost-effective use of storage in your VMware Cloud on AWS SDDC.

Deduplication removes redundant data blocks, whereas compression removes additional redundant data within each data block. These techniques work together to reduce the amount of physical storage required to store the data. VMware vSAN applies applies deduplication followed by compression as it moves data from the cache tier to the capacity tier.

Deduplication occurs inline when data is written back from the cache tier to the capacity tier. The deduplication algorithm uses a fixed block size and is applied within each disk group. Redundant copies of a block within the same disk group are deduplicated.

Storage savings resulting from deduplication and compression are highly dependent on the workload data. On average, storage space savings are in the range of 1.5X to 2X.

Starting with SDDC Version 1.3, deduplication and compression are enabled on all new clusters you create. The feature is automatically enabled for all VMware Cloud on AWS clusters and cannot be turned off.

Clusters created with older SDDC versions have deduplication and compression enabled after they are upgraded to the new SDDC version. During the enablement of deduplication and compression, an extra host is temporarily added to the cluster. You are not charged for this host. The upgrade process moves data out of each disk group, reformats the disk group with deduplication settings, and writes the data back to the disk group.

Enablement of deduplication and compression takes some time. During the process, some clusters might operate in a partial state, with some disk groups having deduplication enabled, and others not. While the upgrade is in progress, you cannot remove hosts from the cluster. During the upgrade process, the cluster may experience some slowness, especially during heavy I/O load.

To speed up the enablement of deduplication and compression, follow these best practices.

- Avoid changes to the cluster configuration as much as possible, including addition of hosts and addition or deletion of VMs.
- Avoid changes to storage policies.
- Maintain current cluster capacity as much as possible.

vSAN Encryption in VMware Cloud on AWS

vSAN encrypts all user data at rest in VMware Cloud on AWS.

Encryption is enabled by default on each cluster deployed in your SDDC, and can't be turned off.

When you deploy a cluster, vSAN uses the AWS Key Management Service (AWS KMS) to generate a Customer Master Key (CMK), which is stored by AWS KMS. vSAN then generates a Key Encryption Key (KEK) and encrypts it using the CMK. The KEK is in turn used to encrypt Disk Encryption Keys (DEKs) generated for each vSAN disk.

You can change KEKs by using either the vSAN API or the vSphere Client UI. This process is known as performing a shallow rekey. Changing the CMK or DEKs is not supported. If you must change the CMK or DEKs, create a new cluster and migrate your VMs and data to it.

Generate New Encryption Keys in VMware Cloud on AWS

You can generate new Key Encryption Keys (KEKs) for your VMware Cloud on AWS cluster if needed.

This process is known as performing a shallow rekey. Changing the CMK or DEKs is not supported. If you must change the CMK or DEKs, create a new cluster and migrate your VMs and data to it

Procedure

1. Log in to the vSphere Client for your cloud SDDC.
2. Navigate to the vSAN cluster.
3. Click the **Configure** tab.
4. Under vSAN, select **Services**.
5. Click **Generate New Encryption Keys**.
6. Click **Apply** to generate a new KEKs.

   The Disk Encryption Keys (DEKs) are re-encrypted with the new KEKs.

Example: Using VMware PowerCLI for this Task

If you know the cloudadmin password, you can use a PowerCLI command like this one to do a shallow re-key for the vSAN service. This example, based the `Vsan-EncryptionRekey.psl` code sample you can download from https://code.vmware.com/samples/2200#code, re-keys the vSan service running on `Cluster-1` of SDDC vCenter `vcenter.sddc-54-200-165-35.vmwarevmc.com`:

```
PS > ./Vsan-EncryptionRekey.psl -vCenter vcenter.sddc-54-200-165-35.vmwarevmc.com -User
      cloudadmin@vmc.local -Password cloudadmin-password -ReKey shallow -ClusterName Cluster-1
```

vSAN Policies

vSAN storage polices define storage requirements for your virtual machines. These policies guarantee the required level of service for your VMs because they determine how storage is allocated to the VM.

VMware Cloud on AWS includes two vSAN datastores, one for the management VMs (vsanDatastore) and one for the workload VMs (WorkloadDatastore). Both datastores share the same underlying storage devices and consume from the same pool of free space.

Each virtual machine deployed to a vSAN datastore is assigned at least one virtual machine storage policy. You can assign storage policies when you create or edit virtual machines.

Note If you do not assign a storage policy to a virtual machine, vSAN assigns a default policy. The default policy has **Primary level of failures to tolerate** set to 1, a single disk stripe per object, and a thin-provisioned virtual disk.

Storage policies have availability attributes and advanced attributes.

Availability Attributes for vSAN VM Storage Policies

Site disaster tolerance

Defines the data redundancy method used by stretched clusters to handle a site failure. This attribute applies to stretched clusters. If you have a standard vSAN cluster, choose None (standard cluster).

The options are:

- None (standard cluster)
- Dual-site monitoring (stretched cluster)
- None-Keep data on primary (stretched cluster)
- None - Keep data on secondary (stretched cluster)

Failures to tolerate

Defines the number of host and device failures that a virtual machine can tolerate. You can choose to have no data redundancy, or select a RAID configuration optimized for either performance (Mirroring) or capacity (Erasure Coding).

Table 8-3. RAID Configurations, FTT, and Host Requirements

RAID Configuration	Failures to Tolerate (FTT)	Minimum Hosts Required
RAID-1 (Mirroring) This is the default setting.	1	3
RAID-5 (Erasure Coding)	1	4
RAID-1 (Mirroring)	2	5
RAID-6 (Erasure Coding)	2	6
RAID-1 (Mirroring)	3	7

The initial number of hosts in a cluster and the way in which hosts are added to or removed from a cluster affects its RAID configuration. For example, a three-host cluster is initially configured with RAID 1. When you add a host, you can reconfigure the cluster for RAID-5, but that reconfiguration is not automatic. A four-host cluster is initially configured with RAID-5. See Storage Capacity and Data Redundancy for details.

Advanced Attributes for vSAN VM Storage Policies

Number of disk stripes per object

Minimum number of capacity devices across which each replica of a virtual machine object is striped. A value higher than 1 might result in better performance, but also results in higher use of system resources. Default value is 1. Maximum value is 12. Change the default value only when recommended by VMware support.

IOPS limit for object

Defines the IOPS limit for an object, such as a VMDK. IOPS is calculated as the number of I/O operations, using a weighted size. If the system uses the default base size of 32 KB, a 64-KB I/O represents two I/O operations.

When calculating IOPS, read and write are considered equivalent, but cache hit ratio and sequentiality are not considered. If a disk’s IOPS exceeds the limit, I/O operations are throttled. If the **IOPS limit for object** is set to 0, IOPS limits are not enforced.

vSAN allows the object to double the rate of the IOPS limit during the first second of operation or after a period of inactivity.

Object space reservation

Percentage of the logical size of the virtual machine disk (vmdk) object that must be reserved, or thick provisioned when deploying virtual machines.

Default value is 0%. Maximum value is 100%.

Flash read cache reservation

Flash capacity reserved as read cache for the virtual machine object. Specified as a percentage of the logical size of the virtual machine disk (vmdk) object. Reserved flash capacity cannot be used by other objects. Unreserved flash is shared fairly among all objects. Use this option only to address specific performance issues.

You do not have to set a reservation to get cache. Setting read cache reservations might cause a problem when you move the virtual machine object because the cache reservation settings are always included with the object.

The Flash Read Cache Reservation storage policy attribute is supported only for hybrid configurations. You must not use this attribute when defining a VM storage policy for an all-flash vSAN cluster.

Default value is 0%. Maximum value is 100%.

Note By default, vSAN dynamically allocates read cache to storage objects based on demand. This feature represents the most flexible and the most optimal use of resources. As a result, typically, you do not need to change the default 0 value for this parameter.

To increase the value when solving a performance problem, exercise caution. Over-provisioned cache reservations across several virtual machines can cause flash device space to be wasted on over-reservations. These cache reservations are not available to service the workloads that need the required space at a given time. This space wasting and unavailability might lead to performance degradation.

Disable object checksum

If the option is set to **No**, the object calculates checksum information to ensure the integrity of its data. If this option is set to **Yes**, the object does not calculate checksum information.

vSAN uses end-to-end checksum to ensure the integrity of data by confirming that each copy of a file is exactly the same as the source file. The system checks the validity of the data during read/write operations, and if an error is detected, vSAN repairs the data or reports the error.

If a checksum mismatch is detected, vSAN automatically repairs the data by overwriting the incorrect data with the correct data. Checksum calculation and error-correction are performed as background operations.

The default setting for all objects in the cluster is **No**, which means that checksum is enabled.

Force provisioning

If the option is set to **Yes**, the object is provisioned even if the **Primary level of failures to tolerate**, **Number of disk stripes per object**, and **Flash read cache reservation** policies specified in the storage policy cannot be satisfied by the datastore. Use this parameter in bootstrapping scenarios and during an outage when standard provisioning is no longer possible.

The default **No** is acceptable for most production environments. vSAN fails to provision a virtual machine when the policy requirements are not met, but it successfully creates the user-defined storage policy.

Storage Policies and SLA Requirements

When working with virtual machine storage policies, it's important to understand how they affect the consumption of storage capacity in the vSAN cluster and whether they meet the requirements defined in the Service Level Agreement for VMware Cloud on AWS (the SLA).

The default vSAN storage policy is initially configured based on the number of hosts in the cluster. For example, a three-host cluster defaults to FTT=1 using the Raid-1 Mirroring policy. A four-host cluster also defaults to FTT=1 but uses the more space-efficient Raid-5 Erasure Coding policy. Clusters with more than six i3.metal hosts in a single AZ default to FTT=2 using the Raid-6 Erasure Coding policy. You can create custom policies that align data availability with the needs of your underlying data, but workload VMs with storage policies that do not meet the requirements set forth in the Service Level Agreement may not qualify for SLA Credits. The VM Storage Policy must be configured with the appropriate level of protection. Ephemeral workloads may use the No Data Redundancy policy to save capacity, foregoing any SLA guarantees of availability.

Important When scaling an i3.metal cluster up from five to six hosts, you must update the underlying policy configuration to FTT=2 using either Mirroring or Erasure Coding to compensate for the larger failure pool. Continued use of FTT=1 for this host configuration means that VMware cannot guarantee availability per the service definition guidance. R5.metal clusters using Elastic vSAN are able to sustain the SLA with FTT=1 regardless of cluster size.

For more information about designing and sizing considerations of storage policies, see the *Administering VMware vSAN* documentation.

About the vSAN Default Storage Policy

vSAN requires that the virtual machines deployed on the vSAN datastores are assigned at least one storage policy. When provisioning a virtual machine, if you do not explicitly assign a storage policy to the virtual machine the vSAN Default Storage Policy is assigned to the virtual machine.

The default policy contains vSAN rule sets and a set of basic storage capabilities, typically used for the placement of virtual machines deployed on vSAN datastores.

Table 8-4. vSAN Default Storage Policy Specifications

Specification	Setting
Primary level of failures to tolerate	1
Number of disk stripes per object	1
Flash read cache reservation, or flash capacity used for the read cache	0
Object space reservation	0 **Note** Setting the Object space reservation to zero means that the virtual disk is thin provisioned, by default.
Force provisioning	No

You can review the configuration settings for the default virtual machine storage policy when you navigate to the **VM Storage Policies** > **vSAN Default Storage Policy** > **Manage** > **Rule-Set 1: VSAN**.

For best results, consider creating and using your own VM storage policies, even if the requirements of the policy are same as those defined in the default storage policy. In some cases, when you scale up a cluster, you must modify the default storage policy to maintain compliance with the requirements of the Service Level Agreement for VMware Cloud on AWS. For information about creating a user-defined VM storage policy, see Define a Virtual Machine Storage Policy for vSAN.

When you assign a user-defined storage policy to a datastore, vSAN applies the settings for the user-defined policy on the specified datastore. At any point, you can assign only one virtual machine storage policy as the default policy to the vSAN datastore.

Characteristics

The following characteristics apply to the vSAN Default Storage Policy.

- The vSAN default storage policy is assigned to all virtual machine objects if you do not assign any other vSAN policy when you provision a virtual machine. The **VM Storage Policy** text box is set to **Datastore default** on the Select Storage page. For more information about using storage policies, see the *vSphere Storage* documentation.

  Note VM swap and VM memory objects receive the vSAN Default Storage Policy with **Force provisioning** set to **Yes**.

- The vSAN default policy only applies to vSAN datastores. You cannot apply the default storage policy to non-vSAN datastores, such as NFS or a VMFS datastore.

- Because the default virtual machine storage policy is compatible with any vSAN datastore in the vCenter Server, you can move your virtual machine objects provisioned with the default policy to any vSAN datastore in the vCenter Server.
- You can clone the default policy and use it as a template to create a user-defined storage policy.
- You can edit the default policy, if you have the StorageProfile.View privilege. You must have at least one vSAN enabled cluster that contains at least one host. Typically you do not edit the settings of the default storage policy.
- You cannot edit the name and description of the default policy, or the vSAN storage provider specification. All other parameters including the policy rules are editable.
- You cannot delete the default policy.
- The default storage policy is assigned when the policy that you assign during virtual machine provisioning does not include rules specific to vSAN.

Assign a Default Storage Policy to vSAN Datastores

You can assign a user-defined storage policy as the default policy to a datastore, to reuse a storage policy that matches your requirements.

Prerequisites

Verify that the VM storage policy you want to assign as the default policy to the vSAN datastore meets the requirements of virtual machines in the vSAN cluster.

Procedure

1. Navigate to the vSAN datastore.
2. Click **Configure**.
3. Under **General**, click the Default Storage Policy **Edit** button, and select the storage policy that you want to assign as the default policy to the vSAN datastore.

   You can choose from a list of storage policies that are compatible with the vSAN datastore, such as the vSAN Default Storage Policy and user-defined storage policies that have vSAN rule sets defined.
4. Select a policy and click **OK**.

   The storage policy is applied as the default policy when you provision new virtual machines without explicitly specifying a storage policy for a datastore.

What to do next

You can define a new storage policy for virtual machines. See Define a Virtual Machine Storage Policy for vSAN.

Define a Virtual Machine Storage Policy for vSAN

You can create a storage policy that defines storage requirements for a VM and its virtual disks. In this policy, you reference storage capabilities supported by the vSAN datastore.

Procedure

1 In the vSphere Client, click **Menu > Policies and Profiles**, then click **VM Storage Policies**.

2 Click **Create VM storage policy**.

3 On the Name and description page:.

   a Leave the vCenter Server selected.

   b Type a name and a description for the storage policy, and click **Next**.

4 On the vSAN page, specify the **Availability** and **Advanced** attributes and click **Next**.

   The defaults are appropriate for many situations. See vSAN Policies for details on each attribute.

5 On the Storage compatibility page, review the list of datastores that match this policy and click **Next**.

   To be eligible, a datastore does not need to satisfy all rule sets within the policy. The datastore must satisfy at least one rule set and all rules within this set. Verify that the vSAN datastore meets the requirements set in the storage policy and that it appears on the list of compatible datastores.

6 On the Ready to complete page, review the policy settings, and click **Finish**.

What to do next

Assign this policy to a virtual machine and its virtual disks. vSAN places the virtual machine objects according to the requirements specified in the policy. For information about applying the storage policies to virtual machine objects, see the *vSphere Storage* documentation.

Assign Storage Policies to Virtual Machines

You can assign a VM storage policy in an initial deployment of a virtual machine or when you perform other virtual machine operations, such as cloning or migrating.

This topic describes how to assign the VM storage policy when you create a virtual machine. For information about other deployment methods that include cloning, deployment from a template, and so on, see the *vSphere Virtual Machine Administration* documentation.

You can apply the same storage policy to the virtual machine configuration file and all its virtual disks. If storage requirements for your virtual disks and the configuration file are different, you can associate different storage policies with the VM configuration file and the selected virtual disks.

Procedure

1 In the vSphere Client, start the virtual machine provisioning process and follow the appropriate steps.

2 Assign the same storage policy to all virtual machine files and disks.

   a On the **Select storage** page, select a storage policy from the **VM Storage Policy** drop-down menu.

      Based on its configuration, the storage policy separates all datastores into compatible and incompatible sets. If the policy references data services offered by a specific storage entity, for example, Virtual Volumes, the compatible list includes datastores that represent only that type of storage.

   b Select an appropriate datastore from the list of compatible datastores.

      The datastore becomes the destination storage resource for the virtual machine configuration file and all virtual disks.

3 Change the VM storage policy for the virtual disk.

   Use this option if requirements for storage placement are different for virtual disks. You can also use this option to enable I/O filter services, such as caching and replication, for your virtual disks.

   a On the **Customize hardware** page, expand the **New hard disk** pane.

   b From the **VM storage policy** drop-down menu, select the storage policy to assign to the virtual disk.

4 Complete the virtual machine provisioning process.

After you create the virtual machine, the **Summary** tab displays the assigned storage policies and their compliance status.

What to do next

If storage placement requirements for the configuration file or the virtual disks change, you can later modify the virtual machine policy assignment.

Workload Networking in VMware Cloud on AWS

9

Networking for workload VMs in your SDDC is routed through the Compute Gateway, which provides north-south network connectivity for virtual machines running in the SDDC.

For more information about SDDC networking, see NSX-T Networking Concepts in *VMware Cloud on AWS Networking and Security*.

This chapter includes the following topics:

- Create a Compute Network Segment
- Attach a VM to or Detach a Workload VM from a Compute Network Segment

Create a Compute Network Segment

A Single Host Starter SDDC is created with a single routed network segment named sddc-cgw-network-1. This network uses CIDR block 192.168.1.0/24, unless that conflicts with the CIDR block you chose for the SDDC management network. In that case, the default network uses CIDR block 172.10.1.0/24.

Multi-host SDDCs are not created with a default network segment, so you must create at least one for your workload VMs. You can use the VMC Console to create additional network segments or delete ones that are no longer in use.

Procedure

1 Create a Network Segment.

  You can skip this step if you're using a single-host starter SDDC and want to use its default segment.

2 Add or Modify Compute Gateway Firewall Rules.

  By default, the compute gateway blocks traffic to all uplinks. You'll need to add Compute Gateway firewall rules to allow your workload VMs to communicate outside the segment to which they are attached.

What to do next

After you have created the logical network segment, you're ready to connect workload VMs to it. If you want, you can configure additional per-segment features such as DNS zones and DHCP relay.

Attach a VM to or Detach a Workload VM from a Compute Network Segment

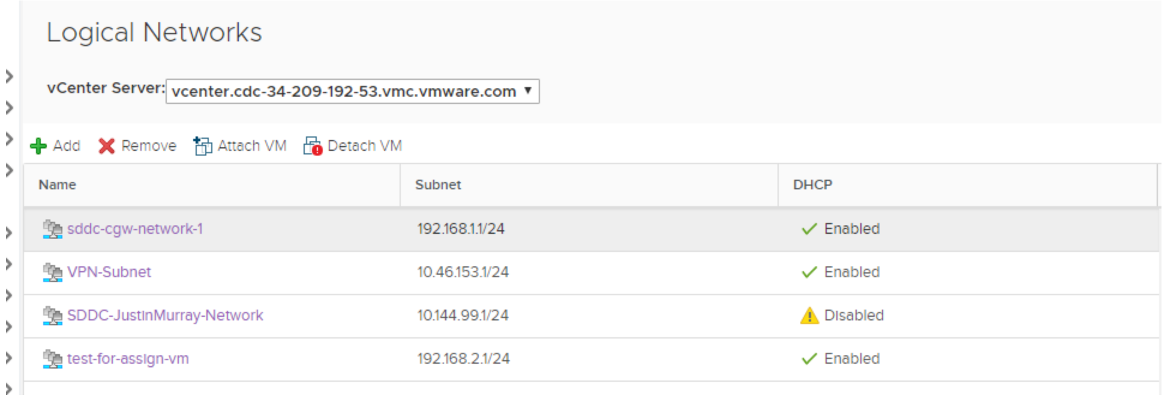
Use the vSphere Web Client to manage attachment of workload VMs to compute network segments.

Prerequisites

Your SDDC compute network must have at least one segment. See Create a Compute Network Segment.

Procedure

1. Log in to the vSphere Client for your SDDC.
2. Select **Menu > Global Inventory Lists**.
3. Select **Logical Networks**.
4. In the **vCenter Server** drop down menu, select the vCenter Server that manages the logical network you want to use.
5. Click next to the logical network name to select it.



vCenter Server: vcenter.cdc-34-209-192-53.vmc.vmware.com ▾

Name	Subnet	DHCP
sddc-cgw-network-1	192.168.1.1/24	✓ Enabled
VPN-Subnet	10.46.153.1/24	✓ Enabled
SDDC-JustinMurray-Network	10.144.99.1/24	⚠ Disabled
test-for-assign-vm	192.168.2.1/24	✓ Enabled

6. Select whether to attach or detach VMs.
   - Click **Attach VM** to attach VMs to the selected network.
   - Click **Detach VM** to detach VMs from the selected network.
7. Select the virtual machine(s) you want to attach or detach, click **>>** to move them to the **Selected Objects** column, and click **Next**.
8. For each VM, select the virtual NIC you want to attach and click **Next**.
9. Click **Finish**.

vSphere Tags and Attributes 10

Tags and attributes allow you to attach metadata to objects in the vSphere inventory to make it easier to sort and search for these objects.

A tag is a label that you can apply to objects in the vSphere inventory. When you create a tag, you assign that tag to a category. Categories allow you to group related tags together. When you define a category, you can specify the object types for its tags, and whether more than one tag in the category can be applied to an object.

For example, if you wanted to tag your virtual machines by guest operating system type, you can create a category called `operating system`. You can specify that it applies to virtual machines only and that only a single tag can be applied to a virtual machine at any time. The tags in this category might be `Windows`, `Linux`, and `Mac OS`.

Tags and categories can span multiple vCenter Server instances:

- If multiple on-premises vCenter Server instances are configured to use Enhanced Linked Mode, tags and tag categories are replicated across all these vCenter Server instances.
- When you use Hybrid Linked Mode, tags and tag categories are maintained across your linked domain. That means the on-premises SDDC and the VMware Cloud on AWS SDDC share tags and tag attributes.

For vSphere Tags and Attributes, VMware Cloud on AWS supports the same set of tasks as an on-premises SDDC.

This chapter includes the following topics:

Create, Edit, or Delete a Tag Category

You use categories to group tags together and define how tags can be applied to objects. You create, edit, and delete a tag category from the vSphere Client.

You can create a tag category explicitly, as explained here, or as part of the tag creation process. Each tag has to belong to at least one tag category.

Prerequisites

The required privilege depends on the task that you want to perform.

Task	Privilege
Create a tag category	**vSphere Tagging.Create vSphere Tag Category** on the root vCenter Server
Edit a tag category	**vSphere Tagging.Edit vSphere Tag Category** on the root vCenter Server.
Delete a tag category	**vSphere Tagging.Delete vSphere Tag Category** on the root vCenter Server.

Procedure

1 In the vSphere Client, click **Menu > Tags & Custom Attributes**.

2 Click the **Tags** tab and click **Categories**.

3 Start the task that you want to perform.

Option	Description
Create a tag category	Click the **New Category** icon.
Edit a tag category	Select a category and click the **Edit Category** icon.
Delete a tag category	Select a category from the list and click the **Delete Category** icon.

4 Edit the category options.

Option	Description
Category Name	The category name must be unique to the currently selected vCenter Server system.
Description	You can provide text in the description to describe the purpose or use of the category.

Option	Description
Tags Per Object	■ If you select **One Tag**, you can apply only one tag from this category to an object. Use this option for categories whose tags are mutually exclusive. For example, if you have a category called Priority with tags High, Medium, and Low, then each object should have only one tag because an object can have only one priority. ■ If you select **Many tags**, you can apply more than one tag from the category to an object. Use this option for categories whose tags are not mutually exclusive. After you have set Tags Per Object, you can change from **One Tag** to **Many Tags**, but not from **Many Tags** to **One Tag**.
Associable Object Types	Select whether tags in this category can be assigned to all objects or only to a specific type of object, such as a virtual machine or a datastore. Changes to the associable object type are limited. ■ If you initially selected a single object type, you can later change the category to work for all object types. ■ If you initially selected All Objects, you cannot restrict the category later.

5. Click **OK** or **Yes** to confirm.

Create, Edit, or Delete a Tag

You use tags to add metadata to inventory objects. You can record information about your inventory objects in tags, and you can use the tags in searches.

Procedure

1. In the vSphere Client, click **Menu > Tags & Custom Attributes**.
2. Click **Tags**.
3. Perform the task.

Option	Description
Create a tag	a Click the New Tag icon. b Specify the **Name** and an optional **Description**. c In the **Category** drop-down menu, select an existing category or create a category. If you select **[New Category]**, the dialog box expands to show the options for creating a category. See Create, Edit, or Delete a Tag Category.
Edit a tag	Select a category and click the **Edit Category** icon.
Delete a tag	Select a category from the list and click the **Delete Category** icon.

4. Click **OK**.

Assign or Remove a Tag

After you have created tags, you can apply or remove them as metadata to objects in the vCenter Server inventory.

Procedure

1. Browse to the object in the vSphere Client inventory.
2. From the Actions menu, select **Tags and Custom Attributes > Assign Tag**.
3. Select a tag from the list and click **Assign**.

   You can later use the same process to remove a tag.
4. From the Actions menu, select **Tags & Custom Attributes > Remove Tag**.
5. Select a tag from the list and click **Remove**.

Permissions on Tag Objects

In the vCenter Server object hierarchy, tag objects are not children of vCenter Server but are created at the vCenter Server root level. In environments with multiple vCenter Server instances, tag objects are shared across vCenter Server instances. Permissions for tag objects work differently than permissions for other objects in the vCenter Server object hierarchy.

Only Global Permissions or Permissions Assigned to the Tag Object Apply

If you grant permissions to a user on a vCenter Server inventory object, such as a virtual machine, that user can perform the tasks associated with the permission. However, the user cannot perform tag operations on the object.

For example, if you grant the **Assign vSphere Tag** privilege to user Dana on host TPA, that permission does not affect whether Dana can assign tags on host TPA. Dana must have the **Assign vSphere Tag** privilege at the root level, that is, a global permission, or must have the privilege for the tag object.

Table 10-1. How Global Permissions and Tag Object Permissions Affect What Users Can Do

Global Permission	Tag-Level Permission	vCenter Server Object-Level Permission	Effective Permission
No tagging privileges assigned.	Dana has **Assign or Unassign vSphere Tag** privileges for the tag.	Dana has **Delete vSphere Tag** privileges on ESXi host TPA.	Dana has **Assign or Unassign vSphere Tag** privileges for the tag.
Dana has **Assign or Unassign vSphere Tag** privileges.	No privileges assigned for the tag.	Dana has **Delete vSphere Tag** privileges on ESXi host TPA.	Dana has **Assign or Unassign vSphere Tag** global privileges. That includes privileges at the tag level.
No tagging privileges assigned.	No privileges assigned for the tag.	Dana has **Assign or Unassign vSphere Tag** privileges on ESXi host TPA.	Dana does not have tagging privileges on any object, including host TPA.

Global Permissions Complement Tag Object Permissions

Global permissions, that is, permissions that are assigned on the root object, complement permissions on tag objects when the permissions on the tag objects are more restrictive. The vCenter Server permissions do not affect the tag objects.

For example, assume that you assign the **Delete vSphere Tag** privilege to user Robin at the root level by using global permissions. For the tag Production, you do not assign the **Delete vSphere Tag** privilege to Robin. In that case, Robin has the privilege for the tag Production because Robin has the global permission. You cannot restrict privileges unless you modify the global permission.

Table 10-2. Global Permissions Complement Tag-Level Permissions

Global Permission	Tag-Level Permission	Effective Permission
Robin has **Delete vSphere Tag** privileges	Robin does not have **Delete vSphere Tag** privileges for the tag.	Robin has **Delete vSphere Tag** privileges.
No tagging privileges assigned	Robin does not have **Delete vSphere Tag** privileges assigned for the tag.	Robin does not have **Delete vSphere Tag** privileges

Tag-Level Permissions Can Extend Global Permissions

You can use tag-level permissions to extend global permissions. That means users can have both a global permission and a tag-level permission on a tag.

Table 10-3. Global Permissions Extend Tag-Level Permissions

Global Permission	Tag-Level Permission	Effective Permission
Lee has **Assign or Unassign vSphere Tag** privilege.	Lee has **Delete vSphere Tag** privilege.	Lee has the **Assign vSphere Tag** privilege and the **Delete vSphere Tag** privilege for the tag.
No tagging privileges assigned.	Lee has **Delete vSphere Tag** privilege assigned for the tag.	Lee has the **Delete vSphere Tag** privilege for the tag.

Add and Edit Custom Attributes

You can create custom attributes in the vSphere Client and associate the attribute with an object, such as a host, virtual machine, cluster, or network. You can also edit custom attributes.

After you create the attributes, set an appropriate value for the attribute on each virtual machine. This value is stored with vCenter Server and not with the virtual machine. Use the new attribute to filter your virtual machines. If you no longer need the custom attribute, remove it. A custom attribute is always a string.

For example, suppose that you have a set of products and you want to sort them by sales representative.

1 Create a Name custom attribute for the sales person's name.

2 Add the Name custom attribute column to one of the list views and add a name to each product entry.

3 You can now click the Name column to sort alphabetically by sales person.

Note Tags and tag categories support a finer-grained mechanism for tagging your object. Consider using tags and tag categories instead of custom attributes.

Procedure

1. In the vSphere Client, select **Menu > Tags and Custom Attributes**.
2. Click **Custom Attributes**.

   All currently defined custom attributes for vCenter Server are displayed.
3. Click **Add**.
4. Enter the values for the custom attribute.
   - a Type the name of the attributes in the **Attribute** text box.
   - b Select the attribute type from the **Type** drop-down menu.
   - c Click **OK**.

   After you have defined an attribute on an object, it is available to all objects of that type in the inventory. However, the value you specify is applied only to the currently selected object.
5. You can later edit a custom attribute.
   - a Select the attribute and click **Edit**.
   - b Change the Name.
   - c Change the type if it's available.
   - d Click **OK**.