LUN Passthrough Tutorial

Version: ZStack 3.10.0

Issue: V3.10.0

ZStack

Copyright Statement

Copyright © 2020 Shanghai Yunzhou Information and Technology Ltd. All rights reserved.

Without its written consent, any organization and any individual do not have the right to extract, copy any part or all of, and are prohibited to disseminate the contents of this documentation in any manner.

Trademark

Shanghai Yunzhou Information and Technology Ltd. reserves all rights to its trademarks, including , but not limited to ZStack and other trademarks in connection with Shanghai Yunzhou Information and Technology Ltd.

Other trademarks or registered trademarks presented in this documentation are owned or controlled solely by its proprietaries.

Notice

The products, services, or features that you purchased are all subject to the commercial contract and terms of Shanghai Yunzhou Information and Technology Ltd., but any part or all of the foregoing displayed in this documentation may not be in the scope of your purchase or use. Unless there are additional conventions, Shanghai Yunzhou Information and Technology Ltd. will not claim any implicit or explicit statement or warranty on the contents of this documentation.

In an event of product version upgrades or other reasons, the contents of this documentation will be irregularly updated and released. Unless there are additional conventions, this documentat ion, considered solely as a using manual, will not make any implicit or explicit warranty on all the statements, information, or suggestions.

Contents

Copyright Statement..... I
1 Overview.....1
2 Preparation.....2
3 Add Host..... 3
4 Attach LUN to VM Instances.....6
5 Typical Usage Scenario.....11
5.1 Microsoft Cluster Server (MSCS) on Windows.....11
5.2 Oracle RAC.....23
Glossary..... 25

1 Overview

ZStack supports the LUN passthrough feature, including FC LUN passthrough and iSCSI LUN passthrough.

- iSCSI LUN passthrough: After you add an iSCSI server, all LUN devices on the iSCSI server are displayed at real time via the data synchronization in the UI and can be passed through to VM instances.
- FC LUN passthrough: After you deploy an FC storage, the FC storage and all LUN devices are displayed at real time in the UI and can be passed through to VM instances.

2 Preparation

- You can use the latest version of ZStack installation package to install the operating system for your servers. In addition, you need to deploy all necessary resources used for creating VM instances.
  - The operating system versions of all servers must be consistent. Specifically, the operating system installations must all be based on c74 ISO or c76 ISO.
  - Specifically, install one of your servers with the **Enterprise Management Node** to serve as a management node.
  - With a management node, you can manage hosts. These hosts can automatically deploy and install LVM, Multipath, and other toolkits.

  For more information about the deployment and installation, see related deployment and installation topics in the *User Guide*.
- You can deploy FC storages or iSCSI storages on hosts, virtually split the FC storages or iSCSI storages into LUNs with certain capacities, and scan and discover LUNs online.
- Multipath accesses can be configured for FC storages or iSCSI storages of more than two link aggregations.

3 Add Host

Note:

ZStack allows you to add hosts via an IP address, IP range, and template. For more information about how to add hosts, see *Add Host* in the User Guide. This topic describes how to add hosts by using an IP address.

In the navigation pane of the ZStack Private Cloud UI, choose **Hardware** > **Host**. On the **Host** page, click **Add Host**. On the displayed **Add Host** page, set the following parameters:

- **Add Method**: Select manual add.
- **Name**: Enter a name for the host.
- **Description**: Optional. Enter a description for the host.
- **Cluster**: Select a cluster where the host belongs.
- **Add IP/IP range**: Select IP.
- **Host IP**: Enter a host IP address.
- **Scan host IOMMU setting**: If not selected, the IOMMU support is disabled.

  Note:

  - If you scan hosts by using an IOMMU device, the physical GPU passthrough feature or the vGPU feature will be supported.
  - If not selected, CPUs on a host will disable IOMMU. This checkbox is not selected by default.
  - If selected, CPUs on a host enable IOMMU, and all available GPU devices on the host will be traversed. Make sure that the BIOS mode on the host has enabled Intel VT-d or AMD IOMMU.
  - If you enable IOMMU for the first time, reboot the host to verify that IOMMU configurations on the CPUs take effect.
- **Disable Intel EPT support**: Disable or enable the Intel EPT support feature.

  Note:

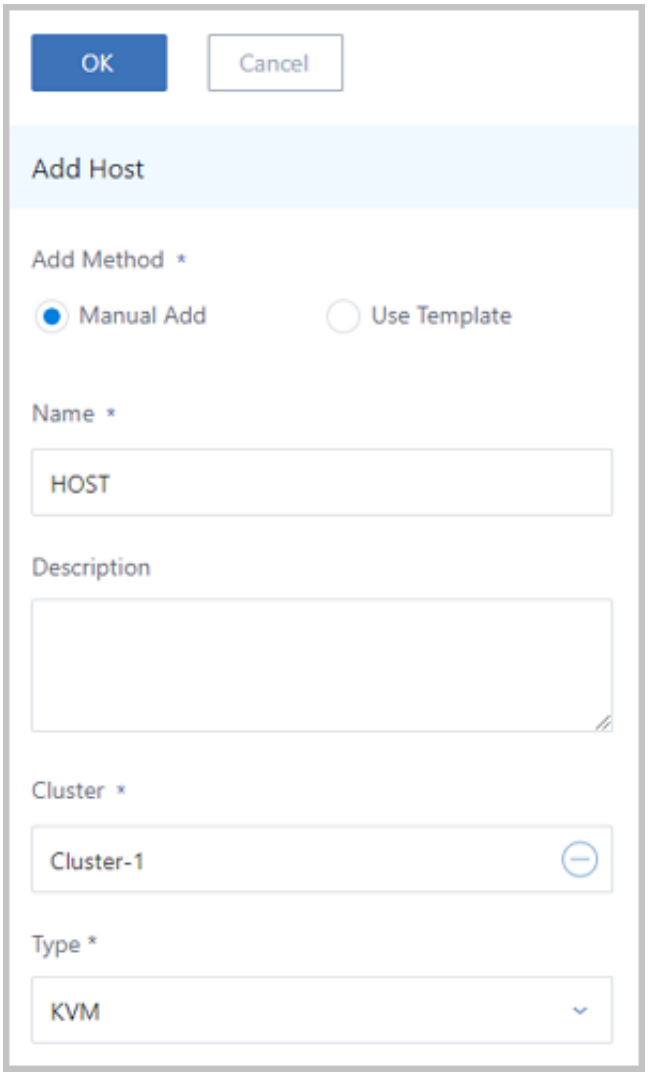
  - This checkbox is not selected by default, indicating that the Intel EPT support is enabled.

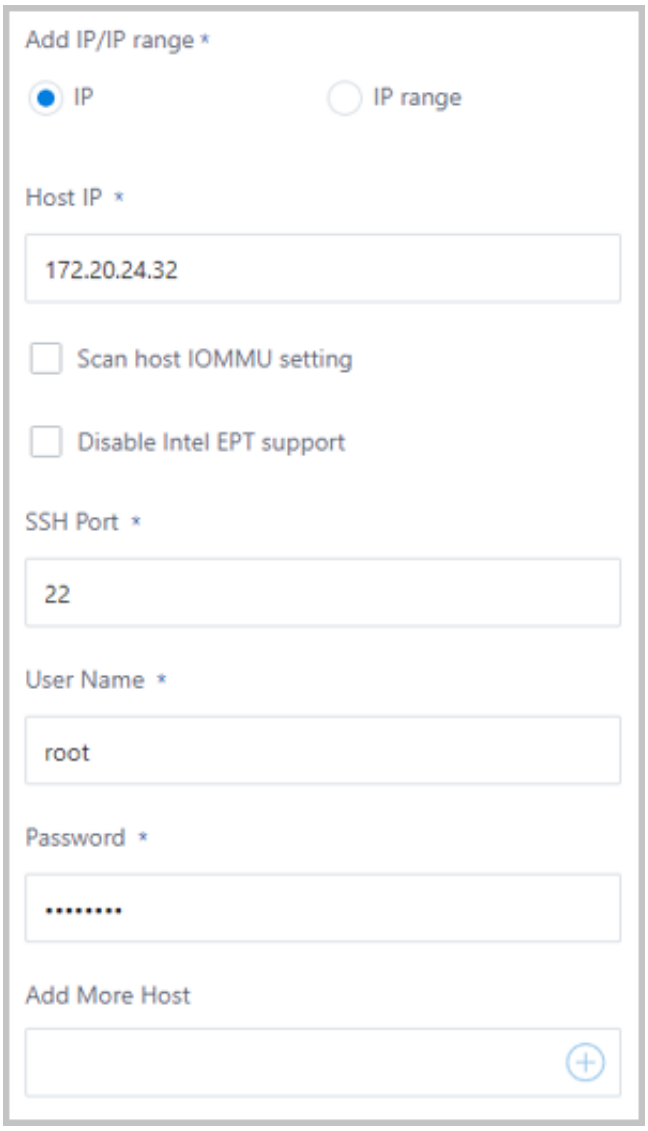
- If CPU models on your servers are too old, you will fail to create VM instances, or cannot open the console UIs of the VM instances. You can select this checkbox to disable the Intel EPT support feature.
- You can change the enabled state or the disabled state of the Intel EPT support feature on the host details page.
- This feature is only applied to Intel CPU.

- **SSH Port**: Set an SSH port for the host. Default value: 22.
- **User Name**: Use either the default user name (root) or enter a regular user name for the host.
- **Password**: Enter the corresponding user password for the host.
- **Add More Host**: Click the plus sign (+) to add more hosts.

Click **OK** to complete adding the host, as shown in *Add Host*.

Figure 3-1: Add Host




Add IP/IP range *
IP
IP range
Host IP *
172.20.24.32
Scan host IOMMU setting
Disable Intel EPT support
SSH Port *
22
User Name *
root
Password *
••••••••
Add More Host

4 Attach LUN to VM Instances

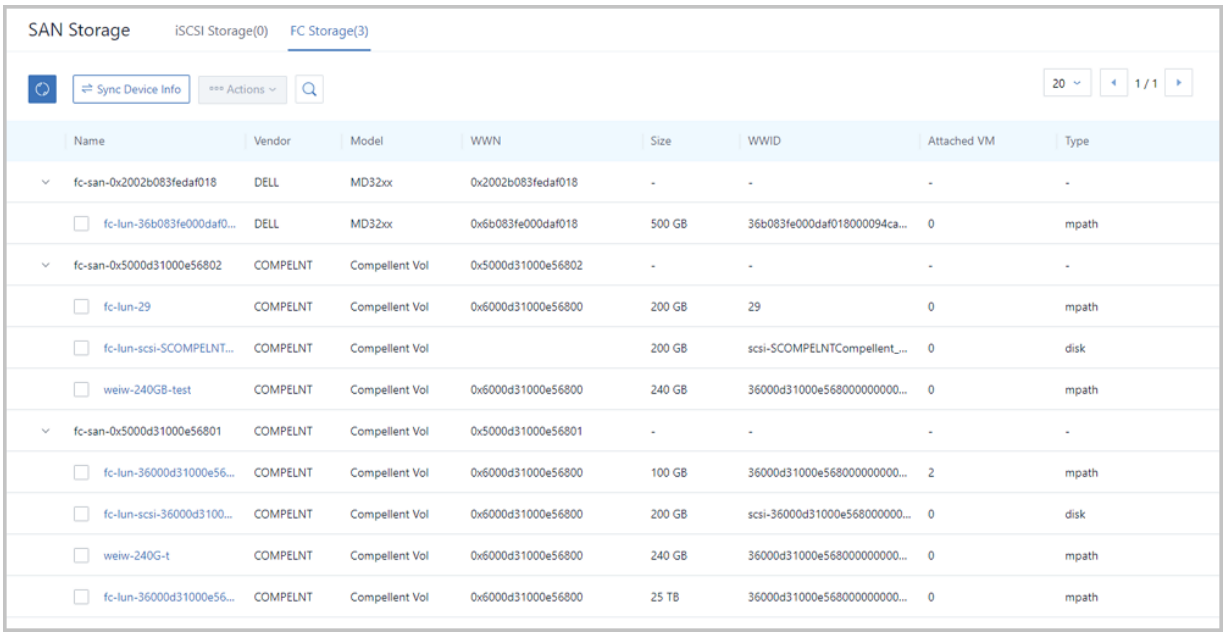
One LUN can be simultaneously attached to multiple VM instances, while one VM instance can attach multiple LUNs. You can attach or detach an LUN to or from a VM instance on the following four UI pages:

- FC storage page: Attach an FC LUN to a VM instance. For more information, see *Attach LUN via FC Storage Page*.
- iSCSI storage page: Attach an iSCSI LUN to a VM instance. For more information, see *Attach LUN via iSCSI Server Page*
- VM details page: Attach an FC LUN and an iSCSI LUN to a VM instance. You can distinguish the storages where LUNs belong via **Type**. For more information, see *Attach LUN via VM Details Page*
- Host details page: Attach an FC LUN and an iSCSI LUN to a VM instance. You can distinguish the storages where LUNs belong via **Type**. For more information, see *Attach LUN via Host Details Page*

Attach LUN via FC Storage Page

In the navigation pane of the ZStack Private Cloud UI, choose **Hardware** > **SAN Storage**. On the **FC Storage** tab page of the **SAN Storage** page, check FC storages and LUNs on the FC storages, as shown in *FC Storage*.

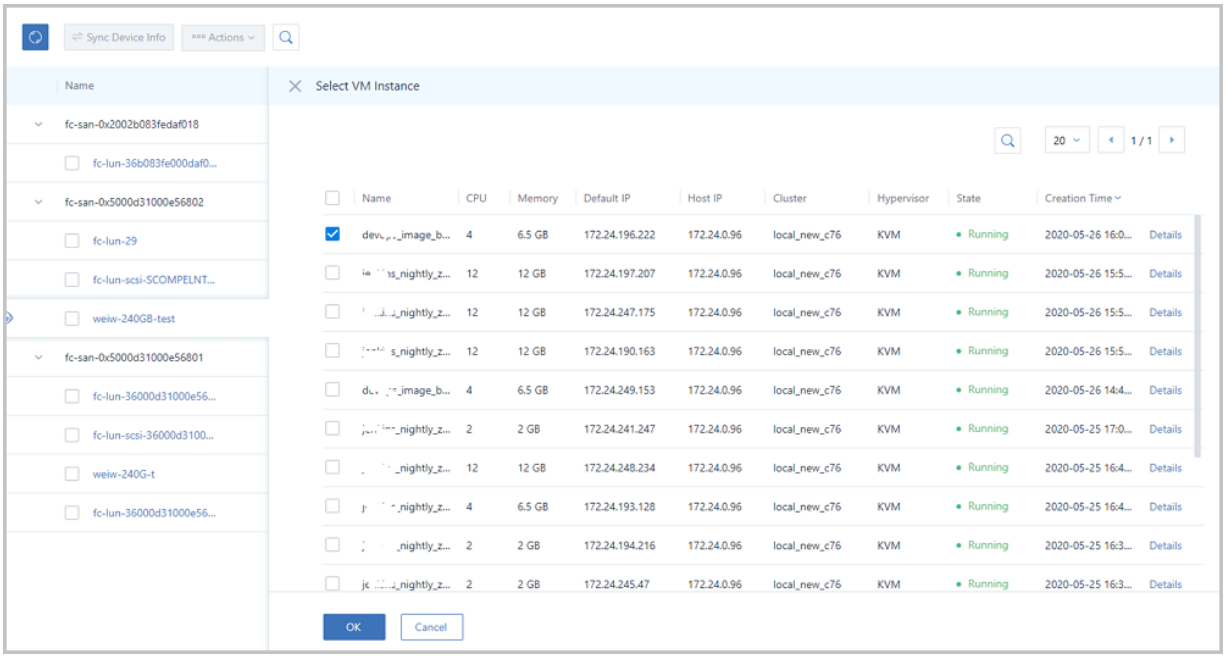
Figure 4-1: FC Storage



After you add a host, if you deploy an FC storage on the host, ZStack will automatically detect and display the host to the FC storage page. Click **Sync Device Info** to manually refresh the list information.

To select an LUN, choose **Actions** > **Attach VM Instance**. On the displayed **Select VM Instance** page, select VM instances that you need to attach, and click **OK** to complete attaching an LUN, as shown in *Attach VM Instance*.

Figure 4-2: Attach VM Instance

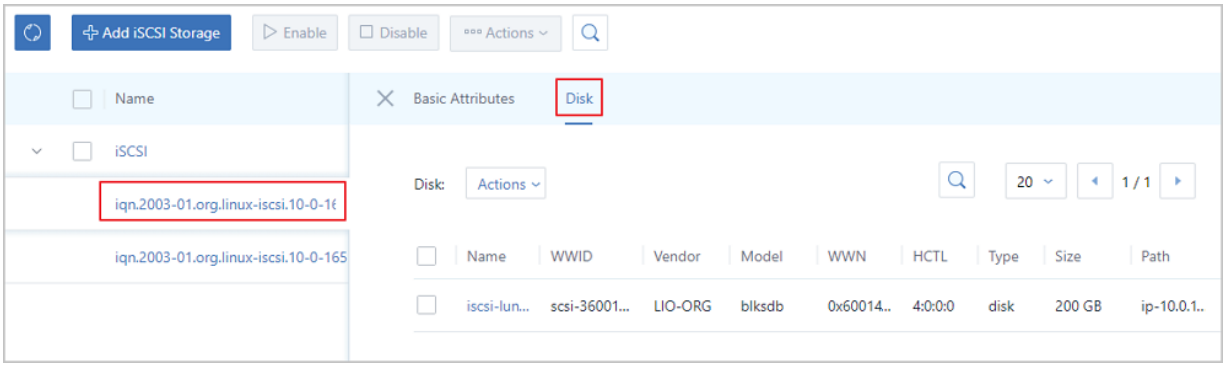


After you attach the LUN to the VM instances, you can choose **Actions** > **Detach VM Instance** to detach the LUN from the VM instances.

Attach LUN via iSCSI Server Page

In the navigation pane of the ZStack Private Cloud UI, choose **Hardware** > **SAN Storage**. On the **iSCSI Storage** tab page, click on the iSCSI qualified name (IQN) to enter the **Disk** tab page, as shown in *Disk Tab Page*.

Figure 4-3: Disk Tab Page



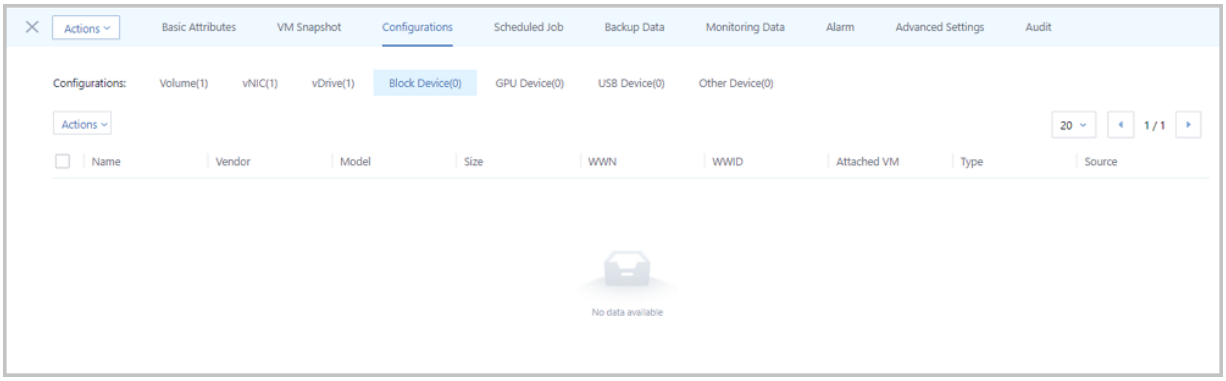
On the **Disk** tab page, select an iSCSI LUN, and choose **Actions** > **Attach LUN to VM Instance**. On the displayed **Select VM Instance** page, select VM instances that you need to attach, and click **OK** to complete attaching the iSCSI LUN.

After you attach the iSCSI LUN to the VM instances, you can choose **Actions** > **Detach LUN from VM Instance** to detach the iSCSI LUN from the VM instances.

Attach LUN via VM Details Page

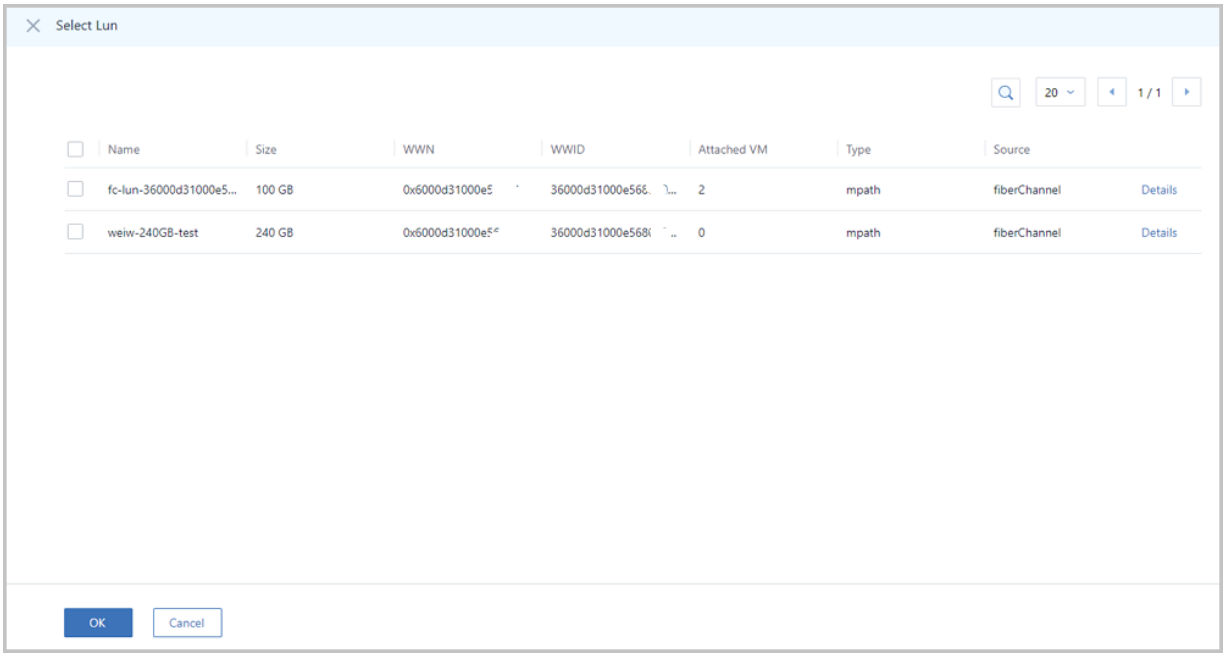
In the navigation pane of the ZStack Private Cloud UI, choose **Resource Pool** > **VM Instance** to enter the **VM Instance** page. You can select a VM instance that needs to attach an LUN, and then select the **Configurations** tab page, as shown in *Configurations Page*.

Figure 4-4: Configurations Page



On the **Block Device** tab page, choose **Actions** > **Attach**. On the **Select LUN** page, select an LUN which will be passed through to the VM instance, and click **OK** to attach the LUN to the VM instance, as shown in *Select LUN*.

Figure 4-5: Select LUN

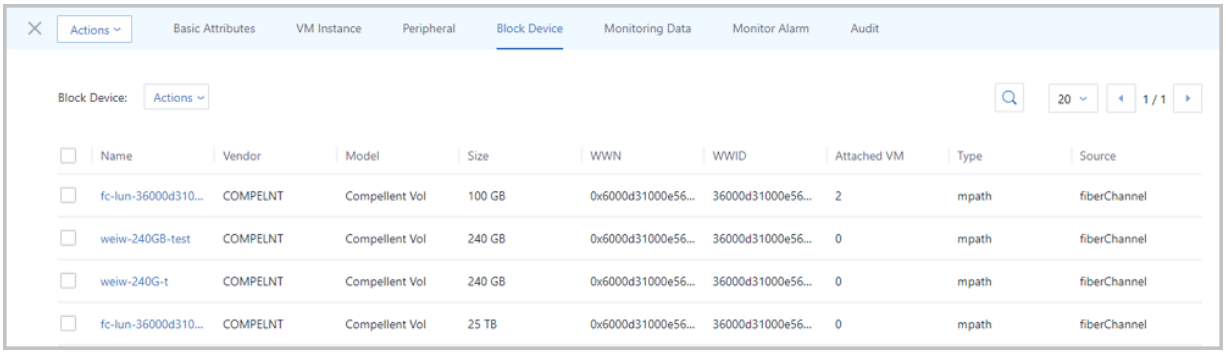


After you attach the LUN to the VM instance, you can choose **Actions** > **Detach** to detach the LUN from the VM instance.

Attach LUN via Host Details Page

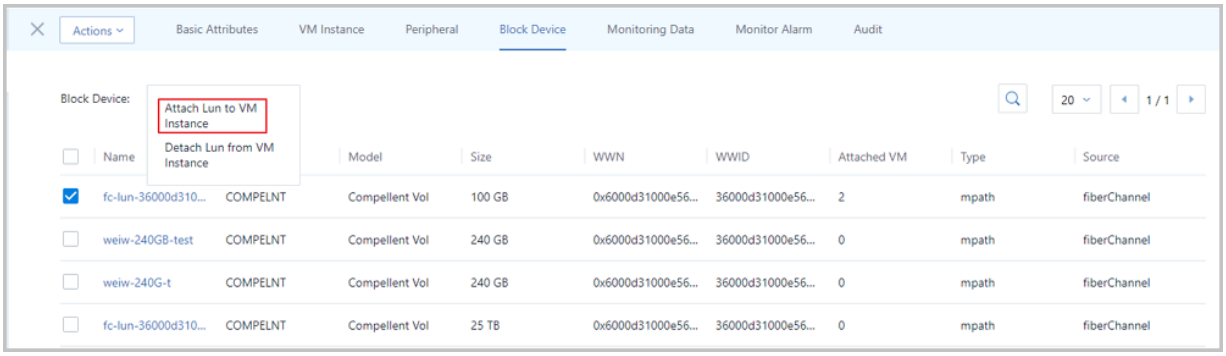
In the navigation pane of the ZStack Private Cloud UI, choose **Hardware** > **Host** to enter the **Host** page. You can select a host that has attached an LUN, and then select the **Block Device** tab page, as shown in *Block Device Page*.

Figure 4-6: Block Device Page



On the **Block Device** tab page, choose **Actions** > **Attach LUN to VM Instance**. On the **Select VM Instance** page, select VM instance to which you need to attach an LUIN, and click **OK** to complete attaching the LUN to the VM instances, as shown in *Attach LUN to VM Instance*.

Figure 4-7: Attach LUN to VM Instance



After you attach the LUN to the VM instances, you can choose **Actions** > **Detach LUN from VM Instance** to detach the LUN from the VM instances.

So far, we have introduced the passthrough feature of LUNs.

5 Typical Usage Scenario

LUN passthrough allows applications of different nodes to consistently access data. Generally, the typical usage scenario for this passthrough is as follows:

- Application softwares or system softwares with sharable data LUN storages, such as Oracle RAC and Microsoft Cluster Server (MSCS).
- Cluster file systems, such as OCFS2, GPFS, and GFS2.

5.1 Microsoft Cluster Server (MSCS) on Windows

Prerequisites

Microsoft Cluster Server (MSCS) is a group of independent servers that work collaboratively to improve high availability and scalability of clustered roles (formerly called application software and services according to *Microsoft Documentation*). Clustered servers (nodes) are connected via physical cables and software. If one or more clustered nodes fail over, other nodes will start to provide services in a process known as failover. Clustered roles will be automatically monitored to make sure whether clustered nodes can work properly. If these clustered nodes do not work, these roles will be restarted or migrated to other clustered nodes.

Microsoft Cluster Server also supports Cluster Shared Volume (CSV) functionality. Servers in failover clusters can share the same storage space. For more information, see *Failover Clustering in Windows Server* in Microsoft Documentation.

Context

Scenario: Assume that you have deployed the latest ZStack Private Cloud environment, and take over an iSCSI-SAN storage to the Cloud. Due to your business requirements, you need to deploy a core business system on the active-backup VM instances, and configure Microsoft Cluster Server for these VM instances. Assume that the virtual IP to access Microsoft Cluster Server is *172.24.20.20*. Besides, as storage planning is required, you need to pass through one LUN to two VM instances, and use this LUN to act as Cluster Shared Volume of Microsoft Cluster Server. Under this scenario, if one VM instance (Master) fails over, another VM instance (Standby) will automatically be started to ensure your business continuity and high availability.

Note:

- Microsoft Cluster Server requires both VM instances to be included in the same anti-affinity group. That is, these VM instances cannot be attached to the same host.

- If these VM instances in Microsoft Cluster Server have been passed through to an LUN, make sure that you must not perform hot migrations for these VM instances. Otherwise, Microsoft Cluster Server will probably be not available. If you do need to perform hot migrations, verify that Microsoft Cluster Server configured on these VM instances is stopped.
- The LUN that is passed through to the VM instances cannot be used for other means.
- If you pass through the LUN to these VM instances for the first time, modify the multi-path configuration file. This can avoid the influence on the used storage.
- We recommend that you use Qlogic HBA to perform the passthrough operation. If you use EMULEX HBA, recompile the driver, attach modules, and restart hosts.

The following configurations are about these two VM instances:

Table 5-1: VM Instance

VM Name	Operating System	Configuration Item	Configuration Information
VM-1 (Master)	Windows Server 2016 Standard	Static IP address	172.24.195.251
		FQDN	node1.example.com
		Domain	example.com
VM-2 (Standby)		Static IP address	172.24.248.71
		FQDN	node2.example.com
		Domain	example.com

To create Microsoft Cluster Server:

1. Pass through an LUN to two VM instances.
2. Initialize the LUN.
3. Add features necessary to Microsoft Cluster Server for the two VM instances.
4. Configure and validate tests.
5. Create a failover cluster.
6. Add disks to Microsoft Cluster Server.
7. Verify whether Microsoft Cluster Server takes effect.

Procedure

1. Pass through an LUN to two VM instances.

In the navigation pane of the ZStack Private Cloud UI, choose **Hardware** > **SAN Storage**. On the **SAN Storage** page, go to the **iSCSI Storage** tab page, and click an iSCSI Qualified Name. On the **Disk** tab page, select an iSCSI LUN. Choose **Actions** > **Attach VM Instance**. On the displayed **Attach VM Instance** page, select two VM instance that you need to attach, and click **OK**, as shown in *Figure 5-1: Pass through LUN*.

Figure 5-1: Pass through LUN

2. Initialize the LUN.

   a) Go to the **Computer Management** page.

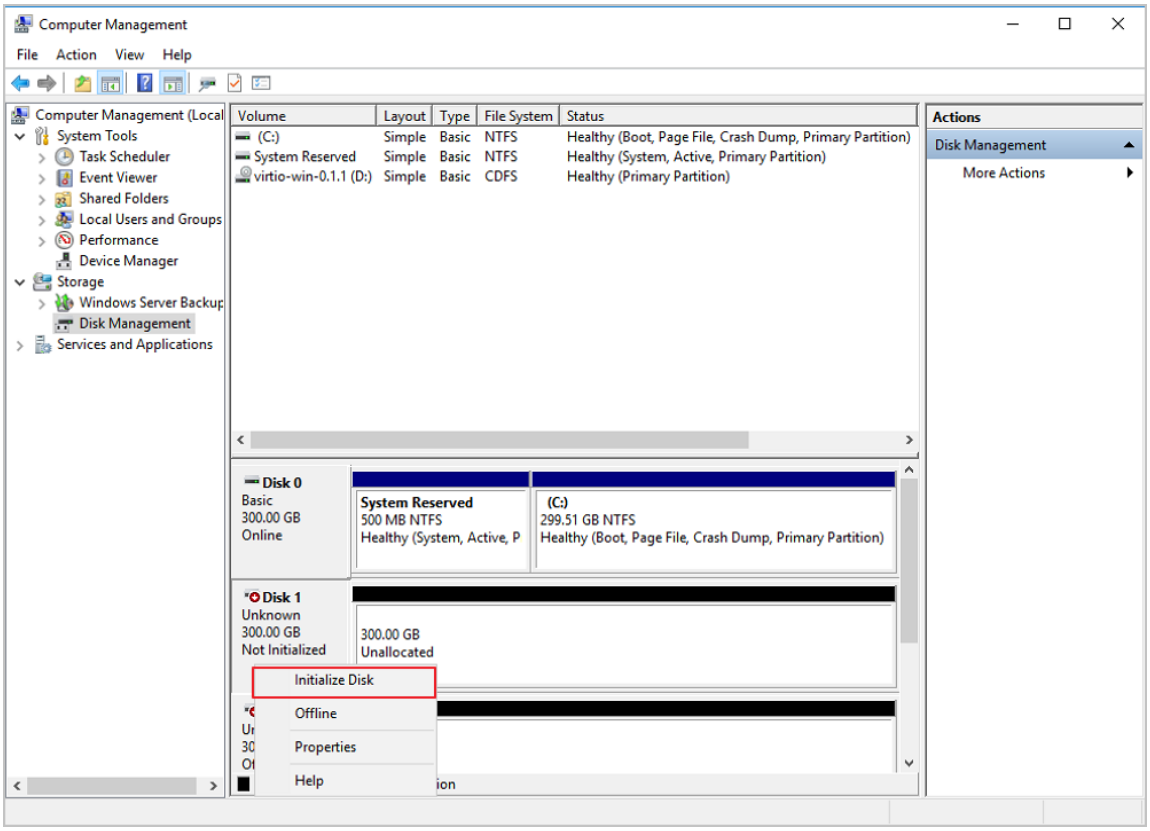
   Open the console of the VM-1, and go to the **Server Manager** page. To go to **Computer Management**, choose **Tools** > **Computer Management**, as shown in *Figure 5-2: Computer Management Page*.

   Figure 5-2: Computer Management Page

   b) Initialize the disk.

   Go to the **Disk Management** details page, right-click **Disk 1**, select **Online**. Right-click **Disk 1** again to select **Initialize Disk**. On the **Initialize Disk** page, click **OK** to complete creating the partition table, as shown in *Figure 5-3: Initialize Disk*.

Figure 5-3: Initialize Disk



3. Add features necessary to Microsoft Cluster Server for these two VM instances.

   a) Add features necessary to Microsoft Cluster Server for the VM-1.

   On the **Server Manager** page, click **Add Roles and Features**. On the displayed **Add Roles and Features Wizard** page, click **Next**. On the **Select features** page, select **Failover Clustering**. On the displayed **Add features that are required for Failover Clustering?** page, click **Add Features**, and then click **Next**. On the **Confirm installation selections** page, click **Install**, and wait until the installation is completed. Finally, click **Close** to complete adding Microsoft Cluster Server.

   The confirmation to add features for the VM instance is shown in *Figure 5-4: Confirmation Page*.

   Figure 5-4: Confirmation Page

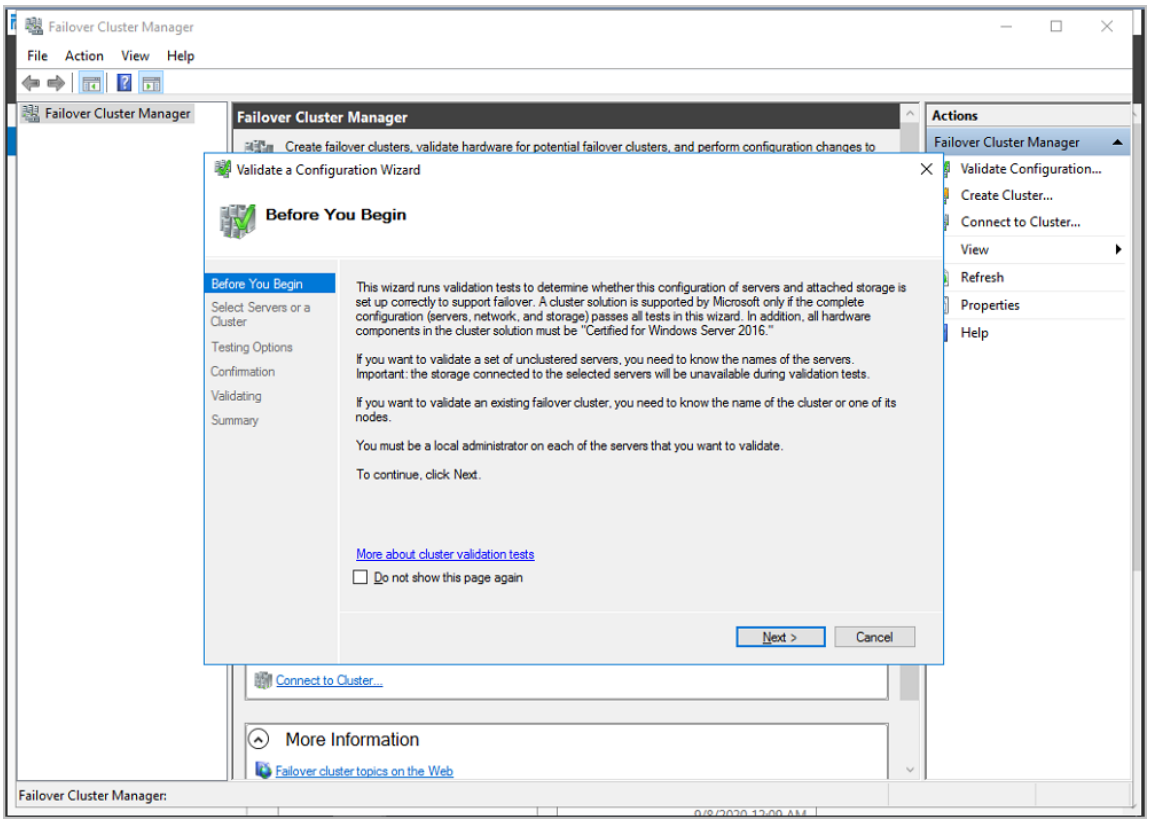
   b) Similarly, add features necessary to Microsoft Cluster Server for the VM-2.

4. Configure and validate tests.

a) Go to the test and configuration wizard page.

On the **Server Manager** page, choose **Tools** > **Failover Cluster Manager**. On the **Failover Cluster Manager** page, to go to **Validate Configuration Wizard** page, click **Validate Configuration**, as shown in *Validate Configuration Wizard*.

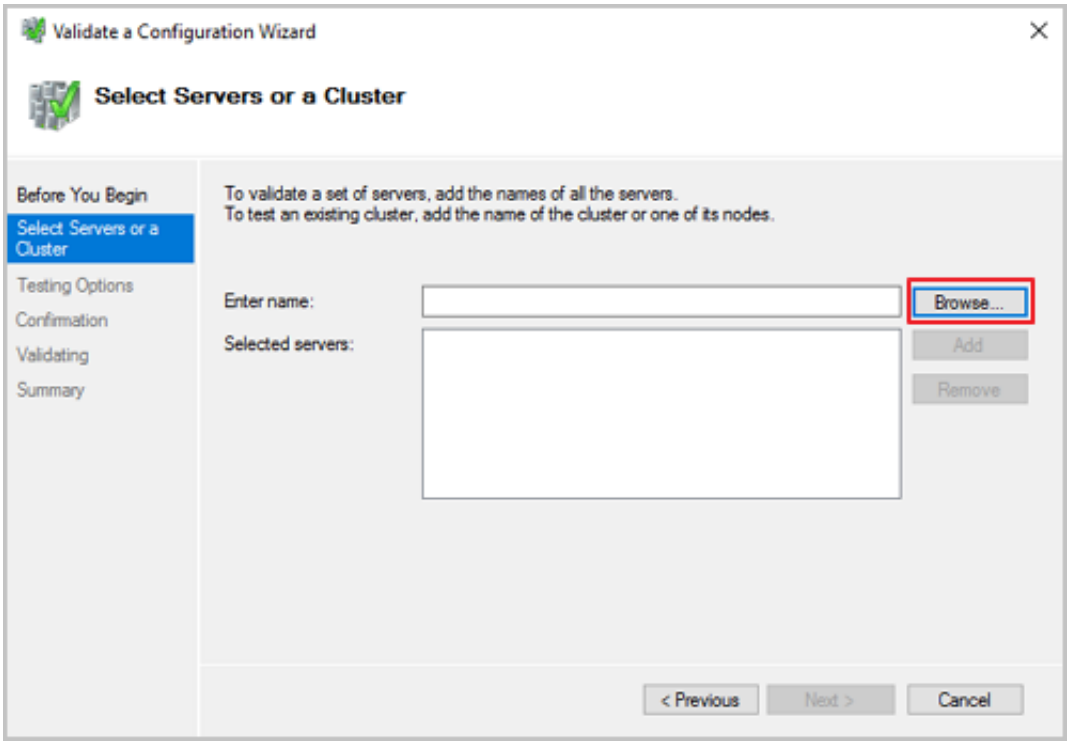
Figure 5-5: Validate Configuration Wizard

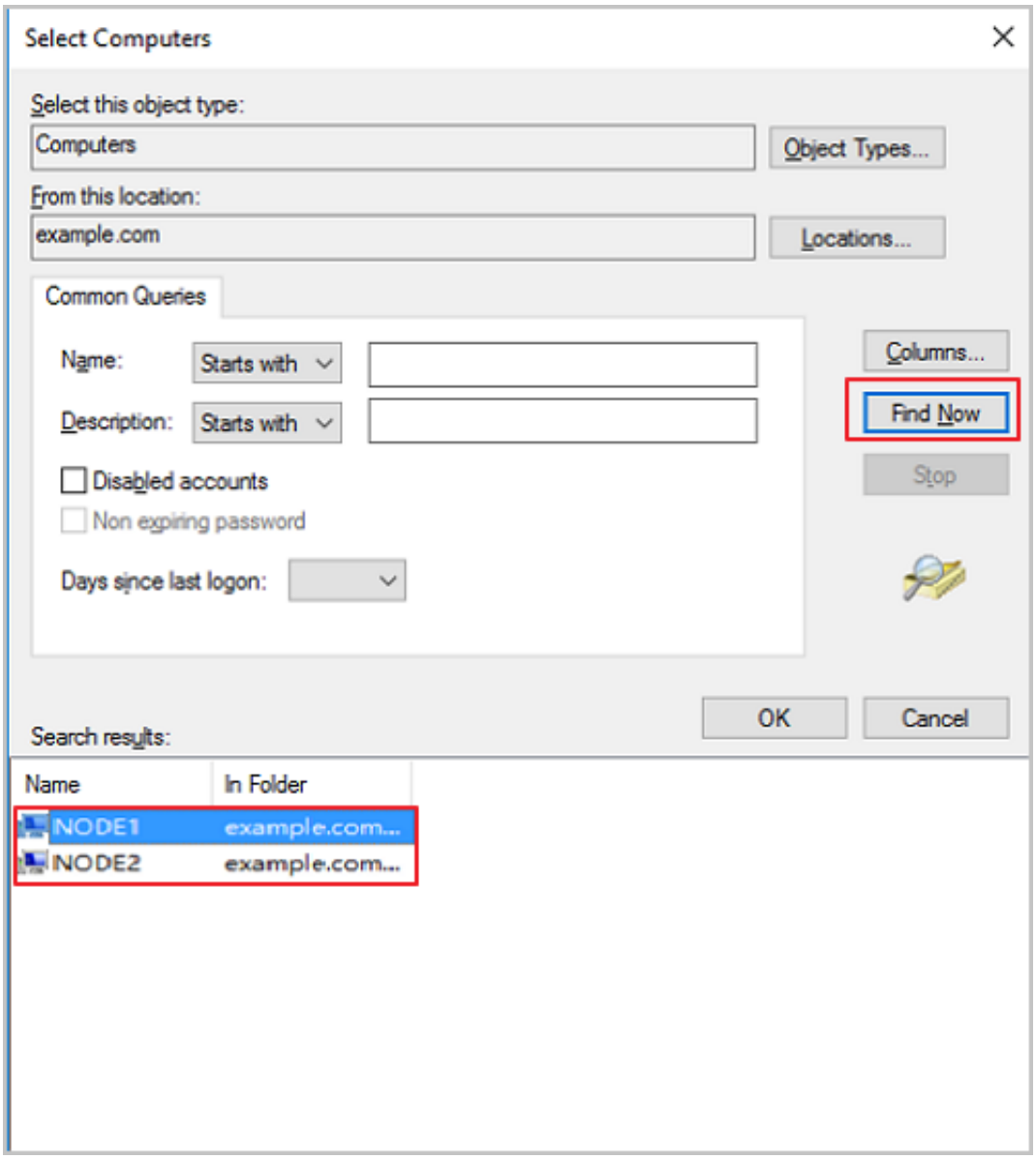


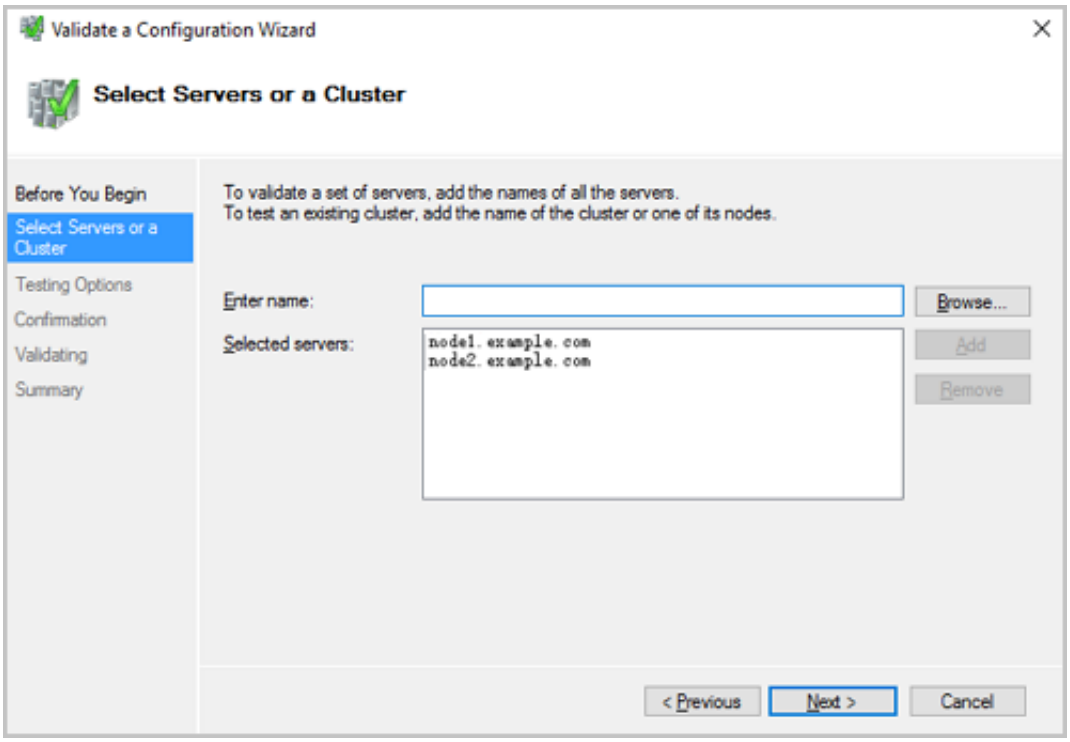
b) Select the subject that you need to test.

On the **Validate Configuration Wizard** page, click **Next**. On the **Select Servers or a Cluster** page, choose **Browse** > **Advanced** > **Find Now**. Select the subject that you need to test, and click **OK**, as shown in *Select Subject to be Validated*.

Figure 5-6: Select Subject to be Validated



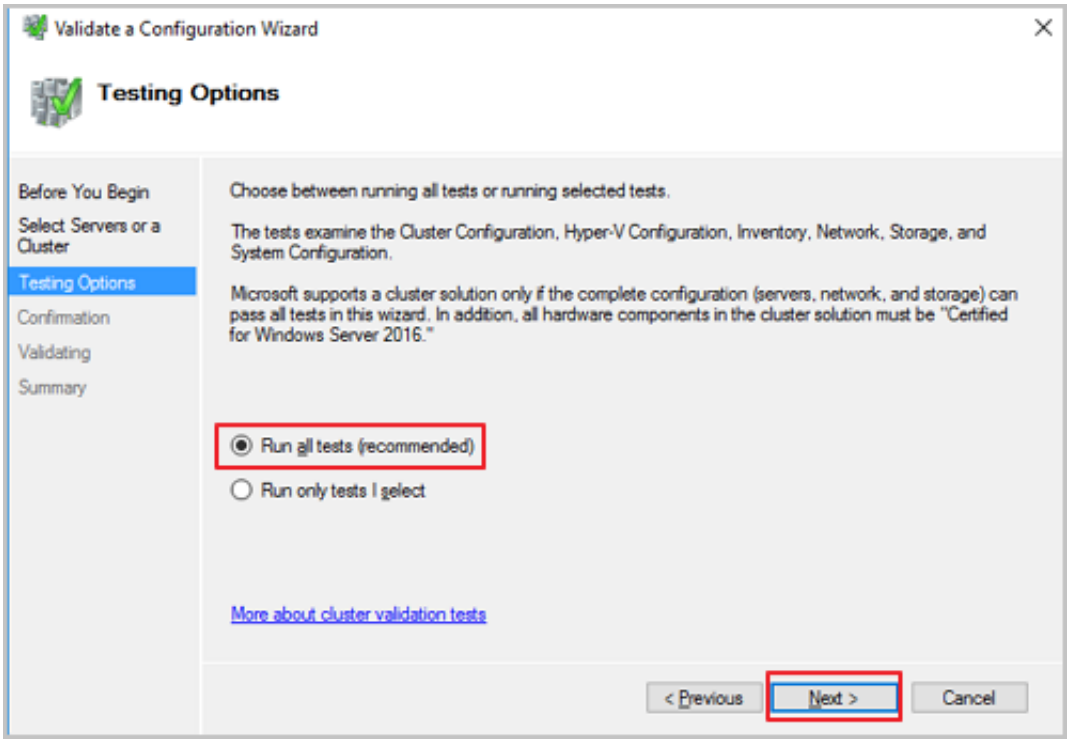

Select Computers
Select this object type:
Computers
Object Types...
From this location:
example.com
Locations...
Common Queries
Name:
Starts with
Description:
Starts with
Disabled accounts
Non expiring password
Days since last logon:
Columns...
Find Now
Stop
OK
Cancel
Search results:
Name
In Folder
NODE1
example.com...
NODE2
example.com...



c) Select the test method.

After you select the subject, click **Next**. On the **Testing Options** page, select **Run all tests (recommended)**, and choose **Next** > **Next** to start testing the subject. Only if all related configuration and verification tests are prepared well, Microsoft will support the Microsoft Cluster Server service. After the test is completed, click **Finish**, as shown in *Figure 5-7: Testing Options*.

Figure 5-7: Testing Options



Note:

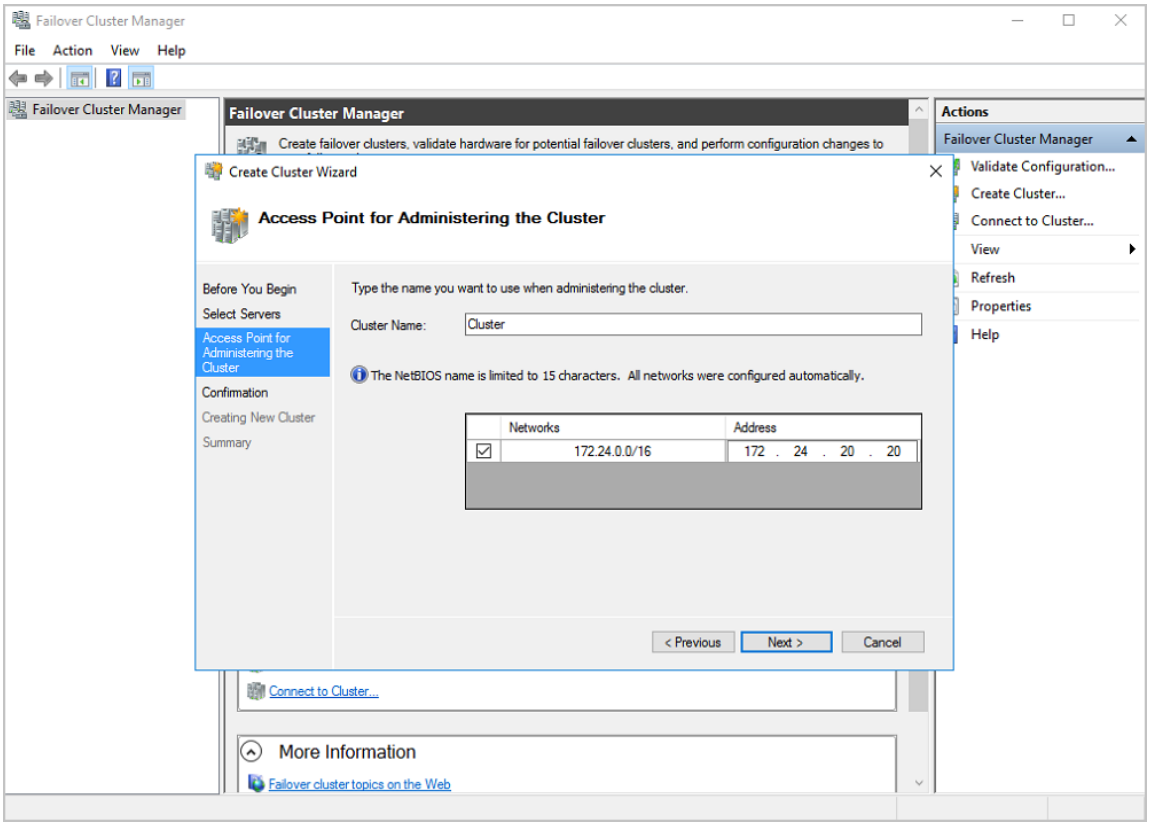
- We recommend that you select **Run all tests (recommended)**.
- If the test result returns errors, verify that subsequent operations can be performed after ell errors are resolved. Otherwise, Microsoft Cluster Server is abnormal.

5. Create failover clusters.

   a) Set parameters for failover clusters.

   On the **Failover Cluster Manager** page, click **Create Cluster**. On the **Create Cluster Wizard** page, click **Next**. On the **Select Servers** page, add the subject that has passed the test. Click **Next** to set the names for the failover clusters and the virtual IP address used to access Microsoft Cluster Server, as shown in *Figure 5-8: Set Cluster Parameters*.

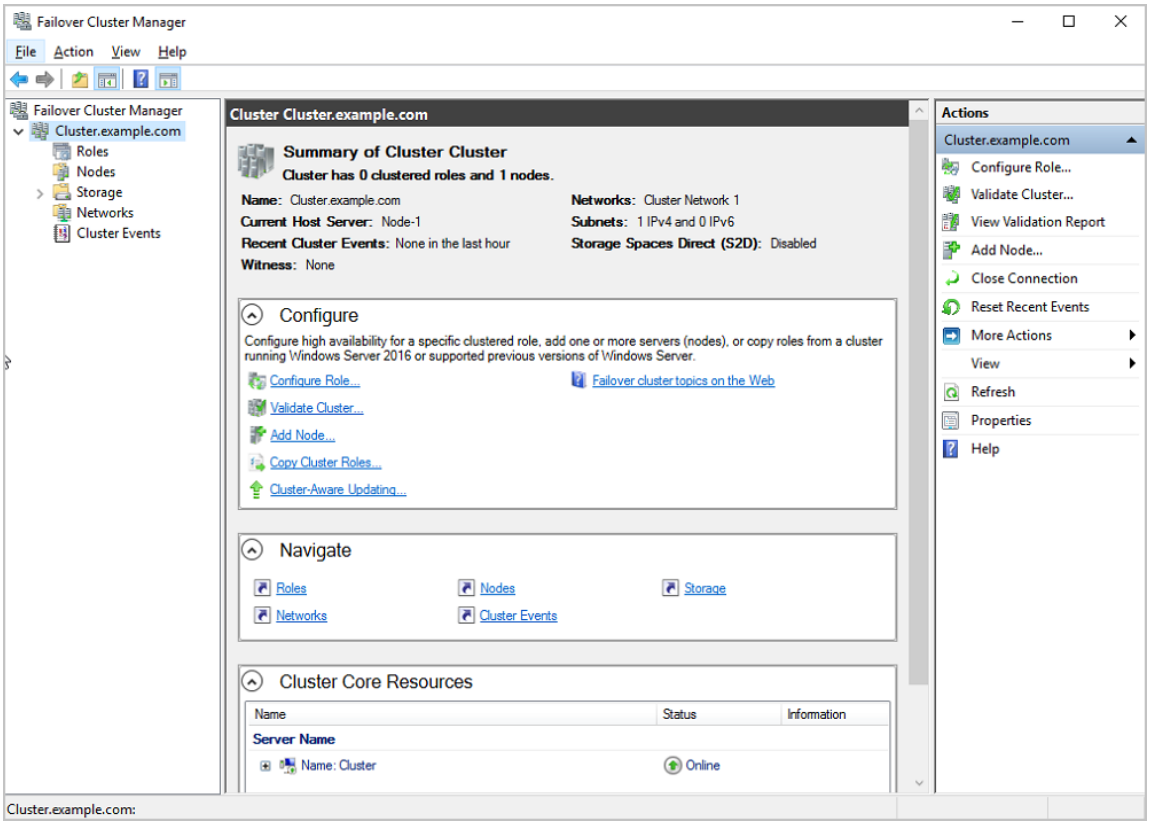
Figure 5-8: Set Cluster Parameters



b) Preview configurations for the failover clusters to be created.

After you set the parameters for the failover clusters, click **Next** to preview the configurations of the failover cluster. After confirmation, click **Next**, and start creating the failover cluster. After you create the failover cluster, click **Finish**, as shown in *Figure 5-9: Preview Configuration of Cluster to Be Created*.

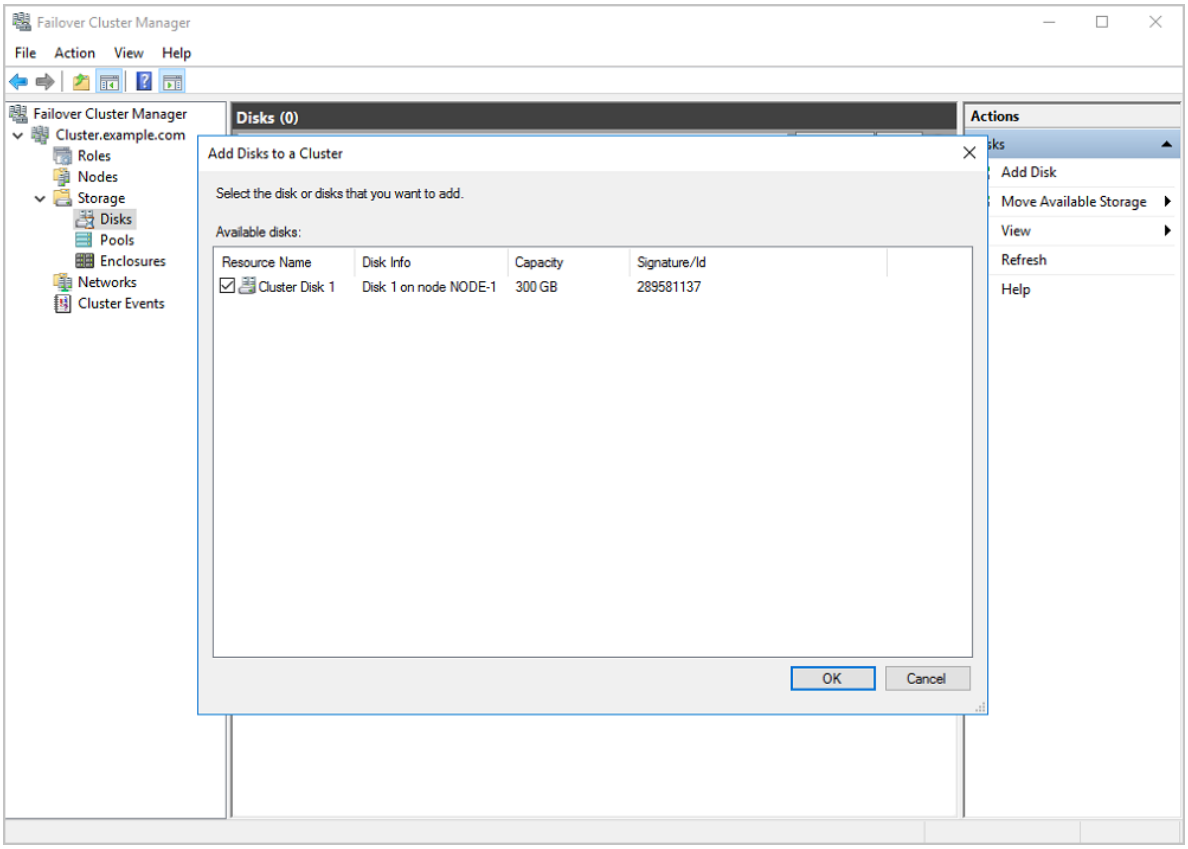
Figure 5-9: Preview Configuration of Cluster to Be Created

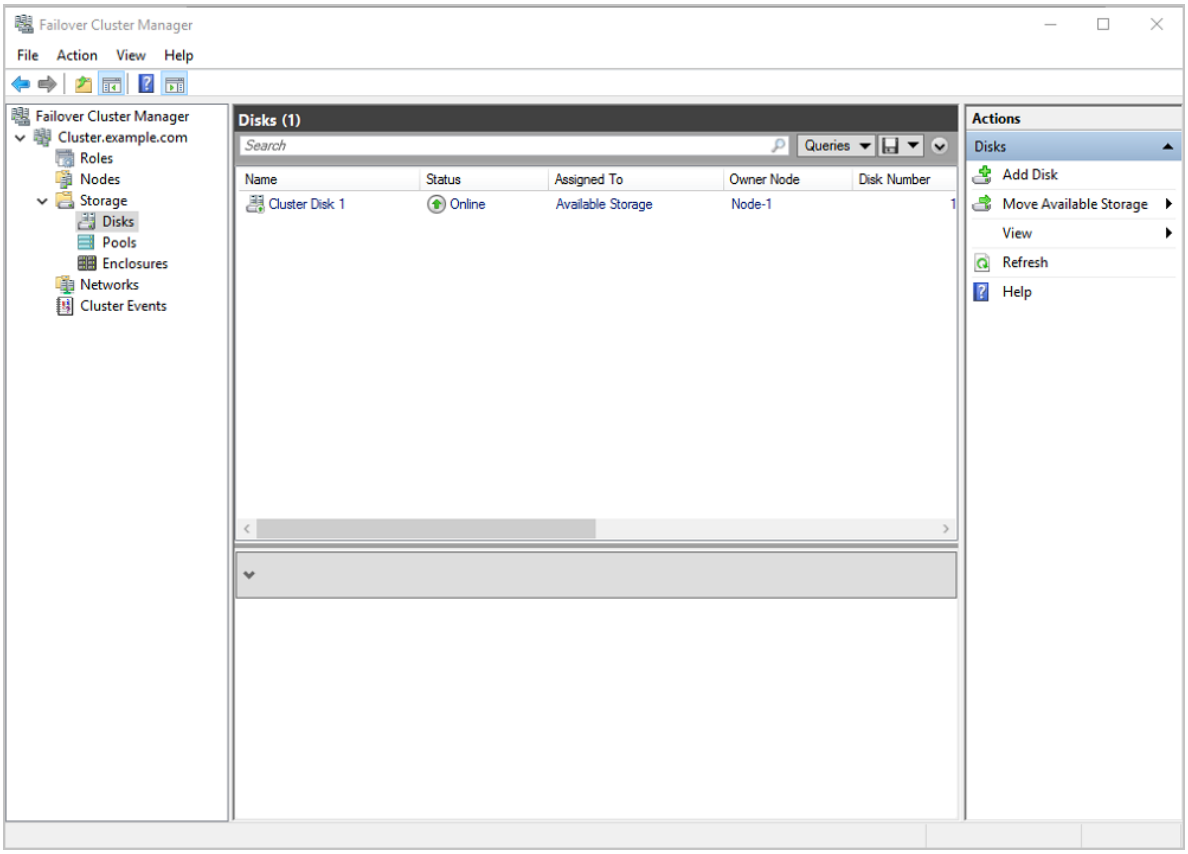


6. Add disks to Microsoft Cluster Server.

   On the **Failover Cluster Manager** page, choose **cluster.example.com** > **Storage** > **Disk**. On the **Storage** details page, right-click **Add Disk**. On the displayed **Add Disks to a Cluster** page, select the disk that you need to add, and click **OK**, as shown in *Figure 5-10: Add Disk*.

Figure 5-10: Add Disk





7. Verify whether Microsoft Cluster Server takes effect.

   Run **ping** to validate whether Microsoft Cluster Server takes effect.

   Expected result:

   1. When VM-1 (Master) runs normally, the network can be connected by running **ping**.
   2. When the VM-2 (Standby) fails over, the network will temporarily be disconnected.
   3. After the Microsoft Cluster Server mechanism triggers, VM-2 (Standby) continues to ensure connectivities.

   Actual result:

```
[root@10-0-73-238 ~]# ping 192.168.200.200
PING 172.24.20.20 (172.24.20.20) 56(84) bytes of data.
64 bytes from 172.24.20.20: icmp_seq=1 ttl=126 time=26.3 ms
64 bytes from 172.24.20.20: icmp_seq=2 ttl=126 time=13.1 ms
64 bytes from 172.24.20.20: icmp_seq=3 ttl=126 time=2.88 ms
64 bytes from 172.24.20.20: icmp_seq=4 ttl=126 time=1.73 ms
64 bytes from 172.24.20.20: icmp_seq=5 ttl=126 time=2.27 ms
64 bytes from 172.24.20.20: icmp_seq=6 ttl=126 time=1.11 ms
Request timeout for icmp_seq 7
Request timeout for icmp_seq 8
Request timeout for icmp_seq 9
Request timeout for icmp_seq 10
Request timeout for icmp_seq 11
64 bytes from 172.24.20.20: icmp_seq=12 ttl=126 time=125.2 ms
64 bytes from 172.24.20.20: icmp_seq=13 ttl=126 time=11.3 ms
64 bytes from 172.24.20.20: icmp_seq=14 ttl=126 time=2.67 ms
^C
--- 172.24.20.20 ping statistics ---
14 packets transmitted, 9 received, 35.7% packet loss, time 15053ms
```

What's next

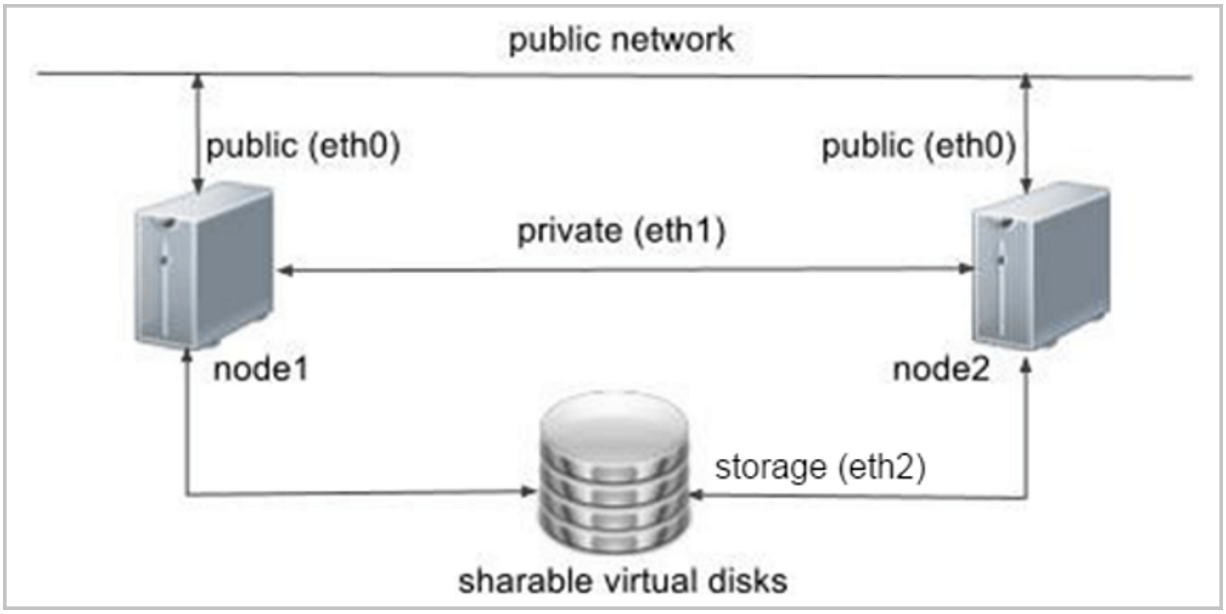
So far, we have introduced the usage of Microsoft Cluster Server.

5.2 Oracle RAC

Oracle Real Application Clusters (Oracle RAC) is initially designed to provide better database services. After years of development, Oracle RAC is now based on a comprehensive high availability system. This high availability system can either act as the base of a database cloud system, or serve as a shared infrastructure to provide higher availability, scalability, flexibility, and agility for all application softwares in a data center. Oracle RAC adopts a shared disk architecture. In this regard, a volume manager and a file system used for storing database data must support cluster identifying. Oracle Automatic Storage Management (Oracle ASM) is a volume (cluster) manager recommended for Oracle database.

Oracle ASM can automatically manage shared disks. Therefore, you can provide bare devices for RAC nodes via FC LUN passthrough, iSCSI LUN passthrough, or shared, virtual disks. These RAC nodes can manage databases without deploying additional shared file systems, as shown in *RAC Workflow*.

Figure 5-11: RAC Workflow



For more information about Oracle RAC, see *Oracle Official Documentation*.

Glossary

Zone

A zone is a logical group of resources such as clusters, L2 networks, and primary storages. Zone is the largest resource scope defined in ZStack.

Cluster

A cluster is a logical group of analogy hosts (compute nodes). Hosts in the same cluster must be installed with the same operating system, have the same network configuration, and be able to access the same primary storage. In a real data center, a cluster usually maps to a rack.

Management Node

A management node is a host with operating system installed to provide UI management and Cloud deployment.

Compute Node

A compute node is a physical server (also known as a host) that provides VM instances with compute, network, and storage resources.

Primary Storage

A primary storage is a storage server used to store disk files in VM instances. Local storage, NFS, Ceph, Shared Mount Point, and Shared Block are supported.

Backup Storage

A backup storage is a storage server used to store image template files. ImageStore, SFTP (Community Edition), and Ceph are supported. We recommend that you deploy backup storage separately.

ImageStore

ImageStore is a type of backup storage. You can use ImageStore to create images for VM instances that are in the running state and manage image version updates and release. ImageStore allows you quickly upload, download, export images, and create image snapshots as needed.

VM Instance

A VM instance is a virtual machine instance running on a host. A VM instance has its own IP address to access public network and run application services.

Image

An image is an image template used by a VM instance or volume. Image templates include system volume images and data volume images.

Volume

A volume can either be a data volume or a root volume. A volume provides storage to a VM instance. A shared volume can be attached to one or more VM instances.

Instance Offering

An instance offering is a specification of the VM instance CPU and memory, and defines the host allocator strategy, disk bandwidth, and network bandwidth.

Disk Offering

A disk offering is a specification of a volume, which defines the size of a volume and how the volume will be created.

L2 Network

An L2 network is a layer 2 broadcast domain used for layer 2 isolation. Generally, L2 networks are identified by names of devices on the physical network.

L3 Network

An L3 network is a collection of network configurations for VM instances, including the IP range, gateway, and DNS.

Public Network

A public network is generally allocated with a public IP address by Network Information Center (NIC) and can be connected to IP addresses on the Internet.

Private Network

A private network is the internal network that can be connected and accessed by VM instances.

L2NoVlanNetwork

L2NoVlanNetwork is a network type for creating an L2 network. If L2NoVlanNetwork is selected, VLAN settings are not used for host connection.

L2VlanNetwork

L2VlanNetwork is a network type for creating an L2 network. If L2VlanNetwork is selected, VLAN settings are used for host connection and need to be configured on the corresponding switches in advance.

VXLAN Pool

A VXLAN pool is an underlay network in VXLAN. You can create multiple VXLAN overlay networks (VXLAN) in a VXLAN pool. The overlay networks can operate on the same underlay network device.

VXLAN

A VXLAN network is a L2 network encapsulated by using the VXLAN protocol. A VXLAN network belongs to a VXLAN pool. Different VXLAN networks are isolated from each other on the L2 network.

vRouter

A vRouter is a custom Linux VM instance that provides various network services.

Security Group

A security group provides L3 network firewall control over the VM instances. It can be used to set different security rules to filter IP addresses, network packet types, and the traffic flow of network packets.

EIP

An elastic IP address (EIP) is a method to access a private network through a public network.

Snapshot

A snapshot is a point-in-time capture of data status in a disk. A snapshot can be either an automatic snapshot or a manual snapshot.