VMware vCenter Server™ 6.0 Availability Guide

TECHNICAL MARKETING DOCUMENTATION

V 2.0/NOVEMBER 2015/MOHAN POTHERI, G. BLAIR FRITZ, PUNEET GUPTA

vmware®

Table of Contents

Overview

A correctly architected and highly available solution provides applications with the largest amount of acceptable operational uptime by countering the impact of unplanned downtime. Although downtime can also be planned—for maintenance and patching, for example—the unplanned outages have the greatest effect on production uptime. This paper will discuss the requirements of defining high availability for VMware vCenter Server™ on Microsoft Windows and VMware vCenter™ Server Appliance™, with recommendations and best practices for providing acceptable levels of protection.

Availability Considerations for vCenter Server 6.0

vCenter Server can be a single point of failure in the environment. Many VMware® solutions, such as VMware Horizon® Suite, VMware vRealize™ Automation™, and so on, are layered on top of vCenter Server; its availability impacts the usability of these components. Loss of vCenter Server curtails the ability of these solutions to perform many critical functions such as making changes and creating new virtual machines (VMs).

VMware vCenter comprises multiple components. To provide reliable availability options, an understanding of what these components entail, including dependencies, is required.
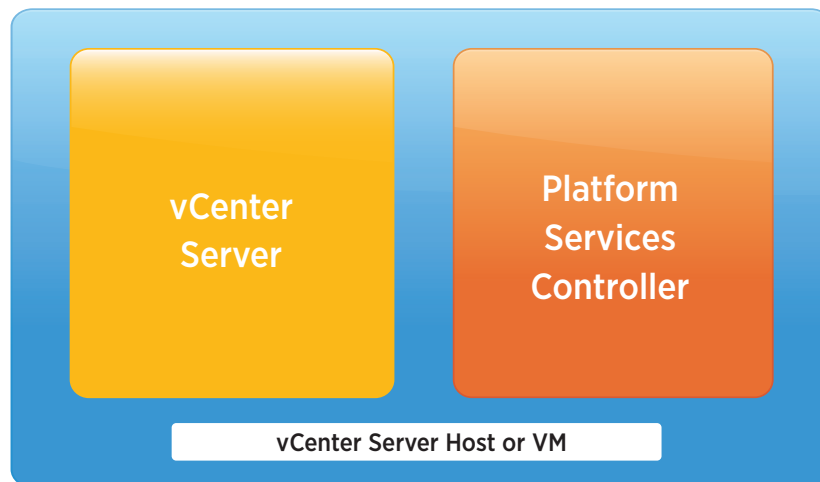

Figure 1. vCenter Server 6.0 Components

The primary vCenter components are vCenter Server and the Platform Services Controller (PSC).

vCenter Server

Many of the traditional services have been consolidated into the vCenter Server node. The following are vCenter Server services:

• VMware vCenter core components
• VMware vSphere® Web Client
• VMware vCenter Inventory Service
• VMware vSphere Storage Profile drive storage
• VMware vSphere Auto Deploy™
• VMware vSphere Syslog Collector
• VMware vSphere ESXi™ Network Dump Collector

Platform Services Controller

The PSC supports data and state replication and includes the following set of common infrastructure services that can be used by vCenter Server, VMware vRealize Suite, and other solutions:

• Single sign-on (SSO)
• Licensing
• Certificate authority
• Global permissions
• Other common services

vCenter Server Components and Their Availability Characteristics

• vCenter Server – vCenter Server services that encompass a majority of its functions, excluding those performed by the PSC, are run on a standalone server that can be a single point of failure. Failure of this server causes loss of these vCenter Server services. Greater availability can be achieved by ensuring that these services are monitored and any failure is remediated.

• Database services – vCenter Server depends on database services that can be internal or external to the server. This is also a single point of failure. The database server must be monitored and protected from all types of failures.

• Platform Services Controller – There can be multiple PSC instances. They replicate state information between themselves, so it is not a single point of failure if there is more than one per site.

vCenter Server High-Availability Solutions and Configurations

There are multiple solutions for high availability. Many of these options can be combined to provide different levels of availability. We will look at all available solutions and discuss their configuration, pros and cons, and applicability to vCenter Server.

VMware vSphere High Availability

VMware vSphere High Availability (vSphere HA) is the time-tested high-availability solution from VMware that is widely used for production environments. If configured with no single points of failure, it can protect the workloads from hardware failures. In the event of any hardware failure, the protected workload is automatically restarted in the remaining nodes of the cluster. There is an outage for the workload for the duration of the detection of the failure and the restart of the VM and the application. vSphere HA is very easy to set up and manage and is the simplest high-availability solution available for protecting virtual workloads.

Protecting vCenter Server with vSphere HA

VMware recommends leveraging vSphere HA and VMware vSphere Distributed Resource Scheduler™ (vSphere DRS) if licensing permits.

Prerequisites
- Verify that the vCenter Server system and its configuration files reside on shared storage.
- Verify that the hosts are configured to access the shared storage so you can power on the VMs by using different hosts in the cluster.
- Verify that hosts are configured to have access to the VM network.
- Verify that you are using redundant management network connections for vSphere HA. For information about setting up network redundancy, see "Best Practices for Networking."
- Verify that you have configured hosts with at least two datastores to provide redundancy for vSphere HA datastore heartbeating.
- Connect vSphere Web Client to vCenter Server using an account with cluster administrator permissions.

Procedure for New Cluster
1. In vSphere Web Client, browse to the data center where you want the cluster to reside and click **Create a Cluster**.
2. Complete the New Cluster Wizard. Do not turn on vSphere HA or vSphere DRS.
3. Click **OK** to close the wizard and create an empty cluster.
4. Based on your plan for the resources and networking architecture of the cluster, use vSphere Web Client to add hosts to the cluster.
5. Browse to the cluster and enable vSphere HA.
   a. Click the **Manage** tab and click **Settings**.
   b. Select **vSphere HA** and click **Edit**.
   c. Select **Turn On vSphere HA**.

6. Select **Host Monitoring**. Enabling **Host Monitoring** enables hosts in the cluster to exchange network heartbeats and enables vSphere HA to take action when it detects failures.

7. Choose a setting for **Virtual Machine Monitoring**. Select **VM Monitoring Only** to restart individual VMs if their heartbeats are not received within a set time. You can also select **VM and Application Monitoring** to enable application monitoring.

8. Click **OK**.

vSphere HA is now turned on and can help reduce downtime for the vCenter Server system during hardware failures. As part of the configuration, enable Host Hardware Monitoring with protection against storage connectivity loss. This will help in situations such as All Paths Down (APD) and PDL (Permanent Device Loss). In the event of a storage failure, the VM is restarted on a healthy host.

vSphere HA and vSphere DRS Affinity Rules

vSphere HA failover can be specified for the following two types of rules:

• VM antiaffinity rules force specified VMs to remain apart during failover actions.

• VM host affinity rules place specified VMs on a particular host or a member of a defined group of hosts during failover actions.

When editing a vSphere DRS affinity rule, select the checkbox or checkboxes that enforce the preferred failover action for vSphere HA.

• vSphere HA is subject to VM antiaffinity rules during failover. If VMs with this rule are placed together, the failover is aborted.

• vSphere HA is subject to VM-to-host affinity rules during failover. vSphere HA attempts to place VMs with this rule on the specified hosts if possible.

vSphere DRS affinity rules help protect vCenter Server and database components by providing appropriate separation between the primary and standby servers and also between multiple PSCs.

Affinity rules can be used to keep the vCenter Server system within a small group of hosts because they can help detect the host location of the system itself in the event of its failure. If the vCenter Server system uses a dedicated external database server, affinity rules can be used to keep them together on the same physical host.

Additional vSphere HA Considerations

• Enable and configure proper admission control for the cluster.

• Set the VM restart priority to High for the VM or VMs that are hosting the vCenter Server system and database.
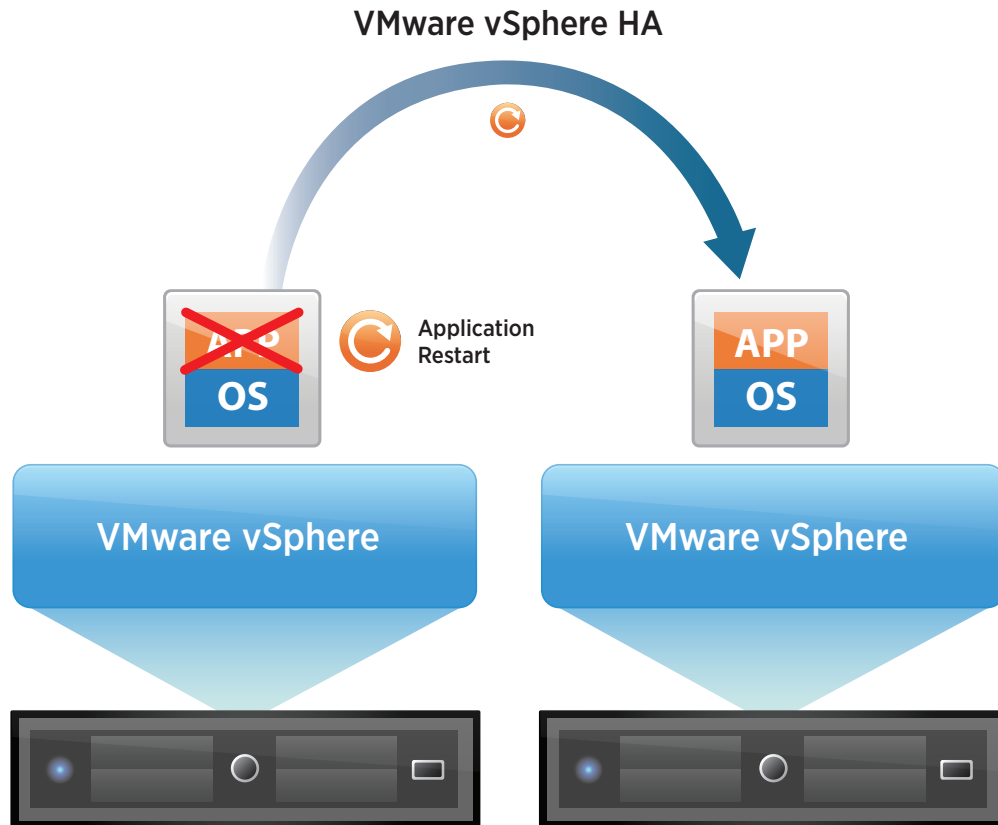

Figure 2. vSphere HA to Protect vCenter Server Combined with Watchdog for Application Restart

Watchdog Protection for vCenter Server

Watchdog is a process that monitors and protects vCenter Server services. It is available on both the vCenter Server Appliance and the Microsoft Windows–based vCenter Server versions and is enabled by default. If any services fail, Watchdog attempts to restart them. If it cannot restart the service because of a host failure, vSphere HA restarts the VM, running the service on a new host. Watchdog can provide better availability by using vCenter Server processes (PID Watchdog) or the vCenter Server API (API Watchdog). Each vCenter Server process has a separate Watchdog process associated with it. The following is an example with vCenter Server Appliance:

```
ps -ef | grep vmware-watchdog
root 7398 1 0 Mar27 ? 00:00:00 /bin/sh /usr/bin/vmware-watchdog -s rhttpproxy -u 30 -q 5
/usr/sbin/rhttpproxy -r /etc/vmware-rhttpproxy/config.xml -d /etc/vmware-rhttpproxy
root 11187 1 0 Mar27 ? 00:00:00 /bin/sh /usr/bin/vmware-watchdog -s vws -u 30 -q 5
/usr/lib/vmware-vws/bin/vws.sh
root 12041 1 0 Mar27 ? 00:09:58 /bin/sh /usr/bin/vmware-watchdog -s syslog -u 30 -q 5 -b
/var/run/rsyslogd.pid /sbin/rsyslogd -c 5 -f /etc/vmware-rsyslog.conf
root 12520 1 0 Mar27 ? 00:09:56 /bin/sh /usr/bin/vmware-watchdog -b
/storage/db/vpostgres/postmaster.pid -u 300 -q 2 -s vmware-vpostgres su -s /bin/bash vpostgres
root 29201 1 0 Mar27 ? 00:00:00 /bin/sh /usr/bin/vmware-watchdog -a -s vpxd -u 3600 -q 2
/usr/sbin/vpxd
```

Starting with vSphere 6.x, a Python daemon called API Watchdog checks the status of APIs for the vpxd service. If the APIs are not running, API Watchdog attempts twice to restart the service. If that still does not resolve the issue, API Watchdog then reboots the VM. API Watchdog starts running immediately after deployment of vCenter Server Appliance. On vCenter Server for Windows, however, vCenter Server must be rebooted once before API Watchdog starts working.

VMware vSphere Fault Tolerance

In the event of server failures, VMware vSphere Fault Tolerance (vSphere FT) provides continuous availability for applications with as many as four virtual CPUs. It does so by creating a live shadow instance of a VM that is always up to date with the primary VM. In the event of a hardware outage, vSphere FT automatically triggers failover, ensuring zero downtime and preventing data loss. Like vSphere HA, it protects against hardware failure but completely eliminates downtime with instantaneous cutover and recovery. After failover, vSphere FT automatically creates a new, secondary VM to deliver continuous protection for the application.

vSphere FT offers the following benefits:

• Protects mission-critical, high-performance applications regardless of operating system (OS)

• Provides continuous availability, for zero downtime and zero data loss with infrastructure failures

• Delivers a fully automated response

When virtualizing vCenter Server, technologies such as vSphere FT can help protect the vCenter management server from hardware failures. Compared to vSphere HA, vSphere FT can provide instantaneous protection, but the following limitations must be considered:

• The vCenter Server system is limited to four vCPUs.

• vSphere FT protects against hardware failures but not against application failures.

• vSphere FT cannot reduce downtime for patching-related outages.

• vSphere FT has resource requirements that can create additional overhead.

Because vSphere FT is suitable for workloads with a maximum of four vCPUs and 64GB of memory, it can be used in "tiny" and "small" vCenter Server deployments.
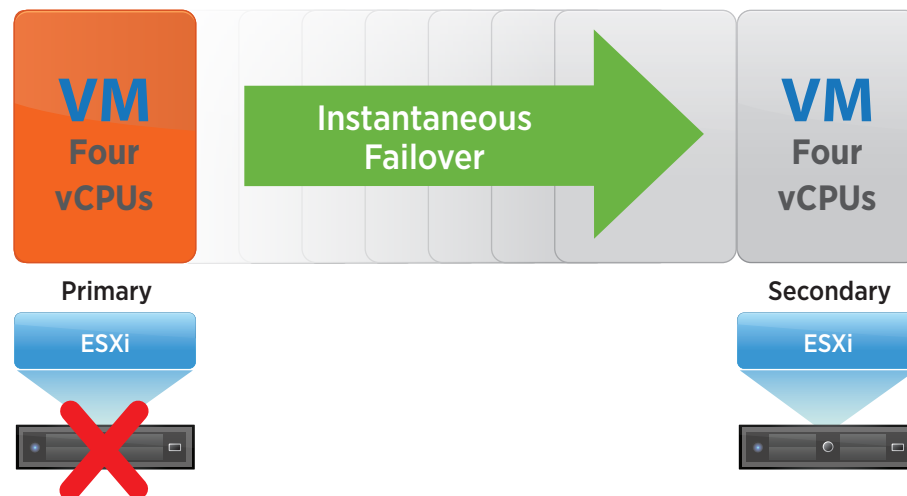

Figure 3. vSphere Fault Tolerance for vCenter Server Protection

vSphere FT can be turned on through vSphere Web Client.

When vSphere FT is turned on, vCenter Server resets the VM's memory limit and sets the memory reservation to the memory size of the VM. While vSphere FT remains turned on, memory reservation, size, and limit as well as number of vCPUs and shares cannot be changed. In addition, disks for the VM cannot be added or removed.

Connect vSphere Web Client to vCenter Server using an account with cluster administrator permissions.

Prerequisites for Using vSphere FT
All hosts with vSphere FT enabled require a dedicated 10Gbps low-latency VMkernel interface for vSphere FT logging traffic.

The option to turn on vSphere FT is unavailable (dimmed) if any of these conditions apply:

- The VM resides on a host that does not have a license for the feature.
- The VM resides on a host that is in maintenance mode or standby mode.
- The VM is disconnected or orphaned—that is, its VMX file cannot be accessed.
- The user does not have permission to turn the feature on.

Configuration Procedure
1. In vSphere Web Client, browse to the VM for which you want to turn on vSphere FT.

2. Right-click the VM representing vCenter Server and select **Fault Tolerance** > **Turn On Fault Tolerance**.

3. Click **Yes**.

4. Select a datastore on which to place the secondary VM configuration files. Then click **Next**.

5. Select a host on which to place the secondary VM. Then click **Next**.

6. Review your selections and then click **Finish**.

The specified vCenter Server VM is designated as a primary VM; a secondary VM is established on another host. The primary vCenter Server VM is now fault tolerant.

Third-Party Tools Leveraging the VMware Application Monitoring API

Third-party tools such as Symantec ApplicationHA provide monitoring capabilities for applications running inside VMs. vSphere HA has an API that enables third-party vendors to develop agents that can monitor the health of an application running within the guest OS and inform vSphere HA when a problem is detected. Symantec is an example of a partner to have developed an agent for providing application awareness within a vSphere cluster. Symantec ApplicationHA runs inside the application and is fully integrated with vSphere HA and vCenter Server.

An agent is installed in the VM to monitor the health of the application resources. If a configured application instance or associated services become unavailable, the agent automatically detects it and tries to start the application services for a configurable number of attempts. If the application services fail to start, the agent interprets this as an application failure and reports the status to vSphere HA. vSphere HA can then restart the VM. After the VM restarts, the agent starts the application services and brings the configured resources online on the system.

This solution does not use a standby VM; instead, it restarts the services or, if all else fails, restarts the VM itself. Planned downtimes relating to patching the OS or application are not protected by this solution. The configuration of third-party tools is outside the scope of this document.

Virtual Machine Guest Operating System Clustering Solutions

Several applications use clustering, including stateless applications such as Web servers and applications with built-in recovery features such as database servers. A typical clustering setup includes disks that are shared between nodes. A shared disk is required as a quorum disk. In a cluster of VMs across physical hosts, the shared disk must be on a Fibre Channel (FC) SAN, FCoE, or iSCSI. A private heartbeat network is required between the nodes.

Figure 5 shows two VMs that run clustering software at the OS level with application monitoring and remediation. The VMs share a private heartbeat and a public network connection backed by shared storage.
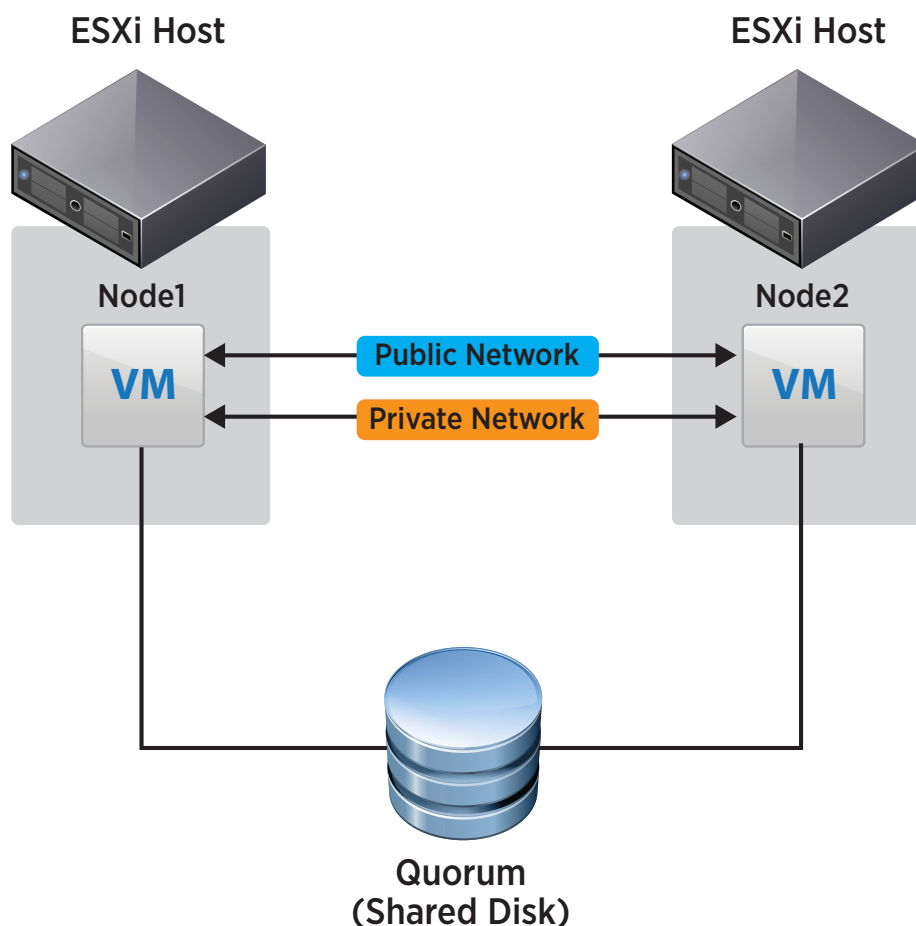
Figure 4. In-Guest Clustering Setup

This solution uses a primary and a standby VM for a particular application being protected. The cluster framework monitors the health of the application resources. If a configured application instance or associated services become unavailable, the cluster services fail over the services to the standby node. Planned downtimes relating to patching the OS are protected by this solution because the application can be failed over to the standby VM during OS patching and the downtime can be minimized.

Windows Server Failover Clustering (WSFC) is one method of guest OS clustering that can be used with the Windows version of vCenter Server 6.0. This method is described in "Appendix A."

Summary: vCenter Server 6.0 Availability Options

vCenter Server can leverage all the solutions for high availability previously discussed. The actual solutions used depend on customer requirements. Table 1 shows the various solutions, their features, and their cost and complexity ratings. The more solutions that are used, the more expensive and complex they are. The choice should be dictated by minimum customer requirements for vCenter Server availability.

HA SOLUTION	VM RESTARTING	APPLICATION MONITORING	FAILOVER TIME	HARDWARE MAINTENANCE DOWNTIME	OS MAINTENANCE DOWNTIME	COST AND COMPLEXITY
vSphere HA + Watchdog (WD)	YES	YES	MED	NONE	HIGH	LOW
vSphere HA + WD + vSphere FT	YES	YES	NONE	NONE	HIGH	MED
vSphere HA + Third-Party Monitoring	YES	YES	MED	NONE	HIGH	MED
Guest Failover Cluster	NO	YES	LOW	MED	LOW	HIGH
vSphere HA + Guest Failover Cluster	YES	YES	LOW	NONE	LOW	HIGH

Table 1. Availability Solutions for vCenter Server and Their Features

High Availability for the Platform Services Controller

Multiple external PSCs can be deployed at a single site serving one or more vCenter Server system. A load balancer is required to front-end the PSC instances. By having more than one PSC instance behind the load balancer, the PSC can be made highly available.
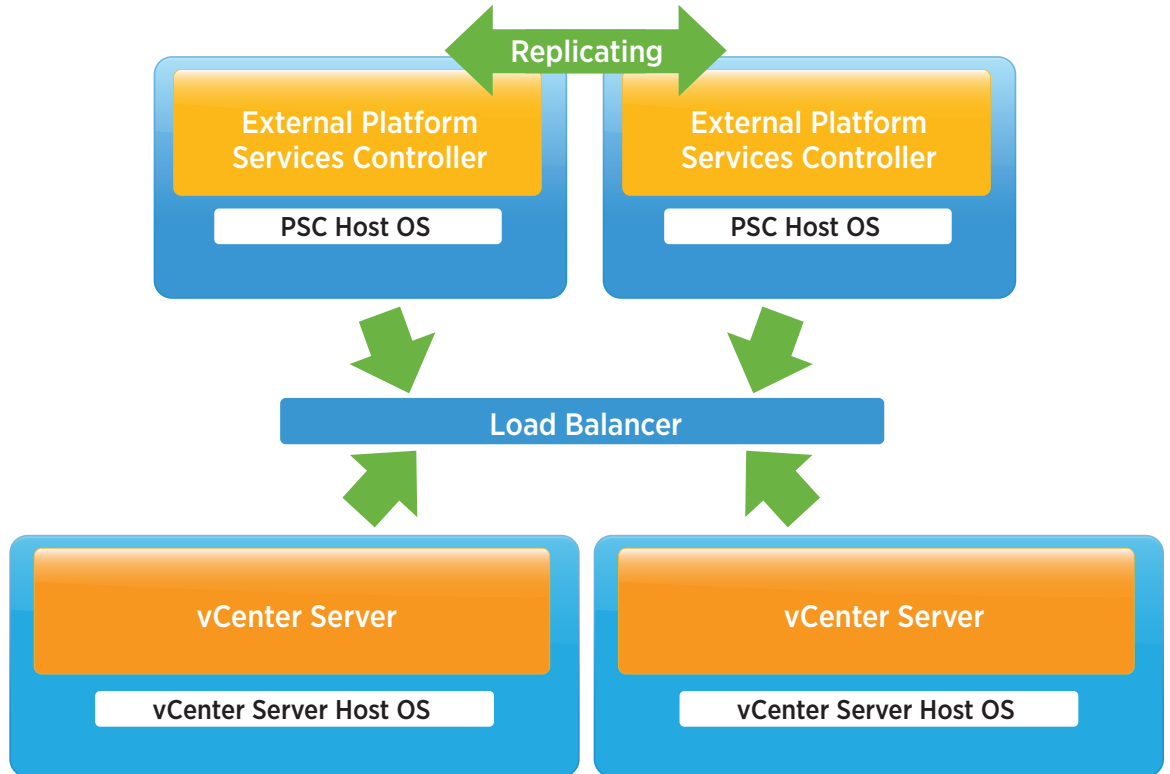
Figure 5. PSC High-Availability Configuration

Deployment Modes for vCenter Server and PSC

Local vCenter Server and PSC High Availability

This model protects the PSC service by having multiple instances of PSCs locally behind a load balancer. Failure of a PSC does not impact the usage of the infrastructure. The PSCs replicate state information between each other and should be separated from one another physically using antiaffinity rules. The clustered vCenter Server systems interact with the PSCs through a load balancer.
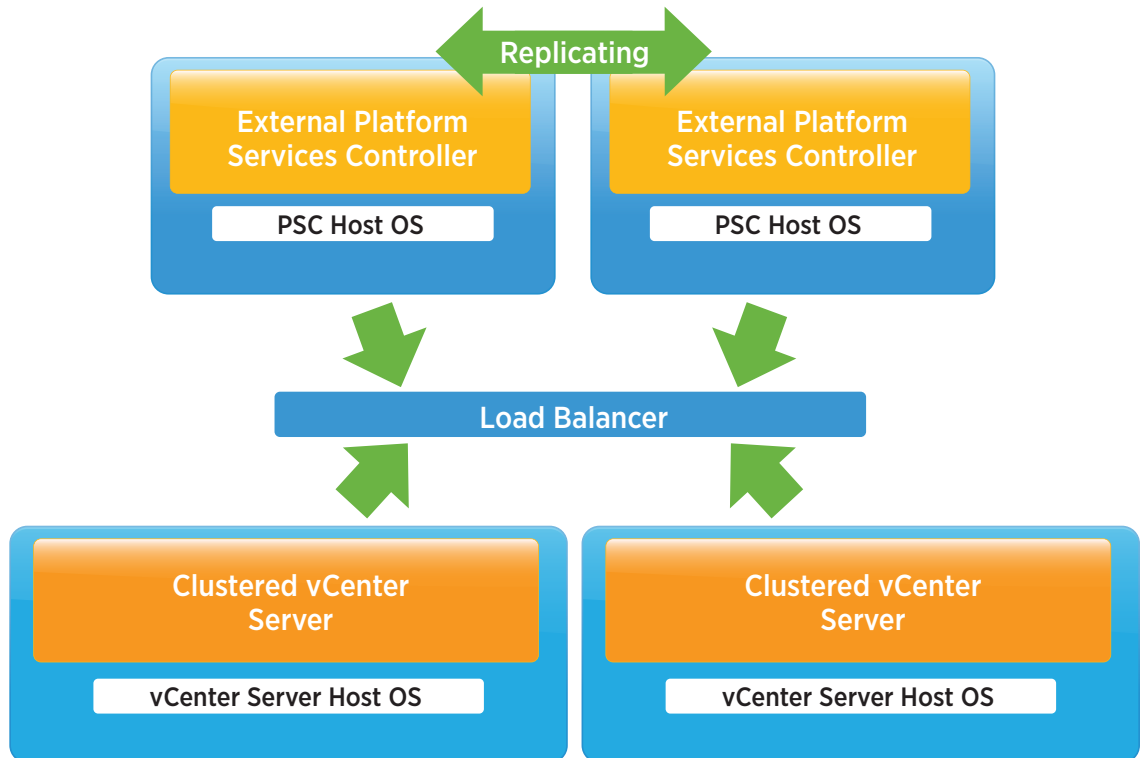
Figure 6. Local vCenter Server and PSC High Availability

Multisite vCenter Server and PSC Basic Architecture

In this configuration, each site is independent, with PSC replication between sites. The vCenter Server system detects site topologies and uses the local PSC under normal circumstances. Users can seamlessly move the vCenter Server systems between PSCs when necessary. This topology enables Enhanced Linked Mode, which is facilitated by the PSC. Enhanced Linked Mode provides for a single point of management for all vCenter Server systems in the same vSphere domain. In vSphere 6.0, the Windows-based and virtual appliance–based vCenter Server systems have the same operational maximums and can belong to the same Linked Mode configuration. The configuration replicates all licenses, global permissions, tags, and roles across all sites.
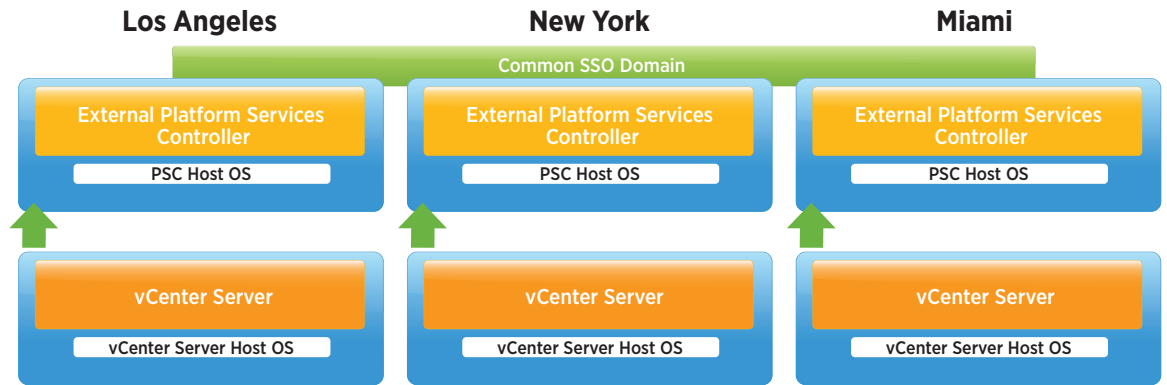
Figure 7. Multisite vCenter Server and PSC Basic Architecture

Multisite vCenter Server and PSC with High Availability

The following factors enable this solution to combine high availability at a local site with a multisite configuration:

• Each site is populated with at least two PSCs.

• vCenter Server nodes are individually clustered.
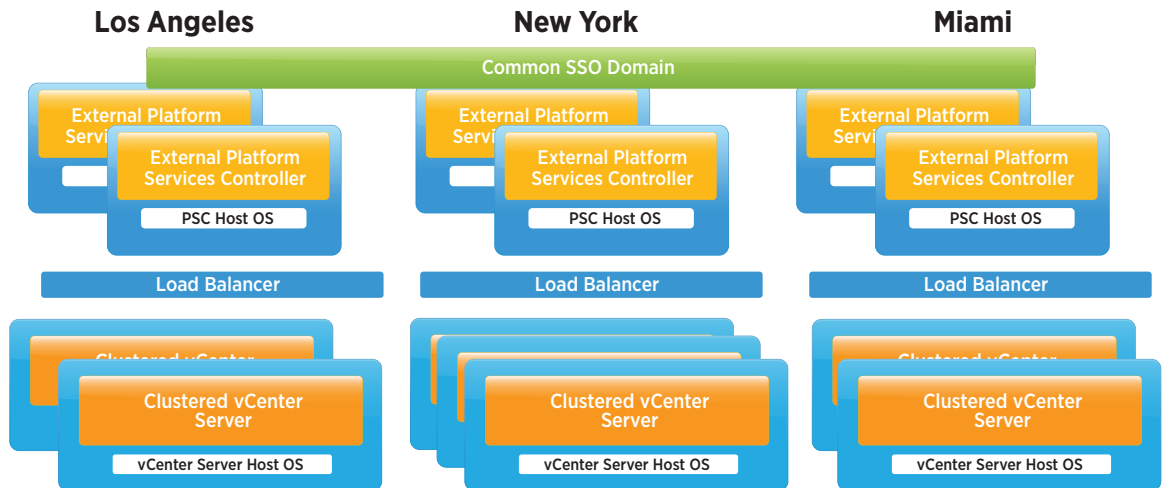
Figure 8. Multisite vCenter Server and PSC High-Availability Architecture

Recovery Options

The availability options discussed protect vCenter Server from minor disasters. Some disasters are caused by other factors such as data loss, corruption, and so on. There are recovery solutions that protect against these types of failures. The following technologies and best practices provide an acceptable level of vCenter Server recoverability, regardless of whether there is an SLA for vCenter Server specifically or vCenter Server is part of a workload SLA.

VMware vSphere Replication

VMware vSphere Replication™ can replicate VMs within a site or across sites to add another layer of protection. To perform a VM recovery, vSphere Replication requires that vCenter Server and vSphere Web Client be online. Even with this requirement, it might still be possible to utilize vSphere Replication to protect vCenter Server. For example, with multiple vCenter Server systems—one that is still available can be used to recover the failed vCenter Server system—this might be in the same physical location or in a different one.

In this scenario, the VM running vCenter Server will be replicated to a remote site designated as the recovery site. vCenter Server can be recovered in the replicated site and be usable as long as its IP address or subnet is available and reachable in the recovery site.

VMware vSphere Data Protection

VMware vSphere Data Protection™ is a backup-and-recovery solution included with all vSphere 6.0 editions. It is deployed as a virtual appliance and is based on industry-leading EMC® Avamar® technology. vSphere Data Protection is an agentless solution that utilizes VM snapshots to back up and restore entire VMs, individual virtual machine disk (VMDK) files, and individual files inside the VM.

vSphere Data Protection is managed using vSphere Web Client. If vCenter Server and the corresponding vSphere Web Client server go offline, Emergency Restore can be used to restore VMs, including those running vCenter Server components. Emergency Restore enables direct-to-host recovery of a VM without the need for vCenter Server and vSphere Web Client. This makes it useful for backing up vCenter Server components when they are running in one or more VM.

vSphere Data Protection utilizes the Windows Volume Shadow Copy Service (VSS) provider built into VMware Tools™. When an image-level backup of a Windows VM is performed, applications for which a VSS writer is installed—such as Microsoft SQL Server and the Windows file system—are quiesced just before the VM snapshot for the backup job is created. This results in backups that are application-level and file-level consistent. Backups of Linux-based VMs are considered crash consistent.

Recommendations for Protecting vCenter Server with vSphere Data Protection

• Run all vCenter Server components in one or more VMs.

• Verify that DNS is properly configured for all vCenter Server VMs, vSphere hosts, and vSphere Data Protection virtual appliances in the environment. Name resolution must be possible using both the fully qualified domain name (FQDN) and the host or short name, as well as by reverse lookups, for each VM.

• Deploy vSphere Data Protection to the same cluster where vCenter Server is located.

• Create an image-level (entire VM) backup job for all VMs that contain and support vCenter Server components. Having a backup job for only the vCenter Server VMs makes it easy to run a manual backup job, in addition to the scheduled job, before patching or updating vCenter Server components and VMs.

• Schedule the backup job to run daily, with a retention policy of at least 10 days.

• Schedule the backup job for when vCenter Server utilization is typically low.

• Configure a vCenter Server alarm to notify administrators when a protected VM is running on a snapshot. See VMware Knowledge Base article 1018029.

• Configure a vCenter Server alarm to notify administrators when a protected VM requires consolidation. See VMware Knowledge Base article 2061896.

• Configure email notification to provide information on the status of the vSphere Data Protection appliance and its backup jobs.

• Routinely perform "practice restores" to verify the integrity of the backups.

Conclusion

There are multiple high-availability options for VMware vCenter Server 6.0. VMware vSphere High Availability, VMware vSphere Fault Tolerance, and watchdog processes can all be leveraged to protect vCenter Server services. Multiple instances of Platform Services Controllers behind a load balancer provide high availability. Windows Server Failover Clustering (WSFC) can be used to further improve availability and protect a vCenter Server environment. Based on customer requirements, multiple deployment modes can be leveraged for availability in local and multisite configurations.

References

1. Disk Quorum and Clustering Requirement

2. vCenter Deployment and System Requirement

3. WSFC in VMware Environments

4. WSFC Failover Knowledge Base

5. vCenter Installation

6. Database Full Recovery Model for vCenter

7. Troubleshooting Guide for vSphere and vCenter 6.0

8. Creating a Windows Server Cluster

9. vCenter Server High Availability with Windows Server Failover Cluster and DFS

10. Estimating Application Availability in ESXi Clusters

11. VMware High Availability Analysis

12. Supported vCenter High Availability Options

Appendix A: Protecting Windows vCenter Server 6.0 with Windows Server Failover Clustering

The following steps provide guidance on setting up Windows Server Failover Clustering (WSFC) to protect the services of the vCenter Server management node.
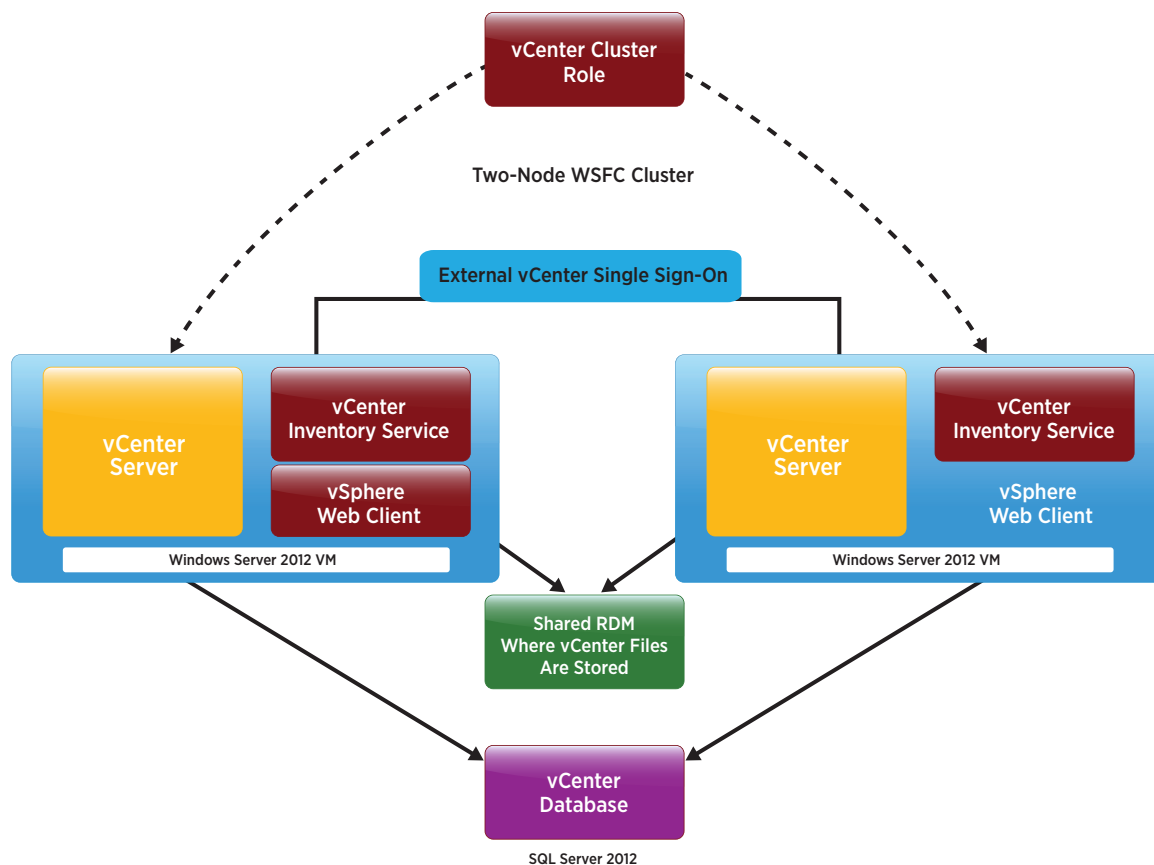

Figure A. vCenter Server Components for Use with Windows Server Failover Clustering

Step-by-Step Guide

1. Create a VM and install Windows GOS.

   Recommended for CISWIN:
   - Windows-2K8-R2-DataCenter
   - Windows-2K12-R2-DataCenter

2. Add two RDM disks to the VM. There is no size requirement for choosing the two RDM disks; generally the disk that acts as a quorum disk is smaller than the one that is used for vCenter Server installation.

   VM > Edit settings > New device > RDM Disk

   Make sure that the RDMs are mounted first.
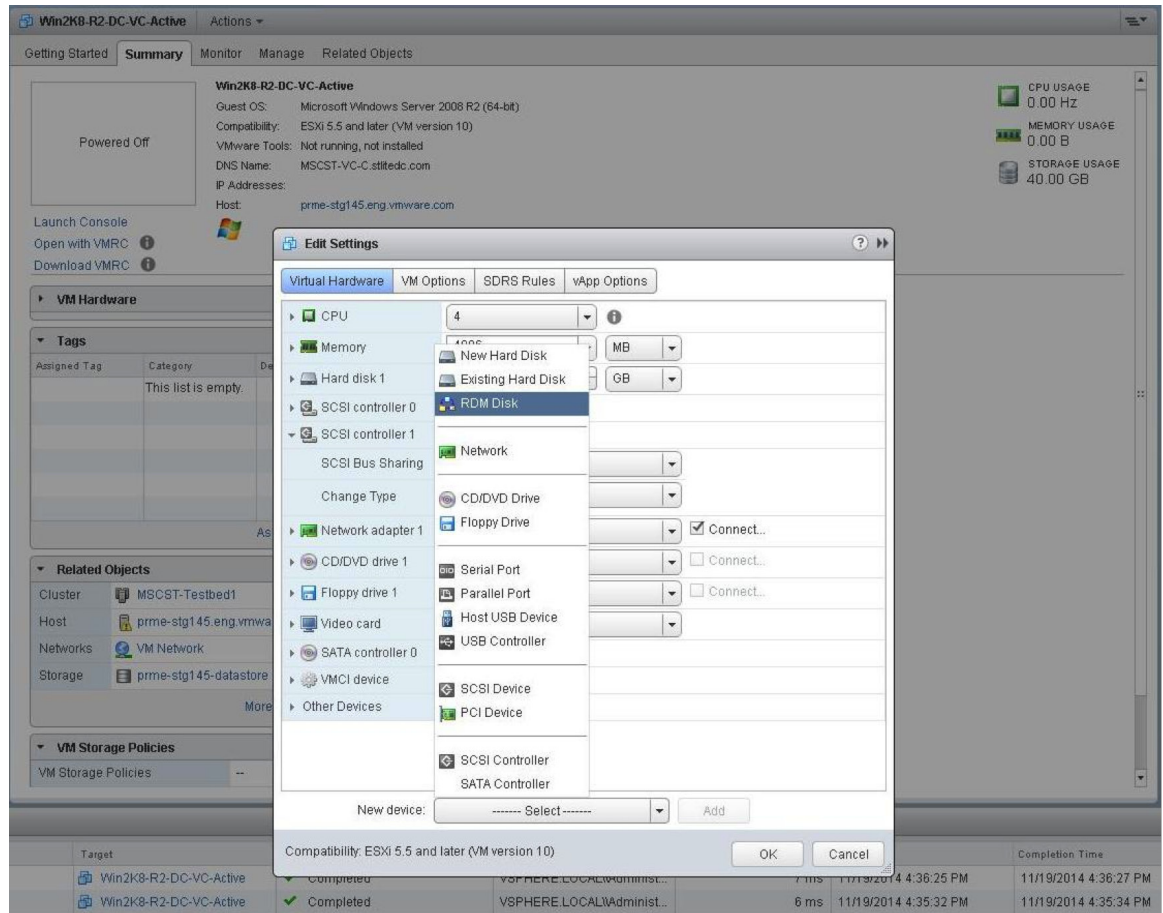
Figure 2.1

While adding the RDM disks, create a separate SCSI controller with the bus sharing option set to physical. Also, ensure that these two disks are independent and persistent (as illustrated in figure 2.2).

VM > Edit settings > New device > RDM Disk

New SCSI control > Bus Sharing option > Physical

Hard disk 1 (RDM 1/Quorum Disk)

• SCSI Controller: New SCSI Controler > SCSI (1:0)

• Compatibility mode: Physical

• Disk Mode: Independent – Persistent

Hard disk 2 (RDM 2/vCenter Server disk)

• SCSI Controller: New SCSI Controler > SCSI (1:1)

• Compatibility mode: Physical

• Disk Mode: Independent – Persistent
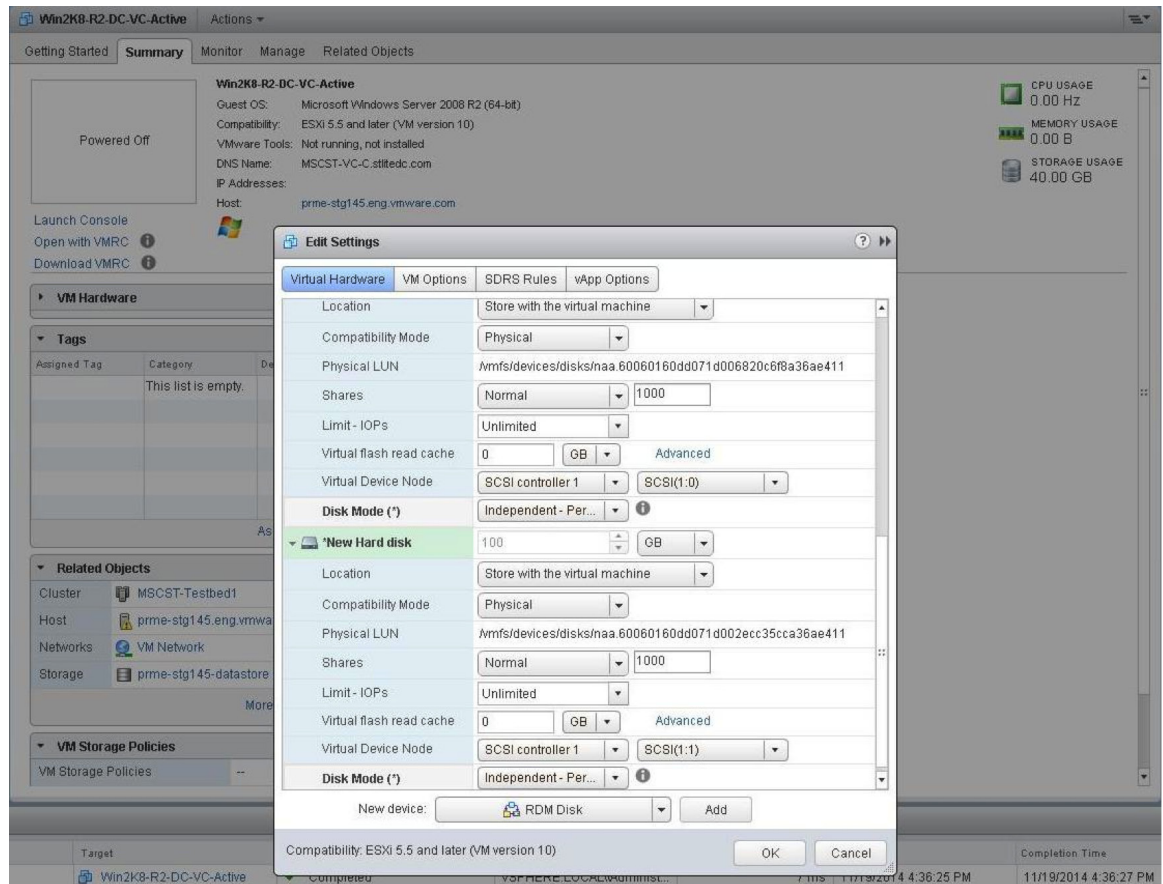
Figure 2.2

3. Power on the VM and format the new disks.

    Windows > Server Manager > Storage > Disk Management

    • Bring both disks online.
    • Convert both disks to MBR, as they are much less than 2 TB. Other disk formats may create problems while validating the cluster. Format the disks and assign drive letters.
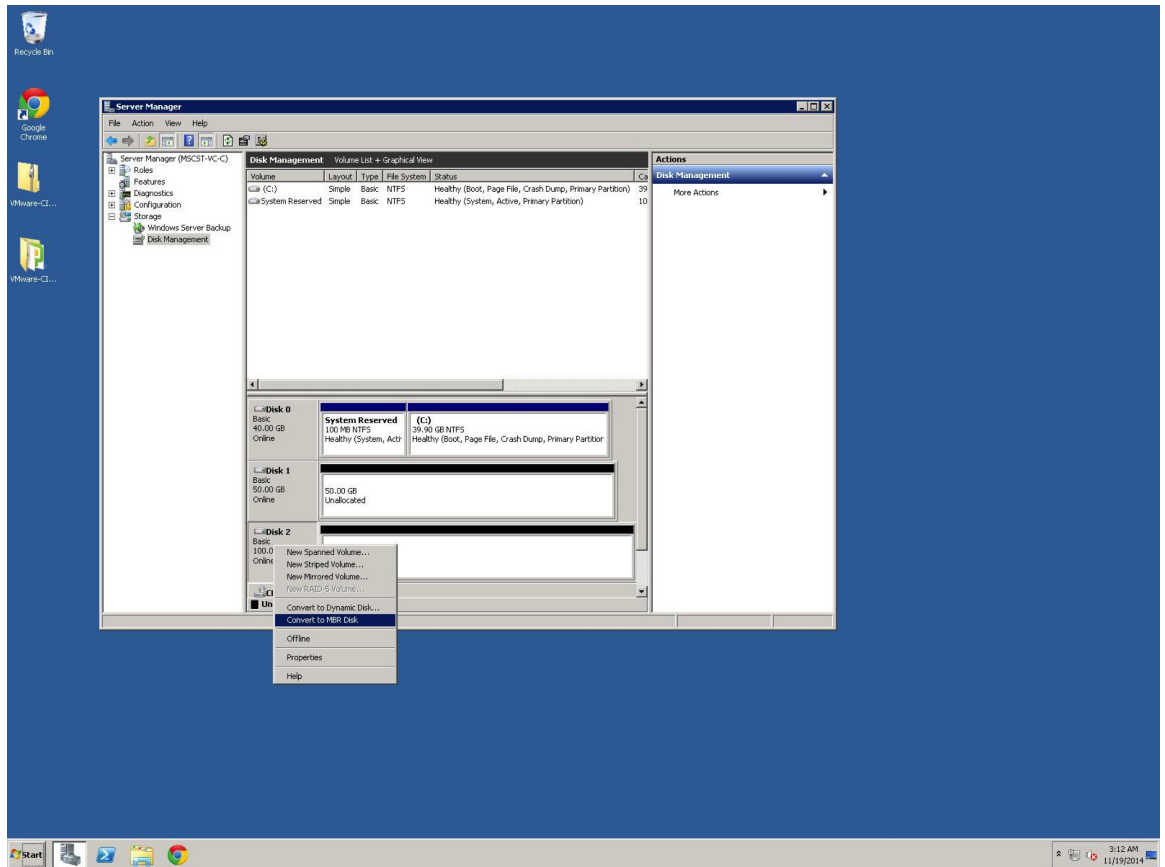
Figure 3.1

4. Install .NET and failover clustering features.

    Windows > Server Manager > Features
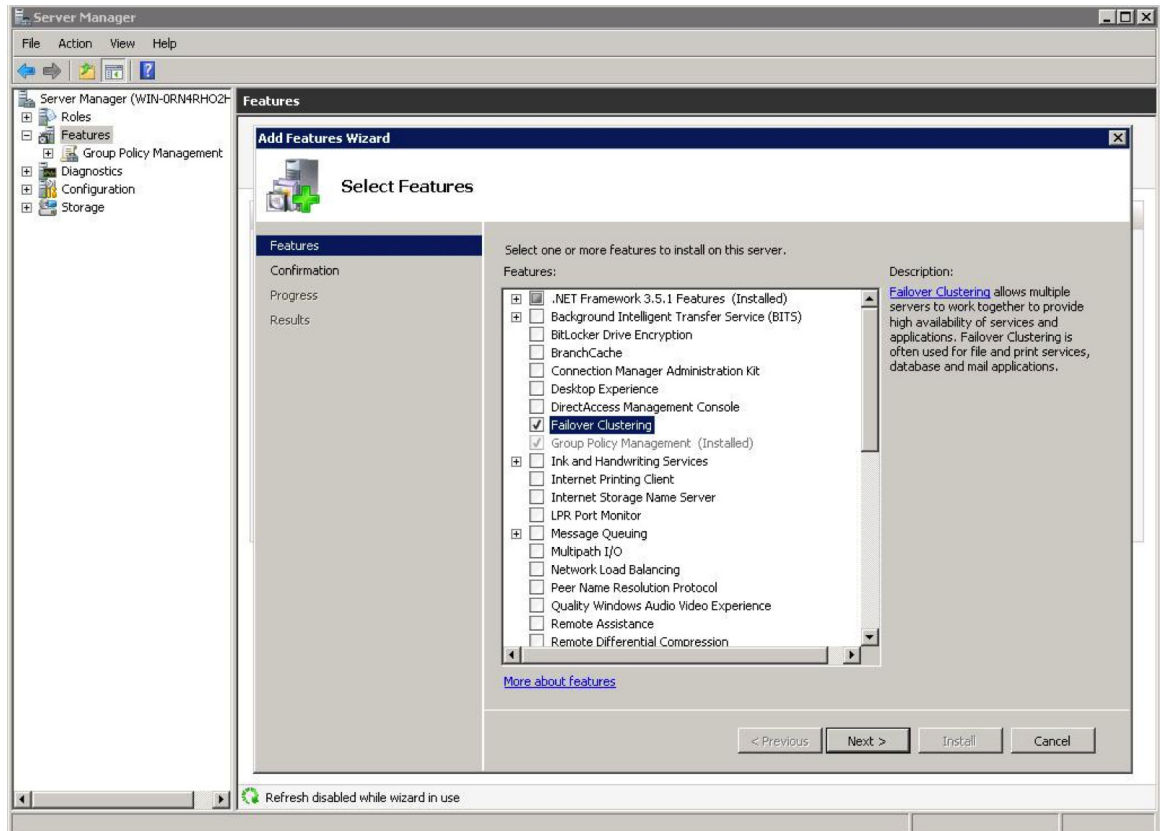
Figure 4.1

5. Install vCenter Server on one of the RDM disks in the guest. Choose the appropriate RDM disk that is more than sufficient for vCenter Server installation, while choosing the custom installation path. *Important: the host name and IP address need to be set to the cluster IP and host name that will be used. This is important to do before installing vCenter Server to make sure the installation uses this information and stores it.*

    vCenter Server Installation > Select the storage location
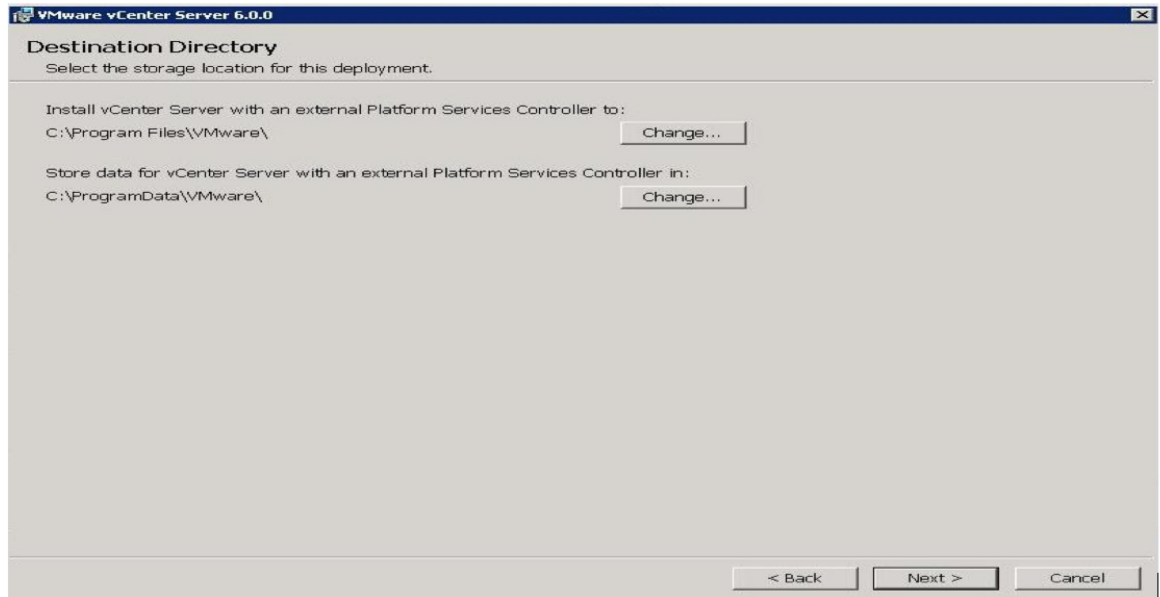
Figure 5.1

6. After successful installation, set all the vCenter Services start options to manual.

   Windows > Control Panel > Services

Figure 6.1

7. Power off the VM.

8. Detach the RDM disks. Do not check the option "delete from disk" while removing the disks from the VM, as illustrated in figure 8.1. The raw mapping disk (vmdk) files should not be deleted from the disk.
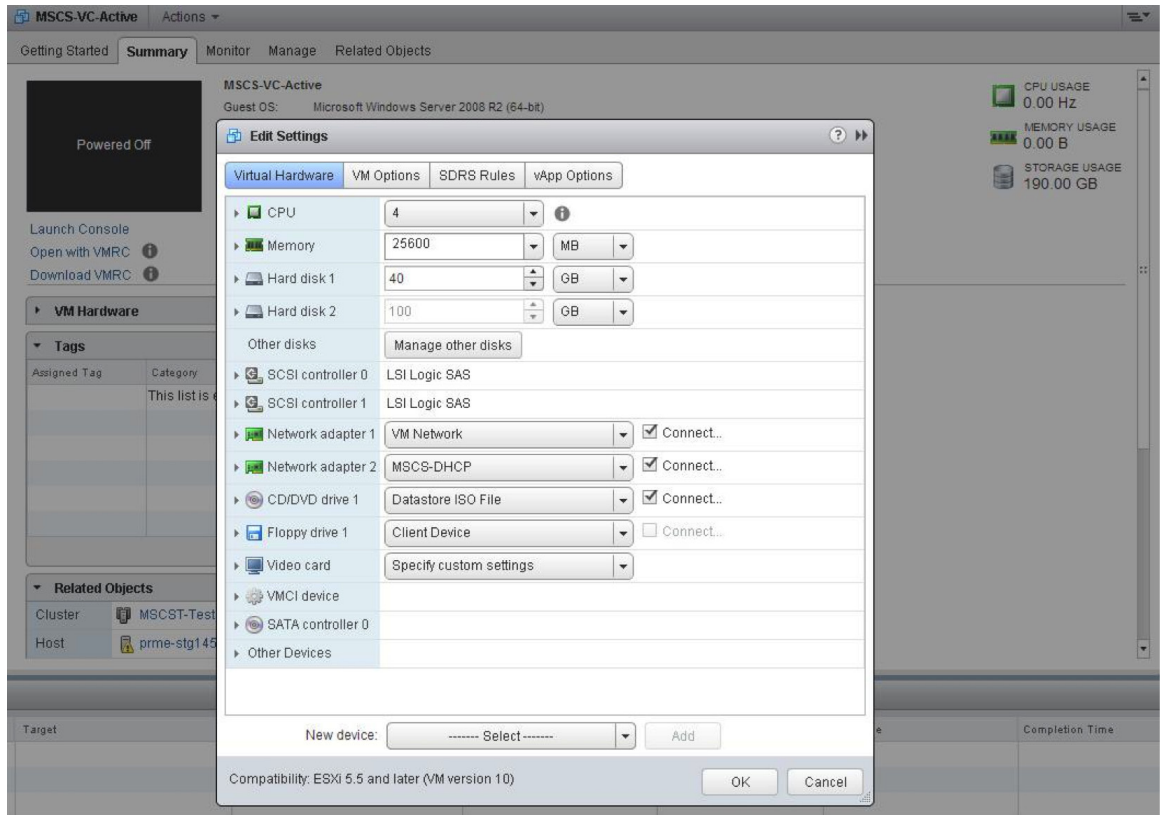

Figure 8.1

9. Clone the VM, and choose "customizing the guest OS" option while cloning to make sure that the clone has a unique identity. This can be done through the default sysprep file, or choose the custom sysprep file while customizing the guest OS during the clone operation.

    Figure 9.1: **New Virtual Machine > Select Clone options > Customize the operating system > New VM Guest Customization Spec**

    Figure 9.2: **New Virtual Machine > Select Clone options > Customize the operating system > New VM Guest Customization Spec > Use Custom Sysprep answer file**
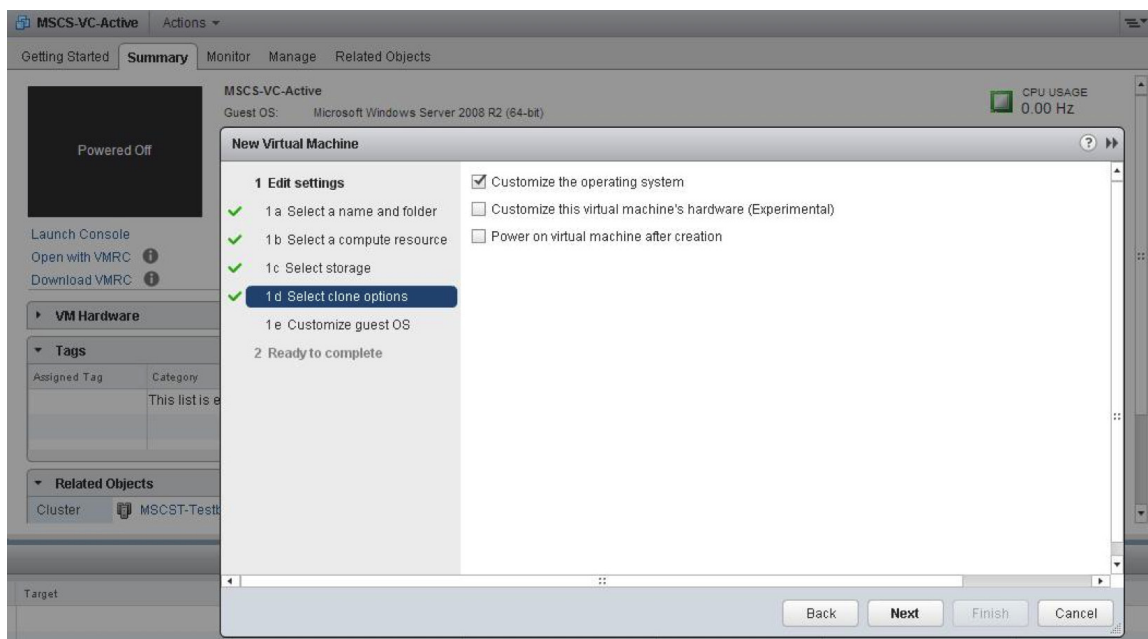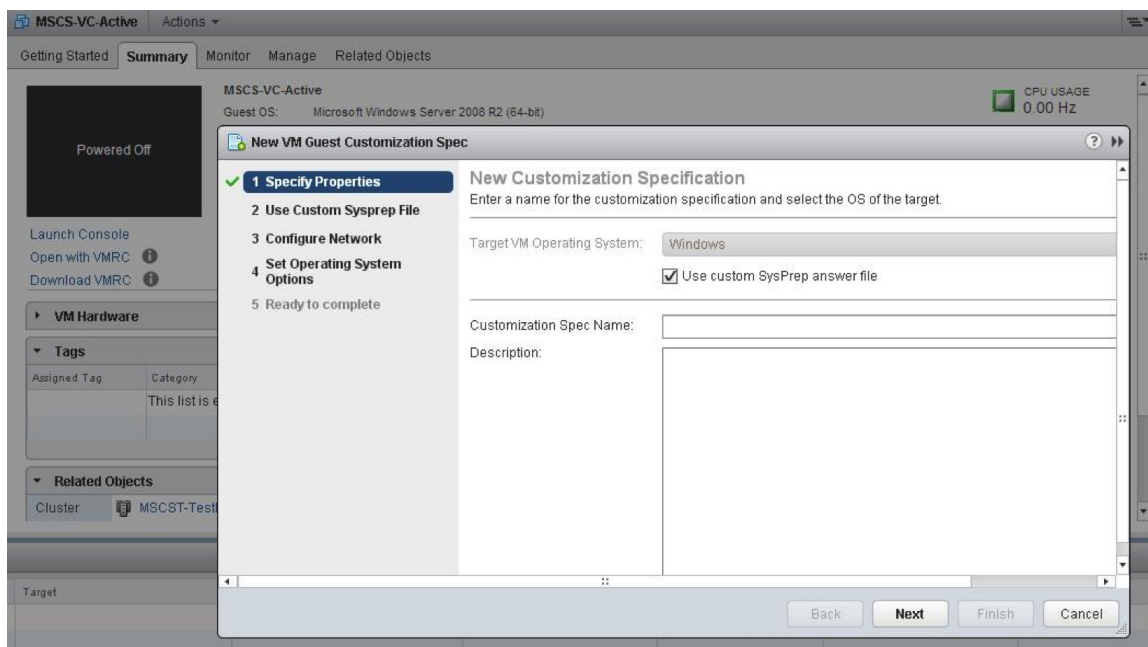
Figure 9.1

Figure 9.2

10. Attach the shared vmdks (RDMs) on both the VMs.

    Add both the shared RDMs again to both the VMs. Note that you are not adding the RDMs this time, as they are already added and mapped to VM-1.

    VM > Edit Settings > New Device > Existing disk > choose the vmdks corresponding to the RDMs that were added to the VM-1 during the VC installation procedure from the VM-1 folder.
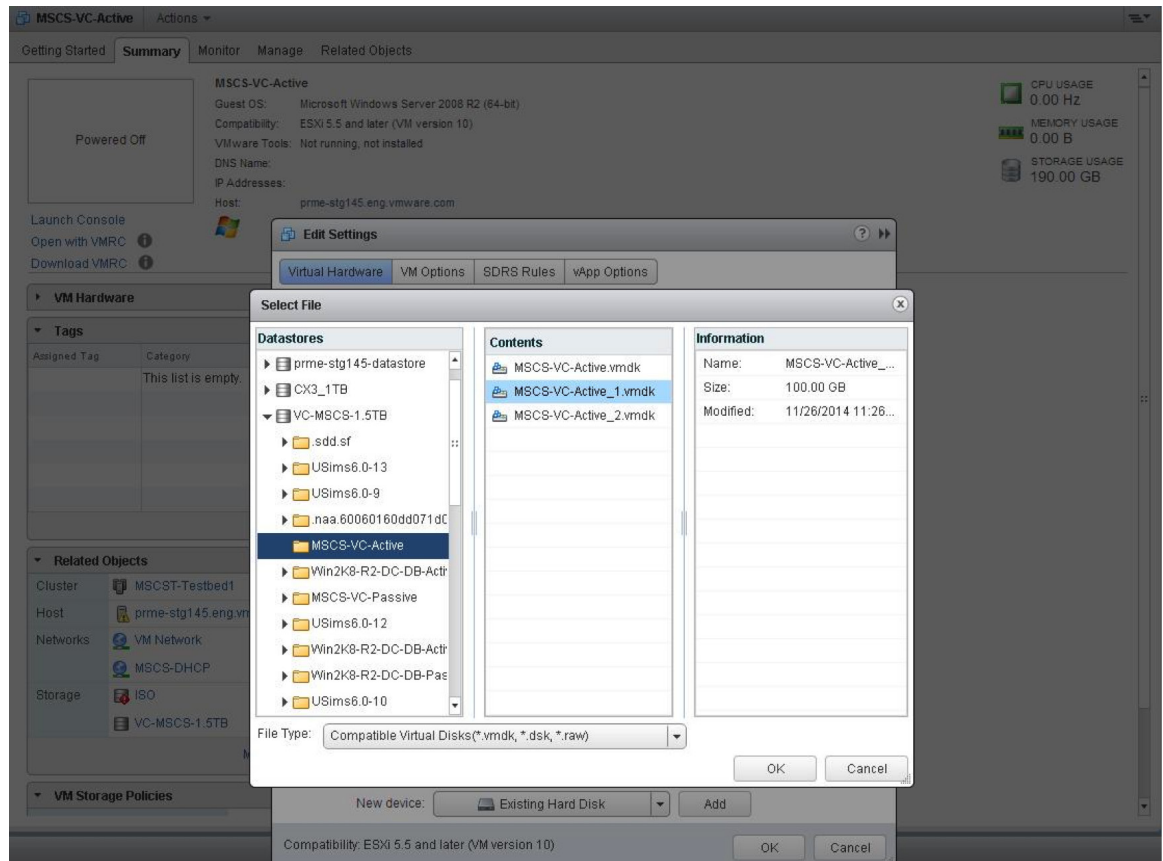
Figure 10.1

11. Power on both the VMs.

12. Change the host name and IP on the first VM. Preserve the original host name and IP for creating the cluster role.

13. Create a MSCS cluster on the first VM by including both nodes to the cluster.

    Windows > Server Manager > Features > Failover Cluster Manager> Create a cluster

    While creating a cluster, follow the cluster creation steps and choose the option to validate the cluster while creating it. This will ensure that the two-node cluster is ready for high availability. Figure 13.1 illustrates choosing the nodes while creating the cluster.
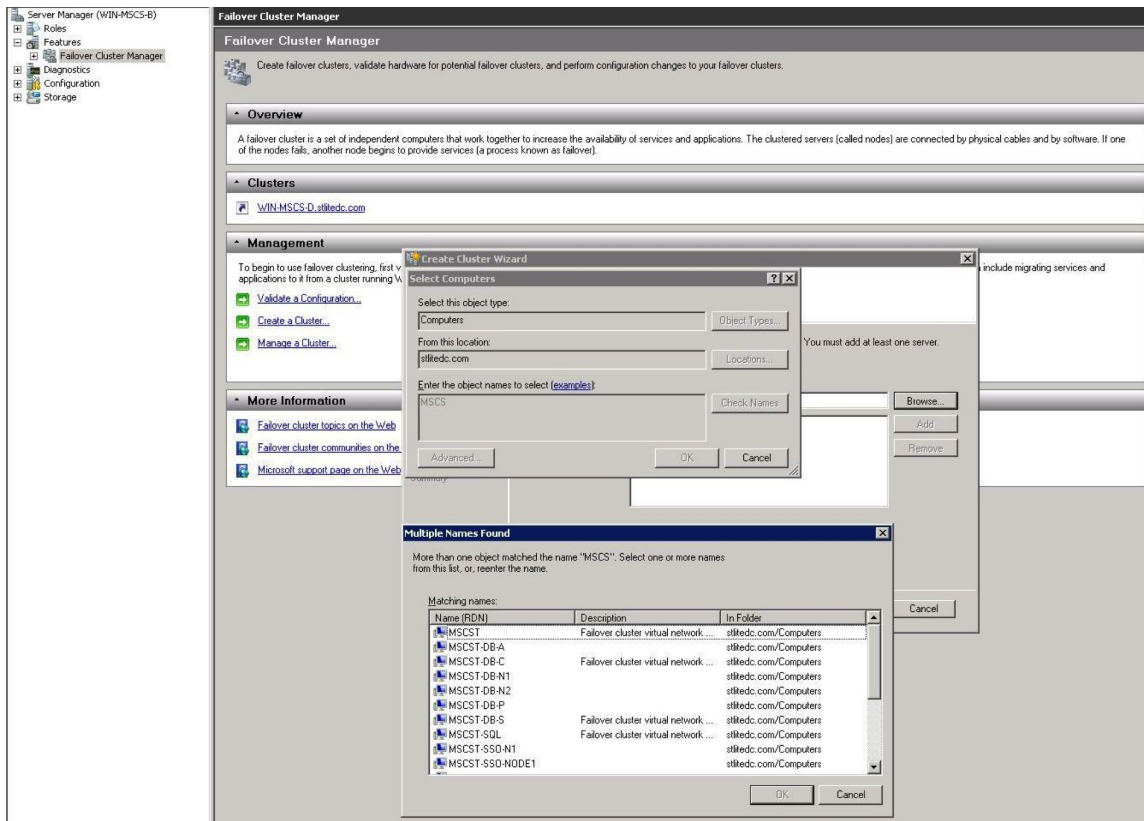
Figure 13.1

14. Create a cluster role by choosing a generic service and "VMware Virtualcenter Server" as role/service.

    While creating the role/service, assign the old IP and host name to this role. The old host name and IPs are the same ones that were preserved during step 10.
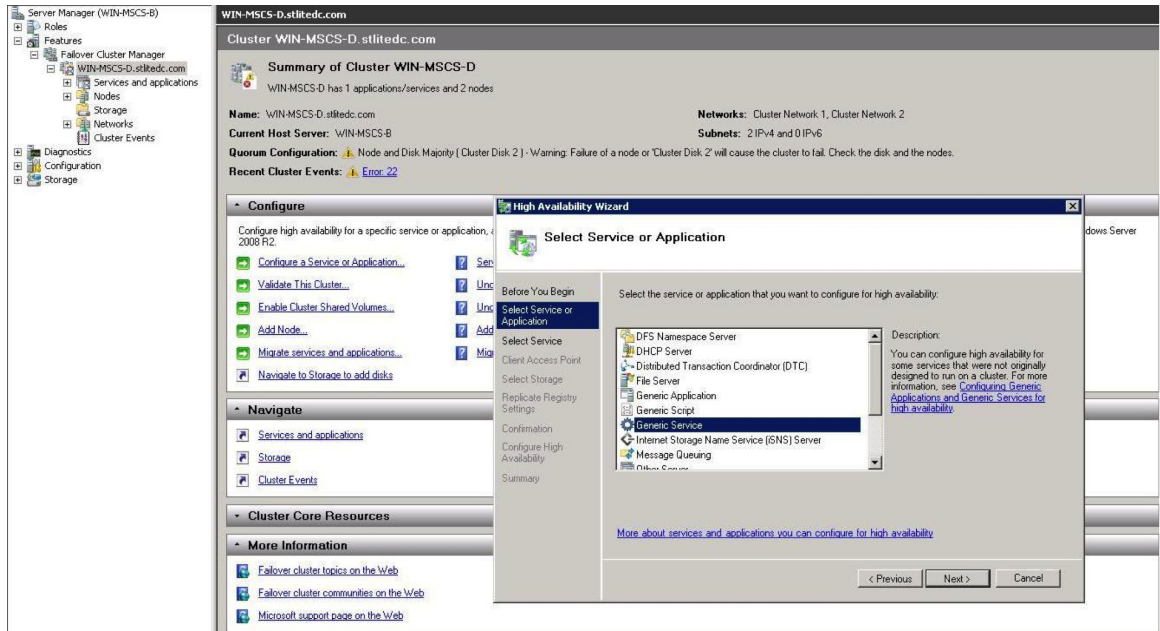
Figure 14.1

15. Add a registry replication after creating the client's access point in the above step:

**Role Creation Wizard > Replicate Registry Settings > Add >
SYSTEM\currentControlSet\Services\VMwareDirectoryService.**

"SYSTEM\currentControlSet\Services\VMwareDirectoryService" from HKLM contains the account details
of VMware directory service. This replication is necessary when changing the vmdir account credentials
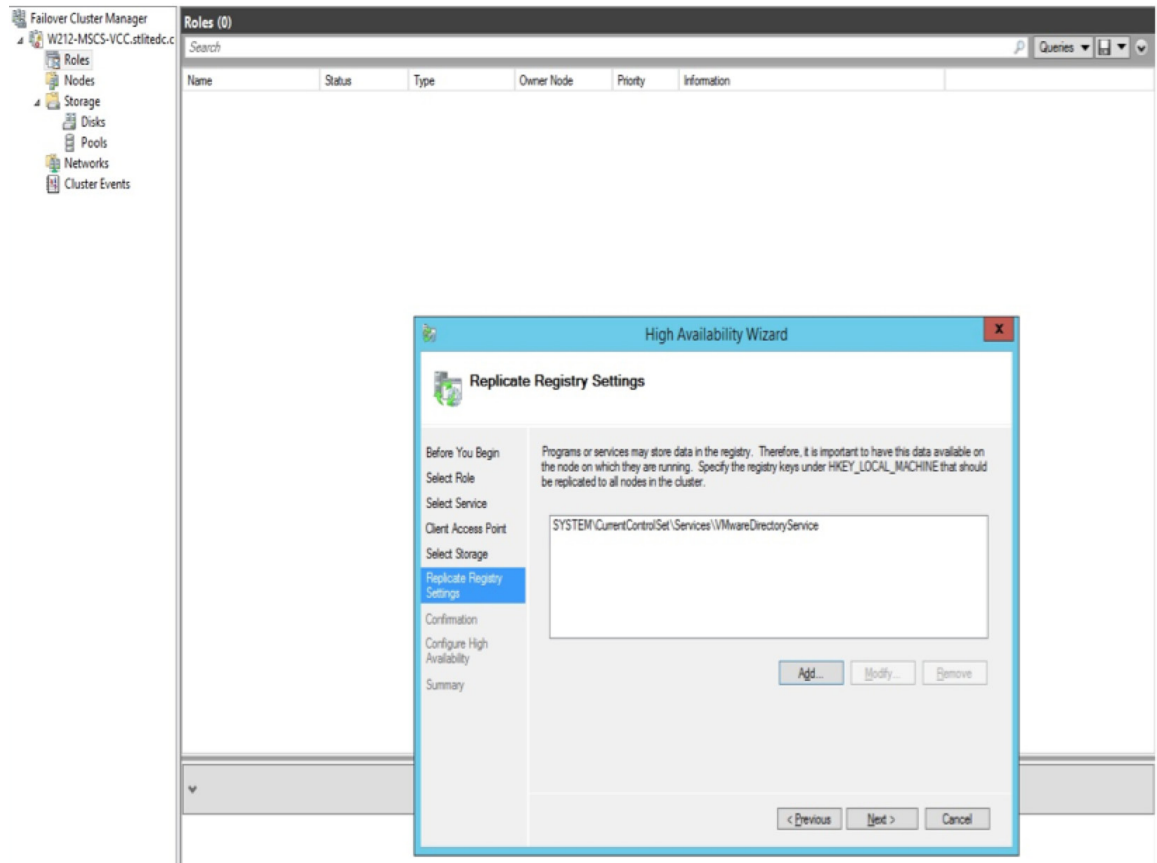of the VC.

Figure 15.1

You can also add the registry replication after the role creation:

Role > Resources tab > Right Click on "VMware Virtualcenter Server" service > Properties > Registry Replication tab > Add > SYSTEM\currentControlSet\Services\VMwareDirectoryService
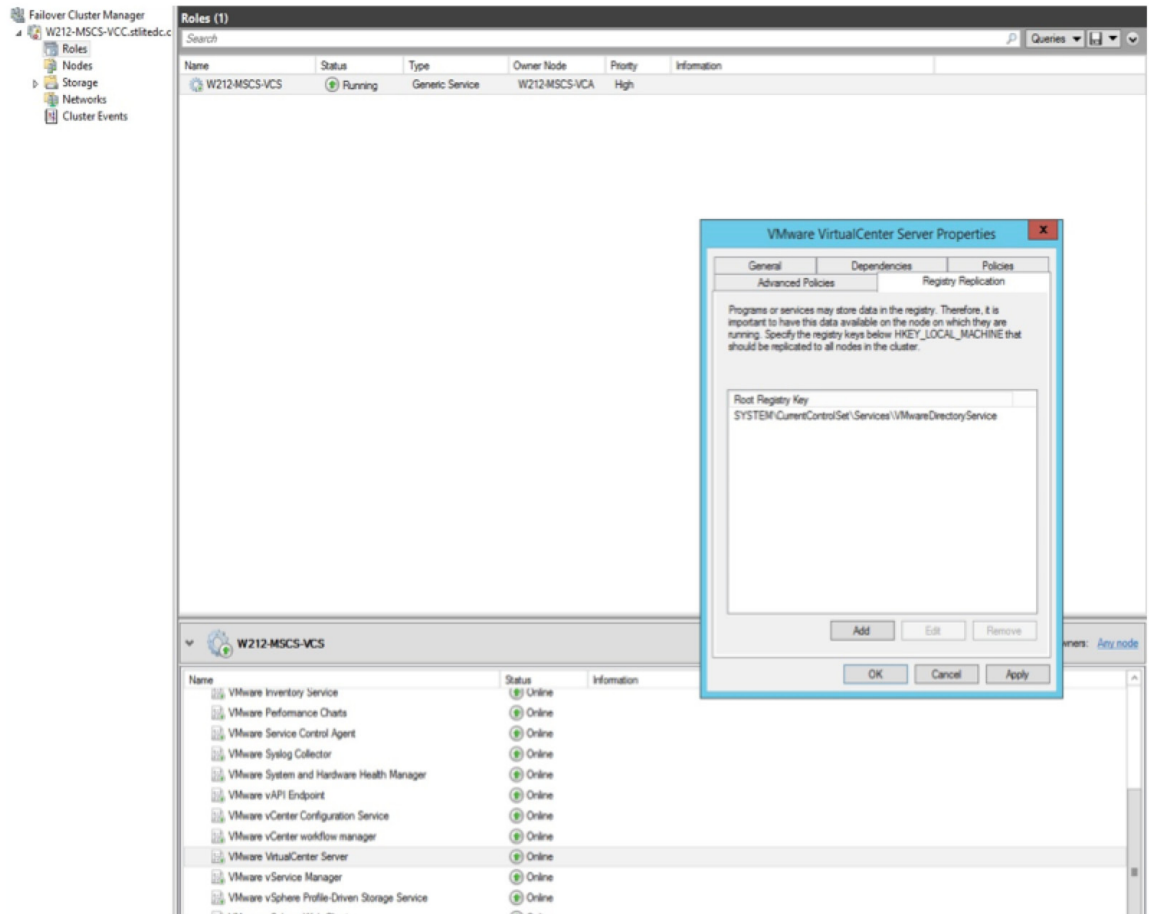
Figure 15.2

16. Add rest of the vCenter services to the created role/service one by one.

    Server Manager > Features > Failover Cluster Manager > New Cluster > Configure a Service or Application > Generic Service > Select rest of the vCenter services one by one to add them to the newly created role.

    Figure 14.1 illustrates adding resources to the role. Figure 14.2 illustrates how it looks after adding all service to the cluster role.
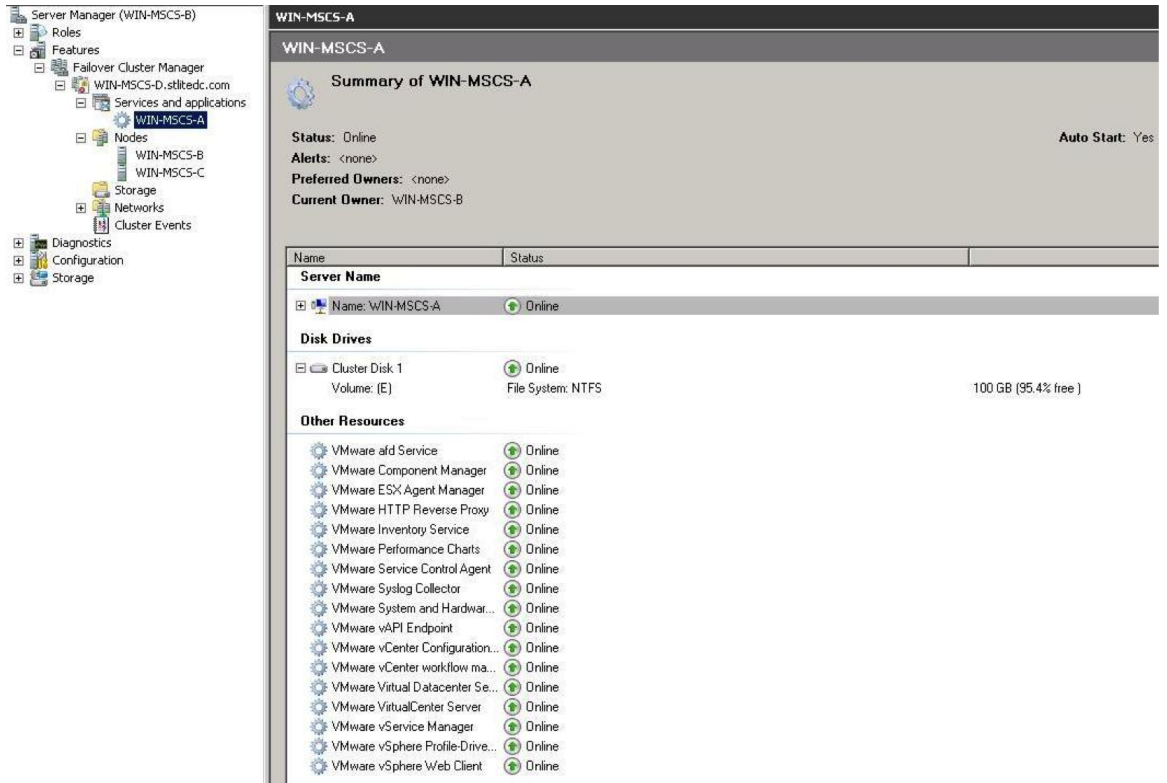
Figure 16.1

17. Verify the failover by powering off the VC VM-1 where the services are currently hosted.

After powering off the VM-1, all the services should be up and running on the VM-2. The expected down time without any load is about 9 minutes.
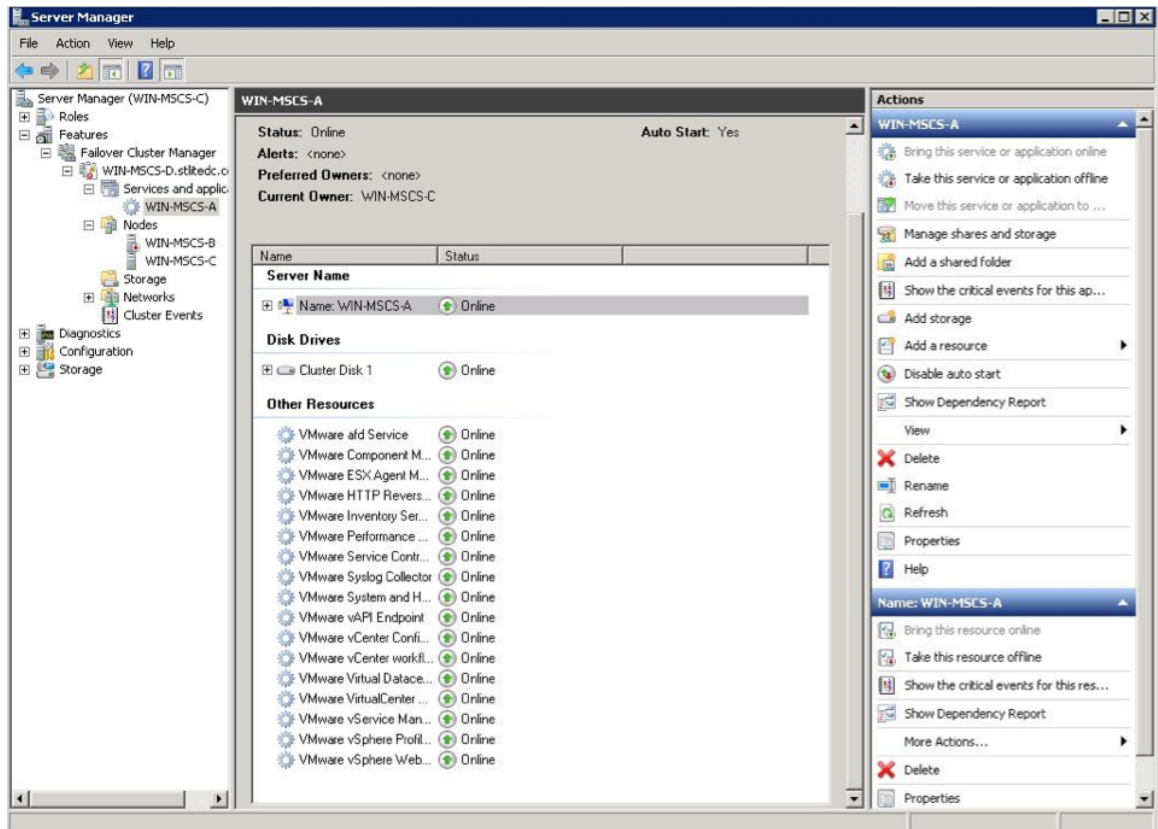
Figure 17.1 illustrates how it looks after the failover.

18. Useful link:

Creating a Microsoft cluster:

http://technet.microsoft.com/en-us/library/cc730647.aspx

Additional Recommendations

1. Network adapter teaming should be used for the public network used by clients to connect with vCenter Server.

2. An antiaffinity rule can be used so both clustered vCenter Server VMs are never on the same host. See VMware Knowledge Base article 1037959.

About the Authors

Mohan Potheri is currently a senior solutions architect at VMware, focusing on VMware vCenter and the virtualization of business-critical applications. He has more than 20 years of experience in IT infrastructure with VMware virtualization, enterprise UNIX, and business-critical applications. Mohan is a CISSP and VMware Certified Design Expert (VCDX#98). He holds master's degrees in electrical engineering and business administration from the University of Houston. Follow @ITVista on Twitter and on the VMware vSphere Blog.

G. Blair Fritz is the global technical lead for VMware vCenter Server and VMware vCenter Single Sign-On™ at VMware. Blair has been working for VMware Support since 2012, focusing on these products. He specializes in improving VMware vSphere features regarding vCenter Single Sign-On and the Platform Services Controller, certificate management, vSphere High Availability, and vCenter Server in general. Blair also has been involved with high-touch, onsite betas with customers around the globe to collect feedback on the next version of vSphere.

Puneet Gupta is currently a staff engineer with VMware, working as a test architect for VMware vSphere virtualization platform system testing. He has more than 12 years of experience in IT with VMware virtualization products, large-scale enterprise deployments, and business-critical applications. Puneet is a VMware Certified Professional (VCP4, VCP3) and EMC Proven Professional (EMCPA). He holds a bachelor's degree in computer science and engineering from the University of Rajasthan in India.

Acknowledgment

The authors would like to thank Duncan Epping and Charu Chaubal for their detailed feedback during the review process.

vmware®