

vSphere Availability

ESXi 6.5
vCenter Server 6.5

This document supports the version of each product listed and supports all subsequent versions until the document is replaced by a new edition. To check for more recent editions of this document, see <http://www.vmware.com/support/pubs>.

EN-002085-01

vmware[®]

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

<http://www.vmware.com/support/>

The VMware Web site also provides the latest product updates.

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

docfeedback@vmware.com

Copyright © 2009–2017 VMware, Inc. All rights reserved. [Copyright and trademark information.](#)

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

Contents

About vSphere Availability	5
Updated Information for vSphere Availability	7
1 Business Continuity and Minimizing Downtime	9
Reducing Planned Downtime	9
Preventing Unplanned Downtime	10
vSphere HA Provides Rapid Recovery from Outages	10
vSphere Fault Tolerance Provides Continuous Availability	11
Protecting the vCenter Server Appliance with vCenter High Availability	12
Protecting vCenter Server with VMware Service Lifecycle Manager	12
2 Creating and Using vSphere HA Clusters	13
How vSphere HA Works	13
vSphere HA Admission Control	21
vSphere HA Interoperability	26
Creating a vSphere HA Cluster	29
Configuring vSphere Availability Settings	31
Best Practices for vSphere HA Clusters	39
3 Providing Fault Tolerance for Virtual Machines	43
How Fault Tolerance Works	43
Fault Tolerance Use Cases	44
Fault Tolerance Requirements, Limits, and Licensing	44
Fault Tolerance Interoperability	45
Preparing Your Cluster and Hosts for Fault Tolerance	47
Using Fault Tolerance	49
Best Practices for Fault Tolerance	53
Legacy Fault Tolerance	55
4 vCenter High Availability	59
Plan the vCenter HA Deployment	60
Configure the Network	65
Configure vCenter HA With the Basic Option	66
Configure vCenter HA With the Advanced Option	67
Manage the vCenter HA Configuration	70
Troubleshoot Your vCenter HA Environment	75
Patching a vCenter High Availability Environment	78

5	Using Microsoft Clustering Service for vCenter Server on Windows High Availability	79
	Benefits and Limitations of Using MSCS	79
	Upgrade vCenter Server in an MSCS Environment	79
	Configure MSCS for High Availability	81
	Index	83

About vSphere Availability

vSphere Availability describes solutions that provide business continuity, including how to establish vSphere[®] High Availability (HA) and vSphere Fault Tolerance.

Intended Audience

This information is for anyone who wants to provide business continuity through the vSphere HA and Fault Tolerance solutions. The information in this book is for experienced Windows or Linux system administrators who are familiar with virtual machine technology and data center operations.

Task instructions in this guide are based on the vSphere Web Client. You can also perform most of the tasks in this guide by using the new vSphere Client. The new vSphere Client user interface terminology, topology, and workflow are closely aligned with the same aspects and elements of the vSphere Web Client user interface. You can apply the vSphere Web Client instructions to the new vSphere Client unless otherwise instructed.

NOTE Not all functionality in the vSphere Web Client has been implemented for the vSphere Client in the vSphere 6.5 release. For an up-to-date list of unsupported functionality, see *Functionality Updates for the vSphere Client Guide* at <http://www.vmware.com/info?id=1413>.

Updated Information for vSphere Availability

vSphere Availability is updated with each release of the product or when necessary.

This table provides the update history of *vSphere Availability*.

Revision	Description
002085-01	<ul style="list-style-type: none">■ Added information on Licensing required for vCenter HA. See, "vCenter HA Hardware and Software Requirements," on page 61.■ Removed requirement to map to FQDN in prerequisite for vCenter HA network. See, "Configure the Network," on page 65.
002085-00	Initial release.

Business Continuity and Minimizing Downtime

1

Downtime, whether planned or unplanned, brings with it considerable costs. However, solutions to ensure higher levels of availability have traditionally been costly, hard to implement, and difficult to manage.

VMware software makes it simpler and less expensive to provide higher levels of availability for important applications. With vSphere, organizations can easily increase the baseline level of availability provided for all applications as well as provide higher levels of availability more easily and cost effectively. With vSphere, you can:

- Provide higher availability independent of hardware, operating system, and applications.
- Reduce planned downtime for common maintenance operations.
- Provide automatic recovery in cases of failure.

vSphere makes it possible to reduce planned downtime, prevent unplanned downtime, and recover rapidly from outages.

This chapter includes the following topics:

- [“Reducing Planned Downtime,”](#) on page 9
- [“Preventing Unplanned Downtime,”](#) on page 10
- [“vSphere HA Provides Rapid Recovery from Outages,”](#) on page 10
- [“vSphere Fault Tolerance Provides Continuous Availability,”](#) on page 11
- [“Protecting the vCenter Server Appliance with vCenter High Availability,”](#) on page 12
- [“Protecting vCenter Server with VMware Service Lifecycle Manager,”](#) on page 12

Reducing Planned Downtime

Planned downtime typically accounts for over 80% of data center downtime. Hardware maintenance, server migration, and firmware updates all require downtime for physical servers. To minimize the impact of this downtime, organizations are forced to delay maintenance until inconvenient and difficult-to-schedule downtime windows.

vSphere makes it possible for organizations to dramatically reduce planned downtime. Because workloads in a vSphere environment can be dynamically moved to different physical servers without downtime or service interruption, server maintenance can be performed without requiring application and service downtime. With vSphere, organizations can:

- Eliminate downtime for common maintenance operations.
- Eliminate planned maintenance windows.
- Perform maintenance at any time without disrupting users and services.

The vSphere vMotion[®] and Storage vMotion functionality in vSphere makes it possible for organizations to reduce planned downtime because workloads in a VMware environment can be dynamically moved to different physical servers or to different underlying storage without service interruption. Administrators can perform faster and completely transparent maintenance operations, without being forced to schedule inconvenient maintenance windows.

Preventing Unplanned Downtime

While an ESXi host provides a robust platform for running applications, an organization must also protect itself from unplanned downtime caused from hardware or application failures. vSphere builds important capabilities into data center infrastructure that can help you prevent unplanned downtime.

These vSphere capabilities are part of virtual infrastructure and are transparent to the operating system and applications running in virtual machines. These features can be configured and utilized by all the virtual machines on a physical system, reducing the cost and complexity of providing higher availability. Key availability capabilities are built into vSphere:

- Shared storage. Eliminate single points of failure by storing virtual machine files on shared storage, such as Fibre Channel or iSCSI SAN, or NAS. The use of SAN mirroring and replication features can be used to keep updated copies of virtual disk at disaster recovery sites.
- Network interface teaming. Provide tolerance of individual network card failures.
- Storage multipathing. Tolerate storage path failures.

In addition to these capabilities, the vSphere HA and Fault Tolerance features can minimize or eliminate unplanned downtime by providing rapid recovery from outages and continuous availability, respectively.

vSphere HA Provides Rapid Recovery from Outages

vSphere HA leverages multiple ESXi hosts configured as a cluster to provide rapid recovery from outages and cost-effective high availability for applications running in virtual machines.

vSphere HA protects application availability in the following ways:

- It protects against a server failure by restarting the virtual machines on other hosts within the cluster.
- It protects against application failure by continuously monitoring a virtual machine and resetting it in the event that a failure is detected.
- It protects against datastore accessibility failures by restarting affected virtual machines on other hosts which still have access to their datastores.
- It protects virtual machines against network isolation by restarting them if their host becomes isolated on the management or Virtual SAN network. This protection is provided even if the network has become partitioned.

Unlike other clustering solutions, vSphere HA provides the infrastructure to protect all workloads with the infrastructure:

- You do not need to install special software within the application or virtual machine. All workloads are protected by vSphere HA. After vSphere HA is configured, no actions are required to protect new virtual machines. They are automatically protected.
- You can combine vSphere HA with vSphere Distributed Resource Scheduler (DRS) to protect against failures and to provide load balancing across the hosts within a cluster.

vSphere HA has several advantages over traditional failover solutions:

Minimal setup	After a vSphere HA cluster is set up, all virtual machines in the cluster get failover support without additional configuration.
Reduced hardware cost and setup	The virtual machine acts as a portable container for the applications and it can be moved among hosts. Administrators avoid duplicate configurations on multiple machines. When you use vSphere HA, you must have sufficient resources to fail over the number of hosts you want to protect with vSphere HA. However, the VMware vCenter Server [®] system automatically manages resources and configures clusters.
Increased application availability	Any application running inside a virtual machine has access to increased availability. Because the virtual machine can recover from hardware failure, all applications that start at boot have increased availability without increased computing needs, even if the application is not itself a clustered application. By monitoring and responding to VMware Tools heartbeats and restarting nonresponsive virtual machines, it protects against guest operating system crashes.
DRS and vMotion integration	If a host fails and virtual machines are restarted on other hosts, DRS can provide migration recommendations or migrate virtual machines for balanced resource allocation. If one or both of the source and destination hosts of a migration fail, vSphere HA can help recover from that failure.

vSphere Fault Tolerance Provides Continuous Availability

vSphere HA provides a base level of protection for your virtual machines by restarting virtual machines in the event of a host failure. vSphere Fault Tolerance provides a higher level of availability, allowing users to protect any virtual machine from a host failure with no loss of data, transactions, or connections.

Fault Tolerance provides continuous availability by ensuring that the states of the Primary and Secondary VMs are identical at any point in the instruction execution of the virtual machine.

If either the host running the Primary VM or the host running the Secondary VM fails, an immediate and transparent failover occurs. The functioning ESXi host seamlessly becomes the Primary VM host without losing network connections or in-progress transactions. With transparent failover, there is no data loss and network connections are maintained. After a transparent failover occurs, a new Secondary VM is respawned and redundancy is re-established. The entire process is transparent and fully automated and occurs even if vCenter Server is unavailable.

Protecting the vCenter Server Appliance with vCenter High Availability

vCenter High Availability (vCenter HA) protects not only against host and hardware failures but also against vCenter Server application failures. Using automated failover from active to passive, vCenter HA supports high availability with minimal downtime.

vCenter HA Deployment Options

vCenter HA protects your vCenter Server Appliance. However, Platform Services Controller provides authentication, certificate management, and licenses for the vCenter Server Appliance. As a result, you have to guarantee high availability of Platform Services Controller. You have these options.

- Deploy an Active node with an embedded Platform Services Controller. As part of the cloning process, the Platform Services Controller and all its services are cloned as well. As part of synchronization from Active node to Passive node, Platform Services Controller on the Passive node is updated.

When failover from the Active node to the Passive node occurs, the Platform Services Controller on the passive node are available and the complete environment is available.

- Deploy at least two Platform Services Controller instances and place them behind a load balancer.

When failover from the Active node to the Passive node occurs, the Passive node continues to point to the load balancer. When one of the Platform Services Controller instances becomes unavailable, the load balancer directs requests to the second Platform Services Controller instance.

See [“vCenter HA Deployment Options,”](#) on page 62.

vCenter HA Configuration Options

You configure vCenter HA from the vSphere Web Client. The configuration wizard provides these options.

Option	Description
Basic	<p>The Basic option clones the Active node to the Passive node and witness node, and configures the nodes for you.</p> <p>You can use this option if your environment meets one of the following requirements.</p> <ul style="list-style-type: none"> ■ Either the vCenter Server Appliance that will become the Active node is managing its own ESXi host and its own virtual machine. This configuration is sometimes called a self-managed vCenter Server. ■ Or the vCenter Server Appliance is managed by another vCenter Server (management vCenter Server) and both vCenter Server instances are in the same vCenter Single Sign-On domain. That means they both use an external Platform Services Controller and both are running vSphere 6.5. <p>See “Configure vCenter HA With the Basic Option,” on page 66.</p>
Advanced	<p>The Advanced option offers more flexibility. You can use this option as long as your environment meets hardware and software requirements.</p> <p>If you select this option, you are responsible for cloning the Active node to the Passive node and the Witness node. You must also perform some networking configuration.</p> <p>See “Configure vCenter HA With the Advanced Option,” on page 67.</p>

Protecting vCenter Server with VMware Service Lifecycle Manager

Availability of vCenter Server is provided by VMware Service Lifecycle Manager.

If a vCenter service fails, VMware Service Lifecycle Manager restarts it. VMware Service Lifecycle Manager monitors the health of services and it takes preconfigured remediation action when it detects a failure. Service does not restart if multiple attempts to remediate fail.

Creating and Using vSphere HA Clusters

2

vSphere HA clusters enable a collection of ESXi hosts to work together so that, as a group, they provide higher levels of availability for virtual machines than each ESXi host can provide individually. When you plan the creation and usage of a new vSphere HA cluster, the options you select affect the way that cluster responds to failures of hosts or virtual machines.

Before you create a vSphere HA cluster, you should know how vSphere HA identifies host failures and isolation and how it responds to these situations. You also should know how admission control works so that you can choose the policy that fits your failover needs. After you establish a cluster, you can customize its behavior with advanced options and optimize its performance by following recommended best practices.

NOTE You might get an error message when you try to use vSphere HA. For information about error messages related to vSphere HA, see the VMware knowledge base article at <http://kb.vmware.com/kb/1033634>.

This chapter includes the following topics:

- “How vSphere HA Works,” on page 13
- “vSphere HA Admission Control,” on page 21
- “vSphere HA Interoperability,” on page 26
- “Creating a vSphere HA Cluster,” on page 29
- “Configuring vSphere Availability Settings,” on page 31
- “Best Practices for vSphere HA Clusters,” on page 39

How vSphere HA Works

vSphere HA provides high availability for virtual machines by pooling the virtual machines and the hosts they reside on into a cluster. Hosts in the cluster are monitored and in the event of a failure, the virtual machines on a failed host are restarted on alternate hosts.

When you create a vSphere HA cluster, a single host is automatically elected as the master host. The master host communicates with vCenter Server and monitors the state of all protected virtual machines and of the slave hosts. Different types of host failures are possible, and the master host must detect and appropriately deal with the failure. The master host must distinguish between a failed host and one that is in a network partition or that has become network isolated. The master host uses network and datastore heartbeating to determine the type of failure.



Sphere HA Clusters (<http://link.brightcove.com/services/player/bcpid2296383276001?bctid=ref:vSphereHAClusters>)

Master and Slave Hosts

When you add a host to a vSphere HA cluster, an agent is uploaded to the host and configured to communicate with other agents in the cluster. Each host in the cluster functions as a master host or a slave host.

When vSphere HA is enabled for a cluster, all active hosts (those not in standby or maintenance mode, or not disconnected) participate in an election to choose the cluster's master host. The host that mounts the greatest number of datastores has an advantage in the election. Only one master host typically exists per cluster and all other hosts are slave hosts. If the master host fails, is shut down or put in standby mode, or is removed from the cluster a new election is held.

The master host in a cluster has a number of responsibilities:

- Monitoring the state of slave hosts. If a slave host fails or becomes unreachable, the master host identifies which virtual machines need to be restarted.
- Monitoring the power state of all protected virtual machines. If one virtual machine fails, the master host ensures that it is restarted. Using a local placement engine, the master host also determines where the restart should be done.
- Managing the lists of cluster hosts and protected virtual machines.
- Acting as vCenter Server management interface to the cluster and reporting the cluster health state.

The slave hosts primarily contribute to the cluster by running virtual machines locally, monitoring their runtime states, and reporting state updates to the master host. A master host can also run and monitor virtual machines. Both slave hosts and master hosts implement the VM and Application Monitoring features.

One of the functions performed by the master host is to orchestrate restarts of protected virtual machines. A virtual machine is protected by a master host after vCenter Server observes that the virtual machine's power state has changed from powered off to powered on in response to a user action. The master host persists the list of protected virtual machines in the cluster's datastores. A newly elected master host uses this information to determine which virtual machines to protect.

NOTE If you disconnect a host from a cluster, all of the virtual machines registered to that host are unprotected by vSphere HA.

Host Failure Types

The master host of a vSphere HA cluster is responsible for detecting the failure of slave hosts. Depending on the type of failure detected, the virtual machines running on the hosts might need to be failed over.

In a vSphere HA cluster, three types of host failure are detected:

- Failure. A host stops functioning.
- Isolation. A host becomes network isolated.
- Partition. A host loses network connectivity with the master host.

The master host monitors the liveness of the slave hosts in the cluster. This communication happens through the exchange of network heartbeats every second. When the master host stops receiving these heartbeats from a slave host, it checks for host liveness before declaring the host failed. The liveness check that the master host performs is to determine whether the slave host is exchanging heartbeats with one of the datastores. See [“Datastore Heartbeating,”](#) on page 19. Also, the master host checks whether the host responds to ICMP pings sent to its management IP addresses.

If a master host cannot communicate directly with the agent on a slave host, the slave host does not respond to ICMP pings. If the agent is not issuing heartbeats, it is viewed as failed. The host's virtual machines are restarted on alternate hosts. If such a slave host is exchanging heartbeats with a datastore, the master host assumes that the slave host is in a network partition or is network isolated.. So, the master host continues to monitor the host and its virtual machines. See [“Network Partitions,”](#) on page 19.

Host network isolation occurs when a host is still running, but it can no longer observe traffic from vSphere HA agents on the management network. If a host stops observing this traffic, it attempts to ping the cluster isolation addresses. If this pinging also fails, the host declares that it is isolated from the network.

The master host monitors the virtual machines that are running on an isolated host. If the master host observes that the VMs power off, and the master host is responsible for the VMs, it restarts them.

NOTE If you ensure that the network infrastructure is sufficiently redundant and that at least one network path is available at all times, host network isolation is less likely to occur.

Proactive HA Failures

A Proactive HA failure occurs when a host component fails, which results in a loss of redundancy or a noncatastrophic failure. However, the functional behavior of the VMs residing on the host is not yet affected. For example, if a power supply on the host fails, but other power supplies are available, that is a Proactive HA failure.

If a Proactive HA failure occurs, you can automate the remediation action taken in the vSphere Availability section of the vSphere Web Client. The VMs on the affected host can be evacuated to other hosts and the host is either placed in Quarantine mode or Maintenance mode.

NOTE Your cluster must be using vSphere DRS for the Proactive HA failure monitoring to work.

Determining Responses to Host Issues

If a host fails and its virtual machines must be restarted, you can control the order in which the virtual machines are restarted with the VM restart priority setting. You can also configure how vSphere HA responds if hosts lose management network connectivity with other hosts by using the host isolation response setting. Other factors are also considered when vSphere HA restarts a virtual machine after a failure.

The following settings apply to all virtual machines in the cluster in the case of a host failure or isolation. You can also configure exceptions for specific virtual machines. See [“Customize an Individual Virtual Machine,”](#) on page 39.

Host Isolation Response

Host isolation response determines what happens when a host in a vSphere HA cluster loses its management network connections, but continues to run. You can use the isolation response to have vSphere HA power off virtual machines that are running on an isolated host and restart them on a non-isolated host. Host isolation responses require that Host Monitoring Status is enabled. If Host Monitoring Status is disabled, host isolation responses are also suspended. A host determines that it is isolated when it is unable to communicate with the agents running on the other hosts, and it is unable to ping its isolation addresses. The host then executes its isolation response. The responses are Power off and restart VMs or Shutdown and restart VMs. You can customize this property for individual virtual machines.

NOTE If a virtual machine has a restart priority setting of Disabled, no host isolation response is made.

To use the Shutdown and restart VMs setting, you must install VMware Tools in the guest operating system of the virtual machine. Shutting down the virtual machine provides the advantage of preserving its state. Shutting down is better than powering off the virtual machine, which does not flush most recent changes to disk or commit transactions. Virtual machines that are in the process of shutting down take longer to fail over while the shutdown completes. Virtual Machines that have not shut down in 300 seconds, or the time specified in the advanced option `das.isolationshutdowntimeout`, are powered off.

After you create a vSphere HA cluster, you can override the default cluster settings for Restart Priority and Isolation Response for specific virtual machines. Such overrides are useful for virtual machines that are used for special tasks. For example, virtual machines that provide infrastructure services like DNS or DHCP might need to be powered on before other virtual machines in the cluster.

A virtual machine "split-brain" condition can occur when a host becomes isolated or partitioned from a master host and the master host cannot communicate with it using heartbeat datastores. In this situation, the master host cannot determine that the host is alive and so declares it dead. The master host then attempts to restart the virtual machines that are running on the isolated or partitioned host. This attempt succeeds if the virtual machines remain running on the isolated/partitioned host and that host lost access to the virtual machines' datastores when it became isolated or partitioned. A split-brain condition then exists because there are two instances of the virtual machine. However, only one instance is able to read or write the virtual machine's virtual disks. VM Component Protection can be used to prevent this split-brain condition. When you enable VMCP with the aggressive setting, it monitors the datastore accessibility of powered-on virtual machines, and shuts down those that lose access to their datastores.

To recover from this situation, ESXi generates a question on the virtual machine that has lost the disk locks for when the host comes out of isolation and cannot reacquire the disk locks. vSphere HA automatically answers this question, allowing the virtual machine instance that has lost the disk locks to power off, leaving just the instance that has the disk locks.

Virtual Machine Dependencies

You can create dependencies between groups of virtual machines. To do so, you must first create the VM groups in the vSphere Web Client by going to the **Configure** tab for the cluster and selecting **VM/Host Groups**. Once the groups have been created, you can create restart dependency rules between the groups by browsing to **VM/Host Rules** and in the Type drop-down menu, select **Virtual Machines to Virtual Machines**. These rules can specify that certain VM groups cannot be restarted until other, specified VM groups have been Ready first.

Factors Considered for Virtual Machine Restarts

After a failure, the cluster's master host attempts to restart affected virtual machines by identifying a host that can power them on. When choosing such a host, the master host considers a number of factors.

File accessibility	Before a virtual machine can be started, its files must be accessible from one of the active cluster hosts that the master can communicate with over the network
Virtual machine and host compatibility	If there are accessible hosts, the virtual machine must be compatible with at least one of them. The compatibility set for a virtual machine includes the effect of any required VM-Host affinity rules. For example, if a rule only permits a virtual machine to run on two hosts, it is considered for placement on those two hosts.
Resource reservations	Of the hosts that the virtual machine can run on, at least one must have sufficient unreserved capacity to meet the memory overhead of the virtual machine and any resource reservations. Four types of reservations are considered: CPU, Memory, vNIC, and Virtual flash. Also, sufficient network ports must be available to power on the virtual machine.

Host limits	In addition to resource reservations, a virtual machine can only be placed on a host if doing so does not violate the maximum number of allowed virtual machines or the number of in-use vCPUs.
Feature constraints	If the advanced option has been set that requires vSphere HA to enforce VM to VM anti-affinity rules, vSphere HA does not violate this rule. Also, vSphere HA does not violate any configured per host limits for fault tolerant virtual machines.

If no hosts satisfy the preceding considerations, the master host issues an event stating that there are not enough resources for vSphere HA to start the VM and tries again when the cluster conditions have changed. For example, if the virtual machine is not accessible, the master host tries again after a change in file accessibility.

VM and Application Monitoring

VM Monitoring restarts individual virtual machines if their VMware Tools heartbeats are not received within a set time. Similarly, Application Monitoring can restart a virtual machine if the heartbeats for an application it is running are not received. You can enable these features and configure the sensitivity with which vSphere HA monitors non-responsiveness.

When you enable VM Monitoring, the VM Monitoring service (using VMware Tools) evaluates whether each virtual machine in the cluster is running by checking for regular heartbeats and I/O activity from the VMware Tools process running inside the guest. If no heartbeats or I/O activity are received, this is most likely because the guest operating system has failed or VMware Tools is not being allocated any time to complete tasks. In such a case, the VM Monitoring service determines that the virtual machine has failed and the virtual machine is rebooted to restore service.

Occasionally, virtual machines or applications that are still functioning properly stop sending heartbeats. To avoid unnecessary resets, the VM Monitoring service also monitors a virtual machine's I/O activity. If no heartbeats are received within the failure interval, the I/O stats interval (a cluster-level attribute) is checked. The I/O stats interval determines if any disk or network activity has occurred for the virtual machine during the previous two minutes (120 seconds). If not, the virtual machine is reset. This default value (120 seconds) can be changed using the advanced option `das.iostatsinterval`.

To enable Application Monitoring, you must first obtain the appropriate SDK (or be using an application that supports VMware Application Monitoring) and use it to set up customized heartbeats for the applications you want to monitor. After you have done this, Application Monitoring works much the same way that VM Monitoring does. If the heartbeats for an application are not received for a specified time, its virtual machine is restarted.

You can configure the level of monitoring sensitivity. Highly sensitive monitoring results in a more rapid conclusion that a failure has occurred. While unlikely, highly sensitive monitoring might lead to falsely identifying failures when the virtual machine or application in question is actually still working, but heartbeats have not been received due to factors such as resource constraints. Low sensitivity monitoring results in longer interruptions in service between actual failures and virtual machines being reset. Select an option that is an effective compromise for your needs.

The default settings for monitoring sensitivity are described in [Table 2-1](#). You can also specify custom values for both monitoring sensitivity and the I/O stats interval by selecting the **Custom** checkbox.

Table 2-1. VM Monitoring Settings

Setting	Failure Interval (seconds)	Reset Period
High	30	1 hour
Medium	60	24 hours
Low	120	7 days

After failures are detected, vSphere HA resets virtual machines. The reset ensures that services remain available. To avoid resetting virtual machines repeatedly for nontransient errors, by default, virtual machines will be reset only three times during a certain configurable time interval. After virtual machines have been reset three times, vSphere HA makes no further attempts to reset the virtual machines after subsequent failures until after the specified time has elapsed. You can configure the number of resets using the **Maximum per-VM resets** custom setting.

NOTE The reset statistics are cleared when a virtual machine is powered off then back on, or when it is migrated using vMotion to another host. This causes the guest operating system to reboot, but is not the same as a 'restart' in which the power state of the virtual machine is changed.

If a virtual machine has a datastore accessibility failure (either All Paths Down or Permanent Device Loss), the VM Monitoring service suspends resetting it until the failure has been addressed.

VM Component Protection

If VM Component Protection (VMCP) is enabled, vSphere HA can detect datastore accessibility failures and provide automated recovery for affected virtual machines.

VMCP provides protection against datastore accessibility failures that can affect a virtual machine running on a host in a vSphere HA cluster. When a datastore accessibility failure occurs, the affected host can no longer access the storage path for a specific datastore. You can determine the response that vSphere HA will make to such a failure, ranging from the creation of event alarms to virtual machine restarts on other hosts.

NOTE When you use the VM Component Protection feature, your ESXi hosts must be version 6.0 or higher.

Types of Failure

There are two types of datastore accessibility failure:

PDL	PDL (Permanent Device Loss) is an unrecoverable loss of accessibility that occurs when a storage device reports the datastore is no longer accessible by the host. This condition cannot be reverted without powering off virtual machines.
APD	APD (All Paths Down) represents a transient or unknown accessibility loss or any other unidentified delay in I/O processing. This type of accessibility issue is recoverable.

Configuring VMCP

VM Component Protection is configured in the vSphere Web Client. Go to the **Configure** tab and click **vSphere Availability** and **Edit**. Under **Failures and Responses** you can select **Datastore with PDL** or **Datastore with APD**. The storage protection levels you can choose and the virtual machine remediation actions available differ depending on the type of database accessibility failure.

PDL Failures Under **Datastore with PDL**, you can select **Issue events** or **Power off and restart VMs**.

APD Failures The response to APD events is more complex and accordingly the configuration is more fine-grained. You can select **Issue events**, **Power off and restart VMs--conservative restart policy**, or **Power off and restart VMs--aggressive restart policy**

NOTE If either the Host Monitoring or VM Restart Priority settings are disabled, VMCP cannot perform virtual machine restarts. Storage health can still be monitored and events can be issued, however.

Network Partitions

When a management network failure occurs for a vSphere HA cluster, a subset of the cluster's hosts might be unable to communicate over the management network with the other hosts. Multiple partitions can occur in a cluster.

A partitioned cluster leads to degraded virtual machine protection and cluster management functionality. Correct the partitioned cluster as soon as possible.

- **Virtual machine protection.** vCenter Server allows a virtual machine to be powered on, but it can be protected only if it is running in the same partition as the master host that is responsible for it. The master host must be communicating with vCenter Server. A master host is responsible for a virtual machine if it has exclusively locked a system-defined file on the datastore that contains the virtual machine's configuration file.
- **Cluster management.** vCenter Server can communicate with the master host, but only a subset of the slave hosts. As a result, changes in configuration that affect vSphere HA might not take effect until after the partition is resolved. This failure could result in one of the partitions operating under the old configuration, while another uses the new settings.

Datastore Heartbeating

When the master host in a vSphere HA cluster can not communicate with a slave host over the management network, the master host uses datastore heartbeating to determine whether the slave host has failed, is in a network partition, or is network isolated. If the slave host has stopped datastore heartbeating, it is considered to have failed and its virtual machines are restarted elsewhere.

vCenter Server selects a preferred set of datastores for heartbeating. This selection is made to maximize the number of hosts that have access to a heartbeating datastore and minimize the likelihood that the datastores are backed by the same LUN or NFS server.

You can use the advanced option `das.heartbeatdsperhost` to change the number of heartbeat datastores selected by vCenter Server for each host. The default is two and the maximum valid value is five.

vSphere HA creates a directory at the root of each datastore that is used for both datastore heartbeating and for persisting the set of protected virtual machines. The name of the directory is `.vSphere-HA`. Do not delete or modify the files stored in this directory, because this can have an impact on operations. Because more than one cluster might use a datastore, subdirectories for this directory are created for each cluster. Root owns these directories and files and only root can read and write to them. The disk space used by vSphere HA depends on several factors including which VMFS version is in use and the number of hosts that use the datastore for heartbeating. With vmfs3, the maximum usage is approximately 2GB and the typical usage is approximately 3MB. With vmfs5 the maximum and typical usage is approximately 3MB. vSphere HA use of the datastores adds negligible overhead and has no performance impact on other datastore operations.

vSphere HA limits the number of virtual machines that can have configuration files on a single datastore. See *Configuration Maximums* for updated limits. If you place more than this number of virtual machines on a datastore and power them on, vSphere HA protects a number of virtual machines only up to the limit.

NOTE A Virtual SAN datastore cannot be used for datastore heartbeating. Therefore, if no other shared storage is accessible to all hosts in the cluster, there can be no heartbeat datastores in use. However, if you have storage that can be reached by an alternate network path that is independent of the Virtual SAN network, you can use it to set up a heartbeat datastore.

vSphere HA Security

vSphere HA is enhanced by several security features.

Select firewall ports opened	vSphere HA uses TCP and UDP port 8182 for agent-to-agent communication. The firewall ports open and close automatically to ensure they are open only when needed.
Configuration files protected using file system permissions	vSphere HA stores configuration information on the local storage or on ramdisk if there is no local datastore. These files are protected using file system permissions and they are accessible only to the root user. Hosts without local storage are only supported if they are managed by Auto Deploy.
Detailed logging	<p>The location where vSphere HA places log files depends on the version of host.</p> <ul style="list-style-type: none"> ■ For ESXi 5.x hosts, vSphere HA writes to syslog only by default, so logs are placed where syslog is configured to put them. The log file names for vSphere HA are prepended with <code>fdm</code>, fault domain manager, which is a service of vSphere HA. ■ For legacy ESXi 4.x hosts, vSphere HA writes to <code>/var/log/vmware/fdm</code> on local disk, as well as syslog if it is configured. ■ For legacy ESX 4.x hosts, vSphere HA writes to <code>/var/log/vmware/fdm</code>.
Secure vSphere HA logins	vSphere HA logs onto the vSphere HA agents using a user account, <code>vpxuser</code> , created by vCenter Server. This account is the same account used by vCenter Server to manage the host. vCenter Server creates a random password for this account and changes the password periodically. The time period is set by the vCenter Server <code>VirtualCenter.VimPasswordExpirationInDays</code> setting. Users with administrative privileges on the root folder of the host can log in to the agent.
Secure communication	All communication between vCenter Server and the vSphere HA agent is done over SSL. Agent-to-agent communication also uses SSL except for election messages, which occur over UDP. Election messages are verified over SSL so that a rogue agent can prevent only the host on which the agent is running from being elected as a master host. In this case, a configuration issue for the cluster is issued so the user is aware of the problem.
Host SSL certificate verification required	vSphere HA requires that each host have a verified SSL certificate. Each host generates a self-signed certificate when it is booted for the first time. This certificate can then be regenerated or replaced with one issued by an authority. If the certificate is replaced, vSphere HA needs to be reconfigured on the host. If a host becomes disconnected from vCenter Server after its certificate is updated and the ESXi or ESX Host agent is restarted, then vSphere HA is automatically reconfigured when the host is reconnected to vCenter Server. If the disconnection does not occur because vCenter Server host SSL certificate verification is disabled at the time, verify the new certificate and reconfigure vSphere HA on the host.

vSphere HA Admission Control

vSphere HA uses admission control to ensure that sufficient resources are reserved for virtual machine recovery if a host fails.

Admission control imposes constraints on resource usage. Any action that might violate these constraints is not permitted. Actions that might be disallowed include the following examples:

- Powering on a virtual machine
- Migrating a virtual machine
- Increasing the CPU or memory reservation of a virtual machine

The basis for vSphere HA admission control is how many host failures your cluster is allowed to tolerate and still guarantee failover. The host failover capacity can be set in three ways:

- Cluster resource percentage
- Slot policy
- Dedicated failover hosts

NOTE vSphere HA admission control can be disabled. However, without it you have no assurance that the expected number of virtual machines can be restarted after a failure. Do not permanently disable admission control.

Regardless of the admission control option chosen, a VM resource reduction threshold also exists. You use this setting to specify the percentage of resource degradation to tolerate, but it is not available unless vSphere DRS is enabled.

The resource reduction calculation is checked for both CPU and memory. It takes into account a virtual machine's reserved memory and memory overload to decide whether to permit it to power on, migrate, or have reservation changes. The actual memory consumed by the virtual machine is not considered in the calculation because the memory reservation does not always correlate with the actual memory utilization of the virtual machine. If the actual utilization is more than reserved memory, insufficient failover capacity is available, resulting in performance degradation on failover.

Setting a performance reduction threshold enables you to specify the occurrence of a configuration issue. For example:

- The default value is 100%, which produces no warnings.
- If you reduce the threshold to 0%, a warning is generated as soon as cluster utilization exceeds the available capacity.
- If you reduce the threshold to 20%, the performance reduction that can be tolerated is calculated as $\text{performance reduction} = \text{current utilization} * 20\%$. When the current utilization minus the performance reduction exceeds the available capacity, a configuration notice is issued.

Cluster Resources Percentage Admission Control

You can configure vSphere HA to perform admission control by reserving a specific percentage of cluster CPU and memory resources for recovery from host failures.

With this type of admission control, vSphere HA ensures that a specified percentage of aggregate CPU and memory resources are reserved for failover.

With the cluster resources percentage option, vSphere HA enforces admission control as follows:

- 1 Calculates the total resource requirements for all powered-on virtual machines in the cluster.
- 2 Calculates the total host resources available for virtual machines.

- 3 Calculates the Current CPU Failover Capacity and Current Memory Failover Capacity for the cluster.
- 4 Determines if either the Current CPU Failover Capacity or Current Memory Failover Capacity is less than the corresponding Configured Failover Capacity (provided by the user).

If so, admission control disallows the operation.

vSphere HA uses the actual reservations of the virtual machines. If a virtual machine does not have reservations, meaning that the reservation is 0, a default of 0MB memory and 32MHz CPU is applied.

NOTE The cluster resources percentage option for admission control also checks that there are at least two vSphere HA-enabled hosts in the cluster (excluding hosts that are entering maintenance mode). If there is only one vSphere HA-enabled host, an operation is not allowed, even if there is a sufficient percentage of resources available. The reason for this extra check is that vSphere HA cannot perform failover if there is only a single host in the cluster.

Computing the Current Failover Capacity

The total resource requirements for the powered-on virtual machines is comprised of two components, CPU and memory. vSphere HA calculates these values.

- The CPU component by summing the CPU reservations of the powered-on virtual machines. If you have not specified a CPU reservation for a virtual machine, it is assigned a default value of 32MHz (this value can be changed using the `das.vmcpumimhz` advanced option.)
- The memory component by summing the memory reservation (plus memory overhead) of each powered-on virtual machine.

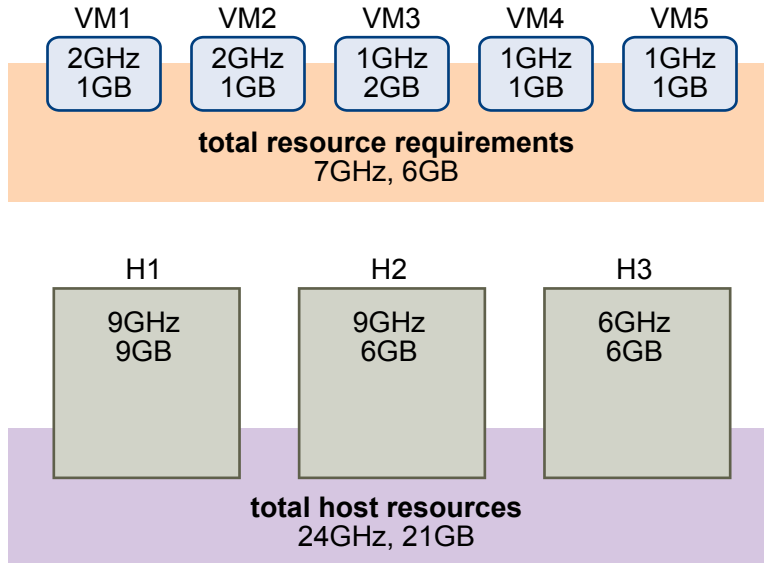
The total host resources available for virtual machines is calculated by adding the hosts' CPU and memory resources. These amounts are those contained in the host's root resource pool, not the total physical resources of the host. Resources being used for virtualization purposes are not included. Only hosts that are connected, not in maintenance mode, and have no vSphere HA errors are considered.

The Current CPU Failover Capacity is computed by subtracting the total CPU resource requirements from the total host CPU resources and dividing the result by the total host CPU resources. The Current Memory Failover Capacity is calculated similarly.

Example: Admission Control Using Cluster Resources Percentage

The way that Current Failover Capacity is calculated and used with this admission control policy is shown with an example. Make the following assumptions about a cluster:

- The cluster is comprised of three hosts, each with a different amount of available CPU and memory resources. The first host (H1) has 9GHz of available CPU resources and 9GB of available memory, while Host 2 (H2) has 9GHz and 6GB and Host 3 (H3) has 6GHz and 6GB.
- There are five powered-on virtual machines in the cluster with differing CPU and memory requirements. VM1 needs 2GHz of CPU resources and 1GB of memory, while VM2 needs 2GHz and 1GB, VM3 needs 1GHz and 2GB, VM4 needs 1GHz and 1GB, and VM5 needs 1GHz and 1GB.
- The Configured Failover Capacity for CPU and Memory are both set to 25%.

Figure 2-1. Admission Control Example with Percentage of Cluster Resources Reserved Policy

The total resource requirements for the powered-on virtual machines is 7GHz and 6GB. The total host resources available for virtual machines is 24GHz and 21GB. Based on this, the Current CPU Failover Capacity is 70% $((24\text{GHz} - 7\text{GHz})/24\text{GHz})$. Similarly, the Current Memory Failover Capacity is 71% $((21\text{GB} - 6\text{GB})/21\text{GB})$.

Because the cluster's Configured Failover Capacity is set to 25%, 45% of the cluster's total CPU resources and 46% of the cluster's memory resources are still available to power on additional virtual machines.

Slot Policy Admission Control

With the slot policy option, vSphere HA admission control ensures that a specified number of hosts can fail and sufficient resources remain in the cluster to fail over all the virtual machines from those hosts.

Using the slot policy, vSphere HA performs admission control in the following way:

- 1 Calculates the slot size.

A slot is a logical representation of memory and CPU resources. By default, it is sized to satisfy the requirements for any powered-on virtual machine in the cluster.

- 2 Determines how many slots each host in the cluster can hold.
- 3 Determines the Current Failover Capacity of the cluster.

This is the number of hosts that can fail and still leave enough slots to satisfy all of the powered-on virtual machines.

- 4 Determines whether the Current Failover Capacity is less than the Configured Failover Capacity (provided by the user).

If it is, admission control disallows the operation.

NOTE You can set a specific slot size for both CPU and memory in the admission control section of the vSphere HA settings in the vSphere Web Client.

Slot Size Calculation



vSphere HA Slot Size and Admission Control
http://link.brightcove.com/services/player/bcpid2296383276001?bctid=ref:video_vsphere_slot_admission_control

Slot size is comprised of two components, CPU and memory.

- vSphere HA calculates the CPU component by obtaining the CPU reservation of each powered-on virtual machine and selecting the largest value. If you have not specified a CPU reservation for a virtual machine, it is assigned a default value of 32MHz. You can change this value by using the `das.vmcpuminhz` advanced option.)
- vSphere HA calculates the memory component by obtaining the memory reservation, plus memory overhead, of each powered-on virtual machine and selecting the largest value. There is no default value for the memory reservation.

If your cluster contains any virtual machines that have much larger reservations than the others, they will distort slot size calculation. To avoid this, you can specify an upper bound for the CPU or memory component of the slot size by using the `das.slotcpuinmhz` or `das.slotmeminmb` advanced options, respectively. See “[vSphere HA Advanced Options](#),” on page 37.

You can also determine the risk of resource fragmentation in your cluster by viewing the number of virtual machines that require multiple slots. This can be calculated in the admission control section of the vSphere HA settings in the vSphere Web Client. Virtual machines might require multiple slots if you have specified a fixed slot size or a maximum slot size using advanced options.

Using Slots to Compute the Current Failover Capacity

After the slot size is calculated, vSphere HA determines each host's CPU and memory resources that are available for virtual machines. These amounts are those contained in the host's root resource pool, not the total physical resources of the host. The resource data for a host that is used by vSphere HA can be found on the host's **Summary** tab on the vSphere Web Client. If all hosts in your cluster are the same, this data can be obtained by dividing the cluster-level figures by the number of hosts. Resources being used for virtualization purposes are not included. Only hosts that are connected, not in maintenance mode, and that have no vSphere HA errors are considered.

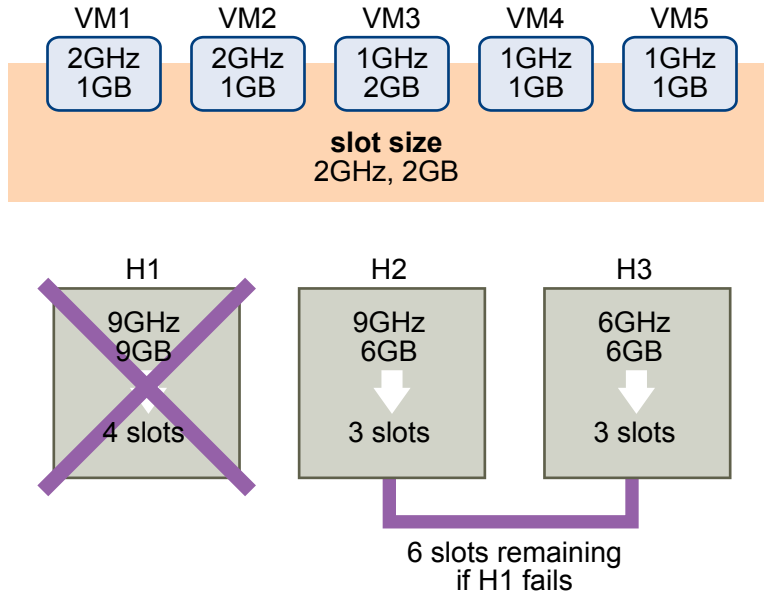
The maximum number of slots that each host can support is then determined. To do this, the host's CPU resource amount is divided by the CPU component of the slot size and the result is rounded down. The same calculation is made for the host's memory resource amount. These two numbers are compared and the smaller number is the number of slots that the host can support.

The Current Failover Capacity is computed by determining how many hosts (starting from the largest) can fail and still leave enough slots to satisfy the requirements of all powered-on virtual machines.

Example: Admission Control Using Slot Policy

The way that slot size is calculated and used with this admission control policy is shown in an example. Make the following assumptions about a cluster:

- The cluster is comprised of three hosts, each with a different amount of available CPU and memory resources. The first host (H1) has 9GHz of available CPU resources and 9GB of available memory, while Host 2 (H2) has 9GHz and 6GB and Host 3 (H3) has 6GHz and 6GB.
- There are five powered-on virtual machines in the cluster with differing CPU and memory requirements. VM1 needs 2GHz of CPU resources and 1GB of memory, while VM2 needs 2GHz and 1GB, VM3 needs 1GHz and 2GB, VM4 needs 1GHz and 1GB, and VM5 needs 1GHz and 1GB.
- The Host Failures Cluster Tolerates is set to one.

Figure 2-2. Admission Control Example with Host Failures Cluster Tolerates Policy

- 1 Slot size is calculated by comparing both the CPU and memory requirements of the virtual machines and selecting the largest.

The largest CPU requirement (shared by VM1 and VM2) is 2GHz, while the largest memory requirement (for VM3) is 2GB. Based on this, the slot size is 2GHz CPU and 2GB memory.

- 2 Maximum number of slots that each host can support is determined.

H1 can support four slots. H2 can support three slots (which is the smaller of $9\text{GHz}/2\text{GHz}$ and $6\text{GB}/2\text{GB}$) and H3 can also support three slots.

- 3 Current Failover Capacity is computed.

The largest host is H1 and if it fails, six slots remain in the cluster, which is sufficient for all five of the powered-on virtual machines. If both H1 and H2 fail, only three slots remain, which is insufficient. Therefore, the Current Failover Capacity is one.

The cluster has one available slot (the six slots on H2 and H3 minus the five used slots).

Dedicated Failover Hosts Admission Control

You can configure vSphere HA to designate specific hosts as the failover hosts.

With dedicated failover hosts admission control, when a host fails, vSphere HA attempts to restart its virtual machines on any of the specified failover hosts. If this is not possible, for example the failover hosts have failed or have insufficient resources, then vSphere HA attempts to restart those virtual machines on other hosts in the cluster.

To ensure that spare capacity is available on a failover host, you are prevented from powering on virtual machines or using vMotion to migrate virtual machines to a failover host. Also, DRS does not use a failover host for load balancing.

NOTE If you use dedicated failover hosts admission control and designate multiple failover hosts, DRS does not attempt to enforce VM-VM affinity rules for virtual machines that are running on failover hosts.

vSphere HA Interoperability

vSphere HA can interoperate with many other features, such as DRS and Virtual SAN.

Before configuring vSphere HA, you should be aware of the limitations of its interoperability with these other features or products.

Using vSphere HA with Virtual SAN

You can use Virtual SAN as the shared storage for a vSphere HA cluster. If enabled, Virtual SAN aggregates the specified local storage disks available on the hosts into a single datastore shared by all hosts.

To use vSphere HA with Virtual SAN, you must be aware of certain considerations and limitations for the interoperability of these two features.

For information about Virtual SAN, see *Administering VMware Virtual SAN*.

NOTE You can use vSphere HA with Virtual SAN stretched clusters.

ESXi Host Requirements

You can use Virtual SAN with a vSphere HA cluster only if the following conditions are met:

- All the cluster's ESXi hosts must be version 5.5 or later.
- The cluster must have a minimum of three ESXi hosts.

Networking Differences

Virtual SAN has its own network. If Virtual SAN and vSphere HA are enabled for the same cluster, the HA interagent traffic flows over this storage network rather than the management network. vSphere HA uses the management network only if Virtual SAN is disabled. vCenter Server chooses the appropriate network if vSphere HA is configured on a host.

NOTE You can enable Virtual SAN only if vSphere HA is disabled.

If you change the Virtual SAN network configuration, the vSphere HA agents do not automatically pick up the new network settings. To make changes to the Virtual SAN network, you must take the following steps in the vSphere Web Client:

- 1 Disable Host Monitoring for the vSphere HA cluster.
- 2 Make the Virtual SAN network changes.
- 3 Right-click all hosts in the cluster and select **Reconfigure for vSphere HA**.
- 4 Re-enable Host Monitoring for the vSphere HA cluster.

[Table 2-2](#) shows the differences in vSphere HA networking whether or not Virtual SAN is used.

Table 2-2. vSphere HA Networking Differences

	Virtual SAN Enabled	Virtual SAN Disabled
Network used by vSphere HA	Virtual SAN storage network	Management network
Heartbeat datastores	Any datastore mounted to > 1 host, but not Virtual SAN datastores	Any datastore mounted to > 1 host
Host declared isolated	Isolation addresses not pingable and Virtual SAN storage network inaccessible	Isolation addresses not pingable and management network inaccessible

Capacity Reservation Settings

When you reserve capacity for your vSphere HA cluster with an admission control policy, you must coordinate this setting with the corresponding Virtual SAN setting that ensures data accessibility on failures. Specifically, the Number of Failures Tolerated setting in the Virtual SAN rule set must not be lower than the capacity that the vSphere HA admission control setting reserved.

For example, if the Virtual SAN rule set allows for only two failures, the vSphere HA admission control policy must reserve capacity that is equivalent to only one or two host failures. If you are using the Percentage of Cluster Resources Reserved policy for a cluster that has eight hosts, you must not reserve more than 25% of the cluster resources. In the same cluster, with the Host Failures Cluster Tolerates policy, the setting must not be higher than two hosts. If vSphere HA reserves less capacity, failover activity might be unpredictable. Reserving too much capacity overly constrains the powering on of virtual machines and intercluster vSphere vMotion migrations.

Using vSphere HA and DRS Together

Using vSphere HA with Distributed Resource Scheduler (DRS) combines automatic failover with load balancing. This combination can result in a more balanced cluster after vSphere HA has moved virtual machines to different hosts.

When vSphere HA performs failover and restarts virtual machines on different hosts, its first priority is the immediate availability of all virtual machines. After the virtual machines have been restarted, those hosts on which they were powered on might be heavily loaded, while other hosts are comparatively lightly loaded. vSphere HA uses the virtual machine's CPU and memory reservation and overhead memory to determine if a host has enough spare capacity to accommodate the virtual machine.

In a cluster using DRS and vSphere HA with admission control turned on, virtual machines might not be evacuated from hosts entering maintenance mode. This behavior occurs because of the resources reserved for restarting virtual machines in the event of a failure. You must manually migrate the virtual machines off of the hosts using vMotion.

In some scenarios, vSphere HA might not be able to fail over virtual machines because of resource constraints. This can occur for several reasons.

- HA admission control is disabled and Distributed Power Management (DPM) is enabled. This can result in DPM consolidating virtual machines onto fewer hosts and placing the empty hosts in standby mode leaving insufficient powered-on capacity to perform a failover.
- VM-Host affinity (required) rules might limit the hosts on which certain virtual machines can be placed.
- There might be sufficient aggregate resources but these can be fragmented across multiple hosts so that they can not be used by virtual machines for failover.

In such cases, vSphere HA can use DRS to try to adjust the cluster (for example, by bringing hosts out of standby mode or migrating virtual machines to defragment the cluster resources) so that HA can perform the failovers.

If DPM is in manual mode, you might need to confirm host power-on recommendations. Similarly, if DRS is in manual mode, you might need to confirm migration recommendations.

If you are using VM-Host affinity rules that are required, be aware that these rules cannot be violated. vSphere HA does not perform a failover if doing so would violate such a rule.

For more information about DRS, see the *vSphere Resource Management* documentation.

vSphere HA and DRS Affinity Rules

If you create a DRS affinity rule for your cluster, you can specify how vSphere HA applies that rule during a virtual machine failover.

The two types of rules for which you can specify vSphere HA failover behavior are the following:

- VM anti-affinity rules force specified virtual machines to remain apart during failover actions.
- VM-Host affinity rules place specified virtual machines on a particular host or a member of a defined group of hosts during failover actions.

When you edit a DRS affinity rule, select the checkbox or checkboxes that enforce the desired failover behavior for vSphere HA.

- **HA must respect VM anti-affinity rules during failover** -- if VMs with this rule would be placed together, the failover is aborted.
- **HA should respect VM to Host affinity rules during failover** --vSphere HA attempts to place VMs with this rule on the specified hosts if at all possible.

NOTE vSphere HA can restart a VM in a DRS-disabled cluster, overriding a VM-Host affinity rules mapping if the host failure happens soon (by default, within 5 minutes) after setting the rule.

Other vSphere HA Interoperability Issues

To use vSphere HA, you must be aware of the following additional interoperability issues.

VM Component Protection

VM Component Protection (VMCP) has the following interoperability issues and limitations:

- VMCP does not support vSphere Fault Tolerance. If VMCP is enabled for a cluster using Fault Tolerance, the affected FT virtual machines will automatically receive overrides that disable VMCP.
- VMCP does not detect or respond to accessibility issues for files located on Virtual SAN datastores. If a virtual machine's configuration and VMDK files are located only on Virtual SAN datastores, they are not protected by VMCP.
- VMCP does not detect or respond to accessibility issues for files located on Virtual Volume datastores. If a virtual machine's configuration and VMDK files are located only on Virtual Volume datastores, they are not protected by VMCP.
- VMCP does not protect against inaccessible Raw Device Mapping (RDM)s.

IPv6

vSphere HA can be used with IPv6 network configurations, which are fully supported if the following considerations are observed:

- The cluster contains only ESXi 6.0 or later hosts.
- The management network for all hosts in the cluster must be configured with the same IP version, either IPv6 or IPv4. vSphere HA clusters cannot contain both types of networking configuration.
- The network isolation addresses used by vSphere HA must match the IP version used by the cluster for its management network.
- IPv6 cannot be used in vSphere HA clusters that also utilize Virtual SAN.

In addition to the previous restrictions, the following types of IPv6 address types are not supported for use with the vSphere HA isolation address or management network: link-local, ORCHID, and link-local with zone indices. Also, the loopback address type cannot be used for the management network.

NOTE To upgrade an existing IPv4 deployment to IPv6, you must first disable vSphere HA.

Creating a vSphere HA Cluster

vSphere HA operates in the context of a cluster of ESXi (or legacy ESX) hosts. You must create a cluster, populate it with hosts, and configure vSphere HA settings before failover protection can be established.

When you create a vSphere HA cluster, you must configure a number of settings that determine how the feature works. Before you do this, identify your cluster's nodes. These nodes are the ESXi hosts that will provide the resources to support virtual machines and that vSphere HA will use for failover protection. You should then determine how those nodes are to be connected to one another and to the shared storage where your virtual machine data resides. After that networking architecture is in place, you can add the hosts to the cluster and finish configuring vSphere HA.

You can enable and configure vSphere HA before you add host nodes to the cluster. However, until the hosts are added, your cluster is not fully operational and some of the cluster settings are unavailable. For example, the Specify a Failover Host admission control policy is unavailable until there is a host that can be designated as the failover host.

NOTE The Virtual Machine Startup and Shutdown (automatic startup) feature is disabled for all virtual machines residing on hosts that are in (or moved into) a vSphere HA cluster. Automatic startup is not supported when used with vSphere HA.

vSphere HA Checklist

The vSphere HA checklist contains requirements that you must be aware of before creating and using a vSphere HA cluster.

Review this list before you set up a vSphere HA cluster. For more information, follow the appropriate cross reference.

- All hosts must be licensed for vSphere HA.
- A cluster must contain at least two hosts.
- All hosts must be configured with static IP addresses. If you are using DHCP, you must ensure that the address for each host persists across reboots.
- All hosts must have at least one management network in common. The best practice is to have at least two management networks in common. You should use the VMkernel network with the **Management traffic** checkbox enabled. The networks must be accessible to each other and vCenter Server and the hosts must be accessible to each other on the management networks. See [“Best Practices for Networking,”](#) on page 39.
- To ensure that any virtual machine can run on any host in the cluster, all hosts must have access to the same virtual machine networks and datastores. Similarly, virtual machines must be located on shared, not local, storage otherwise they cannot be failed over in the case of a host failure.

NOTE vSphere HA uses datastore heartbeating to distinguish between partitioned, isolated, and failed hosts. So if some datastores are more reliable in your environment, configure vSphere HA to give preference to them.

- For VM Monitoring to work, VMware tools must be installed. See [“VM and Application Monitoring,”](#) on page 17.

- vSphere HA supports both IPv4 and IPv6. See [“Other vSphere HA Interoperability Issues,”](#) on page 28 for considerations when using IPv6.
- For VM Component Protection to work, hosts must have the All Paths Down (APD) Timeout feature enabled.
- To use VM Component Protection, clusters must contain ESXi 6.0 hosts or later.
- Only vSphere HA clusters that contain ESXi 6.0 or later hosts can be used to enable VMCP. Clusters that contain hosts from an earlier release cannot enable VMCP, and such hosts cannot be added to a VMCP-enabled cluster.
- If your cluster uses Virtual Volume datastores, when vSphere HA is enabled a configuration Virtual Volume is created on each datastore by vCenter Server. In these containers, vSphere HA stores the files it uses to protect virtual machines. vSphere HA does not function correctly if you delete these containers. Only one container is created per Virtual Volume datastore.

Create a vSphere HA Cluster

To enable your cluster for vSphere HA, you must first create an empty cluster. After you plan the resources and networking architecture of your cluster, use the vSphere Web Client to add hosts to the cluster and specify the cluster's vSphere HA settings.

A vSphere HA-enabled cluster is a prerequisite for vSphere Fault Tolerance.

Prerequisites

- Verify that all virtual machines and their configuration files reside on shared storage.
- Verify that the hosts are configured to access the shared storage so that you can power on the virtual machines by using different hosts in the cluster.
- Verify that hosts are configured to have access to the virtual machine 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,”](#) on page 39.
- 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 by using an account with cluster administrator permissions.

Procedure

- 1 In the 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 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 the vSphere Web Client to add hosts to the cluster.
- 5 Browse to the cluster and enable vSphere HA.
 - a Click the **Configure** tab.
 - b Select **vSphere Availability** and click **Edit**.

- c Select **Turn ON vSphere HA**.
 - d Select **Turn ON Proactive HA** to allow proactive migrations of VMs from hosts on which a provider has notified a health degradation.
- 6 Under **Failures and Responses** select **Enable Host Monitoring**
- With Host Monitoring enabled, hosts in the cluster can exchange network heartbeats and vSphere HA can take action when it detects failures. Host Monitoring is required for the vSphere Fault Tolerance recovery process to work properly.
- 7 Select a setting for **VM Monitoring**.
- Select **VM Monitoring Only** to restart individual virtual machines 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**.

You have a vSphere HA cluster, populated with hosts.

What to do next

Configure the appropriate vSphere HA settings for your cluster.

- Failures and responses
- Proactive HA Failures and Responses
- Admission Control
- Heartbeat Datastores
- Advanced Options

See [“Configuring vSphere Availability Settings,”](#) on page 31.

Configuring vSphere Availability Settings

When you create a vSphere HA cluster or configure an existing cluster, you must configure settings that determine how the feature works.

In the vSphere Web Client, you can configure following the vSphere HA settings:

Failures and responses	Provide settings here for host failure responses, host isolation, VM monitoring, and VM Component Protection.
Proactive HA failures and responses	Provide settings for how Proactive HA responds when a provider has notified its health degradation to vCenter, indicating a partial failure of that host.
Admission Control	Enable or disable admission control for the vSphere HA cluster and choose a policy for how it is enforced.
Heartbeat Datastores	Specify preferences for the datastores that vSphere HA uses for datastore heartbeating.
Advanced Options	Customize vSphere HA behavior by setting advanced options.

Configuring Responses to Failures

The **Failure and Responses** pane of the vSphere HA settings allows you to configure how your cluster should function when problems are encountered.

In this part of the vSphere Web Client, you can determine the specific responses the vSphere HA cluster has for host failures and isolation. You can also configure VM Component Protection (VMCP) actions when Permanent Device Loss (PDL) and All Paths Down (APD) situations occur and you can enable VM monitoring.

The following tasks are available:

- 1 [Respond to Host Failure](#) on page 32
You can set specific responses to host failures that occur in your vSphere HA cluster.
- 2 [Respond to Host Isolation](#) on page 33
You can set specific responses to host isolation that occurs in your vSphere HA cluster.
- 3 [Configure VMCP Responses](#) on page 33
Configure the response that VM Component Protection (VMCP) makes when a datastore encounters a PDL or APD failure.
- 4 [Enable VM Monitoring](#) on page 33
You can turn on VM and Application Monitoring and also set the monitoring sensitivity for your vSphere HA cluster.

Respond to Host Failure

You can set specific responses to host failures that occur in your vSphere HA cluster.

This page is editable only if you have enabled vSphere HA.

Procedure

- 1 In the vSphere Web Client, browse to the vSphere HA cluster.
- 2 Click the **Configure** tab.
- 3 Select **vSphere Availability** and click **Edit**.
- 4 Click **Failures and Responses** and then expand **Host Failure Response**.
- 5 Select from the following configuration options.

Option	Description
Failure Response	If you select Disabled , this setting turns off host monitoring and VMs are not restarted when host failures occur. If Restart VMs is selected, VMs are failed over based on their restart priority when a host fails.
Default VM Restart Priority	The restart priority determines the order in which virtual machines are restarted when the host fails. Higher priority virtual machines are started first. If multiple hosts fail, all virtual machines are migrated from the first host in order of priority, then all virtual machines from the second host in order of priority, and so on.
VM Dependency Restart Condition	A specific condition must be selected as well as a delay after that condition has been met, before vSphere HA is allowed to continue to the next VM restart priority.

- 6 Click **OK**.

Your settings for the host failure response take effect.

Respond to Host Isolation

You can set specific responses to host isolation that occurs in your vSphere HA cluster.

This page is editable only if you have enabled vSphere HA.

Procedure

- 1 In the vSphere Web Client, browse to the vSphere HA cluster.
- 2 Click the **Configure** tab.
- 3 Select **vSphere Availability** and click **Edit**.
- 4 Click **Failures and Responses** and expand **Response for Host Isolation**.
- 5 To configure the host isolation response, select **Disabled**, **Shut down and restart VMs**, or **Power off and restart VMs**.
- 6 Click **OK**.

Your setting for the host isolation response takes effect.

Configure VMCP Responses

Configure the response that VM Component Protection (VMCP) makes when a datastore encounters a PDL or APD failure.

This page is editable only if you have enabled vSphere HA.

Procedure

- 1 In the vSphere Web Client, browse to the vSphere HA cluster.
- 2 Click the **Configure** tab.
- 3 Select **vSphere Availability** and click **Edit**.
- 4 Click **Failures and Responses**, and expand either **Datastore with PDL** or **Datastore with APD**.
- 5 If you clicked **Datastore with PDL**, you can set the VMCP failure response for this type of issue, either **Disabled**, **Issue Events**, or **Power off and restart VMs**.
- 6 If you clicked **Datastore with APD**, you can set the VMCP failure response for this type of issue, either **Disabled**, **Issue Events**, **Power off and restart VMs--Conservative restart policy**, or **Power off and restart VMs--Aggressive restart policy**. You can also set **Response recovery**, which is the number of minutes that VMCP waits before taking action.
- 7 Click **OK**.

Your settings for the VMCP failure response take effect.

Enable VM Monitoring

You can turn on VM and Application Monitoring and also set the monitoring sensitivity for your vSphere HA cluster.

This page is editable only if you have enabled vSphere HA.

Procedure

- 1 In the vSphere Web Client, browse to the vSphere HA cluster.
- 2 Click the **Configure** tab.
- 3 Select **vSphere Availability** and click **Edit**.

- 4 Click **Failures and Responses** and expand **VM Monitoring**.
- 5 Select **VM Monitoring** and **Application Monitoring**.
These settings turn on VMware Tools heartbeats and application heartbeats, respectively.
- 6 To set the heartbeat monitoring sensitivity, move the slider between **Low** and **High** or select **Custom** to provide custom settings.
- 7 Click **OK**.

Your monitoring settings take effect.

Configure Proactive HA

You can configure how Proactive HA responds when a provider has notified its health degradation to vCenter, indicating a partial failure of that host.

This page is editable only if you have enabled vSphere DRS.

Procedure

- 1 In the vSphere Web Client, browse to the Proactive HA cluster.
- 2 Click the **Configure** tab.
- 3 Select **vSphere Availability** and click **Edit**.
- 4 Select **Turn on Proactive HA**.
- 5 Click **Proactive HA Failures and Responses**.
- 6 Select from the following configuration options.

Option	Description
Automation Level	<p>Determine whether host quarantine or maintenance mode and VM migrations are recommendations or automatic.</p> <ul style="list-style-type: none"> ■ Manual. vCenter Server will suggest migration recommendations for virtual machines. ■ Automated. Virtual machines will be migrated to healthy hosts and degraded hosts will be entered into quarantine or maintenance mode depending on the configured Proactive HA automation level.
Remediation	<p>Determine what happens to partially degraded hosts.</p> <ul style="list-style-type: none"> ■ Quarantine mode for all failures. Balances performance and availability, by avoiding the usage of partially degraded hosts as long as virtual machine performance is unaffected. ■ Quarantine mode for moderate and Maintenance mode for severe failure (Mixed). Balances performance and availability, by avoiding the usage of moderately degraded hosts as long as virtual machine performance is unaffected. Ensures virtual machines do not run on severely failed hosts. ■ Maintenance mode for all failures. Ensures virtual machines do not run on partially failed hosts. <p>Host.Config.Quarantine and Host.Config.Maintenance privileges are required to put hosts in Quarantine mode and Maintenance mode, respectively.</p>

Select the check boxes to enable Proactive HA providers for this cluster. Providers appear below when their corresponding vSphere Web Client plugin has been installed and the providers monitor every host in the cluster. Click on the edit link to view/edit the failure conditions supported by the provider.

- 7 Click **OK**.

Your settings for Proactive HA response take effect.

Configure Admission Control

After you create a cluster, you can configure admission control to specify whether virtual machines can be started if they violate availability constraints. The cluster reserves resources so that failover can occur for all running virtual machines on the specified number of hosts.

The Admission Control page appears only if you enabled vSphere HA.

Procedure

- 1 In the vSphere Web Client, browse to the vSphere HA cluster.
- 2 Click the **Configure** tab.
- 3 Select **vSphere Availability** and click **Edit**.
- 4 Click **Admission Control** to display the configuration options.
- 5 Select a number for the **Host failures cluster tolerates**. This is the maximum number of host failures that the cluster can recover from or guarantees failover for.
- 6 Select an option for **Define host failover capacity by**.

Option	Description
Cluster resource percentage	Specify a percentage of the cluster's CPU and memory resources to reserve as spare capacity to support failovers.
Slot Policy (powered-on VMs)	Select a slot size policy that covers all powered on VMs or is a fixed size. You can also calculate how many VMs require multiple slots.
Dedicated failover hosts	Select hosts to use for failover actions. Failovers can still occur on other hosts in the cluster if a default failover host does not have enough resources.
Disabled	Select this option to disable admission control and allow virtual machine power ons that violate availability constraints.

- 7 Set the percentage for the **Performance degradation VMs tolerate**.
This setting determines what percentage of performance degradation the VMs in the cluster are allowed to tolerate during a failure.
- 8 Click **OK**.

Your admission control settings take effect.

Configure Heartbeat Datastores

vSphere HA uses datastore heartbeating to distinguish between hosts that have failed and hosts that reside on a network partition. With datastore heartbeating, vSphere HA can monitor hosts when a management network partition occurs and continue to respond to failures.

You can specify the datastores that you want to be used for datastore heartbeating.

Procedure

- 1 In the vSphere Web Client, browse to the vSphere HA cluster.
- 2 Click the **Configure** tab.
- 3 Select **vSphere Availability** and click **Edit**.
- 4 Click **Heartbeat Datastores** to display the configuration options for datastore heartbeating.

- To instruct vSphere HA about how to select the datastores and how to treat your preferences, select from the following options.

Table 2-3.**Datastore Heartbeating Options**

Automatically select datastores accessible from the host

Use datastores only from the specified list

Use datastores from the specified list and complement automatically if needed

- In the Available heartbeat datastores pane, select the datastores that you want to use for heartbeating. The listed datastores are shared by more than one host in the vSphere HA cluster. When a datastore is selected, the lower pane displays all the hosts in the vSphere HA cluster that can access it.
- Click **OK**.

Set Advanced Options

To customize vSphere HA behavior, set advanced vSphere HA options.

Prerequisites

Verify that you have cluster administrator privileges.

NOTE Because these options affect the functioning of vSphere HA, change them with caution.

Procedure

- In the vSphere Web Client, browse to the vSphere HA cluster.
- Click the **Configure** tab.
- Select **vSphere Availability** and click **Edit**.
- Click **Advanced Options**.
- Click **Add** and type the name of the advanced option in the text box.
You can set the value of the option in the text box in the Value column.
- Repeat step 5 for each new option that you want to add and click **OK**.

The cluster uses the options that you added or modified.

What to do next

Once you have set an advanced vSphere HA option, it persists until you do one the following:

- Using the vSphere Web Client, reset its value to the default value.
- Manually edit or delete the option from the `fdm.cfg` file on all hosts in the cluster.

vSphere HA Advanced Options

You can set advanced options that affect the behavior of your vSphere HA cluster.

Table 2-4. vSphere HA Advanced Options

Option	Description
<code>das.isolationaddress[...]</code>	Sets the address to ping to determine if a host is isolated from the network. This address is pinged only when heartbeats are not received from any other host in the cluster. If not specified, the default gateway of the management network is used. This default gateway has to be a reliable address that is available, so that the host can determine if it is isolated from the network. You can specify multiple isolation addresses (up to 10) for the cluster: <code>das.isolationaddressX</code> , where X = 0-9. Typically you should specify one per management network. Specifying too many addresses makes isolation detection take too long.
<code>das.usedefaultisolationaddress</code>	By default, vSphere HA uses the default gateway of the console network as an isolation address. This option specifies whether or not this default is used (true/false).
<code>das.isolationshutdowntimeout</code>	The period of time the system waits for a virtual machine to shut down before powering it off. This only applies if the host's isolation response is Shut down VM. Default value is 300 seconds.
<code>das.slotmeminmb</code>	Defines the maximum bound on the memory slot size. If this option is used, the slot size is the smaller of this value or the maximum memory reservation plus memory overhead of any powered-on virtual machine in the cluster.
<code>das.slotcpuinmhz</code>	Defines the maximum bound on the CPU slot size. If this option is used, the slot size is the smaller of this value or the maximum CPU reservation of any powered-on virtual machine in the cluster.
<code>das.vmmemoryinmb</code>	Defines the default memory resource value assigned to a virtual machine if its memory reservation is not specified or zero. This is used for the Host Failures Cluster Tolerates admission control policy. If no value is specified, the default is 0 MB.
<code>das.vmcpuminhz</code>	Defines the default CPU resource value assigned to a virtual machine if its CPU reservation is not specified or zero. This is used for the Host Failures Cluster Tolerates admission control policy. If no value is specified, the default is 32MHz.
<code>das.iostatsinterval</code>	Changes the default I/O stats interval for VM Monitoring sensitivity. The default is 120 (seconds). Can be set to any value greater than, or equal to 0. Setting to 0 disables the check. NOTE Values of less than 50 are not recommended since smaller values can result in vSphere HA unexpectedly resetting a virtual machine.
<code>das.ignoreinsufficienthbdatastore</code>	Disables configuration issues created if the host does not have sufficient heartbeat datastores for vSphere HA. Default value is false.
<code>das.heartbeatdsperhost</code>	Changes the number of heartbeat datastores required. Valid values can range from 2-5 and the default is 2.

Table 2-4. vSphere HA Advanced Options (Continued)

Option	Description
fdm.isolationpolicydelaysec	The number of seconds system waits before executing the isolation policy once it is determined that a host is isolated. The minimum value is 30. If set to a value less than 30, the delay will be 30 seconds.
das.respectvmantiaffinityrules	Determines if vSphere HA enforces VM-VM anti-affinity rules. Default value is "false", whereby the rules are not enforced. Can also be set to "true" and rules are enforced (even if vSphere DRS is not enabled). In this case, vSphere HA does not fail over a virtual machine if doing so violates a rule, but it issues an event reporting there are insufficient resources to perform the failover. <i>See vSphere Resource Management for more information on anti-affinity rules.</i>
das.maxresets	The maximum number of reset attempts made by VMCP. If a reset operation on a virtual machine affected by an APD situation fails, VMCP retries the reset this many times before giving up
das.maxterminates	The maximum number of retries made by VMCP for virtual machine termination.
das.terminateretryintervalsec	If VMCP fails to terminate a virtual machine, this is the number of seconds the system waits before it retries a terminate attempt
das.config.fdm.reportfailoverfailevent	When set to 1, enables generation of a detailed per-VM event when an attempt by vSphere HA to restart a virtual machine is unsuccessful. Default value is 0. In versions earlier than vSphere 6.0, this event is generated by default.
vpxd.das.completemetadadataupdateintervalsec	The period of time (seconds) after a VM-Host affinity rule is set during which vSphere HA can restart a VM in a DRS-disabled cluster, overriding the rule. Default value is 300 seconds.
das.config.fdm.memreservationmb	By default vSphere HA agents run with a configured memory limit of 250 MB. A host might not allow this reservation if it runs out of reservable capacity. You can use this advanced option to lower the memory limit to avoid this issue. Only integers greater than 100, which is the minimum value, can be specified. Conversely, to prevent problems during master agent elections in a large cluster (containing 6,000 to 8,000 VMs) you should raise this limit to 325 MB. NOTE Once this limit is changed, for all hosts in the cluster you must run the Reconfigure HA task. Also, when a new host is added to the cluster or an existing host is rebooted, this task should be performed on those hosts in order to update this memory setting.
das.reregisterrestartdisabledvms	When vSphere HA is disabled on a specific VM this option ensures that the VM is registered on another host after a failure. This allows you to power-on that VM without needing to re-register it manually. NOTE When this option is used, vSphere HA does not power on the VM, but only registers it.

NOTE If you change the value of any of the following advanced options, you must disable and then re-enable vSphere HA before your changes take effect.

- `das.isolationaddress[...]`
 - `das.usedefaultisolationaddress`
 - `das.isolationshutdowntimeout`
-

Customize an Individual Virtual Machine

Each virtual machine in a vSphere HA cluster is assigned the cluster default settings for VM Restart Priority, Host Isolation Response, VM Component Protection, and VM Monitoring. You can specify specific behavior for each virtual machine by changing these defaults. If the virtual machine leaves the cluster, these settings are lost.

Procedure

- 1 In the vSphere Web Client, browse to the vSphere HA cluster.
- 2 Click the **Configure** tab.
- 3 Under Configuration, select **VM Overrides** and click **Add**.
- 4 Use the + button to select virtual machines to which to apply the overrides.
- 5 Click **OK**.
- 6 (Optional) You can change other settings, such as the **Automation level**, **VM restart priority**, **Response for Host Isolation**, VMCP settings, **VM Monitoring**, or **VM monitoring sensitivity** settings.

NOTE You can view the cluster defaults for these settings by first expanding **Relevant Cluster Settings** and then expanding **vSphere HA**.

- 7 Click **OK**.

The virtual machine's behavior now differs from the cluster defaults for each setting that you changed.

Best Practices for vSphere HA Clusters

To ensure optimal vSphere HA cluster performance, you should follow certain best practices. This section highlights some of the key best practices for a vSphere HA cluster.

You can also refer to the *vSphere High Availability Deployment Best Practices* publication for further discussion.

Best Practices for Networking

Observe the following best practices for the configuration of host NICs and network topology for vSphere HA. Best Practices include recommendations for your ESXi hosts, and for cabling, switches, routers, and firewalls.

Network Configuration and Maintenance

The following network maintenance suggestions can help you avoid the accidental detection of failed hosts and network isolation because of dropped vSphere HA heartbeats.

- When making changes to the networks that your clustered ESXi hosts are on, suspend the Host Monitoring feature. Changing your network hardware or networking settings can interrupt the heartbeats that vSphere HA uses to detect host failures, and this might result in unwanted attempts to fail over virtual machines.

- When you change the networking configuration on the ESXi hosts themselves, for example, adding port groups, or removing vSwitches, suspend Host Monitoring. After you have made the networking configuration changes, you must reconfigure vSphere HA on all hosts in the cluster, which causes the network information to be reinspected. Then re-enable Host Monitoring.

NOTE Because networking is a vital component of vSphere HA, if network maintenance needs to be performed inform the vSphere HA administrator.

Networks Used for vSphere HA Communications

To identify which network operations might disrupt the functioning of vSphere HA, you should know which management networks are being used for heart beating and other vSphere HA communications.

- On legacy ESX hosts in the cluster, vSphere HA communications travel over all networks that are designated as service console networks. VMkernel networks are not used by these hosts for vSphere HA communications. To contain vSphere HA traffic to a subset of the ESX console networks, use the `allowedNetworks` advanced option.
- On ESXi hosts in the cluster, vSphere HA communications, by default, travel over VMkernel networks. With an ESXi host, if you wish to use a network other than the one vCenter Server uses to communicate with the host for vSphere HA, you must explicitly enable the **Management traffic** checkbox.

To keep vSphere HA agent traffic on the networks you have specified, configure hosts so vmkNICs used by vSphere HA do not share subnets with vmkNICs used for other purposes. vSphere HA agents send packets using any pNIC that is associated with a given subnet if there is also at least one vmkNIC configured for vSphere HA management traffic. Consequently, to ensure network flow separation, the vmkNICs used by vSphere HA and by other features must be on different subnets.

Network Isolation Addresses

A network isolation address is an IP address that is pinged to determine whether a host is isolated from the network. This address is pinged only when a host has stopped receiving heartbeats from all other hosts in the cluster. If a host can ping its network isolation address, the host is not network isolated, and the other hosts in the cluster have either failed or are network partitioned. However, if the host cannot ping its isolation address, it is likely that the host has become isolated from the network and no failover action is taken.

By default, the network isolation address is the default gateway for the host. Only one default gateway is specified, regardless of how many management networks have been defined. You should use the `das.isolationaddress[...]` advanced option to add isolation addresses for additional networks. See [“vSphere HA Advanced Options,”](#) on page 37.

Network Path Redundancy

Network path redundancy between cluster nodes is important for vSphere HA reliability. A single management network ends up being a single point of failure and can result in failovers although only the network has failed. If you have only one management network, any failure between the host and the cluster can cause an unnecessary (or false) failover activity if heartbeat datastore connectivity is not retained during the networking failure. Possible failures include NIC failures, network cable failures, network cable removal, and switch resets. Consider these possible sources of failure between hosts and try to minimize them, typically by providing network redundancy.

The first way you can implement network redundancy is at the NIC level with NIC teaming. Using a team of two NICs connected to separate physical switches improves the reliability of a management network. Because servers connected through two NICs (and through separate switches) have two independent paths for sending and receiving heartbeats, the cluster is more resilient. To configure a NIC team for the management network, configure the vNICs in vSwitch configuration for Active or Standby configuration. The recommended parameter settings for the vNICs are:

- Default load balancing = route based on originating port ID
- Failback = No

After you have added a NIC to a host in your vSphere HA cluster, you must reconfigure vSphere HA on that host.

In most implementations, NIC teaming provides sufficient heartbeat redundancy, but as an alternative you can create a second management network connection attached to a separate virtual switch. Redundant management networking allows the reliable detection of failures and prevents isolation or partition conditions from occurring, because heartbeats can be sent over multiple networks. The original management network connection is used for network and management purposes. When the second management network connection is created, vSphere HA sends heartbeats over both management network connections. If one path fails, vSphere HA still sends and receives heartbeats over the other path.

NOTE Configure the fewest possible number of hardware segments between the servers in a cluster. The goal being to limit single points of failure. Additionally, routes with too many hops can cause networking packet delays for heartbeats, and increase the possible points of failure.

Using IPv6 Network Configurations

Only one IPv6 address should be assigned to a given network interface used by your vSphere HA cluster. Assigning multiple IP addresses increases the number of heartbeat messages sent by the cluster's master host with no corresponding benefit.

Best Practices for Interoperability

Observe the following best practices for allowing proper interoperability between vSphere HA and other features.

vSphere HA and Storage vMotion Interoperability in a Mixed Cluster

In clusters where ESXi 5.x hosts and ESX/ESXi 4.1 or prior hosts are present and where Storage vMotion is used extensively or Storage DRS is enabled, do not deploy vSphere HA. vSphere HA might respond to a host failure by restarting a virtual machine on a host with an ESXi version different from the one on which the virtual machine was running before the failure. A problem can occur if, at the time of failure, the virtual machine was involved in a Storage vMotion action on an ESXi 5.x host, and vSphere HA restarts the virtual machine on a host with a version prior to ESXi 5.0. While the virtual machine might power on, any subsequent attempts at snapshot operations could corrupt the vdisk state and leave the virtual machine unusable.

Using Auto Deploy with vSphere HA

You can use vSphere HA and Auto Deploy together to improve the availability of your virtual machines. Auto Deploy provisions hosts when they power up and you can also configure it to install the vSphere HA agent on such hosts during the boot process. See the Auto Deploy documentation included in vSphere Installation and Setup for details.

Upgrading Hosts in a Cluster Using Virtual SAN

If you are upgrading the ESXi hosts in your vSphere HA cluster to version 5.5 or higher, and you also plan to use Virtual SAN, follow this process.

- 1 Upgrade all of the hosts.
- 2 Disable vSphere HA.
- 3 Enable Virtual SAN.
- 4 Re-enable vSphere HA.

Best Practices for Cluster Monitoring

Observe the following best practices for monitoring the status and validity of your vSphere HA cluster.

Setting Alarms to Monitor Cluster Changes

When vSphere HA or Fault Tolerance take action to maintain availability, for example, a virtual machine failover, you can be notified about such changes. Configure alarms in vCenter Server to be triggered when these actions occur, and have alerts, such as emails, sent to a specified set of administrators.

Several default vSphere HA alarms are available.

- Insufficient failover resources (a cluster alarm)
- Cannot find master (a cluster alarm)
- Failover in progress (a cluster alarm)
- Host HA status (a host alarm)
- VM monitoring error (a virtual machine alarm)
- VM monitoring action (a virtual machine alarm)
- Failover failed (a virtual machine alarm)

NOTE The default alarms include the feature name, vSphere HA.

Monitoring Cluster Validity

A valid cluster is one in which the admission control policy has not been violated.

A cluster enabled for vSphere HA becomes invalid when the number of virtual machines powered on exceeds the failover requirements, that is, the current failover capacity is smaller than configured failover capacity. If admission control is disabled, clusters do not become invalid.

In the vSphere Web Client, select **vSphere HA** from the cluster's **Monitor** tab and then select **Configuration Issues**. A list of current vSphere HA issues appears.

DRS behavior is not affected if a cluster is red because of a vSphere HA issue.

Providing Fault Tolerance for Virtual Machines

3

You can utilize vSphere Fault Tolerance for your virtual machines to ensure business continuity with higher levels of availability and data protection than is offered by vSphere HA.

Fault Tolerance is built on the ESXi host platform, and it provides continuous availability by having identical virtual machines run on separate hosts.

To obtain the optimal results from Fault Tolerance you should be familiar with how it works, how to enable it for your cluster and virtual machines, and the best practices for its usage.



Fault Tolerance Protection for Virtual Machines
(http://link.brightcove.com/services/player/bcpid2296383276001?bctid=ref:video_fault_tolerance_protection_vms)

This chapter includes the following topics:

- [“How Fault Tolerance Works,”](#) on page 43
- [“Fault Tolerance Use Cases,”](#) on page 44
- [“Fault Tolerance Requirements, Limits, and Licensing,”](#) on page 44
- [“Fault Tolerance Interoperability,”](#) on page 45
- [“Preparing Your Cluster and Hosts for Fault Tolerance,”](#) on page 47
- [“Using Fault Tolerance,”](#) on page 49
- [“Best Practices for Fault Tolerance,”](#) on page 53
- [“Legacy Fault Tolerance,”](#) on page 55

How Fault Tolerance Works

You can use vSphere Fault Tolerance (FT) for most mission critical virtual machines. FT provides continuous availability for such a virtual machine by creating and maintaining another VM that is identical and continuously available to replace it in the event of a failover situation.

The protected virtual machine is called the Primary VM. The duplicate virtual machine, the Secondary VM, is created and runs on another host. The Secondary VM's execution is identical to that of the Primary VM and it can take over at any point without interruption, thereby providing fault tolerant protection.

The Primary and Secondary VMs continuously monitor the status of one another to ensure that Fault Tolerance is maintained. A transparent failover occurs if the host running the Primary VM fails, in which case the Secondary VM is immediately activated to replace the Primary VM. A new Secondary VM is started and Fault Tolerance redundancy is reestablished automatically. If the host running the Secondary VM fails, it is also immediately replaced. In either case, users experience no interruption in service and no loss of data.

A fault tolerant virtual machine and its secondary copy are not allowed to run on the same host. This restriction ensures that a host failure cannot result in the loss of both VMs.

NOTE You can also use VM-Host affinity rules to dictate which hosts designated virtual machines can run on. If you use these rules, be aware that for any Primary VM that is affected by such a rule, its associated Secondary VM is also affected by that rule. For more information about affinity rules, see the *vSphere Resource Management* documentation.

Fault Tolerance avoids "split-brain" situations, which can lead to two active copies of a virtual machine after recovery from a failure. Atomic file locking on shared storage is used to coordinate failover so that only one side continues running as the Primary VM and a new Secondary VM is respawned automatically.

vSphere Fault Tolerance can accommodate symmetric multiprocessor (SMP) virtual machines with up to four vCPUs. Earlier versions of vSphere used a different technology for Fault Tolerance (now known as legacy FT), with different requirements and characteristics (including a limitation of single vCPUs for legacy FT VMs). If compatibility with these earlier requirements is necessary, you can instead use legacy FT. However, this involves the setting of an advanced option for each VM. See "[Legacy Fault Tolerance](#)," on page 55 for more information.

Fault Tolerance Use Cases

Several typical situations can benefit from the use of vSphere Fault Tolerance.

Fault Tolerance provides a higher level of business continuity than vSphere HA. When a Secondary VM is called upon to replace its Primary VM counterpart, the Secondary VM immediately takes over the Primary VM's role with the entire state of the virtual machine preserved. Applications are already running, and data stored in memory does not need to be re-entered or reloaded. This differs from a failover provided by vSphere HA, which restarts the virtual machines affected by a failure.

This higher level of continuity and the added protection of state information and data informs the scenarios when you might want to deploy Fault Tolerance.

- Applications that need to be available at all times, especially those that have long-lasting client connections that users want to maintain during hardware failure.
- Custom applications that have no other way of doing clustering.
- Cases where high availability might be provided through custom clustering solutions, which are too complicated to configure and maintain.

Another key use case for protecting a virtual machine with Fault Tolerance can be described as On-Demand Fault Tolerance. In this case, a virtual machine is adequately protected with vSphere HA during normal operation. During certain critical periods, you might want to enhance the protection of the virtual machine. For example, you might be executing a quarter-end report which, if interrupted, might delay the availability of mission critical information. With vSphere Fault Tolerance, you can protect this virtual machine prior to running this report and then turn off or suspend Fault Tolerance after the report has been produced. You can use On-Demand Fault Tolerance to protect the virtual machine during a critical time period and return the resources to normal during non-critical operation.

Fault Tolerance Requirements, Limits, and Licensing

Before using vSphere Fault Tolerance (FT), consider the high-level requirements, limits, and licensing that apply to this feature.

Requirements

The following CPU and networking requirements apply to FT.

CPUs that are used in host machines for fault tolerant VMs must be compatible with vSphere vMotion or improved with Enhanced vMotion Compatibility. Also, CPUs that support Hardware MMU virtualization (Intel EPT or AMD RVI) are required. The following CPUs are supported.

- Intel Sandy Bridge or later. Avoton is not supported.
- AMD Bulldozer or later.

Use a 10-Gbit logging network for FT and verify that the network is low latency. A dedicated FT network is highly recommended.

Limits

In a cluster configured to use Fault Tolerance, two limits are enforced independently.

das.maxftvmsperhost	The maximum number of fault tolerant VMs allowed on a host in the cluster. Both Primary VMs and Secondary VMs count toward this limit. The default value is 4.
das.maxftvcpusperhost	The maximum number of vCPUs aggregated across all fault tolerant VMs on a host. vCPUs from both Primary VMs and Secondary VMs count toward this limit. The default value is 8.

Licensing

The number of vCPUs supported by a single fault tolerant VM is limited by the level of licensing that you have purchased for vSphere. Fault Tolerance is supported as follows:

- vSphere Standard and Enterprise. Allows up to 2 vCPUs
- vSphere Enterprise Plus. Allows up to 4 vCPUs

NOTE FT and legacy FT are not supported in vSphere Essentials and vSphere Essentials Plus.

Fault Tolerance Interoperability

vSphere Fault Tolerance faces some limitations concerning the vSphere features, devices, and other features it can interoperate with.

Before configuring vSphere Fault Tolerance, you should be aware of the features and products Fault Tolerance cannot interoperate with.

vSphere Features Not Supported with Fault Tolerance

When configuring your cluster, you should be aware that not all vSphere features can interoperate with Fault Tolerance.

The following vSphere features are not supported for fault tolerant virtual machines.

- Snapshots. Snapshots must be removed or committed before Fault Tolerance can be enabled on a virtual machine. In addition, it is not possible to take snapshots of virtual machines on which Fault Tolerance is enabled.

NOTE Disk-only snapshots created for vStorage APIs - Data Protection (VADP) backups are supported with Fault Tolerance. However, legacy FT does not support VADP.

- Storage vMotion. You cannot invoke Storage vMotion for virtual machines with Fault Tolerance turned on. To migrate the storage, you should temporarily turn off Fault Tolerance, and perform the storage vMotion action. When this is complete, you can turn Fault Tolerance back on.

- Linked clones. You cannot use Fault Tolerance on a virtual machine that is a linked clone, nor can you create a linked clone from an FT-enabled virtual machine.
- VM Component Protection (VMCP). If your cluster has VMCP enabled, overrides are created for fault tolerant virtual machines that turn this feature off.
- Virtual Volume datastores.
- Storage-based policy management.
- I/O filters.

Features and Devices Incompatible with Fault Tolerance

Not all third party devices, features, or products can interoperate with Fault Tolerance.

For a virtual machine to be compatible with Fault Tolerance, the Virtual Machine must not use the following features or devices.

Table 3-1. Features and Devices Incompatible with Fault Tolerance and Corrective Actions

Incompatible Feature or Device	Corrective Action
Physical Raw Disk mapping (RDM).	With legacy FT you can reconfigure virtual machines with physical RDM-backed virtual devices to use virtual RDMs instead.
CD-ROM or floppy virtual devices backed by a physical or remote device.	Remove the CD-ROM or floppy virtual device or reconfigure the backing with an ISO installed on shared storage.
USB and sound devices.	Remove these devices from the virtual machine.
N_Port ID Virtualization (NPIV).	Disable the NPIV configuration of the virtual machine.
NIC passthrough.	This feature is not supported by Fault Tolerance so it must be turned off.
Hot-plugging devices.	<p>The hot plug feature is automatically disabled for fault tolerant virtual machines. To hot plug devices (either adding or removing), you must momentarily turn off Fault Tolerance, perform the hot plug, and then turn on Fault Tolerance.</p> <p>NOTE When using Fault Tolerance, changing the settings of a virtual network card while a virtual machine is running is a hot-plug operation, since it requires "unplugging" the network card and then "plugging" it in again. For example, with a virtual network card for a running virtual machine, if you change the network that the virtual NIC is connected to, FT must be turned off first.</p>
Serial or parallel ports	Remove these devices from the virtual machine.
Video devices that have 3D enabled.	Fault Tolerance does not support video devices that have 3D enabled.
Virtual EFI firmware	Ensure that the virtual machine is configured to use BIOS firmware before installing the guest operating system.
Virtual Machine Communication Interface (VMCI)	Not supported by Fault Tolerance.
2TB+ VMDK	Fault Tolerance is not supported with a 2TB+ VMDK.

Using Fault Tolerance with DRS

You can use vSphere Fault Tolerance with vSphere Distributed Resource Scheduler (DRS) only when the Enhanced vMotion Compatibility (EVC) feature is enabled. This process allows fault tolerant virtual machines to benefit from better initial placement.

When a cluster has EVC enabled, DRS makes the initial placement recommendations for fault tolerant virtual machines and allows you to assign a DRS automation level to Primary VMs (the Secondary VM always assumes the same setting as its associated Primary VM.)

When vSphere Fault Tolerance is used for virtual machines in a cluster that has EVC disabled, the fault tolerant virtual machines are given DRS automation levels of "disabled". In such a cluster, each Primary VM is powered on only on its registered host and its Secondary VM is automatically placed.

If you use affinity rules with a pair of fault tolerant virtual machines, a VM-VM affinity rule applies to the Primary VM only, while a VM-Host affinity rule applies to both the Primary VM and its Secondary VM. If a VM-VM affinity rule is set for a Primary VM, DRS attempts to correct any violations that occur after a failover (that is, after the Primary VM effectively moves to a new host).

Preparing Your Cluster and Hosts for Fault Tolerance

To enable vSphere Fault Tolerance for your cluster, you must meet the feature's prerequisites and you must perform certain configuration steps on your hosts. After those steps are accomplished and your cluster has been created, you can also check that your configuration complies with the requirements for enabling Fault Tolerance.

The tasks you should complete before attempting to set up Fault Tolerance for your cluster include the following:

- Ensure that your cluster, hosts, and virtual machines meet the requirements outlined in the Fault Tolerance checklist.
- Configure networking for each host.
- Create the vSphere HA cluster, add hosts, and check compliance.

After your cluster and hosts are prepared for Fault Tolerance, you are ready to turn on Fault Tolerance for your virtual machines. See [“Turn On Fault Tolerance,”](#) on page 51.

Fault Tolerance Checklist

The following checklist contains cluster, host, and virtual machine requirements that you need to be aware of before using vSphere Fault Tolerance.

Review this list before setting up Fault Tolerance.

NOTE The failover of fault tolerant virtual machines is independent of vCenter Server, but you must use vCenter Server to set up your Fault Tolerance clusters.

Cluster Requirements for Fault Tolerance

You must meet the following cluster requirements before you use Fault Tolerance.

- Fault Tolerance logging and VMotion networking configured. See [“Configure Networking for Host Machines,”](#) on page 48.
- vSphere HA cluster created and enabled. See [“Creating a vSphere HA Cluster,”](#) on page 29. vSphere HA must be enabled before you can power on fault tolerant virtual machines or add a host to a cluster that already supports fault tolerant virtual machines.

Host Requirements for Fault Tolerance

You must meet the following host requirements before you use Fault Tolerance.

- Hosts must use supported processors.
- Hosts must be licensed for Fault Tolerance.
- Hosts must be certified for Fault Tolerance. See <http://www.vmware.com/resources/compatibility/search.php> and select **Search by Fault Tolerant Compatible Sets** to determine if your hosts are certified.
- The configuration for each host must have Hardware Virtualization (HV) enabled in the BIOS.

NOTE VMware recommends that the hosts you use to support FT VMs have their BIOS power management settings turned to "Maximum performance" or "OS-managed performance".

To confirm the compatibility of the hosts in the cluster to support Fault Tolerance, you can also run profile compliance checks as described in [“Create Cluster and Check Compliance,”](#) on page 49.

Virtual Machine Requirements for Fault Tolerance

You must meet the following virtual machine requirements before you use Fault Tolerance.

- No unsupported devices attached to the virtual machine. See [“Fault Tolerance Interoperability,”](#) on page 45.
- Incompatible features must not be running with the fault tolerant virtual machines. See [“Fault Tolerance Interoperability,”](#) on page 45.
- Virtual machine files (except for the VMDK files) must be stored on shared storage. Acceptable shared storage solutions include Fibre Channel, (hardware and software) iSCSI, NFS, and NAS.

Other Configuration Recommendations

You should also observe the following guidelines when configuring Fault Tolerance.

- If you are using NFS to access shared storage, use dedicated NAS hardware with at least a 1Gbit NIC to obtain the network performance required for Fault Tolerance to work properly.
- The memory reservation of a fault tolerant virtual machine is set to the VM's memory size when Fault Tolerance is turned on. Ensure that a resource pool containing fault tolerant VMs has memory resources above the memory size of the virtual machines. Without this excess in the resource pool, there might not be any memory available to use as overhead memory.
- Use a maximum of 16 virtual disks per fault tolerant virtual machine.
- To ensure redundancy and maximum Fault Tolerance protection, you should have a minimum of three hosts in the cluster. In a failover situation, this provides a host that can accommodate the new Secondary VM that is created.

Configure Networking for Host Machines

On each host that you want to add to a vSphere HA cluster, you must configure two different networking switches (vMotion and FT logging) so that the host can support vSphere Fault Tolerance.

To set up Fault Tolerance for a host, you must complete this procedure for each port group option (vMotion and FT logging) to ensure that sufficient bandwidth is available for Fault Tolerance logging. Select one option, finish this procedure, and repeat the procedure a second time, selecting the other port group option.

Prerequisites

Multiple gigabit Network Interface Cards (NICs) are required. For each host supporting Fault Tolerance, a minimum of two physical NICs is recommended. For example, you need one dedicated to Fault Tolerance logging and one dedicated to vMotion. Use three or more NICs to ensure availability.

NOTE The vMotion and FT logging NICs must be on different subnets. If you are using legacy FT, IPv6 is not supported on the FT logging NIC.

Procedure

- 1 In the vSphere Web Client, browse to the host.
- 2 Click the **Configure** tab and click **Networking**.
- 3 Select **VMkernel Network Adapter**.
- 4 Click the **Add host networking** icon.
- 5 Provide appropriate information for your connection type.
- 6 Click **Finish**.

After you create both a vMotion and Fault Tolerance logging virtual switch, you can create other virtual switches, as needed. Add the host to the cluster and complete any steps needed to turn on Fault Tolerance.

What to do next

NOTE If you configure networking to support FT but subsequently suspend the Fault Tolerance logging port, pairs of fault tolerant virtual machines that are powered on remain powered on. If a failover situation occurs, when the Primary VM is replaced by its Secondary VM a new Secondary VM is not started, causing the new Primary VM to run in a Not Protected state.

Create Cluster and Check Compliance

vSphere Fault Tolerance is used in the context of a vSphere HA cluster. After you configure networking on each host, create the vSphere HA cluster and add the hosts to it. You can check to see whether the cluster is configured correctly and complies with the requirements for the enablement of Fault Tolerance.

Procedure

- 1 In the vSphere Web Client, browse to the cluster.
- 2 Click the **Monitor** tab and click **Profile Compliance**.
- 3 Click **Check Compliance Now** to run the compliance tests.

The results of the compliance test appear, and the compliance or noncompliance of each host is shown.

Using Fault Tolerance

After you have taken all of the required steps for enabling vSphere Fault Tolerance for your cluster, you can use the feature by turning it on for individual virtual machines.

Before Fault Tolerance can be turned on, validation checks are performed on a virtual machine.

After these checks are passed and you turn on vSphere Fault Tolerance for a virtual machine, new options are added to the Fault Tolerance section of its context menu. These include turning off or disabling Fault Tolerance, migrating the Secondary VM, testing failover, and testing restart of the Secondary VM.

Validation Checks for Turning On Fault Tolerance

If the option to turn on Fault Tolerance is available, this task still must be validated and can fail if certain requirements are not met.

Several validation checks are performed on a virtual machine before Fault Tolerance can be turned on.

- SSL certificate checking must be enabled in the vCenter Server settings.
- The host must be in a vSphere HA cluster or a mixed vSphere HA and DRS cluster.
- The host must have ESXi 6.x or greater installed (ESX/ESXi 4.x or greater for legacy FT).
- The virtual machine must not have snapshots.
- The virtual machine must not be a template.
- The virtual machine must not have vSphere HA disabled.
- The virtual machine must not have a video device with 3D enabled.

Checks for Powered-On Virtual Machines

Several additional validation checks are performed for powered-on virtual machines (or those that are in the process of being powered on).

- The BIOS of the hosts where the fault tolerant virtual machines reside must have Hardware Virtualization (HV) enabled.
- The host that supports the Primary VM must have a processor that supports Fault Tolerance.
- Your hardware should be certified as compatible with Fault Tolerance. To confirm that it is, use the VMware Compatibility Guide at <http://www.vmware.com/resources/compatibility/search.php> and select **Search by Fault Tolerant Compatible Sets**.
- The configuration of the virtual machine must be valid for use with Fault Tolerance (for example, it must not contain any unsupported devices).

Secondary VM Placement

When your effort to turn on Fault Tolerance for a virtual machine passes the validation checks, the Secondary VM is created. The placement and immediate status of the Secondary VM depends upon whether the Primary VM was powered-on or powered-off when you turned on Fault Tolerance.

If the Primary VM is powered on:

- The entire state of the Primary VM is copied and the Secondary VM is created, placed on a separate compatible host, and powered on if it passes admission control.
- The Fault Tolerance Status displayed for the virtual machine is **Protected**.

If the Primary VM is powered off:

- The Secondary VM is immediately created and registered to a host in the cluster (it might be re-registered to a more appropriate host when it is powered on.)
- The Secondary VM is not powered on until after the Primary VM is powered on.
- The Fault Tolerance Status displayed for the virtual machine is **Not Protected, VM not Running**.
- When you attempt to power on the Primary VM after Fault Tolerance has been turned on, the additional validation checks listed above are performed.

After these checks are passed, the Primary and Secondary VMs are powered on and placed on separate, compatible hosts. The virtual machine's Fault Tolerance Status is tagged as **Protected**.

Turn On Fault Tolerance

You can turn on vSphere Fault Tolerance through the vSphere Web Client.

When Fault Tolerance is turned on, vCenter Server resets the virtual machine's memory limit and sets the memory reservation to the memory size of the virtual machine. While Fault Tolerance remains turned on, you cannot change the memory reservation, size, limit, number of vCPUs, or shares. You also cannot add or remove disks for the VM. When Fault Tolerance is turned off, any parameters that were changed are not reverted to their original values.

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

Prerequisites

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

- The virtual machine resides on a host that does not have a license for the feature.
- The virtual machine resides on a host that is in maintenance mode or standby mode.
- The virtual machine is disconnected or orphaned (its .vmx file cannot be accessed).
- The user does not have permission to turn the feature on.

Procedure

- 1 In the vSphere Web Client, browse to the virtual machine for which you want to turn on Fault Tolerance.
- 2 Right-click the virtual machine 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 virtual machine is designated as a Primary VM, and a Secondary VM is established on another host. The Primary VM is now fault tolerant.

Turn Off Fault Tolerance

Turning off vSphere Fault Tolerance deletes the secondary virtual machine, its configuration, and all history.

Use the **Turn Off Fault Tolerance** option if you do not plan to reenable the feature. Otherwise, use the **Suspend Fault Tolerance** option.

NOTE If the Secondary VM resides on a host that is in maintenance mode, disconnected, or not responding, you cannot use the **Turn Off Fault Tolerance** option. In this case, you should suspend and resume Fault Tolerance instead.

Procedure

- 1 In the vSphere Web Client, browse to the virtual machine for which you want to turn off Fault Tolerance.
- 2 Right-click the virtual machine and select **Fault Tolerance > Turn Off Fault Tolerance**.
- 3 Click **Yes**.

Fault Tolerance is turned off for the selected virtual machine. The history and the secondary virtual machine for the selected virtual machine are deleted.

Suspend Fault Tolerance

Suspending vSphere Fault Tolerance for a virtual machine suspends its Fault Tolerance protection, but preserves the Secondary VM, its configuration, and all history. Use this option to resume Fault Tolerance protection in the future.

Procedure

- 1 In the vSphere Web Client, browse to the virtual machine for which you want to suspend Fault Tolerance.
- 2 Right-click the virtual machine and select **Fault Tolerance > Suspend Fault Tolerance**.
- 3 Click **Yes**.

Fault Tolerance is suspended for the selected virtual machine. Any history and the Secondary VM for the selected virtual machine are preserved and will be used if the feature is resumed.

What to do next

After you suspend Fault Tolerance, to resume the feature select **Resume Fault Tolerance**.

Migrate Secondary

After vSphere Fault Tolerance is turned on for a Primary VM, you can migrate its associated Secondary VM.

Procedure

- 1 In the vSphere Web Client, browse to the Primary VM for which you want to migrate its Secondary VM.
- 2 Right-click the virtual machine and select **Fault Tolerance > Migrate Secondary**.
- 3 Complete the options in the Migrate dialog box and confirm the changes that you made.
- 4 Click **Finish** to apply the changes.

The Secondary VM associated with the selected fault tolerant virtual machine is migrated to the specified host.

Test Failover

You can induce a failover situation for a selected Primary VM to test your Fault Tolerance protection.

This option is unavailable (dimmed) if the virtual machine is powered off.

Procedure

- 1 In the vSphere Web Client, browse to the Primary VM for which you want to test failover.
- 2 Right-click the virtual machine and select **Fault Tolerance > Test Failover**.
- 3 View details about the failover in the Task Console.

This task induces failure of the Primary VM to ensure that the Secondary VM replaces it. A new Secondary VM is also started placing the Primary VM back in a Protected state.

Test Restart Secondary

You can induce the failure of a Secondary VM to test the Fault Tolerance protection provided for a selected Primary VM.

This option is unavailable (dimmed) if the virtual machine is powered off.

Procedure

- 1 In the vSphere Web Client, browse to the Primary VM for which you want to conduct the test.
- 2 Right-click the virtual machine and select **Fault Tolerance > Test Restart Secondary**.
- 3 View details about the test in the Task Console.

This task results in the termination of the Secondary VM that provided Fault Tolerance protection for the selected Primary VM. A new Secondary VM is started, placing the Primary VM back in a Protected state.

Upgrade Hosts Used for Fault Tolerance

Use the following procedure to upgrade hosts used for Fault Tolerance.

Prerequisites

Verify that you have cluster administrator privileges.

Verify that you have sets of four or more ESXi hosts that are hosting fault tolerant virtual machines that are powered on. If the virtual machines are powered off, the Primary and Secondary VMs can be relocated to hosts with different builds.

NOTE This upgrade procedure is for a minimum four-node cluster. The same instructions can be followed for a smaller cluster, though the unprotected interval will be slightly longer.

Procedure

- 1 Using vMotion, migrate the fault tolerant virtual machines off of two hosts.
- 2 Upgrade the two evacuated hosts to the same ESXi build.
- 3 Suspend Fault Tolerance on the Primary VM.
- 4 Using vMotion, move the Primary VM for which Fault Tolerance has been suspended to one of the upgraded hosts.
- 5 Resume Fault Tolerance on the Primary VM that was moved.
- 6 Repeat [Step 1](#) to [Step 5](#) for as many fault tolerant virtual machine pairs as can be accommodated on the upgraded hosts.
- 7 Using vMotion, redistribute the fault tolerant virtual machines.

All ESXi hosts in a cluster are upgraded.

Best Practices for Fault Tolerance

To ensure optimal Fault Tolerance results, you should follow certain best practices.

The following recommendations for host and networking configuration can help improve the stability and performance of your cluster.

Host Configuration

Hosts running the Primary and Secondary VMs should operate at approximately the same processor frequencies, otherwise the Secondary VM might be restarted more frequently. Platform power management features that do not adjust based on workload (for example, power capping and enforced low frequency modes to save power) can cause processor frequencies to vary greatly. If Secondary VMs are being restarted on a regular basis, disable all power management modes on the hosts running fault tolerant virtual machines or ensure that all hosts are running in the same power management modes.

Host Networking Configuration

The following guidelines allow you to configure your host's networking to support Fault Tolerance with different combinations of traffic types (for example, NFS) and numbers of physical NICs.

- Distribute each NIC team over two physical switches ensuring L2 domain continuity for each VLAN between the two physical switches.
- Use deterministic teaming policies to ensure particular traffic types have an affinity to a particular NIC (active/standby) or set of NICs (for example, originating virtual port-id).
- Where active/standby policies are used, pair traffic types to minimize impact in a failover situation where both traffic types will share a vmnic.
- Where active/standby policies are used, configure all the active adapters for a particular traffic type (for example, FT Logging) to the same physical switch. This minimizes the number of network hops and lessens the possibility of oversubscribing the switch to switch links.

NOTE FT logging traffic between Primary and Secondary VMs is unencrypted and contains guest network and storage I/O data, as well as the memory contents of the guest operating system. This traffic can include sensitive data such as passwords in plaintext. To avoid such data being divulged, ensure that this network is secured, especially to avoid 'man-in-the-middle' attacks. For example, you could use a private network for FT logging traffic.

Homogeneous Clusters

vSphere Fault Tolerance can function in clusters with nonuniform hosts, but it works best in clusters with compatible nodes. When constructing your cluster, all hosts should have the following configuration:

- Common access to datastores used by the virtual machines.
- The same virtual machine network configuration.
- The same BIOS settings (power management and hyperthreading) for all hosts.

Run **Check Compliance** to identify incompatibilities and to correct them.

Performance

To increase the bandwidth available for the logging traffic between Primary and Secondary VMs use a 10Gbit NIC, and enable the use of jumbo frames.

You can select multiple NICs for the FT logging network. By selecting multiple NICs, you can take advantage of the bandwidth from multiple NICs even if all of the NICs are not dedicated to running FT.

Store ISOs on Shared Storage for Continuous Access

Store ISOs that are accessed by virtual machines with Fault Tolerance enabled on shared storage that is accessible to both instances of the fault tolerant virtual machine. If you use this configuration, the CD-ROM in the virtual machine continues operating normally, even when a failover occurs.

For virtual machines with Fault Tolerance enabled, you might use ISO images that are accessible only to the Primary VM. In such a case, the Primary VM can access the ISO, but if a failover occurs, the CD-ROM reports errors as if there is no media. This situation might be acceptable if the CD-ROM is being used for a temporary, noncritical operation such as a patch.

Avoid Network Partitions

A network partition occurs when a vSphere HA cluster has a management network failure that isolates some of the hosts from vCenter Server and from one another. See “[Network Partitions](#),” on page 19. When a partition occurs, Fault Tolerance protection might be degraded.

In a partitioned vSphere HA cluster using Fault Tolerance, the Primary VM (or its Secondary VM) could end up in a partition managed by a master host that is not responsible for the virtual machine. When a failover is needed, a Secondary VM is restarted only if the Primary VM was in a partition managed by the master host responsible for it.

To ensure that your management network is less likely to have a failure that leads to a network partition, follow the recommendations in “[Best Practices for Networking](#),” on page 39.

Using Virtual SAN Datastores

vSphere Fault Tolerance can use Virtual SAN datastores, but you must observe the following restrictions:

- A mix of Virtual SAN and other types of datastores is not supported for both Primary VMs and Secondary VMs.
- Virtual SAN metro clusters are not supported with FT.

To increase performance and reliability when using FT with Virtual SAN, the following conditions are also recommended.

- Virtual SAN and FT should use separate networks.
- Keep Primary and Secondary VMs in separate Virtual SAN fault domains.

Legacy Fault Tolerance

By default, vSphere Fault Tolerance can accommodate SMP virtual machines with up to four vCPUs. If your virtual machine has only a single vCPU, however, you can use legacy FT instead for backward compatibility. Unless legacy FT is technically necessary, avoid using it.

To use legacy Fault Tolerance, you must configure an advanced option for the virtual machine. After you complete this configuration, the legacy FT VM is different in some ways from other vSphere FT VMs.

Differences for VMs That Use Legacy FT

VMs that use vSphere FT and VMs that use legacy FT differ in several ways.

Table 3-2. Differences Between Legacy FT and vSphere FT

	Legacy FT	vSphere FT
Extended Page Tables/Rapid Virtualization Indexing (EPT/RVI)	Not supported	Required
IPv6	Not supported for legacy FT logging NICs.	Supported for vSphere FT-logging NICs.
DRS	Fully supported for initial placement, load balancing, and maintenance mode support.	Only power on placement of Secondary VM and maintenance mode are supported.
vStorage APIs - Data Protection backups	Not supported	Supported

Table 3-2. Differences Between Legacy FT and vSphere FT (Continued)

	Legacy FT	vSphere FT
Eager-zeroed thick .vmdk disk files	Required	Not required because vSphere FT supports all disk file types, including thick and thin
.vmdk redundancy	Only a single copy	Primary VMs and Secondary VMs always maintain independent copies, which can be placed on different datastores to increase redundancy.
NIC bandwidth	Dedicated 1-Gb NIC recommended	Dedicated 10-Gb NIC recommended
CPU and host compatibility	Requires identical CPU model and family and nearly identical versions of vSphere on hosts.	CPUs must be compatible with vSphere vMotion or EVC. Versions of vSphere on hosts must be compatible with vSphere vMotion.
Turn on FT on running VM	Not always supported. You might need to power off VM first.	Supported
Storage vMotion	Supported only on powered-off VMs. vCenter Server automatically turns off FT before performing a Storage vMotion action and turns on FT again after the Storage vMotion action completes.	Not supported. User must turn off vSphere FT for the VM before performing the Storage vMotion action and turn on vSphere FT again.
vlance networking drivers	Not supported	Supported

Additional Requirements for Legacy FT

In addition to the differences listed for legacy FT, it also has the following unique requirements.

- ESXi hosts must have access to the same virtual machine datastores and networks.
- Virtual machines must be stored in virtual RDM or virtual machine disk (VMDK) files that are thick provisioned. If a virtual machine is stored in a VMDK file that is thin provisioned and an attempt is made to use fault tolerance, a message appears. It indicates that the VMDK file must be converted. To perform the conversion, you must power off the virtual machine.
- Hosts must have processors from the vSphere FT-compatible processor group. Verify that the hosts' processors are compatible with one another.
- The host that supports the Secondary VM must have a processor that supports fault tolerance and is the same CPU family or model as the host that supports the Primary VM.
- When upgrading hosts that contain fault tolerant VMs, verify that the Primary and Secondary VMs continue to run on hosts with the same FT version number or host build number. This requirement applies to hosts before ESX/ESXi 4.1.

NOTE If you designated a VM to use legacy FT before you upgraded the hosts in the cluster, that VM continues to use legacy FT after the host upgrade.

Upgrading Hosts Involved with Legacy FT

To upgrade your hosts to vSphere 6.5 or later, you must turn off legacy FT on all affected VMs or move those VMs to other hosts. If you do not prepare for the upgrade in this way, VMware vSphere Update Manager blocks the upgrade.

vCenter Server version 6.5 or later can manage existing legacy FT VMs, but you cannot create legacy FT VMs, even on hosts with a version earlier than version 6.5. The following vSphere FT operations can be performed in this scenario:

- Suspend or resume FT
- Test failover
- Restart secondary
- Migrate secondary
- Turn off FT

NOTE Legacy FT VMs can exist only on ESXi hosts that are running on vSphere versions earlier than 6.5.

vCenter High Availability

vCenter High Availability (vCenter HA) protects vCenter Server Appliance against host and hardware failures. The active-passive architecture of the solution can also help you reduce downtime significantly when you patch vCenter Server Appliance.

After some network configuration, you create a three-node cluster that contains Active, Passive, and Witness nodes. Different configuration paths are available. What you select depends on your existing configuration.

1 [Plan the vCenter HA Deployment](#) on page 60

Before you can configure vCenter HA, you have to consider several factors. A vCenter Server Appliance deployment can use an internal or external Platform Services Controller. A brownfield deployment with components that use different versions of vSphere requires different considerations than a greenfield deployment that includes only vSphere 6.5 components. Resource and software requirements and the networking setup must also be considered carefully.

2 [Configure the Network](#) on page 65

Regardless of the deployment option and inventory hierarchy that you select, you have to set up your network before you can start configuration. To set the foundation for the vCenter HA network, you add a port group to each ESXi host, and add a virtual NIC to the vCenter Server Appliance that later becomes the Active node.

3 [Configure vCenter HA With the Basic Option](#) on page 66

When you use the Basic option, the vCenter HA wizard creates and configures a second network adapter on the vCenter Server Appliance, clones the Active node, and configures the vCenter HA network.

4 [Configure vCenter HA With the Advanced Option](#) on page 67

If you configure the vCenter HA cluster with the Advanced option, you have more control over the environment, and you do not have to meet some of the prerequisites for the Basic configuration. However, you are responsible adding a second NIC to the vCenter Server Appliance, cloning the Active node to the Passive and Witness nodes, and configuring the clones.

5 [Manage the vCenter HA Configuration](#) on page 70

After you configure your vCenter HA cluster, you can perform management tasks. These tasks include certificate replacement, replacement of SSH keys, and SNMP setup. You can also edit the cluster configuration to disable or enable vCenter HA, enter maintenance mode, and remove the cluster configuration.

6 [Troubleshoot Your vCenter HA Environment](#) on page 75

In case of problems you can troubleshoot your environment. The task you need to perform depends on the failure symptoms. For additional troubleshooting information, see the VMware Knowledge Base system.

7 [Patching a vCenter High Availability Environment](#) on page 78

You can patch a vCenter Server Appliance which is in a vCenter High Availability cluster by using the `software-packages` utility available in the vCenter Server Appliance shell. For more information, see *vSphere Upgrade*.

Plan the vCenter HA Deployment

Before you can configure vCenter HA, you have to consider several factors. A vCenter Server Appliance deployment can use an internal or external Platform Services Controller. A brownfield deployment with components that use different versions of vSphere requires different considerations than a greenfield deployment that includes only vSphere 6.5 components. Resource and software requirements and the networking setup must also be considered carefully.

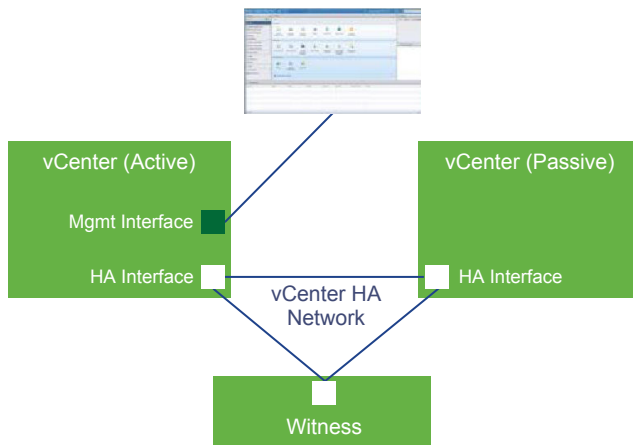
vCenter Architecture Overview

A vCenter HA cluster consists of three vCenter Server Appliance instances. The first instance, initially used as the Active node, is cloned twice to a Passive node and to a Witness node. Together, the three nodes provide an active-passive failover solution.

Deploying each of the nodes on a different ESXi instance protects against hardware failure. Adding the three ESXi hosts to a DRS cluster can further protect your environment.

When vCenter HA configuration is complete, only the Active node has an active management interface (public IP). The three nodes communicate over a private network called vCenter HA network that is set up as part of configuration. The Active node and the Passive node are continuously replicating data.

Figure 4-1. vCenter Three-Node Cluster



All three nodes are necessary for the functioning of this feature. Compare the node responsibilities.

Table 4-1. vCenter HA Nodes

Node	Description
Active	<ul style="list-style-type: none"> ■ Runs the active vCenter Server Appliance instance ■ Uses a public IP address for the management interface ■ Uses the vCenter HA network for replication of data to the Passive node. ■ Uses the vCenter HA network to communicate with the Witness node.
Passive	<ul style="list-style-type: none"> ■ Is initially a clone of the Active node ■ Constantly receives updates from and synchronizes state with the Active node over the vCenter HA network ■ Automatically takes over the role of the Active node if a failure occurs
Witness	<ul style="list-style-type: none"> ■ Is a lightweight clone of the Active node ■ Provides a quorum to protect against a split-brain situations

vCenter HA Hardware and Software Requirements

Before you set up vCenter HA, ensure that you have sufficient memory, CPU, and datastore resources, and ensure that you are using versions of vCenter Server and ESXi that support vCenter HA.

Your environment must meet the following requirements.

Table 4-2. vCenter HA Requirements

Component	Requirements
ESXi	<ul style="list-style-type: none"> ■ ESXi 5.5 or later is required. ■ Three hosts are strongly recommended. Each vCenter HA node can then run on a different host for better protection. ■ Using VMware DRS to protect the set of hosts is recommended. In that case, a minimum of three ESXi hosts is required.
Management vCenter Server (if used)	<p>Your environment can include a management vCenter Server system, or you can set up your vCenter Server Appliance to manage the ESXi host on which it runs (self-managed vCenter Server)</p> <ul style="list-style-type: none"> ■ vCenter Server 5.5 or later is required.
vCenter Server Appliance	<ul style="list-style-type: none"> ■ vCenter Server 6.5 is required. ■ Deployment size Small (4 CPU and 16GB RAM) or bigger is required to meet the RTO. Do not use Tiny in production environments. ■ vCenter HA is supported and tested with VMFS, NFS, and Virtual SAN datastores. ■ Ensure you have enough disk space to collect and store support bundles for all three nodes on the Active node. See “Collecting Support Bundles for a vCenter HA Node,” on page 75.
Network connectivity	<ul style="list-style-type: none"> ■ vCenter HA network latency between Active, Passive, and Witness nodes must be less than 10 ms. ■ The vCenter HA network must be on a different subnet than the management network.
Licensing required for vCenter HA	<ul style="list-style-type: none"> ■ vCenter HA requires a single vCenter Server license. ■ vCenter HA requires a Standard license.

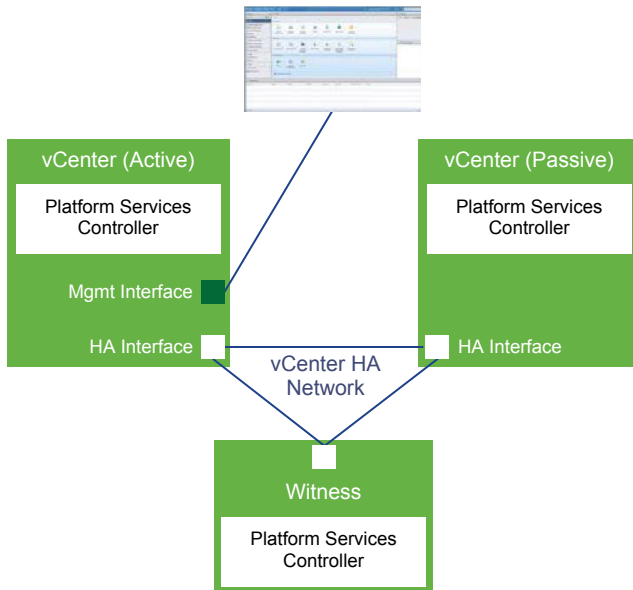
vCenter HA Deployment Options

You can set up your vCenter HA environment with an embedded Platform Services Controller or with an external Platform Services Controller. If you decide to use an external Platform Services Controller, you can place it behind a load balancer for protection in case of Platform Services Controller failure.

vCenter HA with an Embedded Platform Services Controller

When you use vCenter HA with an embedded Platform Services Controller, the environment setup is as follows.

Figure 4-2. vCenter HA with an Embedded Platform Services Controller



- 1 The user provisions the vCenter Server Appliance with an embedded Platform Services Controller.
- 2 Cloning of the vCenter Server Appliance to a Passive and a Witness node occurs.
 - In a Basic configuration, the configuration creates and configures the clones.
 - In an Advanced configuration, the user creates and configures the clones.
- 3 As part of the clone process, Platform Services Controller and all its services are cloned as well.
- 4 When configuration is complete, vCenter HA performs replication to ensure that the Passive node is synchronized with the Active node. The Active node to Passive node replication includes Platform Services Controller data.
- 5 When configuration is complete, the vCenter Server Appliance is protected by vCenter HA. In case of failover, Platform Services Controller and all its services are available on the Passive node.

vCenter HA with an External Platform Services Controller

When you use vCenter HA with an external Platform Services Controller, you must set up an external load balancer to protect the Platform Services Controller. If one Platform Services Controller becomes unavailable, the load balancer directs the vCenter Server Appliance to a different Platform Services Controller.

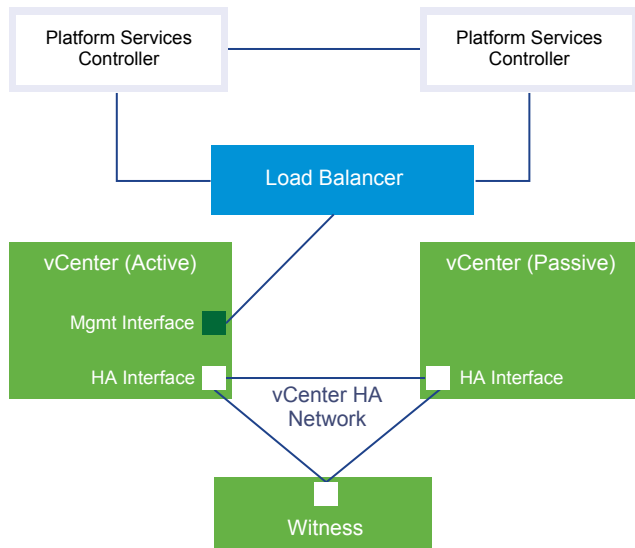
Set up of the external Platform Services Controller is discussed in the following VMware Knowledge Base articles.

- [2147014](#): Configuring Netscaler Load Balancer for use with vSphere Platform Services Controller (PSC) 6.5

- [2147038](#) Configuring F5 BIG-IP Load Balancer for use with vSphere Platform Services Controller (PSC) 6.5
- [2147046](#) Configuring NSX Edge Load Balancer for use with vSphere Platform Services Controller (PSC) 6.5

The environment setup is as follows.

Figure 4-3. vCenter HA with External Platform Services Controller



- 1 The user sets up at least two external Platform Services Controller instances. These instances replicate vCenter Single Sign-On information and other Platform Services Controller information, for example, licensing.
- 2 During provisioning of the vCenter Server Appliance, the user selects an external Platform Services Controller.
- 3 The user sets up the vCenter Server Appliance to point to a load balancer that provides high availability for Platform Services Controller.
- 4 The user or the Basic configuration clones the first vCenter Server Appliance to create the Passive node and Witness node.
- 5 As part of the clone process, the information about the external Platform Services Controller and the load balancer is cloned as well.
- 6 When configuration is complete, the vCenter Server Appliance is protected by vCenter HA.
- 7 If the Platform Services Controller instance becomes unavailable, the load balancer redirects requests for authentication or other services to the second Platform Services Controller instance.

Configuration Workflow Overview

You can select a Basic or Advanced configuration option. The Basic option automatically creates the Passive and Witness nodes as part of vCenter HA configuration. With the Advanced option, you are responsible for manually cloning the Active node to create the Passive and Witness nodes.

The configuration option that you select depends on your environment. The Basic configuration requirements are stricter, but more of the configuration is automated. The Advanced configuration is possible if your environment meets hardware and software requirements, and it offers more flexibility. However, Advanced configuration requires that you create and configure the clones of the Active node.

Basic Configuration Workflow

Basic configuration automatically clones the Active node. You must meet one of the following requirements to perform Basic configuration.

- Either the vCenter Server Appliance that will become the Active node is managing its own ESXi host and its own virtual machine. This configuration is sometimes called a self-managed vCenter Server.
- Or the vCenter Server Appliance is managed by another vCenter Server (management vCenter Server) and both vCenter Server instances are in the same vCenter Single Sign-On domain. That means they both use an external Platform Services Controller and both are running vSphere 6.5.

If you meet the requirements the Basic workflow is as follows.

- 1 The user deploys the first vCenter Server Appliance, which will become the Active node.
- 2 The user adds a second network (port group) for vCenter HA traffic on each ESXi host.
- 3 The user starts the vCenter HA configuration, selects Basic and supplies the IP addresses, the target ESXi host or cluster, and the datastore for each clone.
- 4 The system clones the Active node and creates a Passive node with precisely the same settings, including the same host name.
- 5 The system clones the Active node again and creates a more light-weight Witness node.
- 6 The system sets up the vCenter HA network on which the three nodes communicate, for example, by exchanging heartbeats and other information.

For step-by-step instructions, see [“Configure vCenter HA With the Basic Option,”](#) on page 66.

Advanced Configuration Workflow

If you cannot select the Basic option or you want more control over your deployment, you can perform Advanced configuration. With this option, you are responsible for cloning the Active node yourself as part of vCenter HA setup. If you select this option and remove the vCenter HA configuration later, you are responsible for deleting the nodes that you created.

For the Advanced option, the workflow is as follows.

- 1 The user deploys the first vCenter Server Appliance, which will become the Active node.
- 2 The user adds a second network (port group) for vCenter HA traffic on each ESXi host.
- 3 The user adds a second network adapter (NIC) to the Active node
- 4 The user logs in to the vCenter Server Appliance (Active node) with the vSphere Web Client.
- 5 The user starts the vCenter HA configuration, selects Advanced, and supplies IP address and subnet information for the Passive and Witness nodes. Optionally, the user can override the failover management IP addresses.
- 6 The user logs in to the management vCenter Server and creates two clones of the vCenter Server Appliance (Active node).
- 7 The user returns to the configuration wizard on the vCenter Server Appliance and completes the configuration process.
- 8 The system sets up the vCenter HA network on which the three nodes exchange heartbeats and replication information.
- 9 The vCenter Server Appliance is protected by vCenter HA.

See [“Configure vCenter HA With the Advanced Option,”](#) on page 67 for details.

Configure the Network

Regardless of the deployment option and inventory hierarchy that you select, you have to set up your network before you can start configuration. To set the foundation for the vCenter HA network, you add a port group to each ESXi host, and add a virtual NIC to the vCenter Server Appliance that later becomes the Active node.

After configuration is complete, the vCenter HA cluster has two networks, the management network on the first virtual NIC and the vCenter HA network on the second virtual NIC.

Management network	The management network serves client requests (public IP). The management network IP addresses must be static.
vCenter HA network	<p>The vCenter HA network connects the Active, Passive, and Witness nodes and replicates the appliance state. It also monitors heartbeats.</p> <ul style="list-style-type: none"> ■ The vCenter HA network IP addresses for the Active, Passive, and Witness nodes must be static. ■ The vCenter HA network must be on a different subnet than the management network. The three nodes can be on the same subnet or on different subnets. ■ Network latency between the Active, Passive, and Witness nodes must be less than 10 milliseconds. ■ You must not add a default gateway entry for the cluster network.

Prerequisites

- The vCenter Server Appliance that later becomes the Active node, is deployed.
- You can access and have privileges to modify that vCenter Server Appliance and the ESXi host on which it runs.
- During network setup, you need static IP addresses for the management network. The management and cluster network addresses must be IPv4 or IPv6. They cannot be mixed

Procedure

- 1 Log in to the management vCenter Server and find the ESXi host on which the Active node is running.
- 2 Add a port group to the ESXi host.

This port group can be on an existing virtual switch or, for improved network isolation, you can create a new virtual switch. It must be on a different subnet than the management network on Eth0.

- 3 If your environment includes the recommended three ESXi hosts, add the port group to each of the hosts.

What to do next

What you do next depends on the type of configuration you select.

- With a Basic configuration, the wizard creates the vCenter HA virtual NIC on each clone and sets up the vCenter HA network. When configuration completes, the vCenter HA network is available for replication and heartbeat traffic.
- With an Advanced configuration.
 - You have to first create and configure a second NIC on the Active node. See [“Create and Configure a Second NIC on the vCenter Server Appliance,”](#) on page 67.
 - When you perform the configuration, the wizard prompts for the IP addresses for the Passive and Witness nodes.

- The wizard prompts you to clone the Active node. As part of the clone process, you perform additional network configuration.

See [“Configure vCenter HA With the Advanced Option,”](#) on page 67.

Configure vCenter HA With the Basic Option

When you use the Basic option, the vCenter HA wizard creates and configures a second network adapter on the vCenter Server Appliance, clones the Active node, and configures the vCenter HA network.

Prerequisites

- Deploy vCenter Server Appliance that you want to use as the initial Active node.
 - The vCenter Server Appliance must have a static IP address mapped to an FQDN.
 - SSH must be enabled on the vCenter Server Appliance.
- Verify that your environment meets one of the following requirements.
 - Either the vCenter Server Appliance that will become the Active node is managing its own ESXi host and its own virtual machine. This configuration is sometimes called a self-managed vCenter Server.
 - Or the vCenter Server Appliance is managed by another vCenter Server (management vCenter Server) and both appliances are in the same vCenter Single Sign-On domain. That means they both use an external Platform Services Controller and both are running vSphere 6.5.

If your environment does not meet one of these requirements, perform an Advanced configuration. See [“Configure vCenter HA With the Advanced Option,”](#) on page 67.

- Set up the infrastructure for the vCenter HA network. See [“Configure the Network,”](#) on page 65.
- Determine which static IP addresses to use for the two vCenter Server Appliance nodes that will become the Passive node and Witness node.

Procedure

- 1 Log in to the Active node with the vSphere Web Client.
- 2 Right-click the vCenter Server object in the inventory and select **vCenter HA Settings**.
- 3 Click **Configure**.
- 4 Select the **Basic** configuration option and click **Next**.

This option is available only if your environment meets prerequisites for the Basic option.

- 5 Specify the IP address, subnet mask for the Active node and the port group to connect to the vCenter HA network and click **Next**.
- 6 Provide the vCenter HA network IP address and subnet mask for the Passive node and the Witness node and click **Next**.

The configuration wizard needs the addresses to create the vCenter HA network and to connect the three nodes.

- 7 (Optional) Click **Advanced** if you want to override the failover management IP address for the Passive node.
- 8 Review the information for the Passive and Witness nodes, click **Edit** to make changes, and click **Next**.

If you are not using a DRS cluster, select different hosts and datastores for the Passive and Witness nodes if possible.

- 9 Click **Finish**.

The Passive and Witness nodes are created. When vCenter HA configuration is complete, vCenter Server Appliance has high availability protection.

What to do next

See [“Manage the vCenter HA Configuration,”](#) on page 70 for a list of cluster management tasks.

Configure vCenter HA With the Advanced Option

If you configure the vCenter HA cluster with the Advanced option, you have more control over the environment, and you do not have to meet some of the prerequisites for the Basic configuration. However, you are responsible adding a second NIC to the vCenter Server Appliance, cloning the Active node to the Passive and Witness nodes, and configuring the clones.

Procedure

- 1 [Create and Configure a Second NIC on the vCenter Server Appliance](#) on page 67
Before you can start the Advanced configuration, you have to create and configure a second NIC on the vCenter Server Appliance that will become the active node. This NIC will be used for vCenter HA traffic. You perform this task after you set up the network, but before you start the configuration process.
- 2 [Start the Advanced Configuration Process](#) on page 68
After you configure the network and you add a second NIC to the vCenter Server Appliance, you can start the vCenter HA configuration process.
- 3 [Create and Configure the Clones of the Active Node](#) on page 68
As part of the Advanced configuration, you have to clone the Active node to create the Passive and Witness nodes. Do not exit the Configure vCenter HA wizard while you perform the cloning tasks.
- 4 [Complete the vCenter HA Advanced Configuration](#) on page 69
After you create the Passive and Witness nodes, return to the configuration wizard on the Active node to complete the configuration.

Create and Configure a Second NIC on the vCenter Server Appliance

Before you can start the Advanced configuration, you have to create and configure a second NIC on the vCenter Server Appliance that will become the active node. This NIC will be used for vCenter HA traffic. You perform this task after you set up the network, but before you start the configuration process.

Prerequisites

- Set up the infrastructure for the vCenter HA network. See [“Configure the Network,”](#) on page 65.
- Deploy vCenter Server Appliance that you want to use as initial Active node.
 - The vCenter Server Appliance must have a static IP address mapped to an FQDN.
 - SSH must be enabled on the vCenter Server Appliance.

Procedure

- 1 Log in to the management vCenter Server with the vSphere Web Client.
- 2 Select the vCenter Server Appliance virtual machine (Active node), add a second network adapter, and attach it to the vCenter HA portgroup that you created.

- 3 Log in to the vCenter Server Appliance that will initially become the Active node, directly.

Interface	Action
vCenter Server Appliance	Go to <code>https://appliance-IP-address-or-FQDN:5480</code>
vSphere Web Client	<ol style="list-style-type: none"> Go to <code>https://appliance-IP-address-or-FQDN/vsphere-client</code> Select Administration > System Configuration

- 4 Configure the IP settings for the second network adapter.

Start the Advanced Configuration Process

After you configure the network and you add a second NIC to the vCenter Server Appliance, you can start the vCenter HA configuration process.

Prerequisites

- Deploy vCenter Server Appliance that you want to use as initial Active node.
 - The vCenter Server Appliance must have a static IP address mapped to an FQDN.
 - SSH must be enabled on the vCenter Server Appliance.
- Configure the network. See [“Configure the Network,”](#) on page 65.
- Determine which static IP addresses to use for the two vCenter Server Appliance nodes that will become the Passive node and Witness node.

Procedure

- 1 Log in to the Active node with the vSphere Web Client.
- 2 Right-click the vCenter Server object in the inventory and select **vCenter HA Settings**.
- 3 Click **Configure**.
- 4 Select the **Advanced** configuration option and click **Next**.
- 5 Provide the IP address and subnet mask for the Passive and Witness nodes click **Next**.
You have to specify these IP addresses now even though the nodes do not exist yet. You can no longer change these IP addresses after you click **Next**.
- 6 (Optional) Click **Advanced** if you want to override the failover management IP address for the Passive node.
- 7 Leave the wizard window open and perform the cloning tasks.

What to do next

[“Create and Configure the Clones of the Active Node,”](#) on page 68.

Create and Configure the Clones of the Active Node

As part of the Advanced configuration, you have to clone the Active node to create the Passive and Witness nodes. Do not exit the Configure vCenter HA wizard while you perform the cloning tasks.

Procedure

- 1 Log in to the management vCenter Server, right-click the vCenter Server Appliance virtual machine (Active node), and select **Clone > Clone to Virtual Machine**.

- 2 For the first clone, which will become the Passive node, enter the following values.

Option	Value
New Virtual Machine Name	Name of the Passive node. For example, use vcsa-peer.
Select Compute Resource Select Storage	Use a different target host and datastore than for the Active node if possible.
Clone Options	Select the Customize the operating system and Power on virtual machine after creation check boxes and click the New Customization Spec icon on the next page. In the New Customization Spec wizard that appears specify the following. <ol style="list-style-type: none"> Use the same host name as the Active node. Ensure the timezone is consistent with the Active node. On the Configure Network page, specify the IP settings for NIC1 and NIC2, which map to the management interface and the vCenter HA interface. Leave the NIC2 Default Gateway blank.

- 3 After the first clone has been created, clone the Active node again for the Witness node.

Option	Value
New Virtual Machine Name	Name of the Witness node. For example, use vcsa-witness.
Select Compute Resource Select Storage	Use a different target host and datastore than for the Active and Passive nodes if possible.
Clone Options	Select the Customize the operating system and Power on virtual machine after creation check boxes and click the New Customization Spec icon on the next page. In the New Customization Spec wizard that appears specify the following. <ol style="list-style-type: none"> Use the host name of your choice. Ensure the timezone is consistent with the Active node. On the Configure Network page, specify the IP settings for NIC2, which maps to the vCenter HA interface. Leave the NIC2 Default Gateway blank.

- 4 Ensure that the clone process completes and the virtual machines are powered on.

What to do next

Return to the vCenter HA wizard on the Active node to complete the setup. See [“Complete the vCenter HA Advanced Configuration,”](#) on page 69.

Complete the vCenter HA Advanced Configuration

After you create the Passive and Witness nodes, return to the configuration wizard on the Active node to complete the configuration.

Prerequisites

Complete the process of cloning the Active node to a Passive node and a Witness node.

Procedure

- Return to the configuration wizard and click **Finish**.
- Wait for vCenter HA setup to complete.

Manage the vCenter HA Configuration

After you configure your vCenter HA cluster, you can perform management tasks. These tasks include certificate replacement, replacement of SSH keys, and SNMP setup. You can also edit the cluster configuration to disable or enable vCenter HA, enter maintenance mode, and remove the cluster configuration.

- [Set Up SNMP Traps](#) on page 71
You can set up Simple Network Management Protocol (SNMP) traps to receive SNMP notifications for your vCenter HA cluster.
- [Set Up Your Environment to Use Custom Certificates](#) on page 71
The machine SSL certificate on each node is used for cluster management communication and for encryption of replication traffic. If you want to use custom certificates, you have to remove the vCenter HA configuration, delete the Passive and Witness nodes, provision the Active node with the custom certificate, and reconfigure the cluster.
- [Manage vCenter HA SSH Keys](#) on page 72
vCenter HA uses SSH keys for password-less authentication between the Active, Passive, and Witness nodes. The authentication is used for heartbeat exchange and file and data replication. To replace the SSH keys in the nodes of a vCenter HA cluster, you disable the cluster, generate new SSH keys on the Active node, transfer the keys to the passive node, and enable the cluster.
- [Initiate a vCenter HA Failover](#) on page 72
You can manually initiate a failover and have the Passive node become the Active node.
- [Edit the vCenter HA Cluster Configuration](#) on page 72
When you edit the vCenter HA cluster configuration, you can disable or enable the cluster, place the cluster in maintenance mode, or remove the cluster.
- [Perform Backup and Restore Operations](#) on page 73
For additional security, you can back up the Active node in the vCenter HA cluster. You can then restore the node in case of catastrophic failure.
- [Remove a vCenter HA Configuration](#) on page 74
You can remove a vCenter HA configuration from the vSphere Web Client. If you are using an Advanced configuration, or if any of the nodes are not discoverable, you might have to perform additional cleanup steps.
- [Reboot All vCenter HA Nodes](#) on page 74
If you have to shut down and reboot all nodes in the cluster, you must follow a specific shutdown order to prevent the Passive node from assuming the role of Active node.
- [Change the Appliance Environment](#) on page 75
When you deploy a vCenter Server Appliance, you select an environment. For vCenter HA, Small, Medium, Large, and X-Large are supported for production environments. If you need more space and want to change the environment, you have to delete the Passive node virtual machine before you change the configuration.
- [Collecting Support Bundles for a vCenter HA Node](#) on page 75
Collecting a support bundle from all the nodes in a vCenter HA cluster helps with troubleshooting.

Set Up SNMP Traps

You can set up Simple Network Management Protocol (SNMP) traps to receive SNMP notifications for your vCenter HA cluster.

The traps default to SNMP version 1.

Set up SNMP traps for the Active node and the Passive node. You tell the agent where to send related traps, by adding a target entry to the `snmpd` configuration.

Procedure

- 1 Log in to the Active node by using the Virtual Machine Console or SSH.
- 2 Run the `vicfg-snmp` command, for example:

```
vicfg-snmp -t 10.160.1.1@1166/public
```

In this example, `10.160.1.1` is the client listening address, `1166` is the client listening port, and `public` is the community string.

- 3 Enable the SNMP agent (`snmpd`) by running the following command.

```
vicfg-snmp -e
```

What to do next

You also might find these commands useful.

- To view the complete help for the command, run `vicfg-snmp -h`.
- To disable the SNMP agent, run `vicfg-snmp -D`.
- To show the SNMP agent's configuration, run `vicfg-snmp -s`.
- To reset the configuration to the default, run `vicfg-snmp -r`.

Set Up Your Environment to Use Custom Certificates

The machine SSL certificate on each node is used for cluster management communication and for encryption of replication traffic. If you want to use custom certificates, you have to remove the vCenter HA configuration, delete the Passive and Witness nodes, provision the Active node with the custom certificate, and reconfigure the cluster.

If possible, replace certificates in the vCenter Server Appliance that will become the Active node before you clone the node.

Procedure

- 1 Edit the cluster configuration and select **Remove**.
- 2 Delete the Passive node and the Witness node.
- 3 On the Active node, which is now a standalone vCenter Server Appliance, replace the machine SSL Certificate with a custom certificate.

See the *Platform Services Controller Administration* documentation.

- 4 Reconfigure the cluster.

Manage vCenter HA SSH Keys

vCenter HA uses SSH keys for password-less authentication between the Active, Passive, and Witness nodes. The authentication is used for heartbeat exchange and file and data replication. To replace the SSH keys in the nodes of a vCenter HA cluster, you disable the cluster, generate new SSH keys on the Active node, transfer the keys to the passive node, and enable the cluster.

Procedure

- 1 Edit the cluster and change the mode to **Disabled**.
- 2 Log in to the Active node by using the Virtual Machine Console or SSH.
- 3 Enable the bash shell.


```
bash
```
- 4 Run the following command to generate new SSH keys on the Active node.


```
/usr/lib/vmware-vmtoolsd/scripts/resetSshKeys.py
```
- 5 Use SCP to copy the keys to the Passive node and Witness node.


```
scp /vcha/.ssh/*
```
- 6 Edit the cluster configuration and set the vCenter HA cluster to **Enabled**.

Initiate a vCenter HA Failover

You can manually initiate a failover and have the Passive node become the Active node.

A vCenter HA cluster supports two types of failover.

Automatic failover	The Passive node attempts to take over the active role in case of an Active node failure.
Manual failover	The user can force a Passive node to take over the active role by using the Initiate Failover action.

Initiate a manual failover for troubleshooting and testing.

Procedure

- 1 Log in to the Active node vCenter Server Appliance with the vSphere Web Client and click **Configure**.
- 2 Under **Settings** select **vCenter HA** and click **Initiate Failover**.
- 3 Click **Yes** to start the failover.

A dialog offers you the option to force a failover without synchronization. In most cases, performing synchronization first is best.

- 4 After the failover, you can verify that the Passive node has the role of the Active node in the vSphere Web Client.

Edit the vCenter HA Cluster Configuration

When you edit the vCenter HA cluster configuration, you can disable or enable the cluster, place the cluster in maintenance mode, or remove the cluster.

The operating mode of a vCenter Server Appliance controls the failover capabilities and state replication in a vCenter HA cluster.

A vCenter HA cluster can operate in one of the following modes.

Table 4-3. vCenter HA Cluster Modes of Operation

Mode	Automatic Failover	Manual Failover	Replication	
Enabled	Yes	Yes	Yes	This default mode of operation protects the vCenter Server Appliance from hardware and software failures by performing automatic failover.
Maintenance	No	Yes	Yes	Used for some maintenance tasks. For other tasks, you have to disable vCenter HA.
Disabled	No	No	No	If the Passive or Witness nodes are lost or recovering from a failure, a vCenter HA configuration can be disabled. The Active node continues as a standalone vCenter Server Appliance.

NOTE If the cluster is operating in either Maintenance or Disabled mode, an Active node can continue serving client requests even if the Passive and Witness nodes are lost or unreachable.

Prerequisites

Verify that the vCenter HA cluster is deployed and contains the Active, Passive, and Witness nodes.

Procedure

- 1 Log in to the Active node vCenter Server Appliance with the vSphere Web Client and click **Configure**.
- 2 Under **Settings** select **vCenter HA** and click **Edit**.
- 3 Select one of the options.

Option	Result
Enable vCenter HA	Enables replication between the Active and Passive nodes. If the cluster is in a healthy state, your Active node is protected by automatic failover from the Passive node.
Maintenance Mode	In maintenance mode, replication still occurs between the Active and Passive nodes. However, automatic failover is disabled.
Disable vCenter HA	Disables replication and failover. Keeps the configuration of the cluster. You can later enable vCenter HA again.
Remove vCenter HA cluster	Removes the cluster. Replication and failover no longer are provided. The Active node continues to operate as a standalone vCenter Server Appliance. See “Remove a vCenter HA Configuration,” on page 74 for details.

- 4 Click **OK**.

Perform Backup and Restore Operations

For additional security, you can back up the Active node in the vCenter HA cluster. You can then restore the node in case of catastrophic failure.

NOTE Remove the cluster configuration before you restore the Active node. Results are unpredictable if you restore the Active node and the Passive node is still running or other cluster configuration is still in place.

Prerequisites

Verify the interoperability of vCenter HA and the backup and restore solution. One solution is vCenter Server Appliance file-based restore.

Procedure

- 1 Back up the Active node.
Do not back up the Passive node and Witness node.
- 2 Before you restore the cluster, power off and delete all vCenter HA nodes.
- 3 Restore the Active node.
The Active node is restored as a standalone vCenter Server Appliance.
- 4 Reconfigure the vCenter HA.

Remove a vCenter HA Configuration

You can remove a vCenter HA configuration from the vSphere Web Client. If you are using an Advanced configuration, or if any of the nodes are not discoverable, you might have to perform additional cleanup steps.

Procedure

- 1 Log in to the Active node vCenter Server Appliance and click **Configure**.
- 2 Under **Settings** select **vCenter HA** and click **Edit**.
- 3 Select **Remove vCenter HA cluster**.
 - The vCenter HA cluster's configuration is removed from the Active, Passive, and Witness nodes.
 - The Active node continues to run as a standalone vCenter Server Appliance.
 - You cannot reuse the Passive and Witness nodes in a new vCenter HA configuration.
 - If you performed configuration using the Advanced options, or if the Passive and Witness nodes are not discoverable, you must delete these nodes explicitly.
 - Even if the second virtual NIC was added by the configuration process, the removal process does not remove the virtual NIC.

Reboot All vCenter HA Nodes

If you have to shut down and reboot all nodes in the cluster, you must follow a specific shutdown order to prevent the Passive node from assuming the role of Active node.

Procedure

- 1 Shut down the nodes in this order.
 - Passive node
 - Active node
 - Witness node
- 2 Restart each node.
You can restart nodes in any order.
- 3 Verify that all nodes join the cluster successfully, and that the previous Active node resumes that role.

Change the Appliance Environment

When you deploy a vCenter Server Appliance, you select an environment. For vCenter HA, Small, Medium, Large, and X-Large are supported for production environments. If you need more space and want to change the environment, you have to delete the Passive node virtual machine before you change the configuration.

Procedure

- 1 Log in to the Active node with the vSphere Web Client, edit the cluster configuration, and select **Disable**.
- 2 Delete the Passive node virtual machine.
- 3 Change the vCenter Server Appliance configuration for the Active node, for example, from a Small environment to a Medium environment.
- 4 Reconfigure vCenter HA.

Collecting Support Bundles for a vCenter HA Node

Collecting a support bundle from all the nodes in a vCenter HA cluster helps with troubleshooting.

When you collect a support bundle from the Active node in a vCenter HA cluster, the system proceeds as follows.

- Collects support bundle information from the Active node itself.
- Collects support bundles from Passive and Witness nodes and places them in the `commands` directory on the Active node support bundle.

NOTE The collection of support bundles from the Passive and Witness nodes is a best effort and happens if the nodes are reachable.

Troubleshoot Your vCenter HA Environment

In case of problems you can troubleshoot your environment. The task you need to perform depends on the failure symptoms. For additional troubleshooting information, see the VMware Knowledge Base system.

- [vCenter HA Clone Operation Fails During Deployment](#) on page 76
If the vCenter HA configuration process does not create the clones successfully, you have to resolve that cloning error.
- [vCenter HA Deployment Fails with an Error](#) on page 76
Deployment failures can be caused by configuration issues, especially problems with the networking setup.
- [Troubleshooting a Degraded vCenter HA Cluster](#) on page 76
For a vCenter HA cluster to be healthy, each of the Active, Passive, and Witness nodes must be fully operational and be reachable over the vCenter HA cluster network. If any of the nodes fails, the cluster is considered to be in a degraded state.
- [Recovering from Isolated vCenter HA Nodes](#) on page 77
If all nodes in a vCenter HA cluster cannot communicate with each other, the Active node stops serving client requests.
- [Resolving Failover Failures](#) on page 78
When a Passive node does not become the Active node during a failover, you can force the Passive node to become the Active node.

vCenter HA Clone Operation Fails During Deployment

If the vCenter HA configuration process does not create the clones successfully, you have to resolve that cloning error.

Problem

Clone operation fails.

Cause

Look for the clone exception. It might indicate one of the following problems.

- You have a DRS-enabled cluster, but do not have three hosts.
- The host or database connection is lost.
- Not enough disk space.
- Other **Clone Virtual Machine** errors

Solution

- 1 Resolve the error that caused the problem.
- 2 Remove the cluster and start configuration again.

vCenter HA Deployment Fails with an Error

Deployment failures can be caused by configuration issues, especially problems with the networking setup.

Problem

You start a vCenter HA cluster configuration, and it fails with an error. The error might show the cause of the problem, for example, you might see an SSH Connection Failed message.

Solution

If deployment fails, take steps to resolve the network issues.

- 1 Verify that the Passive and Witness nodes can be reached from the Active node.
- 2 Verify that routing between the nodes is set up correctly.
- 3 Check network latency.

Troubleshooting a Degraded vCenter HA Cluster

For a vCenter HA cluster to be healthy, each of the Active, Passive, and Witness nodes must be fully operational and be reachable over the vCenter HA cluster network. If any of the nodes fails, the cluster is considered to be in a degraded state.

Problem

If the cluster is in a degraded state, failover cannot occur. For information about failure scenarios while the cluster is in a degraded state, see [“Resolving Failover Failures,”](#) on page 78.

Cause

The cluster can be in a degraded state for a number of reasons.

One of the nodes fails

- If the Active node fails, a failover of the Active node to the Passive node occurs automatically. After the failover, the Passive node becomes the Active node.

At this point, the cluster is in a degraded state because the original Active node is unavailable.

After the failed node is repaired or comes online, it becomes the new Passive node and the cluster returns to a healthy state after the Active and Passive nodes synchronize.

- If the Passive node fails, the Active node continues to function, but no failover is possible and the cluster is in a degraded state.

If the Passive node is repaired or comes online, it automatically rejoins the cluster and the cluster state is healthy after the Active and Passive nodes synchronize.

- If the Witness node fails, the Active node continues to function and replication between Active and Passive node continues, but no failover can occur.

If the Witness node is repaired or comes online, it automatically rejoins the cluster and the cluster state is healthy.

Database replication fails

If replication fails between the Active and Passive nodes, the cluster is considered degraded. The Active node continues to synchronize with the Passive node. If it succeeds, the cluster returns to a healthy state. This state can result from network bandwidth problems or other resource shortages.

Configuration file replication issues

If configuration files are not properly replicated between the Active and Passive nodes, the cluster is in a degraded state. The Active node continues to attempt synchronization with the Passive node. This state can result from network bandwidth problems or other resource shortages.

Solution

How you recover depends on the cause of the degraded cluster state. If the cluster is in a degraded state, events, alarms, and SNMP traps show errors.

If one of the nodes is down, check for hardware failure or network isolation. Check whether the failed node is powered on.

In case of replication failures, check if the vCenter HA network has sufficient bandwidth and ensure network latency is 10 ms or less.

Recovering from Isolated vCenter HA Nodes

If all nodes in a vCenter HA cluster cannot communicate with each other, the Active node stops serving client requests.

Problem

Node isolation is a network connectivity problem.

Solution

- 1 Attempt to resolve the connectivity problem. If you can restore connectivity, isolated nodes rejoin the cluster automatically and the Active node starts serving client requests.
- 2 If you cannot resolve the connectivity problem, you have to log in to Active node's console directly.
 - a Power off and delete the Passive node and the Witness node virtual machines.
 - b Log in to the Active node by using SSH or through the Virtual Machine Console.
 - c To enable the Bash shell, enter **shell** at the **appliance\$** prompt.

- d Run the following command to remove the vCenter HA configuration.
`destroy-vcha -f`
- e Reboot the Active node.
The Active node is now a standalone vCenter Server Appliance.
- f Perform vCenter HA cluster configuration again.

Resolving Failover Failures

When a Passive node does not become the Active node during a failover, you can force the Passive node to become the Active node.

Problem

The Passive node fails while trying to assume the role of the Active node.

Cause

A vCenter HA failover might not succeed for these reasons.

- The Witness node becomes unavailable while the Passive node is trying to assume the role of the Active node.
- An appliance state synchronization issue between the nodes exists.

Solution

You recover from this issue as follows.

- 1 If the Active node recovers from the failure, it becomes the Active node again.
- 2 If the Witness node recovers from the failure, follow these steps.
 - a Log in to the Passive node through the Virtual Machine Console.
 - b To enable the Bash shell, enter **shell** at the `appliance$` prompt.
 - c Run the following command.
`vcha-reset-primary`
 - d Reboot the Passive node.
- 3 If both Active node and Witness node cannot recover, you can force the Passive node to become a standalone vCenter Server Appliance.
 - a Delete the Active node and Witness node virtual machines.
 - b Log in to the Passive node through the Virtual Machine Console.
 - c To enable the Bash shell, enter **shell** at the `appliance$` prompt.
 - d Run the following command.
`destroy-vcha`
 - e Reboot the Passive node.

Patching a vCenter High Availability Environment

You can patch a vCenter Server Appliance which is in a vCenter High Availability cluster by using the `software-packages` utility available in the vCenter Server Appliance shell. For more information, see *vSphere Upgrade*.

Using Microsoft Clustering Service for vCenter Server on Windows High Availability

5

When you deploy vCenter Server, you must build a highly available architecture that can handle workloads of all sizes.

Availability is critical for solutions that require continuous connectivity to vCenter Server. To avoid extended periods of downtime, you can achieve continuous connectivity for vCenter Server by using a Microsoft Cluster Service (MSCS) cluster.

This chapter includes the following topics:

- [“Benefits and Limitations of Using MSCS,”](#) on page 79
- [“Upgrade vCenter Server in an MSCS Environment,”](#) on page 79
- [“Configure MSCS for High Availability,”](#) on page 81

Benefits and Limitations of Using MSCS

vCenter Server 5.5 update 3.x supports Microsoft Cluster Service (MSCS) as an option for providing vCenter Server availability.

Multiple instances of vCenter Server are in an MSCS cluster, but only one instance is active at a time. Use this solution to perform maintenance, such as operating system patching or upgrades, excluding vCenter Server patching or upgrades. You perform maintenance on one node in the cluster without shutting down the vCenter Server database.

Another potential benefit of this approach is that MSCS uses a type of “shared-nothing” cluster architecture. The cluster does not involve concurrent disk accesses from multiple nodes. In other words, the cluster does not require a distributed lock manager. MSCS clusters typically include only two nodes and they use a shared SCSI connection between the nodes. Only one server needs the disks at any given time. No concurrent data access occurs. This sharing minimizes the impact if a node fails.

Unlike the vSphere HA cluster option, the MSCS option works only for Windows virtual machines. The MSCS option does not support vCenter Server Appliance.

NOTE This configuration is supported only when vCenter Server is running as a VM, not on a physical host.

Upgrade vCenter Server in an MSCS Environment

If you are running vCenter Server 6.0, you must upgrade to vCenter Server 6.5 to set up an MSCS high availability environment.

vCenter Server 6.0.x has 18 services, assuming that the PSC server is running on a different host. vCenter Server 6.5 has 3 services and the names have changed. An MSCS cluster configuration created to set up high availability for vCenter Server 6.0 becomes invalid after an upgrade to vCenter Server 6.5.

The process for vCenter Server high availability in an MSCS environment is as follows.

- 1 Remove the MSCS configuration for vCenter Server.
- 2 Upgrade the vCenter Server from version 6.0 to version 6.5.
- 3 Configure MSCS to make vCenter Server highly available.

Prerequisites

- Verify that you are not deleting the primary node VM.
- Verify that the primary node is the current active node.
- Verify that all the services of vCenter Server 6.0 are running on the primary node.
- Verify that the Platform Services Controller node upgrade is finished and running vCenter Server 6.5.
- Collect the inventory database backup.

Procedure

- 1 Power off the secondary node and wait for all the vCenter Server services to be started on the primary node.
- 2 Remove the role name.
- 3 Destroy the MSCS cluster. Bring the RDM disks online again before changing the startup type.
- 4 Open the Service Management view and change the startup type for vCenter Server services from manual to automatic.
- 5 Before upgrading to vCenter Server 6.5, change the IP and host name to the IP and host name used for the role.

You must restart the host and ensure that vCenter Server is accessible.

- 6 Mount the vCenter Server 6.5 ISO and start the installation.
- 7 After the installation finishes, open the Service Management view and verify that the new services are installed and running.
- 8 Set up the MSCS cluster configuration again and set the startup type of all vCenter Server services to manual.
- 9 Shut down the primary node and detach the RDM disks, but do not delete them from the datastore.
- 10 After the reconfiguration is complete, select **VM > Clone > Clone to Template**, clone the secondary node, and change its IP and host name.
- 11 Keep the secondary node powered off and add both RDM disks to the primary node. Then power on the primary node and change its IP and host name.
- 12 Add both RDM disks to the secondary node. Then power on the secondary node.
- 13 Open Failover Cluster Manager and configure the MSCS cluster.

You must use the cluster role IP and host name.

What to do next

When configuring the MSCS cluster, you must add vCenter Server services such as the VMware AFD service and the VMware vCenter Configuration service to the role as resources.

Configure MSCS for High Availability

Use the following steps to set up Microsoft Cluster Service (MSCS) as an availability solution for vCenter Server.

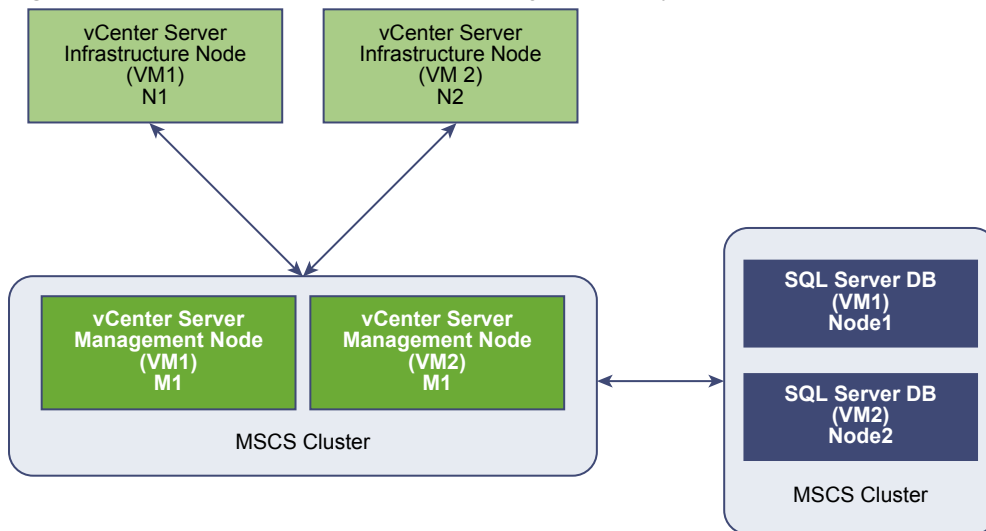
Prerequisites

- Create a virtual machine (VM) with one of the following guest operating systems:
 - Windows 2008 R2 Datacenter
 - Windows 2012 R2 Datacenter
- Add two raw device mapping (RDM) disks to this VM. These disks must be mounted when they are added and the RDM disks must also be independent and persistent.
- Create a separate SCSI controller with the bus sharing option set to physical.

NOTE Since this configuration uses a SCSI controller with the Bus Sharing option set to Physical, backup and restore is not supported. You must use a host-based agent for backup or restore.

- Set the Virtual Device Node to newly created SCSI controller 1.
- Open the MSCS drive and create two folders: one for VC data and another for VC installation.
- Install a Platform Services Controller instance before you install vCenter Server and provide its FQDN during the installation.

Figure 5-1. MSCS Cluster for vCenter Server High Availability



NOTE MSCS as an availability solution for vCenter Server is provided only for management nodes of vCenter Server (M node). For infrastructure nodes, customers must deploy multiple N nodes for high availability. You cannot have M and N nodes on the same VM for MSCS protection.

Procedure

- 1 Power on the VM.
- 2 Format the two RDM disks, assign them drive letters, and convert them to MBR.
- 3 Using **Windows > Server Manager > Features**, install .net.
- 4 Install vCenter Server on one of the RDM disks and set the start option to manual.

- 5 Power off the VM.
- 6 Detach the RDM disks.
Detaching the RDM disks is not a permanent deletion. Do not select **Delete from disk** and do not delete the vmdk files.
- 7 Clone the VM and select the **Customize the operating system** option, so that the clone has a unique identity.
Create a unique identity through either the default sysprep file or the custom sysprep file.
- 8 Attach the shared RDMs to both VMs and power them on.
- 9 Change the host name and IP address on the first VM (VM1).
Note the original IP address and host name that were used at the time of the installation of vCenter Server on VM1. This information is used to assign a cluster role IP.
- 10 Install failover clustering on both nodes.
- 11 To create an MSCS cluster on VM1, include both nodes in the cluster. Also select the validation option for the new cluster.
- 12 To start configuring roles, select **Generic Service** and click **Next**.
- 13 Select **VMware Service Lifecycle Manager** from the listed services and click **Next**.
- 14 Enter the host name and IP used for the VM1. Then assign the RDM to the role.
- 15 In the Replicate Registry Settings wizard, add the registry key **SYSTEM\CurrentControlSet\Services\VMwareDirectoryService** and click **Next**.
- 16 Using Add Resource, add the VMware AFD and VMware vCenter Configuration services to the role.
- 17 Stop and restart the role.

You have created an MSCS cluster that can support vCenter Server availability.

What to do next

After you have created the MSCS cluster, verify that failover is occurring by powering off the VM hosting vCenter Server (VM1). In a few minutes, verify that the services are running on the other VM (VM2).

Index

A

- Active node, functioning **60**
- admission control
 - configuring **35**
 - vSphere HA **21**
- Advanced configuration, vCenter HA **67**
- affinity rules **43, 47**
- anti-affinity rules **43**
- APD **18**
- application monitoring **33**
- Application Monitoring **14, 17**
- Auto Deploy **41**

B

- best practices
 - Fault Tolerance **53**
 - vSphere HA clusters **39**
 - vSphere HA networking **39**
- business continuity **9**

C

- c **74**
- cloning nodes **68**
- Cluster Operational Status **42**
- cluster settings **30**
- cluster validity **42**
- complete Advanced configuration, vCenter HA **69**
- compliance check, Fault Tolerance **49**
- configure second NIC, vCenter HA **67**
- configured failover capacity **23**
- Configured Failover Capacity **21**
- configuring vSphere HA advanced options **36**
- creating a vSphere HA cluster **29**
- current failover capacity **23**
- Current Failover Capacity **21**
- current failover hosts **25**

D

- das.config.fdm.memreservationmb **37**
- das.config.fdm.reportfailoverfailevent **37**
- das.heartbeatdsperhost **19, 37**
- das.ignoreinsufficienthbdatastore **37**
- das.iostatsinterval **17, 37**
- das.isolationaddress **37, 39**
- das.isolationshutdowntimeout **15, 37**

- das.maxftvcpusperhost **44**
- das.maxftvmsperhost **44**
- das.maxresets **37**
- das.maxterminates **37**
- das.reregisterrestartdisabledvms **37**
- das.reservationrequestretryintervalsec **37**
- das.respectvmvmantiaffinityrules **37**
- das.slotcpuinmhz **23, 37**
- das.slotmeminmb **23, 37**
- das.terminateretryintervalsec **37**
- das.usedefaultisolationaddress **37**
- das.vmcpcuminmhz **21, 37**
- das.vmmemoryminmb **37**
- datastore with APD **33**
- datastore heartbeating **14, 19**
- datastore with PDL **33**
- dedicated failover hosts **25**
- default gateway **39**
- degraded vCenter HA cluster **76**
- Distributed Power Management (DPM) **27**
- Distributed Resource Scheduler (DRS)
 - using with vSphere Fault Tolerance **47**
 - using with legacy Fault Tolerance **55**
 - using with vSphere HA **27, 28**
- DNS lookup **29**
- downtime
 - planned **9**
 - unplanned **10**
- DRS Affinity Rules **28**

E

- Enhanced vMotion Compatibility **47**
- error messages
 - Fault Tolerance **43**
 - vSphere HA **13**
- EVC **47**
- events and alarms, setting **42**
- Extended Page Tables (EPT) **46, 55**

F

- failover hosts **25**
- Fault Tolerance
 - anti-affinity rules **43**
 - best practices **53**
 - checklist **47**

- compliance check **49**
- continuous availability **11**
- error messages **43**
- interoperability **45**
- logging **48**
- migrate secondary **52**
- networking configuration **48**
- options **49**
- overview **43**
- preparing for **47**
- prerequisites **47**
- restrictions for turning on **50**
- suspending **52**
- test failover **52**
- test restart secondary **53**
- turning off **51**
- turning on **51**
- use cases **44**
- validation checks **50**
- version **47**
- vSphere configuration **47**
- Fault Tolerance licensing **44**
- Fault Tolerance limits **44**
- Fault Tolerance requirements **44**
- fdm.isolationpolicydelaysec **37**
- firewall ports **20, 39**

H

- Hardware Virtualization (HV) **47, 50**
- host failover capacity **21**
- host failure, vSphere HA response **32**
- host isolation, vSphere HA response **33**
- Host Isolation Response setting **15**
- Host Monitoring feature **30, 39**
- hosts
 - maintenance mode **14, 27**
 - network isolation **14**
- how vCenter HA works **60**

I

- I/O stats interval **17**
- intended audience **5**
- interoperability, Fault Tolerance **45**
- IPv4 **28, 29, 46, 55**
- IPv6 **28, 29, 46, 48, 55**
- iSCSI SAN **47**
- ISO images **53**

L

- legacy FT **43, 48, 55**
- Lifecycle Manager Service **12**
- log files **20**

M

- management network **29, 39**
- master host election **14**
- Maximum per-VM resets **17**
- Microsoft Cluster Service (MSCS) **79**
- migrate secondary, Fault Tolerance **52**
- minimizing downtime **9**
- modifying cluster settings **30**
- monitoring sensitivity **17**
- monitoring vSphere HA **42**
- MSCS **81**

N

- N_Port ID Virtualization (NPIV) **46**
- network isolation address **39**
- network labels **39**
- network partition **14, 19, 53**
- networking configuration, Fault Tolerance **48**
- networking differences with Virtual SAN **26**
- NIC teaming **39**

O

- On-Demand Fault Tolerance **44**

P

- paravirtualization **46**
- Passive node, functioning **60**
- Patching a vCenter High Availability Environment **78**
- PDL **18**
- Percentage of Cluster Resources Reserved **21**
- planned downtime **9**
- planning a vSphere HA cluster **13**
- port group names **39**
- PortFast **39**
- prerequisites, Fault Tolerance **47**
- Proactive HA **14, 34**

R

- Rapid Virtualization Indexing (RVI) **46, 55**
- RDM **46, 47**

S

- set up MSCS for vCenter Server availability **81**
- slot **23**
- slot policy admission control **23**
- slot size calculation **23**
- snapshots **45**
- SSL Certificates **20**
- storage
 - iSCSI **47**

- NAS **47**
 - NFS **47**
 - Storage DRS **41**
 - Storage vMotion **9, 41, 45**
 - suspending, Fault Tolerance **52**
 - Symmetric multiprocessor (SMP) **46**
 - symmetric multiprocessor (SMP) virtual machines **55**
- T**
- TCP port **20**
 - test failover, Fault Tolerance **52**
 - test restart secondary, Fault Tolerance **53**
 - tolerating host failures **23**
 - transparent failover **11, 43**
 - turning off, Fault Tolerance **51**
- U**
- UDP port **20**
 - unplanned downtime **10**
 - updated information **7**
 - upgrading for MSCS high availability **79**
 - upgrading hosts with FT virtual machines **53**
 - use cases, Fault Tolerance **44**
- V**
- VADP backups **55**
 - validation checks **50**
 - vCenter HA
 - advanced deployment **68**
 - backup **73**
 - basic deployment **66**
 - certificate setup **71**
 - clone operation problem **76**
 - cluster failures **77**
 - configuration **61**
 - deployment failure **76**
 - edit a cluster configuration **72**
 - initiate a failover **72**
 - operating modes **62**
 - planning deployment **60**
 - reboot all nodes **74**
 - restore **73**
 - small environment **75**
 - SNMP traps **71**
 - SSH keys **72**
 - support bundles **75**
 - troubleshooting **75**
 - vCenter Server Appliance high availability **63**
 - vCenter HA configuration
 - advanced **67**
 - managing **70**
 - vCenter HA network configuration **65**
 - vCenter High Availability Overview **12**
 - vCenter Server high availability **59**
 - virtual machine overrides **39**
 - virtual machine protection **14, 19**
 - virtual machine restarts **15**
 - Virtual Machine Startup and Shutdown feature **29**
 - Virtual SAN **19, 26, 28, 41**
 - Virtual SAN datastores **53**
 - Virtual SAN stretched clusters **26**
 - VM Component Protection **18, 28–30, 33, 45**
 - VM monitoring **33**
 - VM Monitoring **14, 17**
 - VM resource reduction threshold **21**
 - VM-VM affinity rules **25**
 - vm.uselegacyft **55**
 - VMCP **18, 28–30, 33, 45**
 - VMDK **47, 55**
 - VMFS **19, 39**
 - VMware Tools **17**
 - vpxd.das.completemetadataupdateintervalsec **37**
 - vpxuser user account **20**
 - vSphere HA
 - advantages **10**
 - checklist **29**
 - cluster settings **29**
 - configuring cluster settings **31**
 - error messages **13**
 - monitoring **42**
 - recovery from outages **10**
 - vSphere HA cluster
 - admission control **21**
 - best practices **39**
 - creating **29, 30, 49**
 - master host **14, 19**
 - planning **13**
 - slave host **14**
 - vSphere HA interoperability **26**
 - vSphere HA architecture **13**
 - vSphere HA heartbeat datastores **35**
 - vSphere HA networking
 - best practices **39**
 - path redundancy **39**
 - vSphere HA response to failures **32**
- W**
- Witness node, functioning **60**

