CHAPTER 10

Backup and High Availability

This chapter covers the following VCP exam topics:

- VMware Consolidated Backup (VCB)
- High Availability (HA)
- Admission Control
- Host Backup
- Cluster-in-a-Box
- Cluster-Across-Boxes
- Physical-to-Virtual cluster
- Fault Tolerance
- Data Recovery
- MSCS Clustering
- Host Isolation

(For more information on the VCP exam topics, see "About the VCP Exam" in the introduction.)

What good is an environment without a good backup strategy? This chapter explores the different options by which you can back up your VMware Infrastructure. The chapter also explores VMware High Availability (HA) and the ability to sustain host failures and ensure that your critical virtual machines (VMs) can be restarted on other hosts that are online.

Backup Scenarios

- **Host Backup**
- **VMware Consolidated Backup (VCB)**
- **Data Recovery**

CramSaver

If you can correctly answer this question before going through this section, save time by skimming the Exam Alerts in this section and then completing the Cram Quiz at the end of the section.

1. What are the two types of virtual machine backups? (Select all that apply.)

○ **A.** Differential

○ **B.** File-level

○ **C.** Shadow-copy

○ **D.** Image-level

Answers

1. **B** and **D** are correct. The two types of VM level backups are file-level and image-level backups; therefore, answers A and C are incorrect.

You should think of and treat backup strategies for VMs and their ESX/ESXi host the same way you would approach physical machines. The same best practices and methodologies apply. The only different scenario covered is VMware Consolidated Backup (VCB).

Virtual Machine Backup Options

As far as virtual machines are concerned, you can back them up using any of the following methods:

- Installing a backup agent inside the guest operating system. This is the same as when you install a backup agent inside a physical machine's guest operating system. You can then do file-level backups as frequently as your data changes or as frequently as your environment requires.
- Backing up the virtual machine files. Because a virtual machine is encapsulated inside regular files, you can back up or archive these files. When using this method, you can power off the VM and back up the files. Alternatively, if you need the VM to continue to be online, you can take a snapshot of it and back up the VMDK files.

When considering backup of virtual machine data, make sure that you have your application files stored on separate drives. This makes the process of backing up the data easier. This also concentrates the backup process on the data rather than the operating system files, which should be backed up infrequently because they do not change and you don't want to have to back them up repeatedly. The system drive should be backed up in the event that you want to restore the entire virtual machine to the state it was in when you made the backup. The point to keep in mind here is that you are backing up the system drive for the Registry and the application-specific files that are installed with the virtual machine.

Host Backup Options

The ESX/ESXi host is primarily the Service Console. Because the Service Console is used for command-line advanced options, its files rarely change and most of the configurations you make in your VMware Infrastructure are stored in the vCenter database. That being said, backing up the Service Console is not really worth the time and effort involved. You can easily reinstall ESX and make the changes rather than deal with executing a backup and restore operation of the Service Console.

However, if you want to back up your Service Console, you can do so using one of the following two methods:

- Install a backup agent inside the Service Console and back up the files accordingly. This would be the same traditional agent backup approach that you would take with any physical machine running any guest operating system.
- Use third-party software to create a complete image of the ESX server and then use this same third-party software to restore the entire image and return the state of the ESX/ESXi host to the point when you took the image originally.

Exam**Alert**

Pay attention to the host backup section because you are sure to get a question on this topic on the VCP exam.

VMware Consolidated Backup

VMware Consolidated Backup (VCB) is an alternative method of backing up and restoring virtual machines at the file level or image level. VCB runs on a

Windows server and makes a snapshot of the VM, and for file-level backups, it mounts the VMDK as a disk on the Windows backup server. Once the VMDK is mounted, you can see the contents of the VMDK appear under a specified directory on the VCB server. Now your file-level backup agent can access these files and write them to tape or any other backup destination.

When you want to make a full image backup, VCB also makes a snapshot but then copies the full VMDK as a file to a special backup LUN, which should be at least the size of the VMDK. After copying the VMDK to the LUN, the snapshot of the VM is released and your file-level backup agent can pick up the VMDK from the backup LUN.

To be able to do this, the VCB server has to have a direct connection to the LUNs on which the VMs reside. Keep in mind that VCB is not a backup product itself, but just a tool to help your traditional backup product access your VMs.

Consequently, you have taken the network out of the equation all the way to the backup server. Now, obviously, depending on what type of backup system you are using and how the backup server is connected to your backup robot, the network can remain out of the equation or can be used to move the files from the backup server to the backup tape library.

By using VCB, you have the following advantages:

- Because you are dealing with snapshot-level backups, there is no need for a backup window because no downtime is required to back up the VM. The VM is backed up while it is powered on.
- Backup load is moved away from the ESX/ESXi host because you are taking a snapshot and moving this snapshot to another location to be backed up directly by the VCB proxy server. By doing so, you offload all the processing requirements needed from the ESX/ESXi host to the backup server.
- The backup agent is optional. When using VCB, you take advantage of the VMware Tools that are installed inside the VM that allow for VCB to take place, thereby giving you the option to use the backup agent only if you want to restore directly to this virtual machine. Instead of restoring directly to a VM, you may opt to install a backup agent on a select few VMs and then restore any files to these VMs. At that point, you can copy the restored files to their final destination. This would save money and management of backup agents on multiple virtual machines.

When running VCB, you have full support for file-level backup of Microsoft Windows guest operating systems and image-level backups of any guest operating system.

ExamAlert

Knowing the capabilities and limitations of VCB is critical because you will surely be quizzed on this subject on the exam. VCB is a most critical component of the VI suite.

Data Recovery

Data Recovery is a new backup and recovery tool for VMs feature that is introduced with vSphere 4, aimed at small to medium-sized companies. Data Recovery is a Linux-based appliance that can be imported into vCenter and controlled through a vSphere Client plug-in.

The Data Recovery appliance is an agentless backup to disk type solution. Your destination backup location can be to VMware VMFS Datastores on local disk, iSCSI, FC, or it can be on an NFS Datastore. You can even back up using Data Recovery to Windows Common Internet File System (CIFS) shares.

Tip

Destination Virtual Disks or RDMs would have to manually be added to the Data Recovery Appliance.

A single Data Recovery Appliance can support the following:

- Up to 100 VMs.
- Up to 100 backup jobs.
- Each backup job can have a maximum of one destination.
- Each VM configured in a backup job is backed up once every 24 hours.

Note

The configuration of the backup jobs is saved on the Data Recovery appliance. However, after the completion of a successful backup, a copy of the configuration is stored on the destination location as well. This allows for easier restore to a new appliance should the need arise.

How to Configure the Data Recovery Appliance

Setting up the Data Recovery Appliance is straightforward. After obtaining all the files necessary, follow these steps:

1. Deploy the Data Recovery OVF template to vCenter by going to File > Deploy OVF Template and following the wizard.
2. Configure the networking stack to allow for connectivity through the appliance's console.
3. Configure the appropriate time zone settings through the appliance's console.
4. Add destination storage to the appliance. To accomplish this, you would add a virtual disk to this appliance the same way you would any other VM by going to Edit Settings.
5. Install the Data Recovery plug-in for the vSphere Client.

Once you have installed and configured the appliance, you can access it via your vSphere client from the Home screen in the Solutions and Applications.

Exam**Alert**

The ESX/ESXi host that the appliance will be deployed on and the host that carries the VMs that will be backed up both need to be licensed for Data Recovery.

Tip

After you configure the networking settings of the appliance, you can access its web interface from a supported browser by pointing to its IP address. From there, you can configure many of the settings you can configure via console, such as the time zone.

Backup Process

Backing up VMs using the Data Recovery Appliance is easy and wizard-driven. Once you initiate the wizard, it leads you step by step to a successful backup job creation. The backup job wizard is capable of enumerating all objects in the vCenter inventory, which means you can back up any VM regardless of its logical grouping or location.

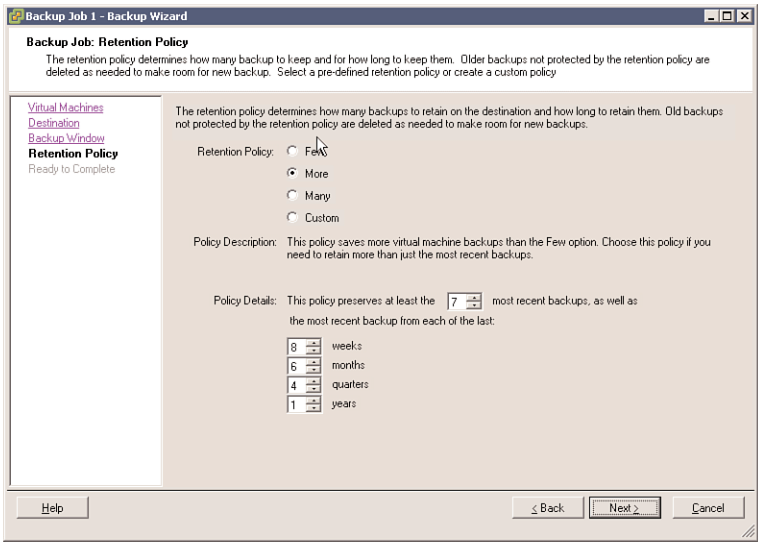
When you run the backup wizard, you are warned if

- You have selected more than 100 VMs for backup.
- If a selected VM is on a nonlicensed host.

To start the backup job wizard, point to the Data Recovery Appliance in vCenter and select the Backup tab. You can then either select New at the top right-hand corner or right-click anywhere and click New.

The backup wizard prompts you to

- Select the VM to back up or select certain components of a VM to back up, such as vdisks.
- Select the destination target from your available options.
- The Backup Window is next; specify the times that backup is allowed to run.
- Retention Policy settings are up next, as shown in Figure 10.1. This screen allows you to configure how long you retain data on the destination and how many backups to retain.



Backup Job: Retention Policy

The retention policy determines how many backup to keep and for how long to keep them. Older backups not protected by the retention policy are deleted as needed to make room for new backup. Select a pre-defined retention policy or create a custom policy

The retention policy determines how many backups to retain on the destination and how long to retain them. Old backups not protected by the retention policy are deleted as needed to make room for new backups.

Retention Policy: ○ Few ● More ○ Many ○ Custom

Policy Description: This policy saves more virtual machine backups than the Few option. Choose this policy if you need to retain more than just the most recent backups.

Policy Details: This policy preserves at least the 7 most recent backups, as well as the most recent backup from each of the last:

8 weeks
6 months
4 quarters
1 years

FIGURE 10.1 **Retention Policy.**

Restore VMs and Files

The restore process using the Data Recovery Appliance is just as easy as the backup process and is completely wizard driven. You can

- Restore a single file back to a VM (Windows and Linux support only).
- Restore a VM to a different host, datastore, or resource pool.
- Restore a VM due to deletion or corruption.
- Restore a VM to an earlier point in time.
- Restore a VM's virtual disks.

To initiate the restore wizard, find the Data Recovery Appliance in your vCenter inventory and click on the Restore tab. You can then click on the Restore link in the top right-hand corner. This initiates the restore wizard, which prompts you to select what to restore. Based on the guidelines we discussed earlier, you can select anything from a full VM restore all the way to a specific vdisk restore.

Now to restore a file back into a VM, additional software needs to be downloaded from the VMware website. You need to download the restore client for Windows or Linux. This restore client is then installed in the VM allowing you to mount a restore point from the Data Recovery Appliance and select the appropriate file for restore directly to the VM.

Cram Quiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading the section again.

1. True or false: It is a recommended and crucial task that you back up the Service Console to preserve the ESX/ESXi host configuration.

   ○ **A.** True

   ○ **B.** False

2. True or false: VMware Consolidated Backup supports file-level backup for all guest operating systems.

   ○ **A.** True

   ○ **B.** False

Cram Quiz Answers

1. **B**, False, is correct. Although backing up the Service Console is an option, it is really not required because most of the ESX/ESXi host configuration is stored in the vCenter database. The Service Console files rarely change, and those that do don't merit a backup. It is thereby easier to reinstall ESX than it is worth backing up the SC.
2. **B**, False, is correct. VMware Consolidated Backup supports only file-level backups on Windows-based systems and image-level backups for all guest operating systems.

High Availability

- **High Availability (HA)**
- **Admission Control**
- **Cluster-in-a-Box**
- **Cluster-Across-Boxes**
- **Physical-to-Virtual Cluster**
- **Host Isolation**
- **MSCS Clustering**
- **Fault Tolerance**

Cram**Saver**

If you can correctly answer this question before going through this section, save time by skimming the Exam Alerts in this section and then completing the Cram Quiz at the end of the section.

1. What are the two VMware HA clusterwide settings that you can configure? (Select all that apply.)

❍ **A.** Host Failures
❍ **B.** DRS
❍ **C.** Admission Control
❍ **D.** Fault Tolerance

2. Which outgoing TCP and UDP ports are used for heartbeats and state synchronization? (Choose all that apply.)

❍ **A.** TCP 2150–2250
❍ **B.** UDP 2150–2250
❍ **C.** TCP 2050–2250
❍ **D.** UDP 2050–2250

Answers

1. **A** and **C** are correct. The two settings that you can configure clusterwide for VMware HA are Host Failures and Admission Control; therefore, answers B and D are incorrect.

2. **C** and **D** are correct. The outgoing TCP and UDP ports are 2050 through 2250; therefore, answers A and B are incorrect.

VMware *High Availability (HA)* deals primarily with ESX/ESXi host failure and what happens to the virtual machines that are running on this host. HA can also monitor and restart a VM by checking whether the VMware Tools are still running. When an ESX/ESXi host fails for any reason, all the running VMs also fail. VMware HA ensures that the VMs from the failed host are capable of being restarted on other ESX/ESXi hosts.

Many people mistakenly confuse VMware HA with fault tolerance. VMware HA is not fault tolerant in that if a host fails, the VMs on it also fail. HA deals only with restarting those VMs on other ESX/ESXi hosts with enough resources. Fault tolerance, on the other hand, provides uninterruptible access to resources in the event of a host failure.

Exam**Alert**

The VCP exam is sure to challenge your knowledge on the difference between HA and fault tolerance. Make sure you have a clear understanding of the difference.

VMware HA maintains a communication channel with all the other ESX/ESXi hosts that are members of the same cluster by using a heartbeat that it sends out every 1 second in vSphere 4.0 or every 10 seconds in vSphere 4.1 by default. When an ESX server misses a heartbeat, the other hosts wait 15 seconds for the other host to respond again. After 15 seconds, the cluster initiates the restart of the VMs on the failing ESX/ESXi host on the remaining ESX/ESXi hosts in the cluster. VMware HA also constantly monitors the ESX/ESXi hosts that are members of the cluster and ensures that resources are always available to satisfy requirements in the event of a host failure.

Tip

VMware HA is a reactive system, which means it kicks in to react to a problem; in this case, the problem is a failed host. A reactive system would be perfect if combined with a proactive system, and this is exactly what you get if you enabled VMware HA and VMware DRS on the same cluster. DRS is a proactive system that is constantly busy trying to load-balance resources on ESX/ESXi hosts.

Virtual Machine Failure Monitoring

Virtual Machine Failure Monitoring is technology that is disabled by default. Its function is to monitor virtual machines, which it queries every 20 seconds via a heartbeat. It does this by using the VMware Tools that are installed inside the VM. When a VM misses a heartbeat, VMware HA deems this VM as failed and attempts to reset it. Think of Virtual Machine Failure Monitoring as sort of High Availability for VMs.

> **Note**
>
> Virtual Machine Failure Monitoring can detect whether a virtual machine was manually powered off, suspended, or migrated, and thereby does not attempt to restart it.

HA Configuration Prerequisites

HA requires the following configuration prerequisites before it can function properly:

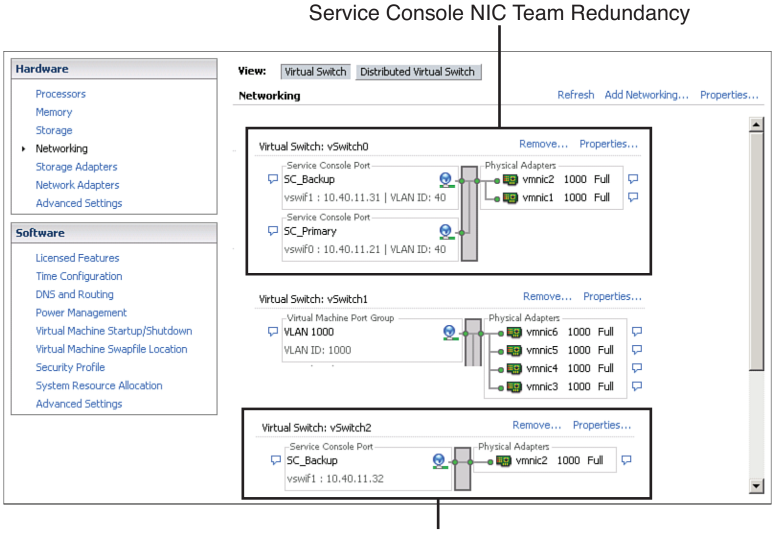
- **vCenter:** Because VMware HA is an enterprise-class feature, it requires vCenter before it can be enabled.
- **DNS resolution:** All ESX/ESXi hosts that are members of the HA cluster must be able to resolve one another using DNS.
- **Access to shared storage:** All hosts in the HA cluster must have access and visibility to the same shared storage; otherwise, they would have no access to the VMs.
- **Access to same network:** All ESX/ESXi hosts must have the same networks configured on all hosts so that when a VM is restarted on any host, it again has access to the correct network.

Service Console Redundancy

Recommended practice dictates that the Service Console have redundancy. VMware HA complains and issues a warning if it detects that the Service Console is configured on a vSwitch with only one vmnic. As Figure 10.2 shows, you can configure Service Console redundancy in one of two ways:

- Create two Service Console port groups, each on a different vSwitch.
- Assign two physical network interface cards (NICs) in the form of a NIC team to the Service Console vSwitch.

Service Console NIC Team Redundancy



Virtual Switch: vSwitch0

Service Console Port
SC_Backup
vswif1 : 10.40.11.31 | VLAN ID: 40

Service Console Port
SC_Primary
vswif0 : 10.40.11.21 | VLAN ID: 40

Physical Adapters
vmnic2 1000 Full
vmnic1 1000 Full

Virtual Switch: vSwitch1

Virtual Machine Port Group
VLAN 1000
VLAN ID: 1000

Physical Adapters
vmnic6 1000 Full
vmnic5 1000 Full
vmnic4 1000 Full
vmnic3 1000 Full

Virtual Switch: vSwitch2

Service Console Port
SC_Backup
vswif1 : 10.40.11.32

Physical Adapters
vmnic2 1000 Full

Service Console second separate vSwitch Redundancy

FIGURE 10.2 **Service Console redundancy.**

In both cases, you need to configure the entire IP stack with IP address, subnet, and gateway. The Service Console vSwitches are used for heartbeats and state synchronization and use the following ports:

- Incoming TCP port 8042
- Incoming UDP port 8045
- Outgoing TCP port 2050
- Outgoing UDP port 2250
- Incoming TCP port 8042–8045
- Incoming UDP port 8042–8045
- Outgoing TCP port 2050–2250
- Outgoing UDP port 2050–2250

Failure to configure SC redundancy results in a warning message when you enable HA. So, to avoid seeing this error message and to adhere to best practice, configure the SC to be redundant.

Exam**Alert**

Service Console redundancy is an important topic and will more than likely be one of the questions on the exam.

Host Failover Capacity Planning

When configuring HA, you have to manually configure the maximum host failure tolerance. This is a task that you should thoughtfully consider during the hardware sizing and planning phase of your deployment. This would assume that you have built your ESX/ESXi hosts with enough resources to run more VMs than planned to be able to accommodate HA. For example, in Figure 10.3, notice that the HA cluster has four ESX hosts and that all four of these hosts have enough capacity to run at least three more VMs. Because they are all already running three VMs, that means that this cluster can afford the loss of two ESX/ESXi hosts because the remaining two ESX/ESXi hosts can power on the six failed VMs with no problem if failure occurs.

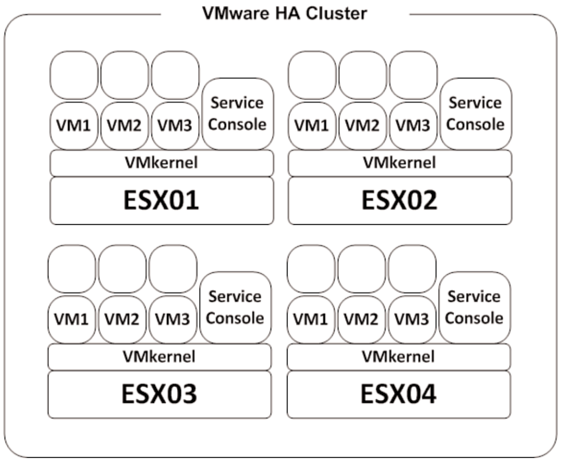


FIGURE 10.3 HA capacity planning.

During the configuration phase of the HA cluster, you are presented with a screen similar to that shown in Figure 10.4 that prompts you to define two clusterwide configurations as follows:

- **Host Monitoring Status:**
    - **Enable Host Monitoring:** This setting enables you to control whether the HA cluster should monitor the hosts for a heartbeat.

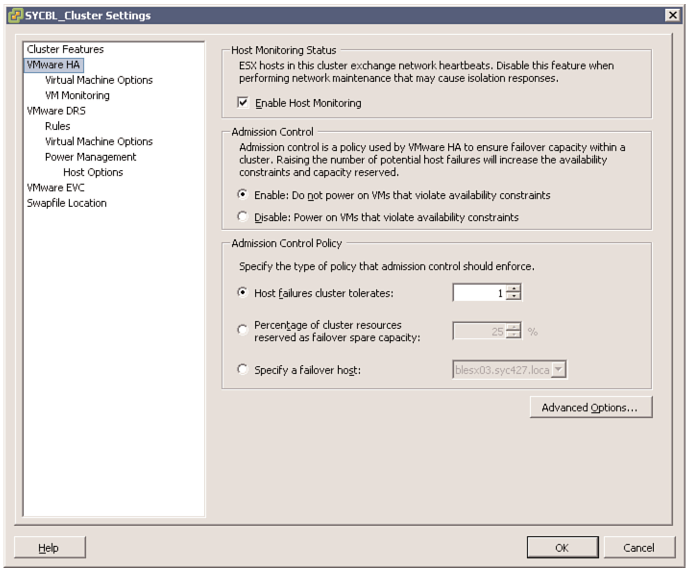
This is the cluster's way of determining whether a host is still active. In some cases, when you are running maintenance tasks on ESX/ESXi hosts, it might be desirable to disable this option to avoid isolating a host.

- **Admission Control:**
    - **Enable: Do not power on VMs that violate availability constraints:** Selecting this option indicates that if no resources are available to satisfy a VM, it should not be powered on.
    - **Disable: Power on VMs that violate availability constraints:** Selecting this option indicates that you should power on a VM even if you have to overcommit resources.

- **Admission Control Policy:**
    - **Host failures cluster tolerates:** This setting enables you to configure how many host failures you want to tolerate. The allowed settings are 1 through 4.
    - **Percentage of cluster resources reserved as failover spare capacity:** Selecting this option indicates that you are reserving a percentage of the total cluster resources in spare for failover. In a four-host cluster, a 25% reservation indicates that you are setting aside a full host for failover. If you want to set aside fewer, you can choose 10% of the cluster resources instead.
    - **Specify a failover host:** Selecting this option indicates that you are selecting a particular host as the failover host in the cluster. This might be the case if you have a spare host or have a particular host that has significantly more compute and memory resources available.

Exam**Alert**

The VCP exam may present you with a scenario and ask you to identify or configure capacity for HA. Make sure you are comfortable with capacity planning.



SYCBL_Cluster Settings

Host Monitoring Status

ESX hosts in this cluster exchange network heartbeats. Disable this feature when performing network maintenance that may cause isolation responses.

☑ Enable Host Monitoring

Admission Control

Admission control is a policy used by VMware HA to ensure failover capacity within a cluster. Raising the number of potential host failures will increase the availability constraints and capacity reserved.

● Enable: Do not power on VMs that violate availability constraints

○ Disable: Power on VMs that violate availability constraints

Admission Control Policy

Specify the type of policy that admission control should enforce.

● Host failures cluster tolerates: 1

○ Percentage of cluster resources reserved as failover spare capacity: 25 %

○ Specify a failover host: blesx03.syc427.loca

FIGURE 10.4 HA clusterwide policies.

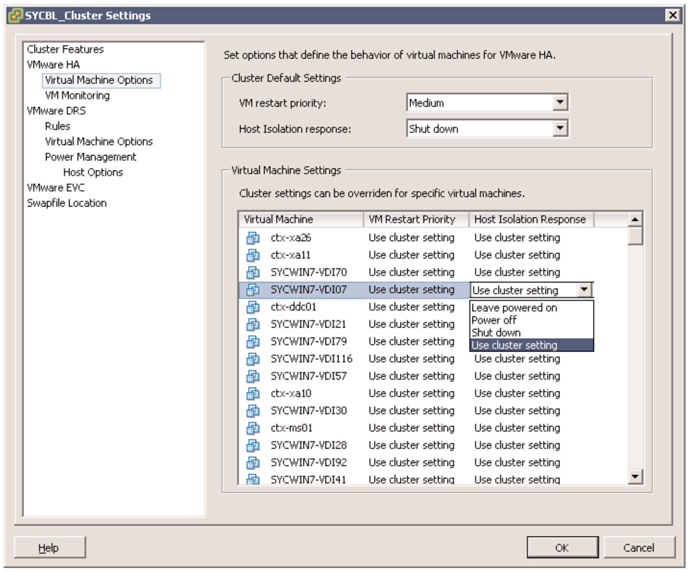
Host Isolation

A network phenomenon known as a *split-brain* occurs when the ESX/ESXi host has stopped receiving a heartbeat from the rest of the cluster. The heartbeat is queried for every 1 second in vSphere 4.0 or 10 seconds in vSphere 4.1. If a response is not received, the cluster thinks the ESX/ESXi host has failed. When this occurs, the ESX/ESXi host has lost its network connectivity on its management interface. The host might still be up and running and the VMs might not even be affected considering they might be using a different network interface that has not been affected. However, vSphere needs to take action when this happens because it believes a host has failed. For that matter, the host isolation response was created. Host isolation response is HA's way of dealing with an ESX/ESXi host that has lost its network connection.

You can control what happens to VMs in the event of a host isolation. To get to the VM Isolation Response screen, right-click the cluster in question and click on Edit Settings. You can then click Virtual Machine Options under the VMware HA banner in the left pane. You can control options clusterwide by setting the host isolation response option accordingly. This is applied to all the VMs on the affected host. That being said, you can always override the cluster settings by defining a different response at the VM level.

As shown in Figure 10.5, your Isolation Response options are as follows:

- **Leave Powered On:** As the label implies, this setting means that in the event of host isolation, the VM remains powered on.
- **Power Off:** This setting defines that in the event of an isolation, the VM is powered off. This is a hard power off.
- **Shut down:** This setting defines that in the event of an isolation, the VM is shut down gracefully using VMware Tools. If this task is not successfully completed within five minutes, a power off is immediately executed. If VMware Tools is not installed, a power off is executed instead.
- **Use Cluster Setting:** This setting forwards the task to the clusterwide setting defined in the window shown previously in Figure 10.5.



Set options that define the behavior of virtual machines for VMware HA.

Cluster Default Settings

VM restart priority: Medium

Host Isolation response: Shut down

Virtual Machine Settings

Cluster settings can be overriden for specific virtual machines.

Virtual Machine	VM Restart Priority	Host Isolation Response
ctx-xa26	Use cluster setting	Use cluster setting
ctx-xa11	Use cluster setting	Use cluster setting
SYCWIN7-VDI70	Use cluster setting	Use cluster setting
SYCWIN7-VDI07	Use cluster setting	Use cluster setting
ctx-ddc01	Use cluster setting	
SYCWIN7-VDI21	Use cluster setting	
SYCWIN7-VDI79	Use cluster setting	
SYCWIN7-VDI116	Use cluster setting	Use cluster setting
SYCWIN7-VDI57	Use cluster setting	Use cluster setting
ctx-xa10	Use cluster setting	Use cluster setting
SYCWIN7-VDI30	Use cluster setting	Use cluster setting
ctx-ms01	Use cluster setting	Use cluster setting
SYCWIN7-VDI28	Use cluster setting	Use cluster setting
SYCWIN7-VDI92	Use cluster setting	Use cluster setting
SYCWIN7-VDI41	Use cluster setting	Use cluster setting

FIGURE 10.5 VM-specific isolation policy.

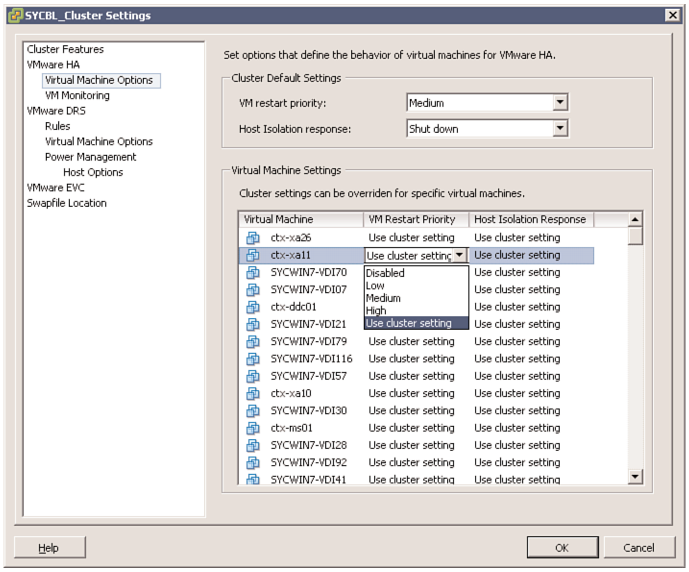
In the event of an isolation, this does not necessarily mean that the host is down. Because the VMs might be configured with different physical NICs and connected to different networks, they might continue to function properly; you therefore have to consider this when setting the priority for isolation. When a host is isolated, this simply means that its Service Console cannot communicate with the rest of the ESX/ESXi hosts in the cluster.

Virtual Machine Recovery Priority

Should your HA cluster not be able to accommodate all the VMs in the event of a failure, you have the ability to prioritize on VMs. The priorities dictate which VMs are restarted first and which VMs are not that important in the event of an emergency. These options are configured on the same screen as the Isolation Response covered in the preceding section. You can configure cluster-wide settings that will be applied to all VMs on the affected host, or you can override the cluster settings by configuring an override at the VM level.

As you can see in Figure 10.6, you can set a VM's restart priority to one of the following:

- **High:** VMs with a high priority are restarted first.
- **Medium:** This is the default setting.
- **Low:** VMs with a low priority are restarted last.
- **Use Cluster Setting:** VMs are restarted based on the setting defined at the cluster level defined in the window shown in Figure 10.6.
- **Disabled:** The VM does not power on.



Set options that define the behavior of virtual machines for VMware HA.

Cluster Default Settings

VM restart priority: Medium

Host Isolation response: Shut down

Virtual Machine Settings

Cluster settings can be overriden for specific virtual machines.

Virtual Machine	VM Restart Priority	Host Isolation Response
ctx-xa26	Use cluster setting	Use cluster setting
ctx-xa11	Use cluster setting	Use cluster setting
SYCWIN7-VDI70		Use cluster setting
SYCWIN7-VDI07		Use cluster setting
ctx-ddc01		Use cluster setting
SYCWIN7-VDI21		Use cluster setting
SYCWIN7-VDI79	Use cluster setting	Use cluster setting
SYCWIN7-VDI116	Use cluster setting	Use cluster setting
SYCWIN7-VDI57	Use cluster setting	Use cluster setting
ctx-xa10	Use cluster setting	Use cluster setting
SYCWIN7-VDI30	Use cluster setting	Use cluster setting
ctx-ms01	Use cluster setting	Use cluster setting
SYCWIN7-VDI28	Use cluster setting	Use cluster setting
SYCWIN7-VDI92	Use cluster setting	Use cluster setting
SYCWIN7-VDI41	Use cluster setting	Use cluster setting

FIGURE 10.6 **VM restart priority.**

The priority should be set based on the importance of the VMs. In other words, you might want to restart domain controllers and not restart print servers. The higher priority virtual machines are restarted first. VMs that can tolerate remaining powered off in the event of an emergency should be configured to remain powered off to conserve resources.

MSCS Clustering

The main purpose of a cluster is to ensure that critical systems remain online at any cost and at all times. Similar to physical machines that can be clustered, virtual machines can also be clustered with ESX using three different scenarios:

- **Cluster-in-a-box:** In this scenario, all the VMs that are part of the cluster reside on the same ESX/ESXi host. As you might have guessed, this immediately creates a single point of failure: the ESX/ESXi host. As far as shared storage is concerned, you can use virtual disks as shared storage in this scenario, or you can use Raw Device Mapping (RDM) in virtual compatibility mode.
- **Cluster-across-boxes:** In this scenario, the cluster nodes (VMs that are members of the cluster) reside on multiple ESX/ESXi hosts, whereby each of the nodes that make up the cluster can access the same storage so that if one VM fails, the other can continue to function and access the same data. This scenario creates an ideal cluster environment by eliminating a single point of failure. Shared storage is a prerequisite in this and must reside on Fiber Channel SAN, You also must use an RDM in Physical or Virtual Compatibility Mode as virtual disks are not a supported configuration for shared storage. Whereby each of the nodes that make up the cluster can access the same storage so that if one VM fails, the other can continue to function and access the same data.
- **Physical-to-virtual cluster:** In this scenario, one member of the cluster is a virtual machine, whereas the other member is a physical machine. Shared storage is a prerequisite in this scenario and must be configured as an RDM in Physical Compatibility Mode.

Tip

When configuring a cluster-across-boxes, it is highly recommended that you create an anti-affinity rule that always separates the VMs that are part of the cluster so that DRS never vMotions them to the same ESX/ESXi host, thereby creating a single point of failure scenario.

Whenever you are designing a clustering solution you need to address the issue of shared storage, which would allow multiple hosts or VMs access to the same data. vSphere offers several methods by which you can provision shared storage as follows:

- **Virtual disks:** You can use a virtual disk as a shared storage area only if you are doing clustering in a box—in other words, only if both VMs reside on the same ESX/ESXi host.
- **RDM in Physical Compatibility Mode:** This mode enables you to attach a physical LUN directly into a VM or physical machine. This mode prevents you from using functionality such as snapshots and is ideally used when one member of the cluster is a physical machine while the other is a VM.
- **RDM in Virtual Compatibility Mode:** This mode enables you to attach a physical LUN directly into a VM or physical machine. This mode gives you all the benefits of virtual disks running on VMFS including snapshots and advanced file locking. The disk is accessed via the hypervisor and is ideal when configuring a cluster-across-boxes scenario where you need to give both VMs access to shared storage.

At the time of this writing, the only VMware-supported clustering service is Microsoft Clustering Services (MSCS). You can consult the VMware white paper "Setup for Failover Clustering and Microsoft Cluster Service" on the topic located at http://www.vmware.com/pdf/vsphere4/r40/vsp_40_mscs.pdf.

Exam**Alert**

Because there are probably many third-party software vendors that advertise support for clustering inside ESX VMs, it is likely that the VCP exam will challenge your knowledge on the official VMware stand on supported clustering services with VMs.

VMware Fault Tolerance

VMware Fault Tolerance (FT) is another form of VM clustering developed by VMware for systems that require extreme uptime. One of the most compelling features of FT is its ease of setup. FT is simply a check box that can be enabled. Compared to traditional clustering that requires specific configurations and in some instances cabling, FT is simple but powerful.

How Does It Work?

When protecting VMs with FT, a secondary VM is created in lockstep of the protected VM, the first VM. FT works by simultaneously writing to the first

VM and the second VM at the same time. Every task is written twice. If you click on the Start menu on the first VM, the Start menu on the second VM will also be clicked. The power of FT is its capability to keep both VMs in sync.

If the protected VM should go down for any reason, the secondary VM immediately takes its place, seizing its identity and its IP address, continuing to service users without an interruption. The newly promoted protected VM then creates a secondary for itself on another host and the cycle restarts.

To clarify, let's see an example. If you wanted to protect an Exchange server, you could enable FT. If for any reason the ESX/ESXi host that is carrying the protected VM fails, the secondary VM kicks in and assumes its duties without an interruption in service.

Exam**Alert**

If the initially protected VM should recover from its failure, it would not assume its previous role or identity; it would simply show up in the inventory as a VM.

Note

VMware FT protects the VM from failure but not the guest operating system or any application running on it. If the primary VM should blue screen, the secondary will blue screen as well. If the primary has an application error, so will the secondary.

Table 10.1 outlines the different High Availability and clustering technologies that you have access to with vSphere and highlights limitations of each.

TABLE 10.1 **vSphere HA and Clustering Support Matrix**

	HA	**FT**	**MSCS**
Availability Type	High Availability	Fault Tolerance	Fault Tolerance
Downtime	Some	None	Some
Supported OS	All supported OS	All supported OS	Only Microsoft supported OS
Supported Hardware	All supported ESX hardware	All supported ESX hardware with CPUs that support FT	Hardware supported by Microsoft
Use Cases	HA for all VMs	FT for critical VMs	FT for critical applications

Fault Tolerance Requirements

Fault Tolerance is no different from any other enterprise feature in that it requires certain prerequisites to be met before the technology can function properly and efficiently. These requirements are outlined in the following list and broken down into the different categories that require specific minimum requirements:

- **Host Requirements:**
  - FT-compatible CPU. Check this VMware KB article for more information: http://kb.vmware.com/kb/1008027.
  - Hardware virtualization must be enabled in the bios.
  - Host's CPU clock speeds must be within 400 MHz of each other.
- **VM Requirements:**
  - VMs must reside on supported shared storage (FC, iSCSI & NFS).
  - VMs must run a supported OS. Check out the supported guest OS http://kb.vmware.com/kb/1008027.
  - VMs must be stored in either a VMDK or a virtual RDM.
  - VMs cannot have thinly provisioned VMDK and must be using an Eagerzeroedthick virtual disk.
  - VMs cannot have more than 1 vCPU configured.
- **Cluster Requirements:**
  - All ESX/ESXi hosts must be same version and same patch level.
  - All ESX/ESXi hosts must have access to the VM datastores and networks.
  - VMware HA must be enabled on the cluster.
  - Each host must have a vMotion and FT Logging NIC configured.
  - Host certificate checking must also be enabled.

Tip

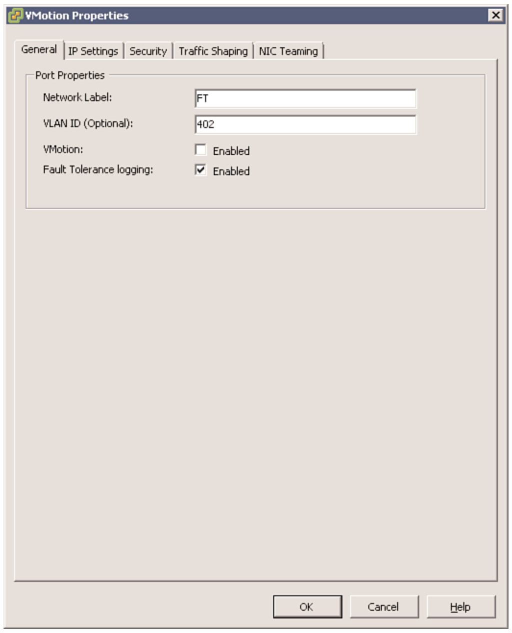
It is highly advisable that in addition to checking processor compatibility with FT, you check your server's make and model compatibility with FT against the VMware Hardware Compatibility List (HCL) here: http://www.vmware.com/resources/compatibility/search.php.

While FT is a great clustering solution, it is important to note that it also has certain limitations. For example, FT VMs cannot be snapshotted, and they cannot be Storage vMotioned. As a matter of fact, these VMs will automatically be flagged DRS-Disabled and will not participate in any dynamic resource load balancing.

How To Enable FT

Enabling FT is not difficult, but it does involve configuring a few different settings. The following settings need to be properly configured for FT to work:

- **Enable Host Certificate Checking:** To enable this setting, log on to your vCenter server and click on Administration from the File menu and click on vCenter Server Settings. In the left pane, click SSL Settings and check the vCenter Requires Verified Host SSL Certificates box.
- **Configure Host Networking:** The networking configuration for FT is easy and follows the same steps and procedures as vMotion, except instead of checking the vMotion box, check the Fault Tolerance Logging box as shown in Figure 10.7.



Port Properties

Network Label: FT
VLAN ID (Optional): 402
VMotion: □ Enabled
Fault Tolerance logging: ☑ Enabled

FIGURE 10.7 FT port group settings.

- **Turning FT On and Off:** Once you have met the preceding requirements, you can now turn FT on and off for VMs. This process is also straightforward: Find the VM you want to protect, right-click it, and select Fault Tolerance > Turn On Fault Tolerance.

Exam**Alert**

FT is an HA technology; therefore, the proper configuration of a VMware HA cluster is an imperative prerequisite for FT.

Note

You cannot turn FT On or Off from the secondary VM. It can only be done on the primary VM.

While FT is a first generation clustering technology, it works impressively well and simplifies overcomplicated traditional methods of building, configuring, and maintaining clusters. FT is an impressive technology for an uptime standpoint and from a seamless failover standpoint.

Cram Quiz

Answer these questions. The answers follow the last question. If you cannot answer these questions correctly, consider reading the section again.

1. When does the phenomenon of split-brain occur?

❍ **A.** When ESX has lost access to its shared storage

❍ **B.** When ESX has stopped receiving a heartbeat from its VMs

❍ **C.** When ESX has stopped receiving a heartbeat from the other nodes in the cluster

❍ **D.** When ESX cannot resolve DNS

2. Which incoming TCP and UDP ports are used for heartbeats and state synchronization? (Choose all that apply.)

❍ **A.** TCP 8042–8045

❍ **B.** UDP 8042–8045

❍ **C.** TCP 8012–8015

❍ **D.** UDP 8012–8015

3. How often does an ESX/ESXi host that is part of an HA cluster send out a heartbeat to the rest of the hosts in the same cluster? (Choose all that apply.)

❍ **A.** 1 second

❍ **B.** 15,000 milliseconds

❍ **C.** 5 seconds

❍ **D.** 20 milliseconds

4. True or false: VMware HA is a fault-tolerance system that allows a VM to have zero downtime in the event that its parent host should fail.

❍ **A.** True

❍ **B.** False

5. Which of the following is not a supported VM cluster in ESX?

❍ **A.** Cluster-in-a-box

❍ **B.** Cluster-across-boxes

❍ **C.** ESX-host-cluster

❍ **D.** Physical-to-virtual cluster

Cram Quiz Answers

1. **C** is correct. This phenomenon occurs when an ESX/ESXi host stops receiving a heartbeat from other members of the HA cluster, and it cannot ping its SC gateway address; therefore, answers A, B, and D are incorrect.
2. **A** and **B** are correct. The incoming TCP and UDP ports are 8042 through 8045; therefore, answers C and D are incorrect.
3. **A** is correct. ESX/ESXi 4.1 hosts that are members of the same HA cluster inform each other that they are still alive via a heartbeat that they broadcast every 1 second; therefore, answers B, C and D are incorrect.
4. **B**, False, is correct. VMware HA is not a fault-tolerance system, and consequently, in the event of a host failure, the VMs running on that host also fail. However, FT can be enabled as part of an HA cluster to extend “never fail” capabilities to VMs in the event that their host should fail. HA natively, however, does not support this feature; it would have to be enabled and configured.
5. **C** is correct. There is no such thing as an ESX-host-cluster other than the vCenter features of DRS and HA; therefore, answers A, B, and D are incorrect.