Chapter 9

Backing Up Your Virtual Environment

Good backups are a must for any environment, physical or virtual. Traditional backup methods that are used in physical environments can still be used in virtual environments, but there are alternatives available that can be more efficient and faster for backing up your VMs. VMware introduced a product called Consolidated Backup in VI3 that was a new approach to backing up ESX hosts. Consolidated Backup leverages a proxy server to back up virtual disks to offload the burden from the VM and to provide networkless backups by backing up the VM's virtual disks over the SAN fabric. Virtual environments allow for some different approaches to backups because of their snapshot abilities and unique architecture. Taking a snapshot of a VM's virtual disk suspends writes to it and allows for backups to occur without the possibility of files being modified during the backup. In addition, some third-party backup products have been specifically developed for ESX hosts and are worth considering for even better integration, efficiency, and backup speed.

You can use several methods to back up your VMs, and they all have advantages and disadvantages. The simplest and least efficient is to use traditional backup agents running inside the guest operating system. Another method is to use backup scripts that run inside the Service Console that snapshot a VM and make a copy of its virtual disk file to an alternative storage device. You can also use VMware's Consolidated Backup product or one of the third-party backup solutions. Let's take a look at each method so that you can decide which one will work best in your environment.

Traditional Backups

Traditional backups are where a backup agent is installed on the guest operating system, and the backup system communicates with the agent to back up the files on the server. This method can cause high disk and network I/O and high CPU utilizations on the servers that are being backed up. Although this might be fine on a physical server, when done in a virtual environment it can affect other VMs running on the same host because they are all competing for the host's resources and a backup on one VM can leave less resources for the others.

Advantages of this method

- Simple to deploy because most environments are already using this method.
- Easier restores of individual files than other methods.
- No additional software is needed and no process changes are required if you are already using this method.

Disadvantages of this method

- Causes excessive resource usage on hosts.
- Slower than other methods.
- No bare-metal restores, requires OS to be installed first.

If your VMs do not see that much usage during your backup windows and the increased resource utilization is not a problem in your environment, you may consider this method. If you do use this method, be sure to monitor your host's performance during backups to ensure you are not encountering resource bottlenecks that could be affecting other VMs on the host. Also, try to ensure that you only back up a single VM concurrently on a particular host, if possible, to avoid bottlenecks, or stagger your backups so that you are not backing up too many VMs on the same host or LUN at the same time.

Backup Scripts

Backup scripts typically work by taking a snapshot of a VM and then copying a VM's large VMDK file to another storage device where it can be backed up to tape or just left there to use in case a file or the VM needs to be restored. These scripts run inside the ESX Service Console and can use the Perl language or other scripting methods (such as bash) that are supported by ESX. This method can be slow and inefficient, and is useful as a complementary method to another backup method. Because the VMDK file as a whole is being backed up, it makes individual file restores more difficult. To restore individual files, the VMDK file needs to be mounted by another VM or utility, and then the file copied to the original server. This method is best suited for restoring a whole VM image, and is useful for disaster recovery scenarios. You can read detailed information about how snapshots work in Chapter 11, "Advanced Topics."

Advantages of this method

- Simple to set up and use.
- No additional software needed.
- Good for whole VM image restores (bare metal).

Disadvantages of this method

- Individual file restores are difficult.
- Requires access to the ESX Service Console.
- May not work with ESXi.
- Can be slower than other methods.

If you choose to use this method either as your primary backup solution or in conjunction with another method, you might want to check out some of the user-developed scripts available from http://vmts.net and www.xtravirt.com.

Consolidated Backup

Chapter 2, "Planning Your Virtual Environment," provided a detailed summary of the Consolidated Backup application that comes with the VI3 Enterprise license, and rather than repeat everything that was written there, this chapter contains just a brief description of the feature and how it works.

VMware Consolidated Backup (VCB) is a Windows-based application that provides a centralized backup facility to back up VMs through a proxy server without affecting the VM itself. It is an alternative to traditional agent-based backup methods and was designed specifically for virtual environments to minimize host server impact during backup operations. VCB is an enablement technology and cannot back up VMs by itself, but instead works with third-party backup products to help offload backup overhead from the host servers. It integrates with most major software backup providers and eliminates the need to back up VMs over the network when using a Fibre Channel SAN. It works by taking a snapshot of a VM, which suspends the writes to the original disk, and then copies the original disk to a proxy server, where the VM's disk is mounted on the backup proxy server. It then backs up the contents of the VM as a virtual disk image or as a set of files and directories, and then unmounts the virtual disk and deletes the snapshot. This backup is performed over the SAN fabric and not over the network, which results in greater backup speeds.

Advantages of using VCB

- Eliminates the need for installing a software backup agent on every VM that you want to back up.
- New version now supports backing up VMs on SAN, iSCSI, NFS, and local disk.

- Supported by most major backup software, including Symantec NetBackup and Backup Exec, CA Arcserve, CommVault Simpana, HP Data Protector, and IBM Tivoli Storage Manager.
- Supports both file-level full and incremental backups for VMs running Windows and image-level backups for VMs running any operating system (Windows and Linux).
- Provides OS-level quiescing of file systems, yielding a file system-consistent backup.
- Integrates with VSS to provide application-level consistency for applications that are VSS aware.
- Offloads CPU, disk, and network resource usage from the host to the proxy server.

Disadvantages of using VCB

- More costly, requires a separate physical server and Fibre Channel card to back up VMs on a SAN. (On newer versions of VCB, you can optionally run VCB inside a VM to avoid this.)
- Limited OS support (Windows and Linux only).
- May require excessive additional space on the SAN to back up VMs.
- Individual file restores require more steps and are more difficult.
- Can be complicated to set up, configure, and maintain.

Consolidated Backup is a great a solution to remove the backup burden from your hosts, and is more suited for larger environments. The first release of Consolidated Backup was rather limited, but VCB has evolved with each new release, with more support and better features, and you should definitely consider using it in your environment if you are using one of the backup applications supported by it. VMware has released several guides that you can use to help implement VCB. These are listed in the "Additional Resources" section at the end of this chapter.

Third-Party VI3-Specific Backup Products

Several third-party backup products have been developed specifically for backing up VI3 environments and are a great alternative to traditional backup methods and provide more options and greater flexibility when backing up and restoring your VMs. These products all work by copying VM virtual disks over the network from a source datastore to a destination disk-based storage device (local, SAN, NFS, iSCSI, and CIFS). In addition, most of these products can integrate with VCB and can do full and differential backups and both whole VM (bare-metal) and individual file restores.

These products all vary in price, performance, and features, and you should evaluate each one to see which one integrates best into your environment and satisfies your requirements.

You should also make sure that any product you look at supports everything in your environment, including VCB, ESXi, guest operating systems, and your storage devices. Make sure to request an evaluation copy of each product and install and use it to make sure it will work properly for you before selecting a backup product.

Vizioncore vRanger Pro

vRanger Pro provides image-level hot backups while VMs are still running. It has an easy to use graphical interface, startup wizards, and VirtualCenter integration. Advanced functionality, such as VCB integration, VSS quiescing, and differential backups, is enabled through its GUI, with no manual RPM installs or complicated scripting. A differential engine backs up only the changes made since the last full backup image, reducing the size of the backup files on disk. Differential characteristics and retention policies can easily be configured to suit customer needs. vRanger Pro supports differential backups via the VCB framework. It also supports backing up and restoring both ESX and ESXi host servers. You can find more information about vRanger Pro at http://vizioncore.com/vRangerPro-features.html.

Veeam Backup

Veeam Backup provides enterprise-ready functionality, including an intuitive user interface with simple backup and replication job management; restoration of a full VM or a single file with just a few mouse clicks; comprehensive statistics, reporting, and email alerting; and integrated ESX file management through Veeam FastSCP. Veeam Backup can work with the VCB framework for backup and replication purposes. Using its proprietary "VCB on-the-fly" technology, Veeam Backup does not require storing VM images on the VCB proxy during backup, allowing for faster backup directly to the target without additional disk space requirements. Veeam Backup also supports incremental backup and replication through VCB, which is not supported natively by VCB. Veeam Backup features limited support for ESXi servers through VCB. You can now back up ESXi servers using the VCB option in the Backup Wizard. File-level recovery is fully supported for guests running on ESXi, whereas full image restore is only supported to ESX 3.x servers. After the guest has been restored to an ESX 3.x server, you can then VMotion the restored machine back to the ESXi server. Replication is also supported with ESXi as a source and ESX 3.x as a target. You can find more information about Veeam Backup at http://veeam.com/vmware-esx-backup.html.

esXpress

esXpress by PHD Technologies makes use of virtual backup appliances (VBAs) to quickly and securely back up VMs without affecting console resources or VM availability. VBAs are just tiny VMs that back up VMs. Unlike dedicated hardware solutions, VBAs are cost-effective and scale with your environment. And because of their distributed nature, VBAs are completely fault tolerant, very low impact, and yet yield incredibly fast speeds. esXpress allows for multiple transport protocols and targets, including FTP, SSH, SMB/CIFS, iSCSI, SAN, local storage, and more. Additional esXpress features include support for both file- and image-level backups, compression for increased backup speeds, and integration with the VI Client. You can find more information about esXpress at http://esxpress.com/index.php.

Did You Know?

Many backup applications use a process called quiescing before backing up servers. This process ensures that the data on the disk is in a state suitable for backups and includes operations such as flushing dirty buffers from the OS memory cache to disk. Different types of quiescing can be done. When it is done at the operating system level, it may not be aware of applications that are running and consequently might not work properly. A better form of quiescing is done at the application level (for example, Exchange, SQL) and is aware of the application and so it can properly write data before backing up the server.

What to Back Up

When it comes to backing up VMs, you have two options: back up the large VMDK virtual disk file (known as image-level backups) at the virtualization datastore level or back up the individual files (known as file-level backups) at the guest operating system level. Both options have advantages and can be used jointly to provide different restoration alternatives for your VMs.

File-level backups allow for individual file restores, which are useful when you have just a few files to restore to a VM and do not need to restore all the data just to restore the individual files. File-level backups can be done using traditional backup methods by running a backup agent inside the guest operating system of a VM. In addition, many backup applications that have support for virtualization can perform file-level backups through nontraditional methods such as using VMware’s Consolidated Backup application. File-level backups are great for restoring a small number of files, but restoring a whole VM can be more time-consuming because you typically need to install a guest operating system and backup agent before you can restore the rest of the VM data.

Image-level backups allow for whole VM or bare-metal restores and are useful when you need to quickly restore a VM to a previous state or for disaster recovery purposes, which require bare-metal restores rather than individual file restores. To perform an image-level backup, you need to use a product or method that supports it rather than traditional backup agents, which work inside the guest operating system and do not have access to the virtualization layer and the VMDK virtual disk file. You can still do individual file restores with image-level backups, it just typically takes a few more steps to do it. To do this, you can restore the disk image and mount it on another VM to access the file and then copy it back to the source VM. In addition, VCB can make use of vCenter Converter to restore individual files.

So, which method should you use? It really depends on the backup method you plan to use. If you are going to use traditional backup methods that will enable you to restore individual files, you might also look to use one of the scripts or third-party products that can copy the image to disk storage so that you have that available as another restore option. This is a good way to make use of local disk on your hosts that you might otherwise not use or a low-cost NFS or iSCSI system or server. You might back up your VMs every day using the file-level method and back them up once a week using the image-level method. If you instead plan to use VCB or a third-party backup product, many of them will already provide you with the ability to do both file- and image-level backups. Image-level backups are great when you want to retire a physical or virtual server so that you have a bare-metal copy of it if you ever need it later.

Another question that is often asked is whether you should install a backup agent inside the ESX Service Console to back it up and the VM files from the VMFS datastores. Generally this is not a good idea for several reasons. The first is that it can have a big impact on the performance of the host when backups are running, which will also affect the VMs on that host. Second, you should avoid installing any software or utilities on the Service Console (with the exception of hardware management agents) because it could cause instability, which could lead to your host crashing. The last reason is that it is easier to rebuild an ESX host if there is a problem with it rather than try to restore it. When you rebuild an ESX host, you have the option to leave all the VMFS datastores intact so that your VMs can be re-registered with the host after it has been rebuilt. You will lose some configuration information (for example, vSwitches), but you can back up certain key files from your host to restore this. If your configuration is fairly simple, you can reconfigure your host instead. It is also a good idea to periodically run the vm-support script on your ESX hosts, which will provide detailed information about the host configuration and much more that you can later use as a guide to reconfigure it.

To back up the key ESX Service Console configuration files, just back up the /etc/vmware directory. You can do this with the `tar` command by typing `tar -cvf_esxhostnamebackup.date.tar /etc/vmware`. This will create a single tar file with the contents of that directory that you can copy to another location for safekeeping. You can do this periodically for each of your hosts to use if you ever need to restore them. VMware has a knowledge base article that covers how to back up and restore ESX host configurations (KB article 1000761, http://kb.vmware.com/kb/1000761). To back up the configuration of ESXi hosts, you can use the vicfg-cfgbackup.pl script that is part of the RCLI.

Attaching Tape Drives to ESX Hosts

Although it is best to avoid doing this, sometimes it is necessary, especially in smaller environments that do not have a dedicated tape backup server. By installing a tape drive in your ESX host, you can use a VM to access it to back up the VMs on your host or other hosts. If you want to do this, you should follow a few guidelines to ensure success:

- Your tape device must use an Adaptec SCSI I/O adapter. (A dedicated one is recommended.) You cannot use an IDE or FC tape device, because even though the host may see it the VM will be unable to because it supports only virtual SCSI controllers.
- Make sure that the Adaptec SCSI controller that you are using is listed on the ESX hardware compatibility list for I/O devices.
- To check whether ESX can see the tape drive, use the following command in the Service Console: `cat /proc/scsi/scsi`. This command will list all attached SCSI devices. When you can see the tape device in the ESX Service Console, install a backup agent that supports ESX. If one is not available, use a backup agent that supports Red Hat Linux Enterprise version 3.
- When selecting a SCSI controller for your VM, make sure to choose LSI Logic rather than BusLogic.

For more information about using a tape device in your ESX host, see VMware knowledge base article 1000024 (http://kb.vmware.com/kb/1000024).

Watch Out!

One risk of attaching a tape drive to an ESX host is that if the tape drive gets hung up, which is not uncommon, you will need to reboot the ESX host to clear it so that you can use it again.

Summary

Backups are important, so make sure you choose a solution that will work for the needs of your environment. Backups are like having good insurance policies: You pay your premium each month, and you hope you never need to use them; but if something happens, you are real glad that you have them. Restoring your data is even more important than backing it up. You need to make sure the solution that you use can easily restore files as needed in your environment. You should also plan your backup strategy around your current or future disaster recovery strategy because good backups are the foundation for any disaster recovery plan. If you plan on doing disaster recovery, look for a backup strategy that supports your disaster recovery requirements for recovering from an event. In addition, you may look into some of the products that can replicate your VMs from your production environment to a disaster recovery environment. When you implement a backup solution in your virtual environment, make sure you test the restore capability so that you don't run into any surprises later when trying to restore critical data.

Additional Resources

For more information about backing up and restoring your environment, check out these resources:

- Virtual Machine Backup Guide
  http://vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_vm_backup.pdf
- Consolidated Backup in VMware Infrastructure 3
  www.vmware.com/pdf/vi3_consolidated_backup.pdf
- Using VMware Infrastructure for Backup and Restore
  www.vmware.com/pdf/esx3_backup_wp.pdf
- VMware Consolidated Backup
  www.vmware.com/files/pdf/vcb_best_practices.pdf
- VMware Consolidated Backup: Improvements in Version 3.5
  www.vmware.com/files/pdf/vcb_35_new.pdf
- VMware Consolidated Backup—Partner Integration Guide
  www.vmware.com/files/pdf/vcb_partner_integration_guide.pdf
- Best Practices for Architecting VCB Enabled Solutions (VMworld 2007 presentation, free registration required)
  http://vmworld.com/vmworld/mylearn?classID=11154