Eric Siebert

THE EXPERT GUIDE TO VMware Data Protection and Disaster Recovery

Sponsored by VEEAM

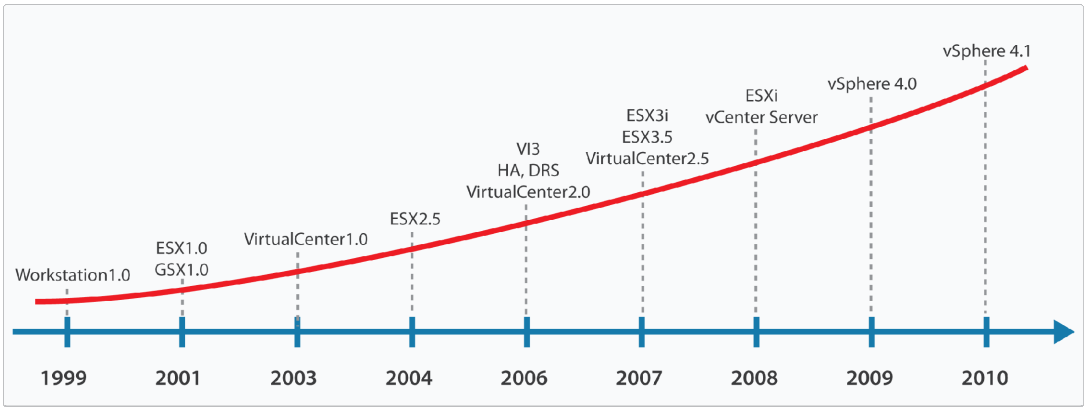
CONTENTS

INTRODUCTION

Virtualization is not just a fad and is definitely here to stay with most companies either currently using it or planning to in the near future. Running servers as virtual machines just makes sense on so many different levels when you factor in all the great benefits that they provide. Virtualization technology continues to mature at a very rapid pace with VMware leading the pack in almost all areas with their advanced features and widespread popularity. In this chapter we will cover some of the key concepts of virtualization and explain the different types of virtualization as well as go into detail into what a virtual machine actually is. We'll also cover how virtual machines differ from traditional physical servers and why managing virtual machines requires a different mindset as well as the usage of applications designed specifically to work with virtual environments.

A BRIEF HISTORY OF VMWARE

VMware basically invented x86 server virtualization with their release of Workstation 1.0 in 1999 followed by ESX (Elastic Sky X) 1.0 and GSX (Ground Storm X) 1.0 in 2001. In 2003 VMware introduced VirtualCenter 1.0 with their groundbreaking new VMotion feature to migrate running virtual machines between hosts without interruption. In 2004 VMware released ESX 2.5 which was another major release for ESX. VMware Infrastructure 3 (VI3) came out in June of 2006 and was a major new release for VMware which introduced big new features like High Availability (HA) and Distributed Resource Scheduler (DRS) as well as a completely overhauled version of VirtualCenter (2.0). In 2007 VMware introduced their new hypervisor platform, ESX 3i that designed to eventually replace ESX. ESXi was designed with a much smaller footprint due to its smaller management console compared to the large Service Console that ESX uses. As a result VMware touted the new feature of being able to boot and run ESXi from a 1GB flash drive. Also in 2007 VMware released their next version of ESX (3.5) and VirtualCenter (2.5) that had more big new features like Distributed Power Management (DPM) and Storage VMotion.



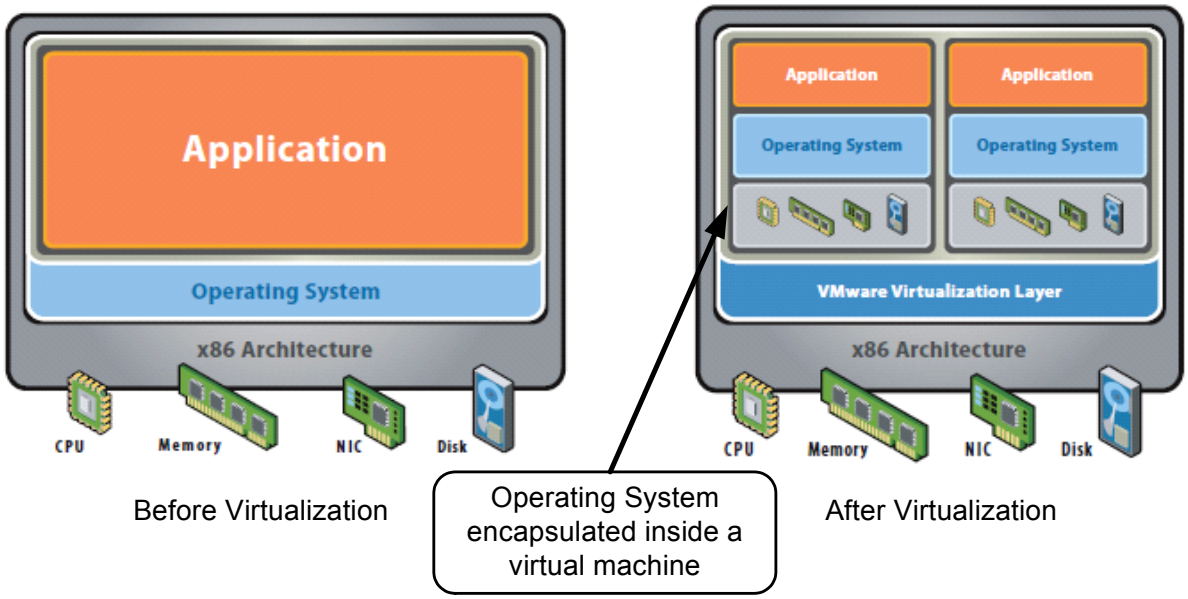
In 2008 VMware renamed ESX 3i to ESXi and announced they were going to give away the feature limited, entry level version of ESXi for free. This was mainly done in response to Microsoft making the move to give Hyper-V away for free. Also in 2008 VMware did some major product renaming and started to embrace the "v" naming convention for many of their products. As a result VirtualCenter was renamed to vCenter Server and the other management and automation products were all given the vCenter name as well (i.e. vCenter Lab Manager). In May of 2009 VMware released their much anticipated and biggest release yet, vSphere 4.0. This release was packed with new features and enhancements including the new Fault Tolerance (FT) feature, vStorage APIs, Distributed vSwitches, Host Profiles and much more. The vSphere 4.0 release was followed up by vSphere 4.1 which was released in July 2010 and as usual VMware packed more great features into it including Storage & Network I/O control as well as enhancements to many existing features. vSphere 4.1 also signaled the end of the road for ESX as VMware announced that 4.1 would be the last major release to include ESX and VMware would instead focus their development efforts on ESXi.

VIRTUALIZATION ARCHITECTURE

Virtualization is a very general term for simulating a physical entity by using software. There are many different forms of virtualization that may be found in a data center including server, network and storage virtualization. When talking about server virtualization there are many unique terms and concepts that you may hear that are part of the technology that makes up server virtualization. In this section we will cover some of those concepts as well as explain some of the other forms of virtualization.

The Hypervisor

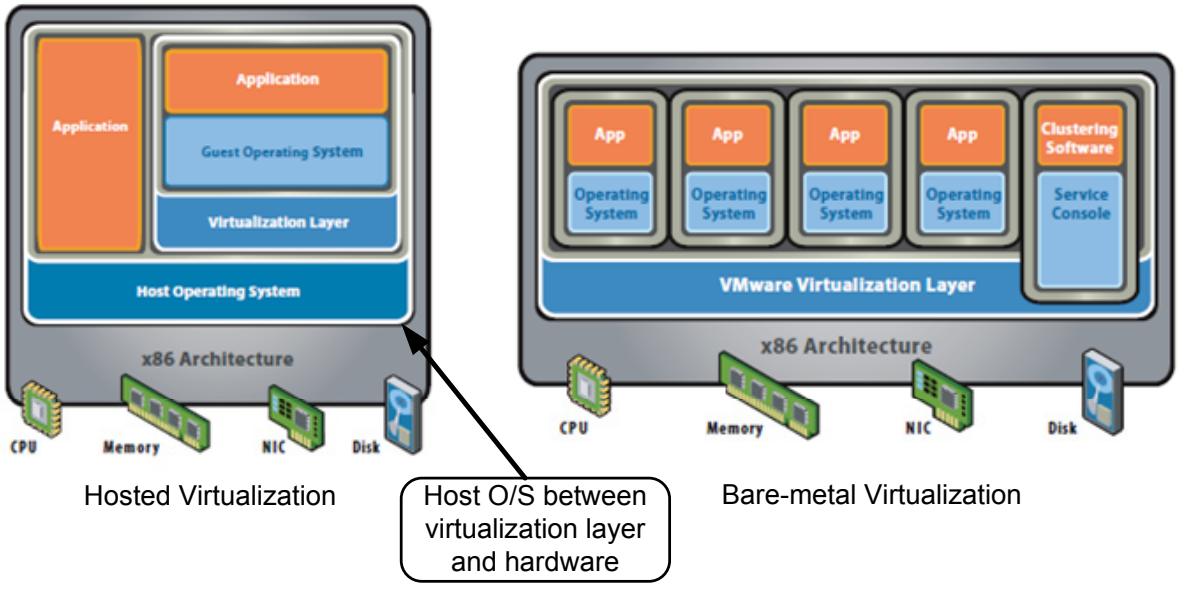
The hypervisor is essentially the brains of server virtualization and in ESX & ESXi it is also known as the VMkernel. A hypervisor is a program that allows multiple operating systems to share a single physical hardware host. Each guest operating system (virtual machine) running on the host appears to have exclusive access to the host's processor, memory and other resources. However, the hypervisor is actually controlling the processor and other host resources and only allocates what is needed to each guest operating system while simultaneously ensuring that each guest operating system is isolated from the others and cannot disrupt each other. In this way each VM is living inside its own small little ecosystem inside the host, blissfully unaware that its home is actually shared with many other tenants. An analogy for this is that while a physical server may be considered a standalone house a virtual host is more of a townhouse building.



There are two different types of hypervisor models, the first are considered bare-metal or type-1 hypervisors and the others are considered hosted or type-2 hypervisors. The characteristics of each type of hypervisor are listed in the below table.

HOSTED HYPERVISOR CHARACTERISTICS	BARE-METAL HYPERVISOR CHARACTERISTICS
Requires a host operating system (Windows/Linux/Mac), installs like an application	Installs directly on the bare metal of a physical server
Virtual machines can use all the hardware resources that the host can see	Virtual machines' resource usage can be limited by advanced resource controls
Maximum hardware compatibility as the operating system supplies all the hardware device drivers	Runs on a narrower range of hardware due to limited driver support
Overhead of a full general-purpose operating system between the virtual machines and the physical hardware results in performance 70-90% of native	Because there is no overhead from a full host operating system performance is 83-98% of native. There is a small bit of overhead from the virtualization layer of the hypervisor
Limited feature support	Many advanced features for resource management, high availability and security

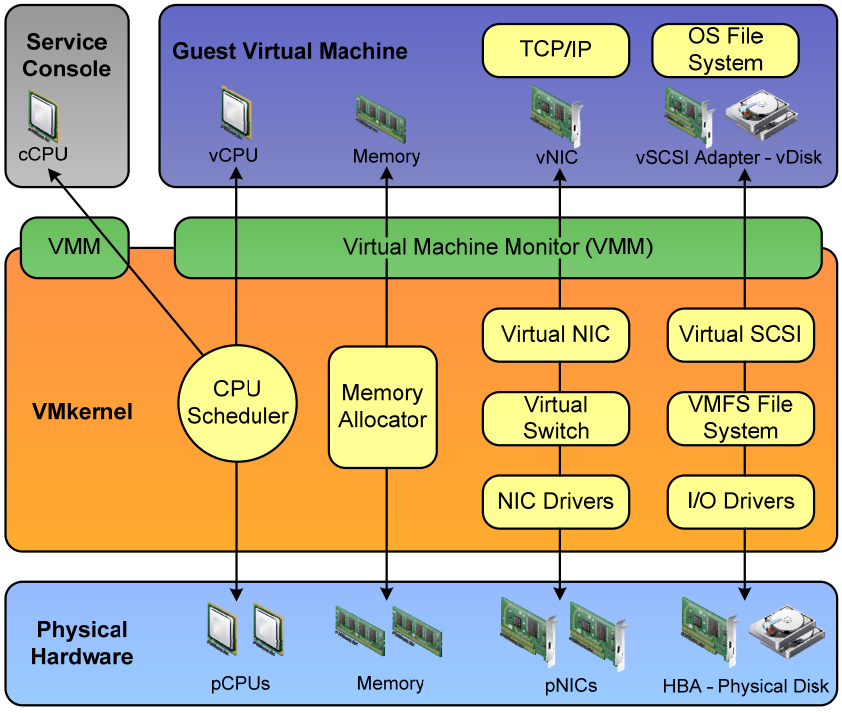
Because of less overhead of the virtualization layer, bare-metal hypervisors support more VMs per physical CPU than hosted hypervisors do. In VMware's line of products their bare-metal hypervisors are ESX & ESXi, and their hosted hypervisors are Workstation, Server, Fusion and Player.



The VMkernel is considered a microkernel as it has the bare minimum of software code to provide the mechanisms to support an operating system. It contains 3 basic outside interfaces: one for hardware, one for guest systems (VMs) and one for the management console OS. The VMkernel directly handles all VM access to CPU and memory resources; access to other hardware resources is done through modules (drivers) that are adapted to work specifically with the VMkernel. These modules dictate what hardware can be used with ESX & ESXi and are the basis for the Hardware Compatibility Guide that VMware publishes. Much like in Windows if a driver for a certain model storage or network adapter is not installed you wouldn't be able to use the device. VMware has included a mechanism for loading additional drivers on to ESX & ESXi hosts but they must be certified by VMware first.

In VI3 the VMkernel was compiled as a 32-bit application and therefore supported both 32-bit and 64-bit server hardware. Despite having a 32-bit VMkernel you could still run 64-bit VMs on a VI3 host because VMs do not run directly on the VMkernel, they instead run on another component called the Virtual Machine Monitor (VMM). The VMM is a process that runs in the VMkernel which is responsible for hardware abstraction and emulating hardware to the guest OS. The VMM decouples the VM from the VMkernel and is responsible for the execution of the VM by partitioning and sharing host resources including CPU, memory, timers & I/O devices. The VMM passes all storage and I/O requests to the VMkernel and passes all other requests go to a special VMX process which handles the emulation of all non-critical hardware resources like CD-ROM and USB ports. Each VM has its own VMM process devoted to it so a host with multiple VMs running will have one VMM for each running VM.

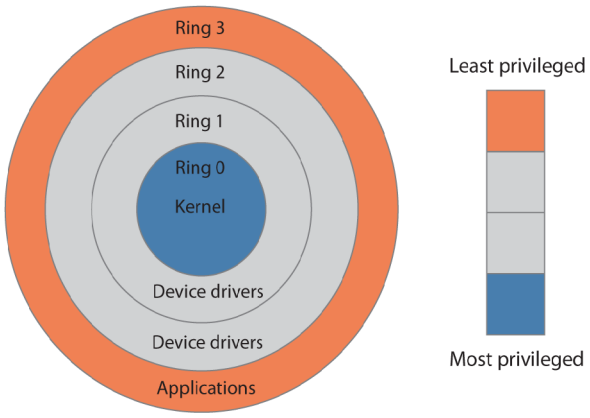
The VMM contains all the logic to run both 32-bit and 64-bit operating systems which is why you could run 64-bit VMs on a 32-bit VMkernel. Another limitation of 32-bit applications is only being able to address up to 4GB of memory; the 32-bit VMkernel was able to get around this limitation by not looking at all the physical memory at once but instead changing the view of the 4GB of memory that it was seeing whenever it needed to. The VMs also have their own mappings to memory on the host so they can see memory that the VMkernel doesn't need to see. In vSphere, VMware changed the VMkernel and made it a 64-bit application.



The VMM emulates the hardware that guest operating systems see and provides the same type of emulated hardware to each VM. The VMM provides the mapping from the emulated hardware provided to the VM to the actual physical hardware devices of the host. This provides the connection from the driver used inside the guest operating system for the emulated hardware to the driver used inside the VMkernel for the actual hardware. VMware Tools that is installed inside the guest operating system contains special paravirtualized device drivers that differ from the standard guest OS drivers as they are aware of the virtualization layer. These drivers can take advantage of the host device capabilities to provide better I/O throughput and reduced CPU overhead. VMware Tools also provides a backdoor connection between the guest OS and the VMkernel for services like power management and time sync.

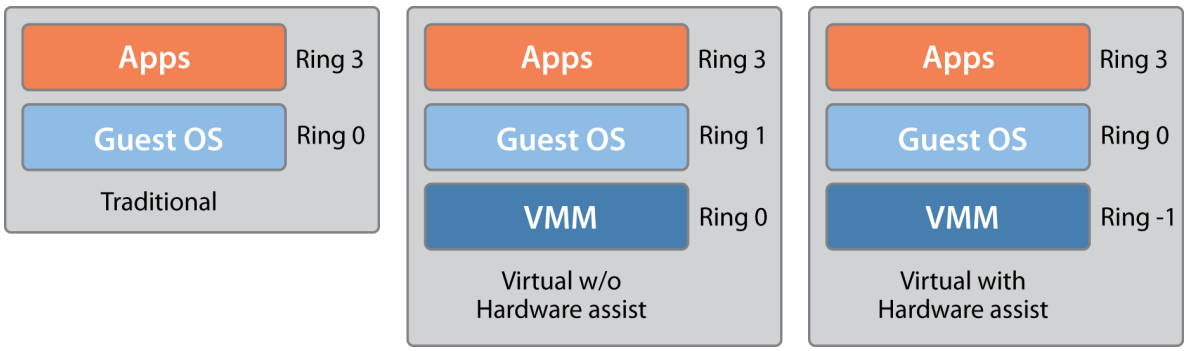
Rings in virtualization

x86 CPUs use a concept called "Rings" which provide a range of privilege levels in which code inside the guest operating system can execute. Rings are designed to protect the operating system from faults that may occur and also to provide security from any malicious code that may try to compromise the guest OS. They provide protection to resources such as memory, I/O devices and certain privileged machine instructions. The enforcement of these protection rings is handled by the CPU hardware of the server which can operate in different privilege levels. Ring privilege levels typically range from Ring 0 which has the highest privileges and interacts directly with the physical hardware to Ring 3 which has the least privileges. The operating system usually runs in Ring 0 and is known as system space, kernel mode or supervisor mode. Device drivers usually run in Ring 1 & 2 and applications usually run in the less privileged Ring 3.

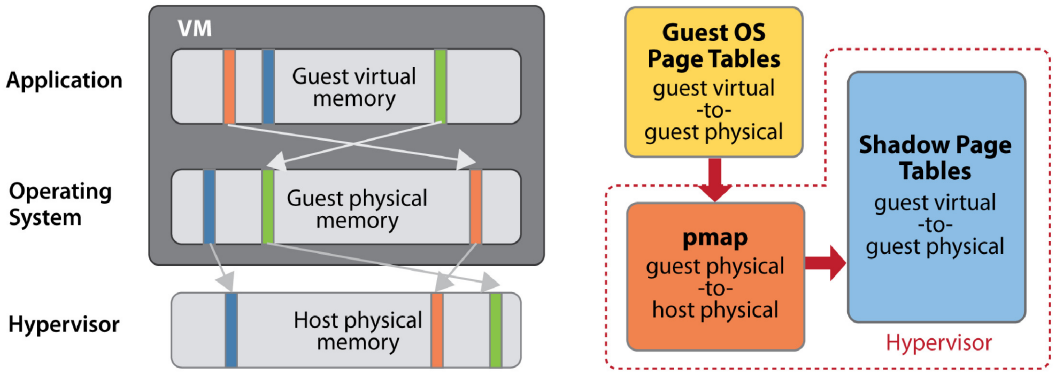


Access between rings is controlled by special gates which allow an application running in Ring 3 to access things like I/O devices by making a special call to the operating system kernel running in Ring 0. In this manner the kernel can directly control exactly what applications running on the server can do and access. With virtualization technology though this model had to change as the VMM had to be inserted between the operating system running in Ring 0 and the physical hardware. This was done using a technique called binary translation (BT) which put the VMM in Ring 0 and moved the operating system from Ring 0 to Ring 1. Operating systems are designed to run in Ring 0 though and to allow the operating system to be able to run without modifications the VMM plays a neat little trick by fooling the operating system into thinking it's running in Ring 0. It does this by trapping all privileged instructions destined for Ring 0 and replacing them with instructions that perform the privileged operations in the VM instead of on the physical hardware. This operation is completely transparent to the VM and essentially encapsulates it into its own little world.

This binary translation does come with a cost though as there is extra CPU & memory overhead when performing this operation compared to a guest OS running natively on the physical hardware. Intel & AMD took note of this and engineered new features directly into their CPUs to eliminate the need for the binary translation operation. What they came up with was to introduce a new privilege level below Ring 0 where the VMM could reside so the operating system could remain in Ring 0 and binary translation was no longer needed. This new ring was called Ring -1 and is referred to as hardware assisted virtualization. The technology from Intel is called VT-x and from AMD is called AMD-V and is available on most modern servers sold today.



Besides managing CPU access the VMM must also manage physical memory access which is normally done by the Memory Management Unit (MMU) of the physical hardware. For this the VMM uses a software technique called Shadow Page Tables to do the mapping from linear, physical, virtual and machine memory addresses. Again this technique comes with overhead so as a result both Intel and AMD developed their own hardware technologies to help the VMM perform memory translations. AMD's implementation of this is called Rapid Virtualization Indexing (RVI) and Intel's is called Extended Page Tables (EPT) and provide better performance than using Shadow Page Tables.



With all these technologies and techniques available to use to help virtualize a guest OS the VMM has the ability to choose which ones to use to provide the best performance for the VM. This is referred to as the Monitor Mode and is determined when a VM is first powered on, the deciding factors of which mode to use are based on the physical CPU features and also the guest OS version and type (32-bit or 64-bit). The available modes that it can choose from are listed below:

- Binary Translation using Shadow Page Tables (BT-swMMU)
- VT-x or AMD-v using Shadow Page Tables (HV-swMMU)
- VT-x using EPT or AMD-V using RVI (HV-hwMMU)

For most 32-bit VMs the BT-swMMU is the chosen monitor mode and for most 64-bit VMs the HV-hwMMU monitor mode is chosen. It is possible to override the MMU that is chosen and force a VM to run using a specific mode but this is not recommended as the MMU usually chooses the best choice for the VM for optimal performance.

CPU Scheduler

The CPU scheduler that is built into the VMkernel has a daunting job; it must organize and handle all the requests for CPU time from the VMs running on the host. This can be quite a job on a busy host and is akin to a maître d' at a very busy restaurant trying to find tables for all the customers, not only must he find a table for each he must find the right size table. This can be considered controlled chaos as a host has limited CPUs and the virtual CPUs can outnumber the physical CPUs anywhere from 4:1 to 8:1. Earlier versions of ESX had a pretty rigid scheduler especially when it came to co-scheduling multiple requests simultaneously. For example when a VM with four vCPUs needed CPU time it had to find four pCPUs available simultaneously to handle the request. On a busy host with a limited number of pCPUs this could be a challenge. To help with this VMware introduced relaxed co-scheduling in VI3 that allowed only a subset of the VM's CPUs to be scheduled simultaneously. This made it much easier for the CPU scheduler to do its job and improved performance and efficiency. In vSphere VMware significantly improved the CPU scheduler and even further relaxed the co-scheduling requirements.

With all those requests coming in for CPU time the scheduler needs to make some intelligent decisions on how to schedule and handle them all. It can't simply tell them to form a single line and handle requests on a first-come, first-served basis. It must schedule requests in a manner that is fair to all VMs so they are not kept waiting too long and it must make the VMs think that they completely own the CPU. To help with this it uses a proportional share based algorithm that uses entitlements that are calculated on a number of factors including VM share, reservation and limit settings. Priorities for CPU time are calculated by using entitlements set by the administrator in combination with how consumed (usage) the CPU resources are. The ratio of the consumed CPU resource to its entitlement is used to determine the priority to give a VM. This allows for fair and balanced CPU scheduling and also allows the settings on the VM to help determine its CPU priority. In addition to scheduling CPU time the scheduler also optimizes the placement of vCPUs onto different sockets to maximize overall cache utilization, and to improve cache affinity by minimizing virtual CPU migrations. The CPU scheduler tries to spread loads across all sockets by default to maximize the aggregate amount of cache available to the running virtual CPUs.

Differences between ESX & ESXi

In the beginning there was only ESX and there were no decisions to make when it came to deploying a bare-metal hypervisor. Then a few years ago ESXi came along and caused people to have to make a decision between the two bare-metal hypervisors. Many people were confused by this as ESX & ESXi had some big differences from each other including a very different architecture as well as some feature differences. As far as architecture goes, ESX & ESXi run the exact same VMkernel code and the difference between the two bare-metal hypervisors lies in their management console. The management consoles are not built into the hypervisor but actually run on top of it as privileged virtual machines. The management console is tightly integrated with the VMkernel and has a direct interface to communicate with it. In this manner it serves as the middle man between administrators and the VMkernel. When you connect to a host using the vSphere Client you are connecting to the management console of the host and not the VMkernel directly. With ESX the console is referred to as the Service Console and basically has a

full operating system installed inside it. The Service Console OS runs a modified Red Hat Linux operating system; in vSphere it uses the 2.6.28 kernel. As a result most traditional Linux commands can be used when logged into the Service Console. In addition VMware has added many custom commands that are specifically used to manage the VMkernel, many of these commands begin with esxcfg-. Because it uses a full operating system the Service Console and its many disk partitions are typically about 9GB in size. Administrators can log in to the Service Console locally or remotely using an SSH client.

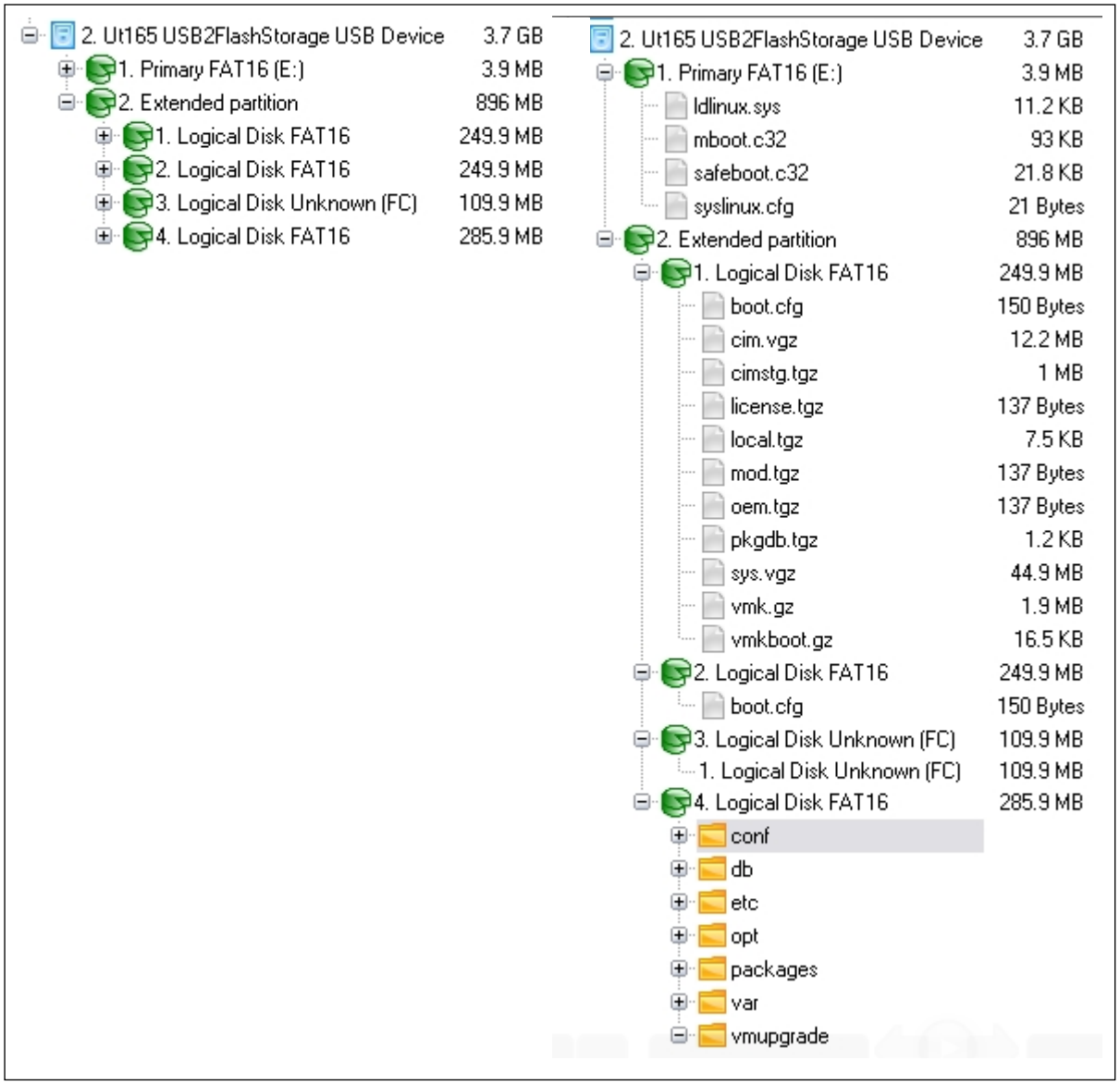
With ESXi the console is simply called the Management Console and is using an ultra-light operating system which dramatically reduces its disk footprint compared to ESX. Instead of a full operating system, ESXi runs POSIX (Portable Operating System Interface for Unix) which is a set of API, shell & utility standards for Unix variants. An application called Busybox runs inside POSIX and is a single binary that contains many applications packaged inside of it. As a result the total disk footprint of ESXi which includes the VMkernel and management console is around 60MB. Because of the compact size the management of ESXi is a bit different than ESX, with ESX you had a full OS that you could log into and perform various tasks. ESXi instead uses what is called the Direct Console User Interface (DCUI) which is a small menu-driven interface where you can perform minimal configuration and administrative tasks.

Because ESXi does have an OS there is also a console that can be acessed that shares some similarities with ESX. This console is referred to as Tech Support Mode (TSM) because VMware intended it to only be used for troubleshooting purposes when instructed to do so by VMware's tech support. Prior to vSphere 4.1 TSM was accessed in a secretive manner by pressing Alt-F1 and typing the word 'unsupported' and then entering the root password. You could also modify a configuration file to enable remote SSH access to TSM. In vSphere 4.1 VMware officially began supporting TSM and made it accessible through the DCUI and added some options for managing it from the vSphere Client as well as made enabling SSH easier. Because the TSM console is much smaller than the ESX Service Console it has a more limited set of commands that can be used inside of it. Some of the basic Linux commands are present in TSM along with many of the VMware specific commands. While not as powerful as the ESX Service Console the TSM console still provides a good way to manage and troubleshoot ESXi hosts.

Because of ESXi's small size it can be installed on to a flash drive and used to boot from on a server eliminating the need to install ESXi on the server's local disk. The total size of an ESXi installation is under 1GB and consists of several disk partitions. There is one 4MB primary partition on the disk that is the bootloader that runs when the host boots up. There is also an 896MB extended partition with 4 logical disks that are used for the following:

1. A 250MB logical disk that contains the core hypervisor code (VMkernel) which is packaged into several files that total up to 60MB. Server manufacturer customizations (i.e. HP, Dell, IBM) are also stored here.
2. A 250MB logical disk that is used to hold a backup copy of the hypervisor code whenever any updates are applied to ESXi. This disk is initially empty except for one 150 byte boot.cfg file until an update to ESXi is applied. Once an update is applied all the files from logical disk 1 are copied to logical disk 2 before the new ESXi image is copied to logical disk 1.
3. A 110MB logical disk that is initially empty but can be used to hold diagnostic core dumps.

4. A 286MB logical disk that contains VMware tool ISO files, the vSphere Client installer, other tools and drivers and runtime storage. These files are not part of the VMkernel but are mostly auxiliary files.



Being able to boot ESXi from a flash drive is a nice feature and is officially supported from VMware as long as you use one of the server vendor model flash drives that are listed on the Hardware Compatibility Guide. Because of the smaller architecture ESXi is much easier to patch as patches are delivered as an entire new version in a single file which replaces the previous ESXi version. With ESX patches had to be applied individually and they often had dependicies on other patches being applied.

VMware has continually stated that ESXi is their future architecture and that ESX will no longer be developed at some point in the future. With the release of vSphere 4.1, VMware has made it official and ESX and its Service Console will not be present in the next major release of vSphere.

WHAT IS A VIRTUAL MACHINE?

You hear the term virtual machine all the time and you may work with them on a day to day basis but have you ever wondered what exactly a VM is and what components it consists of? While VMs live in the RAM of a host they do have a physical presence as well in the form of files that reside on the host. In this section we will cover the various components that make up a virtual machine in detail.

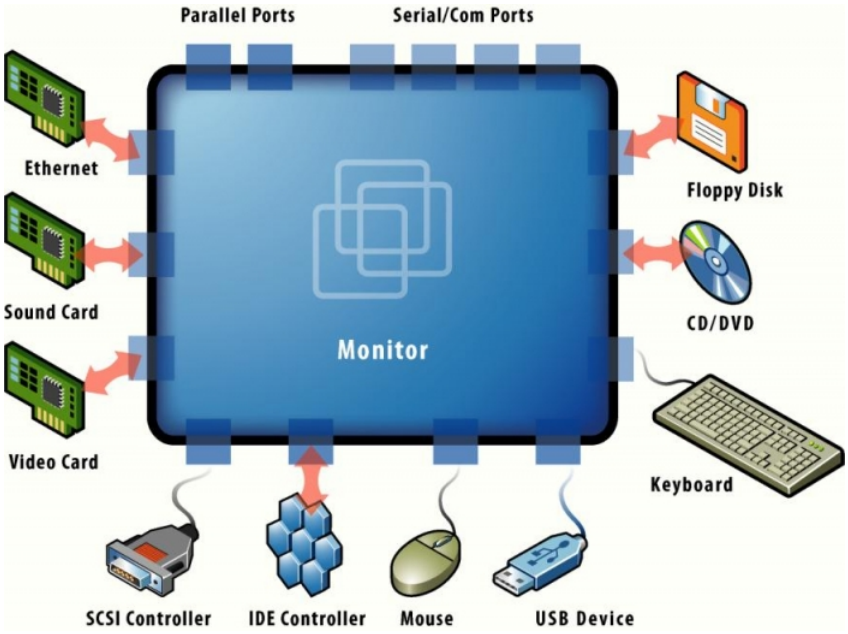
Encapsulation

A virtual machine is encapsulated into a single large virtual disk file which makes it very portable. If you were to move a guest OS running on a physical server to another physical server you would have to copy its files one by one to the new server. A guest OS running on a VM has all of its files encapsulated into the single virtual disk which serves as its own self contained world. A VM can have multiple virtual disks as well with each one being presented to the guest OS in the same manner as a physical disk would be on a physical server. Having this encapsulation has some benefits as you can easily copy a VM from one storage device to another and even carry it in your pocket on a flash drive. This means you can do things like transport a VM on a removable storage device and copy it to another host and power it up.

Encapsulation also provides some additional benefits; VMs always see the same virtual hardware regardless of the underlying physical hardware of the host. As a result you can replace or upgrade the physical host hardware without worrying about causing compatibility problems with the operating systems and applications running on your VMs. It also makes disaster recovery easier as you don't have to worry about matching up physical hardware to the same specifications at both sites as your VMs will see the same virtual hardware regardless of the physical hardware. Encapsulation also provides another important benefit which is critical to a virtual environment, isolation from other VMs and the host so each VM lives in its own world and is not mixed up with other VMs. By having this isolation the VMkernel can ensure that VMs cannot gain access to other VMs or the host itself which is a very important security requirement. While this encapsulation provides some nice benefits it does also mean you must treat and manage VMs differently from physical servers, we will cover this in detail later on in this chapter.

Virtual Machine Hardware

Virtual machines are presented with the same emulated hardware regardless of the physical host hardware. When you create a VM and assign virtual hardware components to it such as CD/DVD drives and SCSI adapters these components will all be seen by the operating system as specific hardware components.



The exception to this is virtual CPUs which are seen by the VM as whatever brand and type is running on the host which is the reason that you cannot VMotion between hosts with different physical CPU families & brands. The standard virtual hardware that a VM sees in vSphere is listed below:

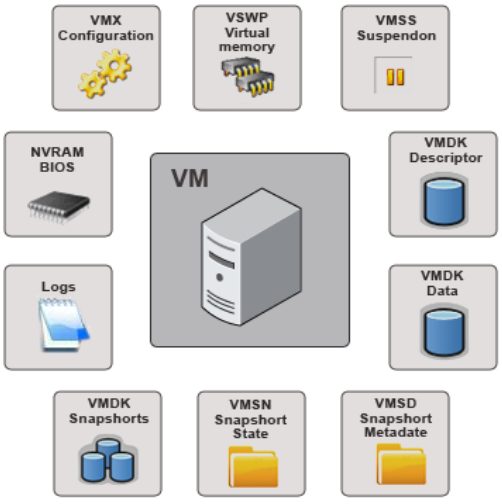
- **System Manufacturer** – VMware
- **BIOS** – Phoenix 6.0
- **Motherboard** – Intel 440BX
- **CPU** – Will vary based on whatever CPU brand & model is in the host servers. VMs running on AMD hosts will see whatever AMD processor is in the host and VMs running on Intel hosts will see whatever Intel processor is in the host. VMs will only see the number of vCPUs that are assigned to it and not the actual physical number that the host has.
- **Memory** - vSphere has 32 virtual memory slots to accommodate the increase of memory that can be assigned to a VM (255GB). The maximum memory module size that can be used is 32GB but ESX will present different size modules based on the amount of memory assigned to a VM. For example a VM with 32GB of memory assigned to it will see two 16GB virtual memory modules; a VM with 8GB of memory assigned to it will see a single 8GB virtual memory module.
- **Video controller** - VMware Standard VGA Graphics Adapter with 4MB video memory is the default. The amount of video memory can be adjusted by editing the VMs settings.
- **CD/DVD-ROM drive** - NEC VMware IDE CDR10

- **Network controller** – There are multiple types of virtual network adapters that you can choose for your VM:
    - Vmxnet3 - displays as a "Vmxnet3 Ethernet Adapter"
    - Vmxnet2 - displays as a "VMware PCI Ethernet Adapter"
    - E1000 - displays as a "Intel PRO/1000 MT Network Connection"
    - Flexible - displays as a "VMware Accelerated AMD PCNet Adapter"
- **IDE controller** – Intel 82371 AB/EB/MB PCI Bus Master EIDE Controller
- **SCSI controller** - There are multiple types of storage adapters that you can choose from for your SCSI virtual disks:
    - LSI Logic Parallel - displays as a "LSI Logic PCI-X Ultra320 SCSI Host Adapter"
    - Buslogic Parallel - displays as a "Buslogic BA80c30 PCI-SCSI MultiMaster"
    - LSI Logic SAS (serial) - displays as a "LSI Adapter, SAS 3000 Series, 8-port with 1068 -StorPort"
    - VMware Paravirtual - displays as a "VMware PVSCSI Controller"

This virtual hardware is what the guest OS on a VM will see regardless of which host it is on. The VMM is what presents this virtual hardware for the guest OS and also handles the mapping from the virtual hardware to the physical host hardware.

Virtual Machine Files

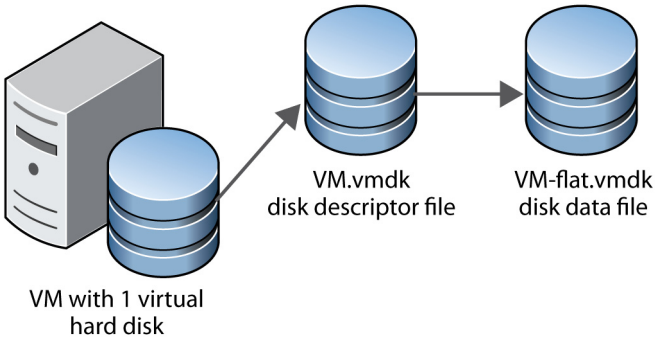
While virtual hardware is just made up in the host's RAM and presented to VM each time it is powered on, a VM also has actual files where all of its data is stored. These files are all stored in its home directory which is located on either a local or shared datastore connected to the host server.



Most of the VM's files will typically have the name of the VM as the first part of the filename followed by different extensions for the various files that are associated with the VM. These files include the VM's virtual disks as well as configuration, BIOS, swap and state files. Depending on the state of the VM you may not see certain files until that state changes. For example the virtual swap file (vswp) of the VM is only created when the VM is powered on and is deleted when it is powered off and the suspended state (vmss) file is only present when a VM is suspended. Below is a listing of all the different files by their extensions that make up a VM:

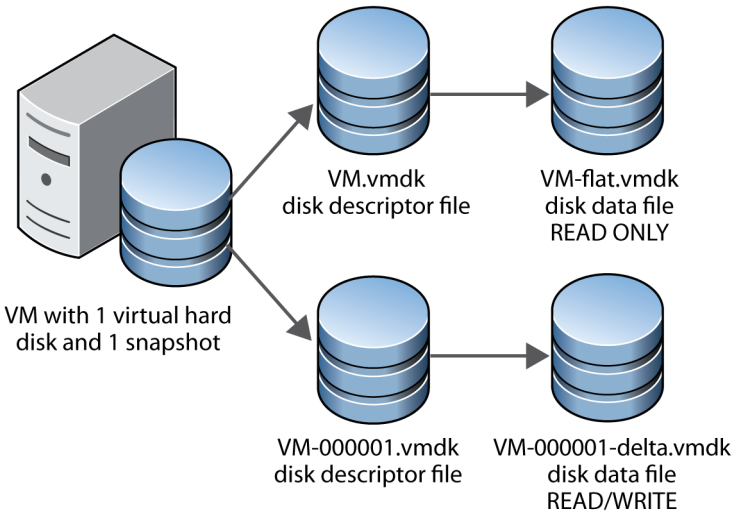
- **.nvram** – This small binary format file (less then 10k) contains the Phoenix BIOS that is used as part of the boot process of the virtual machine. This is the same type of BIOS that is used in a physical server that lets you set hardware configuration options. You can access the BIOS by pressing the F2 key at the black splash screen that is displayed when a VM boots up. Any changes made to the hardware configuration of the VM are saved in the nvram file. If this file is deleted or missing from a VM's directory it is automatically re-created when the VM is powered on.
- **.vmx** – This small text format file contains all of the configuration information and hardware settings of the virtual machine. The data in this file is written when you make changes to the configuration of a VM using the vSphere Client. You can also edit this file directly when a VM is powered off but you should be careful and make a backup copy beforehand. The type of data in this file includes things like specific hardware configuration (i.e. RAM size, NIC and hard drive configurations), resource settings, VMware tools and power management options.
- **.vswp** – This is the memory swap file that is created when a VM is powered on that is used if a host is overcommitted and uses all of its physical memory or when a memory limit is placed on a VM which limits the amount of physical host memory that a VM can use. It is possible to assign more virtual memory to a VM than a host physically has and the vswp file makes up for the physical memory that the host does not have. The size of this file is equal to the amount of memory assigned to a VM minus any memory reservations (default is 0) that a VM may have set on it (i.e. a 4GB VM with a 1GB reservation will have a 3GB vswp file created). A VM will not power on if there is not enough free space available to create this file. These files are deleted when a VM is powered off or suspended.
- **.vmss** – This file is used when virtual machines are suspended and is used to preserve the memory contents of the VM so it can start up again where it left off. This file will be approximately the same size as the amount of memory that is assigned to a VM as even empty memory contents are written to it. When a VM is brought out of a suspended state the contents of this file are written back into the physical memory of a host server. After this file is created it is not automatically deleted until a VM is powered off (not rebooted). If a previous suspend file exists when a VM is suspended again this file is re-used instead of deleted and re-created. If this file is deleted while the VM is suspended then the VM will start normally and not from a suspended state.
- **.vmsd** – This text file is used with snapshots to store metadata and other information about each snapshot that is active on a VM. It is initially 0 bytes in size and is updated with information as snapshots are created or deleted. There will be only one of these files present regardless of the number of snapshots running as they all update this single file. The information in this file consists of the name of the vmdk & vmsn file used by each snapshot, the display name, description and UID of each snapshot. Once you delete snapshots this file still retains old snapshot information but increments the snapshot UID to be used with future snapshots.

- **.vmsn** – This file is similar to the vmss file and is used with snapshots to store the state of a virtual machine when a snapshot is taken. One of these files will be created for every snapshot that is created for a VM and they are automatically deleted when the snapshot is deleted. This file will vary in size depending on whether or not you choose to include the memory state of the VM with the snapshot. If you choose to store the memory state this file will be slightly larger than the amount of memory assigned to the VM as the entire memory contents including empty memory are copied to this file. If you do not choose to store the memory state of the snapshot then this file will be fairly small (under 32KB).
- **.vmdk** – These are the virtual hard disk files; there are two different types of these files that make up a single virtual disk. The first type is a large data file equal to the size of the virtual disk and contains the raw data of the virtual disk. The second type is a small (less then 2k) text disk descriptor file which describes the size and geometry of the virtual disk file and also contains a pointer to the large data file as well as information on the virtual disk's drive sectors, heads, cylinders and disk adapter type. In most cases these files will have the same name as the data file that it is associated with (i.e. veeam1.vmdk and veeam1-flat.vmdk). You can tell which descriptor file is tied to which data file by checking the Extent Description field in this file to see which –flat, -rdm or –delta file is linked to it. There are four different types of virtual disk data files that can be used with VMs which are covered below:
    - **–flat.vmdk** – This is the default large virtual disk data file that is created when you add a virtual hard drive to your VM that is not a RDM. When using thick disks this file will be approximately the same size as what you specify when you create your virtual hard drive. One of these files is created for each virtual hard drive that a VM has configured.

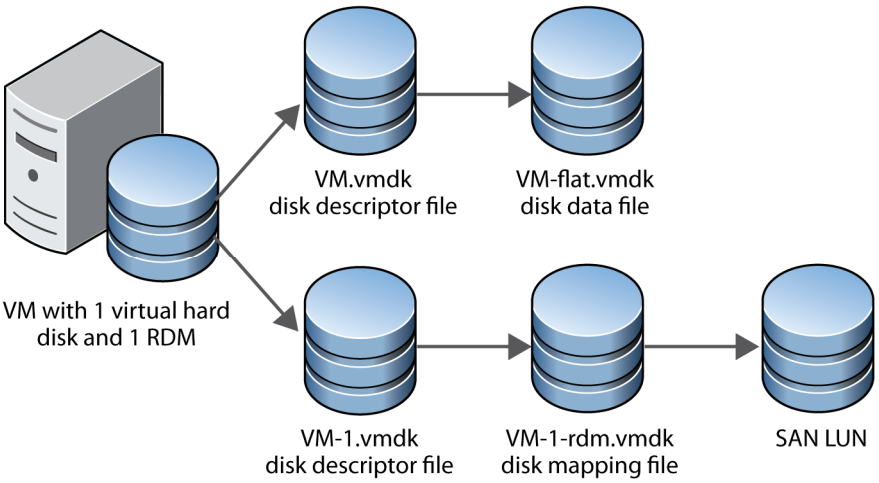


    - **–delta.vmdk** – These virtual disk data files are only used when snapshots are created of a virtual machine. When a snapshot is created all writes to the original –flat.vmdk are halted and it becomes read-only; changes to the virtual disk are then written to these delta files instead. The initial size of these files is 16MB and they are grown as needed in 16MB increments as changes are made to the VM's virtual hard disk. Because these files are a bitmap of the changes made to a virtual disk a single delta.vmdk file cannot exceed the size of the original flat.vmdk file. A delta file will be created for each snapshot that you create for a VM and their filenames will be incremented numerically (i.e. veeam1-000001-delta.vmdk, veeam1-000002-delta.vmdk). These files are automatically deleted when

the snapshot is deleted after they are merged back into the original flat.vmdk file.

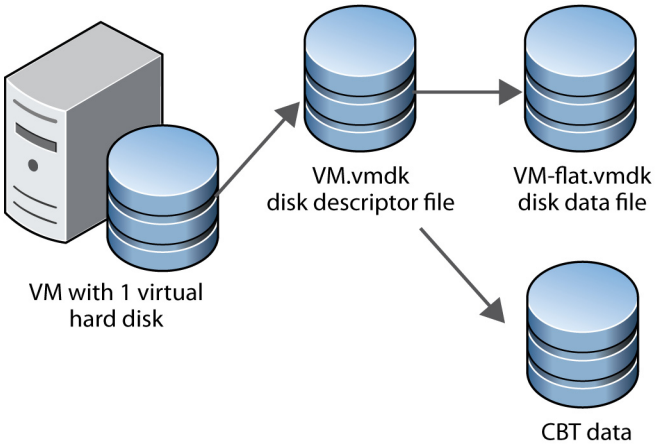


- **–rdm.vmdk** - This is the mapping file for a Raw Device Mapping (RDM) that contains information for the mapping to a SAN LUN. The mapping file is presented to the ESX host as an ordinary disk file and is available for the usual file system operations. However to the virtual machine the storage virtualization layer presents the mapped device as a virtual SCSI device. The metadata in the mapping file includes the location of the mapped device (name resolution) and the locking state of the mapped device. If you do a directory listing you will see that these files will appear to be taking up the same amount of disk space on the VMFS volume as the actual size of the LUN that it is mapped to but in reality they just appear that way and their size is very small. One of these files is created for each RDM that is created for a VM.



- **-ctk.vmdk** - This is the file where the new Changed Block Tracking (CBT) feature stores the information about changed blocks for a virtual disk. These files are created in a VM's home directory for every virtual disk that the CBT feature is enabled on. This size of this file is fixed and does not grow beyond its initial size unless you increase the size of a virtual disk. The size of this file will vary based on the size of a virtual disk which is

approximately .5MB for every 10GB of virtual disk size. Inside this file the state of each block is stored for tracking purposes using sequence numbers that can tell applications if a block has changed or not. One of these files will exist for each virtual disk that CBT is enabled on.



- **.log** – These are the files that are created to log various types of information about the virtual machine. There will be a number of these files present in a VM's directory; the current log file is always named vmware.log and up to 6 older log files will also be retained with a number at the end of their name (i.e. vmware-2.log). A new log file is created either when a VM is powered off and back on or if the log file reaches the maximum defined size limit. The amount of log files that are retained and the maximum size limits are both defined as VM advanced configuration parameters (log.rotateSize & log.keepOld).

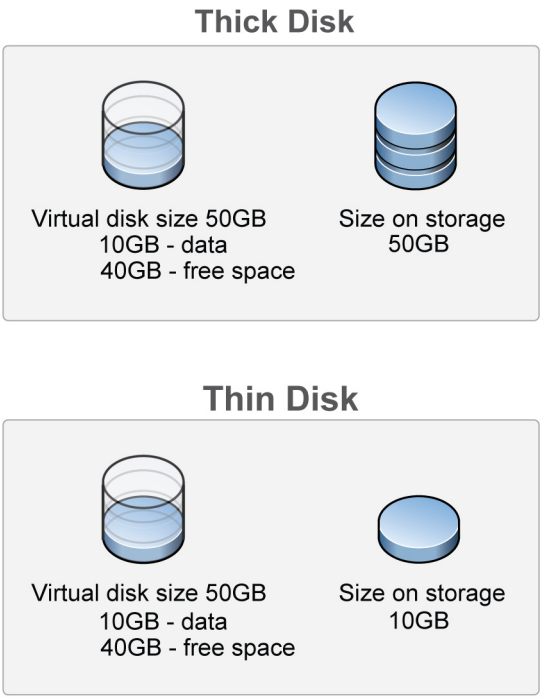
These files are what make up a VM and will typically reside in the VM's home directory. It is possible to change the locations of some of these files so virtual disks, swap files and snapshots can reside on alternate datastores. You can view these files by using the vSphere Client's Datastore Browser or by using a tool like Veeam FastSCP which allows you to access a host's filesystem.

Virtual Disks

While there are different types of .vmdk virtual disk files there are also different types of virtual disk formats. There are 3 main formats (raw, thin & thick) used in vSphere for virtual disk files which are described in detail below.

- **Raw disks** - Also known as Raw Device Mappings (RDMs), raw disks enable a VM to have a virtual disk that is mapped directly to a LUN on a SAN. RDMs have a small mapping file on a VMFS volume in the VMs home directory that contains mapping data to the SAN LUN. When you create an RDM for a VM you can choose one of two modes that it can operate in. The first mode is virtual compatibility mode which virtualizes the mapped device and is mostly transparent to the guest operating system. The other mode is physical compatibility mode which provides minimal SCSI virtualization of the mapped device and the VMkernel passes most SCSI commands directly to the device which allows for closer integration between the VM and the LUN.

- **Thick disks** – With thick disks all space on a datastore is allocated at the time that it is created so the resulting vmdk file will take up as much room on the datastore as the size of the disk that you create. There are several types of thick disks with the main difference between them being how disk blocks are zeroed before they are written to. Disk blocks are normally zeroed as a security feature to erase any data that may already exist from old VMs that may have been deleted. If blocks are not zeroed it is possible to look at that old data from the new VM using a utility that can read data at the block level. The different types of thick disks are listed below:
    - **Non-zeroed thick disks** – Also known as monotholitic preallocated zeroed_out_never. All space is allocated at creation time and may contain stale data on the physical media. Disk blocks are not zeroed before being written to the first time which makes these types of disk less secure because any previously written data has not been cleared from the disk blocks.
    - **Lazy Zeroed thick disks** – Also known as monotholitic preallocated zeroed_out_later. All space is allocated at creation time and wiped clean of any previous data on the physical media. Disk blocks are zeroed out on demand as they are written but only for the first write to a disk block. This type of disk is the default used when creating virtual disks on VMFS volumes using the vSphere Client. These disks have slightly less I/O performance on the first write to a disk block because it must be zeroed before writing to it. Subsequent writes to a disk block have the same optimal performance as the other disk types.
    - **Eager Zeroed thick disks** – Also known as monotholitic preallocated zeroed_out_now. All space is allocated at creation time and wiped clean of any previous data on the physical media. All disk blocks are zeroed out when the disk is created which increases the time it takes to create these disks compared to the other types. These types of disk are the most secure and also offer slightly better performance only on the first write to a disk block because it has been already zeroed. Subsequent writes to a disk block have the same optimal performance as the other disk types. This type of disk is required by the Fault Tolerance feature.
- **Thin disks** – Also known as monotholitic growable. Thin disks are virtual disks that start small and grow as data is written to them. Unlike thick disks where are all space is allocated at the time of disk creation, when a thin disk is created its initial size is 1MB (or up to 8MB depending on the default block size) and it then grows up to the maximum size that was defined when it was created as data is written to it by the guest OS. Thin disks have a slight performance penalty both as they grow and space is allocated on demand and also when the first write to a disk block is zeroed on demand. Once the disk has grown and its blocks have been zeroed they have the same performance as the other disk types. Thin disks are good for conserving disk space on a VMFS volume but can cause problems if you do not monitor their growth to make sure you do not run out of disk space on your VMFS volume. Thin disks are often used by default on NFS datastores based on the allocation policy of the NFS server and not the ESX server.



- **2GB Sparse Disks** - Sparse disks are a special format disk that cannot be used with a running VM on an ESX/ESXi host and is instead often used to import/export virtual disks to and from ESX hosts. This format divides a virtual disk in up to multiple 2GB pieces which is handy when you need to copy a virtual disk to another host or so it can fit on physical media like DVD-ROMs. The 2GB file sizes are used because some older file systems do not support file sizes larger than 2GB. Other VMware products like Workstation and Server can use this format for running VMs but for ESX hosts you must first import these type of disks into a thick or thin format using the vmkfstools utility.

Thin disks have become a popular choice in vSphere because VMware vastly improved their usability by adding management capabilities for them to the vSphere Client. It is possible to change disk formats if needed by using either the Storage VMotion feature or the vmkfstools command line utility.

VIRTUAL MACHINE MANAGEMENT

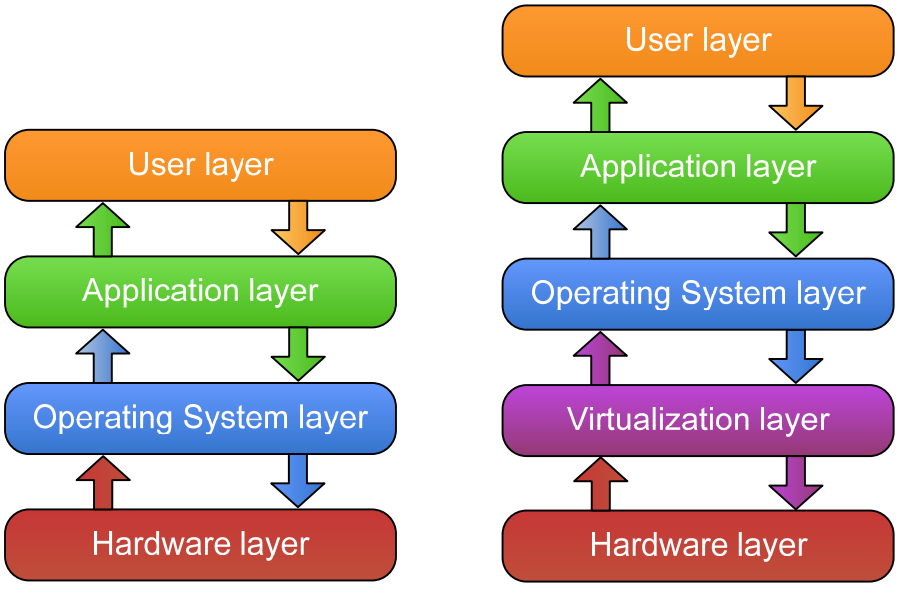
If you're used to managing physical servers you'll need to change your way of thinking when you start managing virtual machines. In this section we will cover some of the differences between virtual machines and physical servers as well as explain why you need to change your operational procedures in virtual environments. We'll also explain why applications used to manage physical servers are often not efficient in virtual environments and why you should use applications designed for virtual environments instead.

The Virtualization Layer

Much like there are different layers that make up our atmosphere, there are also different layers that make up a server; these layers are also referred to as the computing stack because they are stacked on top of each other. A traditional physical server is made up of the following layers:

1. Hardware layer – this is the bottom layer and includes all the hardware components of a server including CPU, memory, storage, networking and other I/O devices.
2. Operating system layer – this installs directly on the hardware and provides an environment for applications to run on. An operating system handles all the interfaces between the users, application and the server hardware.
3. Application layer – applications or software install on the operating system and must go through the OS layer to access the server hardware.

These layers have a hierarchy with applications being at the top and the hardware being at the bottom. With virtualization another layer gets added to this in between the hardware and operating system. The virtualization layer is designed to be transparent to the operating system, but to access physical hardware the OS must traverse through the virtualization layer.



This extra layer adds some additional overhead but with bare-metal hypervisors like ESX & ESXi this overhead is minimal. Because applications and the operating system are un-aware of this extra layer they can often produce inaccurate information when reporting on metrics related to hardware as they can only see the virtual hardware presented to them and not the underlying physical hardware.

How virtual machines differ from physical servers

While virtual machines are designed to do the same jobs as their physical counterparts there are some big key differences between them that you should be aware of. This is especially important when trying to design a virtual environment and select the right hardware for your hosts. You need to really change your way of thinking when designing a virtual environment because you have many virtual machines all competing for the limited resources of a host. Where designing a

physical server is similar to designing a house, designing a vSphere environment is almost like designing a small city; there are lots of inter-related components and many critical design decisions to be made to ensure that all the residents' needs are met properly. If you don't size your water, gas and electricity properly in your city your houses won't have the resources they need for basic services and to handle peak loads. Similarly with designing a virtual environment you need to size your storage, network, CPU and memory resources properly or your virtual machines will not have the resources they need to run applications.

To help understand what the differences between physical servers and virtual machines are we can take a look at the characteristics of each.

Physical Servers	**Virtual Machines**
Stationary location in the datacenter.	Portable, can move between hosts and datacenters.
Hardware upgrades require provisioning new hardware, shutting down server to install which can be disruptive and risky.	Virtual hardware can be easily changed or added at any time, even with the VM powered on in some cases.
Building a new server requires provisioning new hardware, putting it together, configuring infrastructure components and installing OS and other necessary applications.	New VM can be built in minutes with just a few mouse clicks. VMs can be created from templates and cloned from existing VMs.
When a hardware component fails the server is often degraded or un-usable until the defective component is replaced which can take some time.	If a hardware component fails in the physical host the VM can easily be brought back up on another host if it resides on shared storage.
Network and storage connections are fixed and re-configuration can be difficult and require multiple groups to accomplish.	A VM can move from one network to another or from one storage device to another quickly and easily.

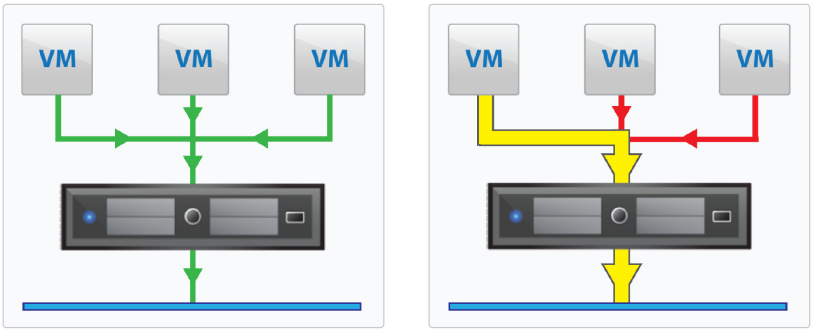
As you can see virtual machines offer many advantages compared to physical servers so besides consolidating server hardware you also get many other benefits from a virtual environment. On the flip side you need to treat VMs differently than you would physical servers because their different architecture requires it. This means common tasks like provisioning, monitoring, troubleshooting and backing up VMs require different methodologies and tools compared to physical servers. If you were to experience a performance problem with a physical server troubleshooting it would be pretty straightforward as you mainly need to only focus on the server itself and nothing else. In a virtual environment you have many more areas that you would need to investigate to pinpoint the cause of the problem. This includes inside the VM with both the apps and operating system, the host, storage devices and networking. A virtual environment is much more complicated than a physical environment with many dependencies and interconnections between various components. As a result you must adapt your normal way of thinking with physical servers to take into account the many other areas inside and outside a host that could impact your VM.

A host has a limited set of resources that all the VMs on the host are competing for, as a result your host can end up like the wild west with VMs battling it out to claim the resources that they need to run their applications. With physical servers you typically do not have that problem as each server has its own resources that are not shared with other servers. To try to maintain control inside your host you typically need to implement resource controls to ensure law and order amongst your VMs and that you do not have rogue VMs hogging all a host's resources. Once you insert that virtualization layer between the hardware and operating system you are changing the way the game is played. The bottom line is across the board if you try to treat and manage a VM like you would a physical server it will be inefficient, problematic and a waste of both administrative and host resources.

Why virtual machines require virtual tools

As we have pointed out already, virtual machines are much different than physical servers and as a result require tools and applications that are designed specifically to work with virtual environments. The problem with using tools designed for physical environments only is they are not aware of the virtualization layer sitting between the hardware and operating system. As a result they are missing out on a big piece of the computing stack and while they may still work they are not seeing the big picture which can result in mis-leading results and inefficiencies. Imagine trying to drive a car with one eye closed, while it is possible, it's also a lot more difficult as you are not seeing everything that you should be seeing and as a result may make mistakes or even have an accident. So when deploying a virtual environment you should look at deploying tools and applications that you will need to manage that environment that have been designed with virtualization in mind. Sure the tools and applications that you use with physical servers will still work, after all a server is a server regardless of whether it is virtualized or not. But in many cases they will be less effective, inaccurate and mis-leading. Let's take a look at some examples of why you should use the right tools and applications for your virtual environment.

Backing up servers can be a very resource intensive operation where a large amount of data must be read from a server in the shortest amount of time possible. With physical servers you typically install a backup agent inside the guest OS and then have your backup server contact each guest through the OS layer to back up its data. This approach is OK for physical servers but in virtual environment this resource intensive activity on a single VM will have a ripple effect on all the other VMs running on the host creating an undesirable performance situation. This effect is similar to using water in a house, there is a fixed size water pipe going to a house with a set maximum pressure. If a faucet is turned on in one room you get full volume and pressure out of it, however if another faucet is turned on at the same time the pressure will be split between them and they will both operate with reduced pressure. The more faucets that are opened will result in reduced performance for all the faucets as they are all competing for the limited water pressure and volume that are available.



The same holds true when backing up VMs, you need to use a backup application that is aware of the virtualization layer and not go through the operating system to back up the VM but instead back it up at the virtualization layer. Doing this will not only greatly reduce the resource usage on your hosts and VMs from backup operations but also result in more efficient and quicker backups.

Another area where physical servers differ from VMs is with networking. With a physical host the network endpoint stops at the physical NIC that is inside the host. With virtualization the network is extended from the physical NIC(s) inside a host to virtual switches that reside inside the host's RAM to the virtual NICs that are connected to your VMs. As a result the physical network is often not aware of much of the network traffic that occurs in the host as in certain vSwitch configurations traffic between VMs never travels over the physical network. As a result traditional network controls and tools are less effective as they cannot monitor traffic they can't see and they cannot protect VMs against threats that may occur from within the host itself. Having network tools that can extend into the virtual host and see and manage the network traffic is critical to maintaining a secure and efficient network.

Performance monitoring & reporting is another big area where you must use tools designed for virtual environments if you want accurate results. Remember the hypervisor is fooling VMs into thinking they have the host all to itself but in reality it has to juggle the host's resources between many VMs. Performance tools designed for physical servers are not aware of the virtualization layer and everything happening behind the scenes. As a result the numbers that they produce may not be an accurate reflection of the true performance of a VM. This is especially true when it comes to CPU & memory statistics, the hypervisor has a lot of tricks up its sleeve that it uses on VMs to help conserve and maximize host memory that the guest OS is unaware of. The same holds true for CPU usage, the CPU scheduler does a lot of fancy footwork to keep VMs going in an efficient manner but the guest OS is blissfully unaware of this as well. There may also be resource controls in play at the virtualization layer that the guest OS is unaware of which can cause misleading reporting results. Having the right tools that know about the virtualization layer will ensure that you get the most accurate reporting statistics from your VMs.

SUMMARY

In this chapter we introduced you to some of the unique concepts of virtualization to help you understand the mechanics behind virtualization so you can more effectively manage and administer virtual machines. If you are new to virtual machines you should prepare to think outside the box if you want to be successful with virtualization. Virtual environments offer many new and more efficient ways for doing traditional tasks so you should try to take advantage of them all by using the right tools. One of the key takeaways from this chapter is to not treat a VM like you would a physical server. Once you understand the key differences between the two you can work to implement the right tools that will provide the most effective management for your virtual environment.

CONTENTS

INTRODUCTION

Backup methods have evolved over the years as new technology has become available, but the introduction of virtualization technology changes everything and introduces more options and flexibility when backing up your servers. While traditional physical backups have evolved over the last few decades, the changes have mostly been with the media being used on the target device. From the original punch card backups, to floppy disks, to magnetic tape, to optical disks and finally to hard disks, the target media has greatly improved over the years. The one area that has not changed that much is with the methods, which mainly involve using an agent installed inside the operating system that a backup server connects to over the network to copy data to the target device. With the introduction of virtualization, this method still works, but it is not as efficient because of the architectural differences in virtual environments. Companies like Veeam recognized the need for new backup and recovery methodologies for virtual environments and developed applications that were made to order for virtualization.

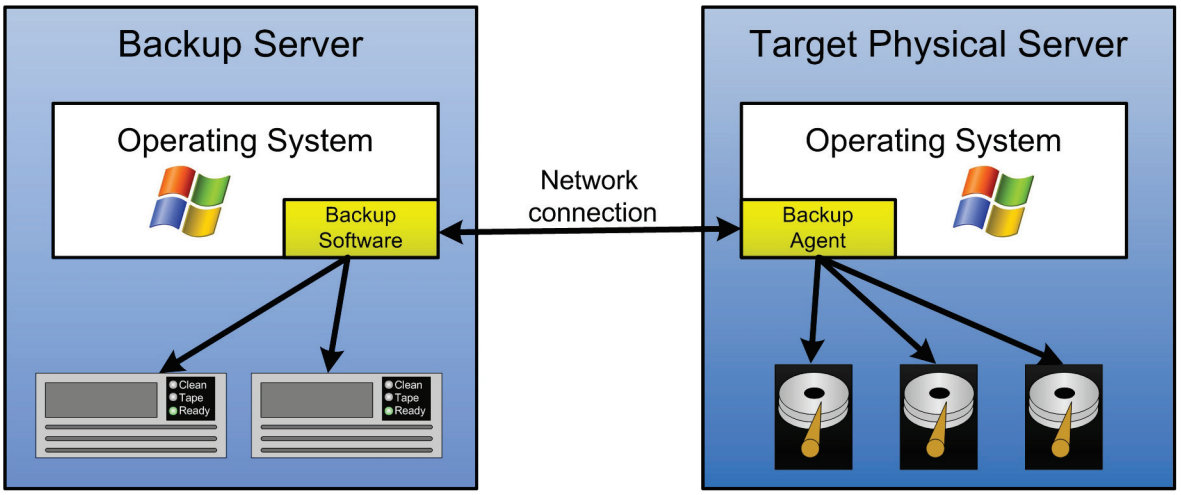
This chapter looks at traditional backup and recovery methods and why they do not work well in virtual environments, as well as the new backup and recovery methods that were developed specifically for virtualization.

TRADITIONAL BACKUP AND RECOVERY METHODS

When we refer to traditional backup and recovery methods we are talking about the methods used with a single guest operating system installed directly on a physical server, which was how things were done before virtualization was introduced by VMware for x86 servers. This section covers the methods that have been used for both backing up and restoring physical servers so you are aware of how these methods relate to those used in virtual environments.

Operating system backup agents

This method relies on a software backup agent that is installed inside the operating system. The backup server contacts the backup agent, which is always running on the server, and establishes a session over a TCP/IP port with the agent. Once communication is established, the agent can copy all the data from the operating system file system to the backup server, which can then write the data to the target media, either a tape or disk system. Because the backup agent is installed inside the operating system, it can read the file systems of all the disks that are configured on the server. It can access the Master File Table of the OS to determine all of the files that exist on the disk partitions so it can copy them one by one to the backup server. This method of backing up a server is considered a file-level backup because files are being copied at the file system level from one device to another, much like you would copy files using Windows Explorer or the command-line COPY command. To support incremental backups, an archive bit value is changed whenever a file is modified by the operating system. Once this value is changed, an incremental backup recognizes that the file was modified since the last backup and includes it in the incremental backup and then clears the archive bit.



Traditional backup method using operating system agents

Since this method works at the file level inside the OS, only active files inside the OS are backed up; all empty disk blocks and deleted files are ignored. This results in a full backup equal in size to the amount of disk space the files take up on disk regardless of how big the disk is. For incremental backups, however, this method is inefficient because if only a small part of a file changes, the entire file is backed up because it has been flagged as modified. For example, if you had a 500MB database on your disk and only a few records totaling 1MB were updated since the last backup, the entire 500MB would be backed up again because the backup agent works at the file level and not the block level. To help with this big limitation, special backup agents are

deployed for certain applications like Microsoft Exchange that know how to back up only the changed records for that application and not the entire database.

Another limitation with this method is with files that are open and in use when the backup runs. These files are typically locked by the OS or application, which means the backup agent cannot access them. To cope with this problem and to ensure that all files are backed up, backup applications can use special open file agents. The Microsoft Volume Shadow-copy Service (VSS) that is built into most new Windows versions is an example of this. VSS works by briefly freezing application write I/O requests while the shadow copy is created. It then flushes the file system buffers and freezes the file system to ensure that all data that is shadow-copied is written in a consistent order. VSS works at the disk block level and creates a clone or shadow copy of the disk volume that is a temporary read-only copy of all the data on the volume. Because this copy is read-only it can be read by the backup application regardless of any applications that may have the file open and are writing to them. All writes to the files happen on the original disk blocks while the shadow copy is active, but before the write happens the original disk blocks are written to a special differences area that is part of the shadow copy. Because of this the creation of the shadow copy is a quick process and does not take up much disk space as only changed disk blocks are written to the shadow copy. Once the backup completes the shadow copy is discarded.

Backups performed within the operating system have worked well for many decades with traditional physical servers. With the introduction of virtualization, however, they have become inefficient, and alternative methods that leverage the strength and architecture of virtualization were needed instead. Companies like Veeam recognized this and developed backup methods specifically designed for virtual environments that bypass the operating system layer and back up VMs at the virtualization layer instead. With virtualization becoming more and more popular the days of backing up through operating system agents is numbered as more efficient methods are used instead.

Scheduling and performing backups

Backing up a server is a resource-intensive operation especially at the network and storage layers. When performing backups of traditional physical servers, the resource usage is limited to the server being backed up. Many servers can be backed up simultaneously with the resource bottleneck typically at the backup server, which has limited tape drives and network bandwidth. Most data centers have backup windows, which is a period of time where usage is low so backups will not affect users who access the applications running on the servers being backed up. Because each server has its own resources, one server being backed up will not have any impact on another, which makes scheduling backups fairly easy.

When backups are performed the backup server connects to each individual server and copies the data from that server to the target backup device, which is typically a tape system. Tapes are a convenient backup media because they can hold a lot of data and can easily be sent offsite for storage. There have been many different types of tape formats over the years such as DLT (Digital Linear Tape), AIT (Advanced Intelligent Tape), DAT (Digtal Audio Tape) and, one of the most popular today, LTO (Linear Tape-Open). The first generation of LTO tapes (LTO-1) was introduced in 2000 and had a maximum native capacity of 100GB with a maximum backup speed of 15 MB/s. Today the 5th generation of LTO tape (LTO-5) has a native capacity of 1.5TB and a maximum backup speed of 140 MB/s.

While tapes are used by many organizations to back up their data, many companies are implementing disk systems as a backup target either in place of or to complement a tape backup system. When using a disk system as a target, data that is backed up to a disk target can be replicated offsite after the backup completes. A disk target can also be used as a holding area for data that is going to be backed up to tape; the data is quickly and easily copied to the disk target during the backup and then later copied to tape without involving the source server. Backing up to disk can also make restores much easier and quicker as you do not have to worry about finding and loading backup tapes. In addition, cheap disk systems can be used as the target storage to provide a relatively high capacity for an inexpensive price.

Backups also have a number of different backup models that can be used to back up data. The different methods are designed to reduce both the amount of time and the amount of space that backups require. Listed below are the different backup methods that are commonly used:

- **Full**—This is where all data on the source server is copied to the backup target. Full backups require the most time to perform and take up the most space. Full backups are the foundation for the other backup types.
- **Synthetic**—With synthetic backups a full backup is only done once and subsequent backups are all incremental backups. This method provides smaller backup windows and less resource consumption than traditional backups as you never have to do periodic full backups. Once an incremental backup takes place it is combined with previous backups to synthesize a full backup. This way an up-to-date full backup copy is always on hand without ever having to perform a full backup. You can still restore older data as well because all changes are backed up and saved as rollback files and historical data is used to calculate reverse increments. This method is also referred to as reverse-delta.
- **Incremental**—This method only backs up data that has changed since either the last full or incremental backup. This method relies on an archive bit that is cleared after the previous backup. Once a file is changed the archive bit is set, which indicates the file has changed since the last backup. The incremental backup then only backs up those files that have the archive bit set and then clears the archive bit once the files are backed up. While this method is quick, it can also make restores more difficult as multiple backup points may have to be used to restore files that have changed since the last full backup.
- **Differential**—This method is similar to incremental backups, but all changed data since the last full backup is backed up every time a differential backup runs. As changed files are backed up with differential backups, the archive bit is not cleared, so any file that is changed since the last full back is backed up with each subsequent differential backup. While this method takes more time and more space on the target, restores are easier since only one backup point is needed to restore files that have changed since the last full backup.

Most companies implement a backup schedule that includes periodic full backups (e.g., once a week), followed by incremental or differential backups until the next full backup occurs, and then the cycle repeats. Backup schedules will vary based on each company's requirements, which include retention timeframes, backup windows, budgets and infrastructure.

Restoring data

When it comes to restoring data from backups, there are two ways to do it: restoring individual files and directories, or restoring all the data on the server. Restoring a select group of files is the most common type of restore, which may result from a scenario where files were accidently

deleted or previous versions of a file are needed. To restore select files back to a server, a point in time is selected from the available backups that exist for the server. Once the files are selected the media that the files are backed up on must be available before the restore can begin. Oftentimes this involves locating and re-calling tapes from offsite storage; if the data is on an existing snapshot it can also be leveraged for this. Once the point-in-time version of the file is located from the restore media, it is copied either back to its original location, overwriting the existing copy or to an alternate location where it can be accessed without disturbing the original file.

Restoring a complete copy of a server is referred to as a bare-metal restore as nothing but the bare metal of a server is needed without any requirement for an operating system being installed on the server. The restore includes the operating system along with all the applications and data that reside on it. This type of restore is typically only needed in the event of a complete hardware failure where all data on the server is lost. It is also used for disaster recovery purposes where a whole copy of a server must be restored because of a failure at the primary site. Because the backup includes the operating system, disk partitions, system state and all of its device drivers, similar server hardware is often needed for the restore to work properly. Performing a bare-metal restore can be complicated and tricky and proper planning needs to be done to ensure a successful restore operation. To complete a bare-metal restore an operating system needs to be installed on the server first so a backup agent can be installed. Once that has been completed the backup server can communicate with the agent to restore all of the files including system state information on top of the installed operating system. Some backup products can also leverage network boot capabilities using PXE to load boot images from a PXE boot server to eliminate the need for an OS to be installed first to perform a bare-metal restore.

Testing and verification of backups

While backups are a critical part of any computing department, being able to successfully restore data when needed is even more critical. Backups are pointless if you can't restore the data—imagine backing up servers for months only to find out when you need to restore some critical data that your backups have not been working properly. Backups are like a good insurance policy: you pay your premiums every month hoping you will never have to use them, but when something bad happens you need to have something to fall back on. Because of this you shouldn't just trust that your backups are working properly—you need to periodically verify this by trying to restore data successfully.

Verification of backups is more than just verifying through the backup application that the backup completed successfully and that the media is error-free. It also involves being able to properly restore files, applications, databases and whole servers when needed. The problem with verifying backups is that it can be a complicated and time-consuming process. While restoring individual files or directories to an alternate location may be simple enough, trying to restore applications or a whole server can be challenging. When you restore an application you need to ensure that it is in a working state, and restoring multi-tier applications can be even more difficult. You need a separate environment in which to perform the restore so you do not affect your production environment. This often requires having one or more physical servers available to perform the restore on; when testing bare-metal restores this can be difficult because hardware similar to the server that was backed up is needed. In data centers that have many model and generations of servers this can be even more difficult. As a result backup verification in traditional physical server environments can be a tedious and difficult process.

VIRTUAL ENVIRONMENT BACKUP METHODS

While traditional methods that are used with physical servers can be used in virtual environments, they are not very efficient. Because the architecture of a virtual environment is drastically different from a physical environment, the traditional methods can create bottlenecks, cause performance problems and take longer to complete. Virtual environments require methods that leverage the strength of the virtualization architecture to perform highly efficient backups with minimal impact on the virtual machines. In this section we detail the different backup and recovery methods that are used in virtual environments that are much different than traditional methods.

Image-level backups

With virtualization it is more efficient to back up the single large virtual disk file (vmdk) at the virtualization layer instead of going through the guest operating system to back up files individually. This type of backup is known as an image-level backup because you are backing up a whole image of the virtual machine's disk file, much like you would if you used an imaging utility like Ghost. While it is more efficient backing up just one large file instead of thousands of smaller files, there is one issue with this. Since image-level backups cannot see inside the operating system and are backing up the whole virtual disk they are also backing up empty disk blocks and deleted files as well. If a VM has an 80GB virtual disk file and only 20GB is in use, 80GB is backed up with an image-level backup; with a file-level backup only the 20GB in use is backed up. To get around this, backup vendors like Veeam that backup at the virtualization layer detect empty disk blocks that have not been written to by the operating system yet and ignore them. While checking for empty disk blocks they also use inline deduplication to detect duplicate blocks and ignore them as well.

You might wonder how an image-level backup can handle open files and avoid corruption from files that change when the backup occurs if it cannot see inside the guest operating system. This is done by first quiescing the VM using a special driver (either VMware Tools or a special driver supplied by the backup vendor) that runs inside the guest operating system that momentarily pauses the running processes on a guest and forces the operating system and applications to write any pending data to disk. Once that is complete a snapshot of the VM is taken at the virtualization layer, which creates a new temporary virtual disk file (delta) for any new disk writes that occur on the VM and prevents the original disk from being written to while the backup is running. Once the backup is completed, the temporary virtual disk file is merged back into the original disk file and the snapshot is deleted.

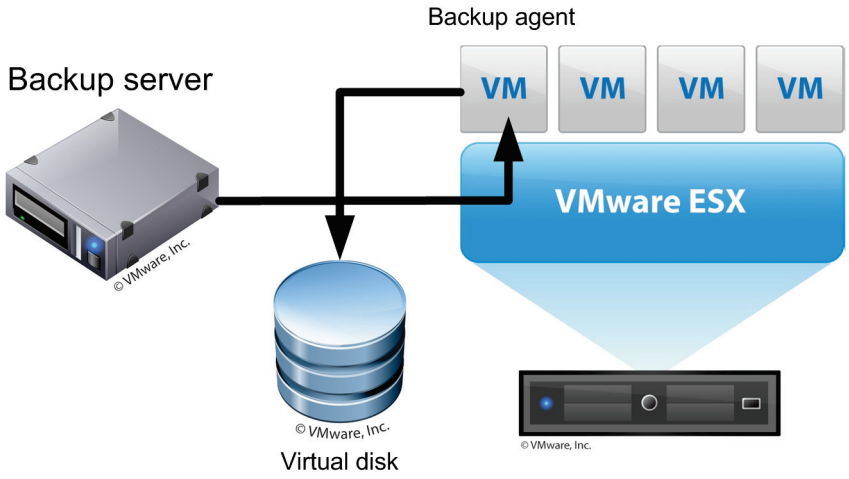
A common misconception with image-level backups is that you cannot do incremental backups because you are only backing up one large file and if any disk block changes the whole file must be backed up again. With traditional file-level backups only the files that have changed are backed up on incremental backups; this is noted by setting a flag called an archive bit that indicates that a file has changed since the last backup. Once the file is backed up the archive bit is cleared until it changes again. With image-level backups, a backup application has to keep track of all the blocks that have changed since the last backup so it knows which ones to back up when doing incremental backups. This process can increase the time of backups as the backup application must calculate a hash for each block and then scan the entire virtual disk and compare it against a hash table to see what has changed since the last backup. To speed up incremental backups, most backup vendors have taken advantage of the new Changed Block Tracking (CBT) feature that is accessible via the vStorage APIs for Data Protection. This allows the

backup application to simply query the VMkernel to find out which disk blocks have changed since the last backup, which greatly speeds up incremental backups.

Image-level backups offer some advantages over traditional file-level backups of physical servers. Having the server encapsulated into one big file makes for easy portability—the virtual disk file can easily be copied to any other storage device. For example, one could easily copy a VM from a host server to another storage device, external hard drive or flash drive for safekeeping. This makes creating ad-hoc backups of VMs a simple process. It also makes performing bare-metal restores much easier as all you need to do is restore the files that make upmakeup the VM to any host and you are up and running.

File-level backups

File-level backups are traditionally done using an agent inside the guest operating system that is aware of all the individual files. While this can be done in a virtual environment, it can cause excessive resource usage on the host server, which can negatively affect the other VMs on that host. Backups that use OS agents on virtual machines must navigate through the virtualization layer to get to the guest operating system layer to back up files as depicted below.



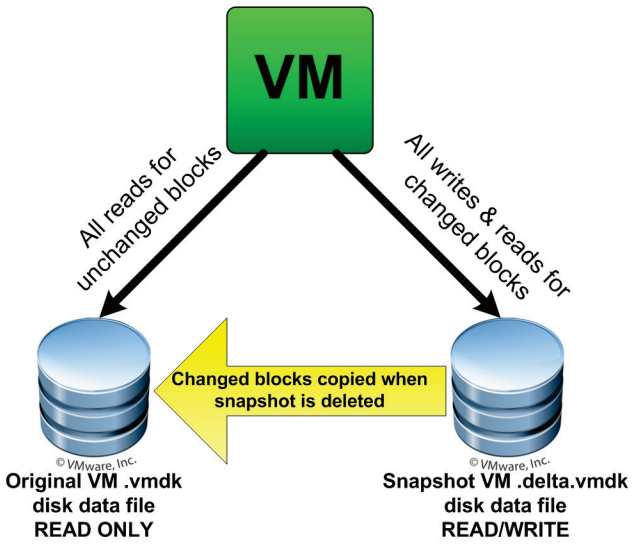
Backing up a VM using a traditional agent inside the OS

As a result, backing up a VM using a backup agent running inside the OS is not very efficient as it increases network and disk I/O as well as CPU utilization on the host while the backup is running. This results in fewer resources for the other VMs on that host. If multiple backups are running on the host, the problem will be even worse and can seriously degrade the performance of the host. This can also have an impact on your consolidation ratios as well as you will be forced to have less VMs per host to stay inside your backup windows.

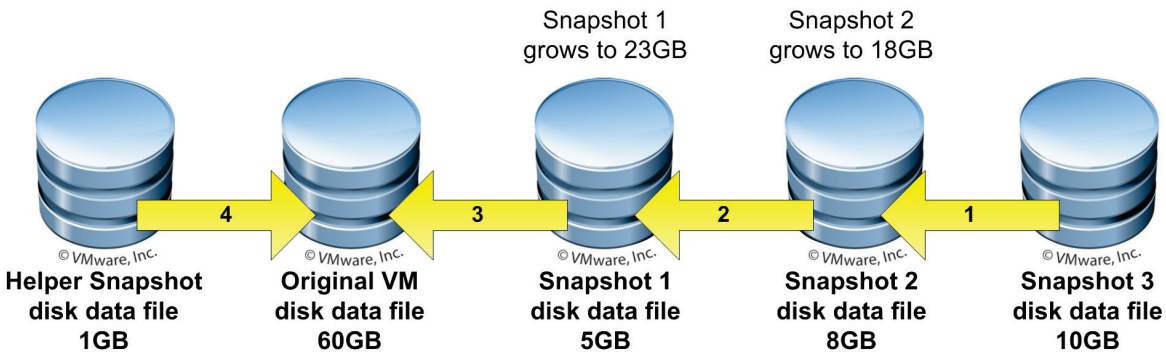
Most virtualization backup tools perform image-level backups at the virtualization layer outside of the guest OS, but they can still restore individual files if needed. This is possible because the backup software can mount the backed up virtual disks and browse the file system to access any file on it. Because of this, file-level backups are no longer needed in a virtual environment as backup applications can provide file-level restore capabilities using image-level backups for most major operating systems.

Virtual machine snapshots

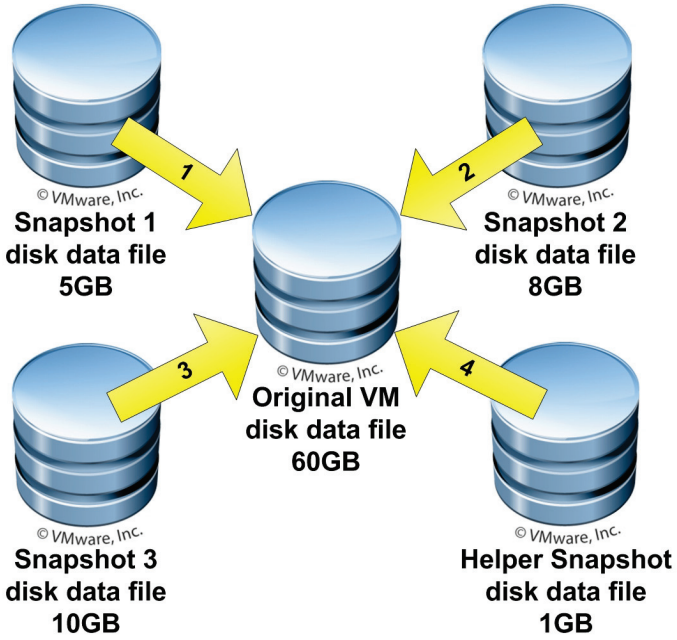
Snapshots are one of the great features that virtualization provides and can be a real lifesaver when upgrading or patching applications and servers. A snapshot is a point-in-time picture of a VM that preserves the disk file system and system memory of the VM. You can create more than one snapshot of a VM so you have multiple restore points available to recover from. When you create a snapshot all writes to the original VM disk files are suspended and they become read-only from that point on. All writes to the original disk files are deflected to new delta virtual disk files that are created when the snapshot is taken. When you delete all snapshots for a VM, all of the delta files that were created are merged back into the original vmdk disk file for the VM and then deleted when completed. If you choose to only delete an individual snapshot, then just that snapshot is merged into its parent snapshot and deleted.



The process for deleting multiple snapshots has changed across vSphere versions. In all versions, when you delete a snapshot before the data is committed back to the original disk file, a helper snapshot is first created to hold any disk writes that are made while the snapshots are being written back to the original disk. In some older 4.0 versions of vSphere, if a VM had 3 snapshots active and you deleted them all, the process would be Snapshot 3 would be copied to Snapshot 2, which would be copied to Snapshot 1, which would be copied to the original disk file and then the helper snapshot would be copied to the original disk file as outlined below.



This process would result in extra disk space being needed as each snapshot would grow as the previous snapshot was added to it. If there was not sufficient free disk space available on the datastore, the snapshots could not be committed. In later vSphere 4.0 versions and in vSphere 4.1, instead of each snapshot being merged into the previous one, each snapshot is merged directly back into the original disk in turn. So if a VM had 3 snapshots active and you deleted them all, the process would be Snapshot 1 would be copied to the original disk, Snapshot 2 would be copied to the original disk, Snapshot 3 would be copied to the original disk and then the helper snapshot would be copied to the original disk file as outlined below.



Because each snapshot is merged back into the original directly and one at a time, no extra disk space is needed except for the helper file. If you choose to revert to a snapshot, the current disk and memory states are discarded and the VM is brought back to the state of the snapshot you reverted to. Whichever snapshot you revert to then becomes the new parent snapshot. The parent snapshot is not always the most recently taken snapshot; if you revert back to an older snapshot it then becomes the parent of the current state of the virtual machine. The parent snapshot is always noted by the "You are here" label under it in Snapshot Manager.

A single snapshot file can never exceed the size of the original disk file; the reason for this is that any time a disk block is written to, it is created once in the delta file and simply updated if changed later on. If you changed every single disk block on your server after taking a snapshot, your snapshot would be the same size as your original disk file. However, the combined space of multiple snapshots could easily exceed the size of the original disk file. Snapshot files will initially be small in size (16MB) and grow larger as writes are made to the VM's disk files. Snapshots grow in 16MB increments to help reduce SCSI reservation conflicts. Whenever a request is made to change a block on the original disk, it is instead changed in the delta file. If the previously changed disk block in a delta file is changed again, it will not increase the size of the delta file because it simply updates the existing block in the delta file. The rate of growth of a snapshot will be determined by how much disk write activity occurs on your server after the snapshot is taken. Servers that have disk write-intensive applications like SQL and Exchange will have their snapshot files grow rapidly. On the other hand, servers with mostly static content and fewer disk writes like web and application servers will grow at a much slower rate. When you create multiple

snapshots, new delta files are created and the previous delta files become read only. With multiple snapshots, each delta file can potentially grow as large as the original disk file.

Virtual machine snapshots should not be considered as a primary method for backing up VMs. They are useful for short-term backups that are needed on the fly when it is necessary to preserve the state of a VM. The reason for this is that snapshots slightly degrade the performance of a VM as they grow and they also consume extra disk space on datastores. Additionally, because a VM's disk is split it creates the potential for problems and can prevent certain VM operations and features from being used. Snapshots should be closely monitored so you do not leave them running longer than necessary. The larger they grow the longer they can take to merge back into the original disk file when they are deleted. The commit operation that occurs when a snapshot is deleted can be very resource intensive on the VM while it is occurring.

While VM snapshots should not be used as a primary backup means, they are a primary enabler for performing image-level backups. Performing a snapshot before doing an image-level backup is necessary to prevent any disk blocks from changing while the virtual disk file is being backed up. The snapshot makes the VM's virtual disk read-only so the backup application can mount it and have exclusive access to it while the disk blocks are copied from it. Once the backup completes, the snapshot is deleted and all changes are written back to the original disk. Almost all virtualization backup applications rely on VM snapshots to perform image-level backups.

Consistent backup states

Whenever a snapshot of a VM is taken prior to backing it up, it is important to ensure it is in a consistent state so data can be properly restored if needed. This is especially important for transaction-sensitive applications like Active Directory, Exchange and SQL Server. When a snapshot of a VM is taken, its disk is frozen so no more writes occur to it while it is being backed up. At the point that the disk is frozen, however, there may be pending transactions or data in memory that has not yet been written to disk. This missing data can cause part of the backup data to be corrupt or incomplete. Because of this, before the snapshot is taken it is important to quiesce the VM, which temporarily halts the operating system and applications while they write any pending data to disk. Quiesce is a term used to describe the operation of pausing a computer while all outstanding writes are flushed to disk. Once this operation completes, the snapshot is taken and the operating system and applications can proceed as usual. This operation guarantees that the now read-only disk is in a state where applications and data can be properly restored if needed. There are several different states that a server can be in when the snapshot is taken, as outlined below:

- **Crash consistent**—this state is the same as if a VM hard crashed or had its power turned off without being properly shut down. All pending transactions and data in the VM's memory are lost and not written to disk. This is the default state if you take a snapshot of a VM without quiescing it first.
- **File system consistent**—in this state the operating system is quiesced, which allows for the operating system to write any pending data to disk before the snapshot is taken. This state does not take into account any applications that may be running and that may need to take additional steps to properly write all data to disk. This state is better than crash consistent but only ensures that the operating system is in a proper state to be backed up and not the applications running on it.

- **Application consistent**—in this state both the operating system and applications are quiesced so the VM is in a state where both the operating system and applications can be properly restored. This is the best possible state a VM can be in to ensure good backups. This type of quiescing only works with applications that specifically support being told to pause and write pending data when necessary. Typically this includes transaction-based applications like databases, email servers and financial systems.

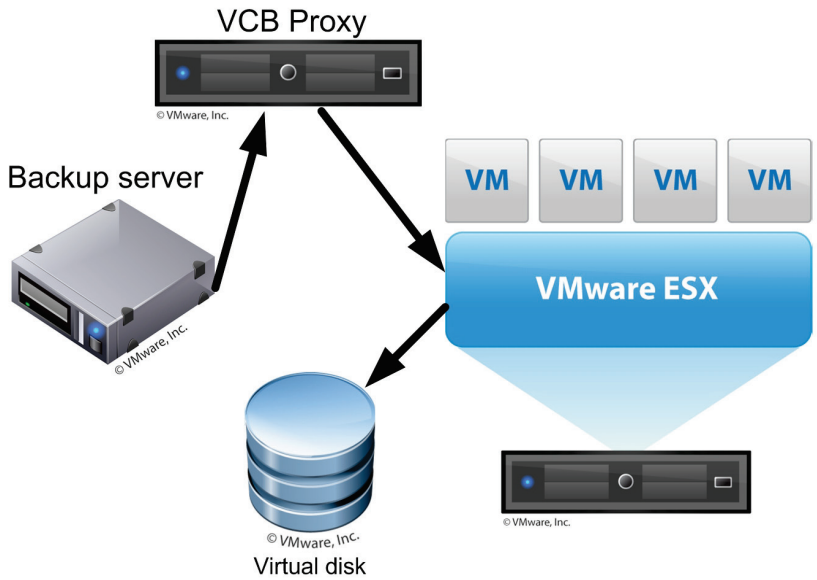
For VMs running Windows the quiescing is done through Microsoft VSS service, which works with the file system and applications to pause them and put them in a proper state for a snapshot and backup. The VSS service has several components that work together to take a shadow copy and put the disk in a consistent state, as outlined below:

- **Requestor**—this is the component that initiates the request to the VSS service. This is typically a backup application like Veeam Backup & Replication. Requestors also work with writers to collect information about the data to be backed up.
- **Writers**—this is the component that is part of applications and services that are designed to work with the VSS service. The writers work with VSS to prepare applications to quiesce their data and to ensure that no data is written until the shadow copy is created. Doing this ensures they are in a proper state to be backed up and all data in memory is written to disk.
- **Providers**—this is the component that does the actual work of creating the shadow copy. Once the writer has done the work to ensure that applications are quiesced, the provider creates and maintains the shadow copies until they are no longer needed. There are different types of providers that can be used, including hardware providers that offload the shadow copies to hardware storage devices, software providersthat work at the software layer to intercept and process I/O requests and write them to any type of storage device, and system providers that are built into the Windows OS and write to a NTFS volume on the system.

Having application-consistent backups is critical to ensure that your data is properly backed up. Most virtualization backup applications leverage the VMware Tools application that is installed inside the VM to quiesce the operating system and applications prior to creating the snapshot for backups. Instead of relying on VMware Tools, some vendors like Veeam also include their own agent that can run inside the VM and work directly with the backup application to quiesce the VM when needed. The reason for this is that VMware can be slow to update the integration with VSS in VMware Tools as changes are made to Windows. By not relying on VMware Tools to handle the quiescing, Veeam can be quicker to respond to changes or new versions in the Windows OS and doesn't have to wait for VMware to update its VSS integration in VMware Tools.

VCB & vStorage APIs

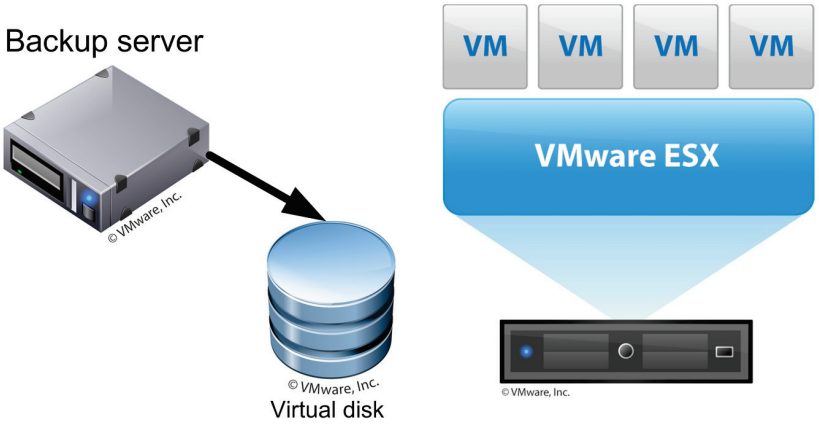
VMware recognized that a more efficient way of doing backups was necessary and developed methods to perform backups at the virtualization layer and never enter the guest operating system. VMware's first attempt at this was the VMware Consolidated Backup (VCB) framework included with VI3. VCB acted as a proxy server to offload backups from the virtual machine by mounting the virtual disk on the VCB server and then doing an image-level backup of it without involving the host or the VM. This shifted the backup overhead from the VM and the host to the proxy server instead. While this was a step in the right direction, it required a middleman between the backup device and the target disk as shown below:



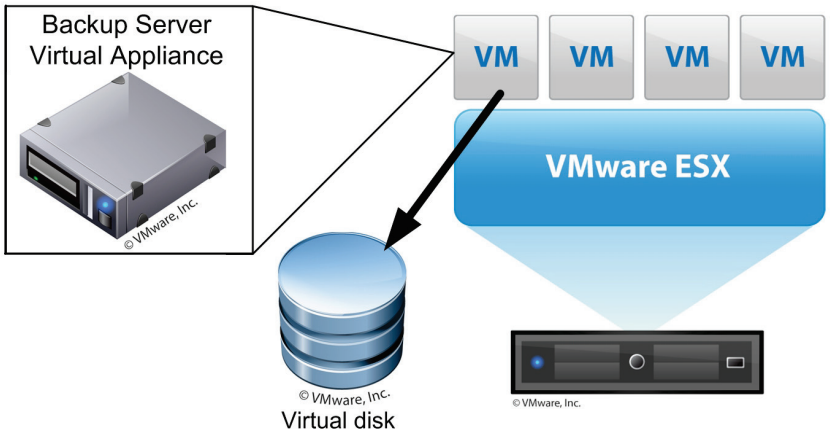
Backing up a VM using a VCB Proxy offload

With the vSphere release VMware eliminated VCB and the proxy that was used and instead leveraged APIs and SDKs so backup vendors could directly connect to virtual storage targets to backup VMs. The new vStorage APIs for Data Protection (VADP) include the functionality that was available in VCB and also added new functionality such as Changed Block Tracking (CBT) and the ability to directly interact with the contents of virtual disks. Doing this was much more efficient and offered more features than using VCB to back up virtual machines.

The two methods for using VADP to back up VMs are having the backup server directly access VM datastores and using a virtual appliance to hot-add a VMs disk file directly from its datastore as shown below:



Backing up a VM using VADP with direct disk access



Backing up a VM using VADP with a virtual appliance

The vStorage APIs existed in VI3 but were referred to as the VCB Backup Framework; however, unlike VCB they are not a separate standalone application and instead are built directly into ESX(i) and require no additional software installation. While the VCB Backup Framework still exists in vSphere and can be used by backup applications, the vStorage APIs are the successor to VCB and will eventually completely replace it.

The vStorage APIs are not really a single API, and the term is basically just a name for a collection of interfaces that can be utilized by 3rd party applications to interact with storage devices in vSphere. These interfaces consist of various SDKs that exist in vSphere and also the Virtual Disk Development Kit (VDDK). The VDDK is a combination API and SDK that enables vendors to develop applications that create and access virtual disk storage. The VDDK is used in conjunction with other vStorage APIs to offer a complete integrated solution for management of storage in vSphere. For example, while VM snapshots can be managed using the SDK functionality, other operations like mounting virtual disks are handled through the VDDK.

The vStorage APIs are broken into 4 groups that have different types of functionality, with the vStorage APIs for Data Protection being most beneficial to backup and replication applications. Changed Block Tracking is the standout feature as it allows 3rd party applications to query the API to find out which disk blocks have changed in a virtual machine's disk file since the last backup operation. Without this feature applications would have to figure this out on their own, which can be time-consuming; now with CBT they can instantly find this out so they know exactly which disk blocks need to be backed up. This enables much faster incremental backups and also allows for near-continuous data protection when replicating virtual disk files to other locations.

CBT is supported on any storage device and datastore in vSphere except physical mode Raw Device Mappings, which includes iSCSI, VMFS, NFS and local disks and it also works with both thin and thick disk types. As CBT is a new feature to vSphere, it does require that the virtual machine hardware be version 7, which is the default in vSphere. CBT is disabled by default as there is a very slight performance penalty that occurs when using it. It can be enabled on select VMs by adding parameters (ctkEnabled=true and scsi#:#.ctkEnabled=true) to the configuration file of the VM; backup applications can also enable it using the SDKs. Once CBT is enabled, a VM must go through something called a stun/unstun cycle for it to take effect; this cycle happens during certain VM operations including power on/off, suspend/resume and create/delete snapshot. During this cycle a VM's disk are re-opened, which allows a change tracking filter to be inserted into the storage stack for that VM.

The CBT feature stores information about changed blocks in a special "-ctk.vmdk" file that is created in each VM's home directory. This file is fixed length and does not grow; the size will vary based on the size of a virtual disk (about .5MB per 10GB of virtual disk size). Inside this file the state of each block is stored for tracking purposes using sequence numbers that can tell applications if a block has changed or not. One of these files will exist for each virtual disk for which CBT is enabled.

The vStorage APIs for Data Protection and the CBT feature make backups quicker and easier in vSphere and are a big improvement over VCB. Veeam was quick to recognize the many advantages that the vStorage APIs offered, and Veeam Backup & Replication was the first backup application to fully embrace and make use of the vStorage APIs.

Scheduling and performing backups

With physical environments servers don't share resources like VMs do so you can typically schedule backups without regard for other backups that may be running at the same time. Scheduling backups in a virtual environment does require some planning though. The reason for this is you don't want to put too much resource strain on hosts and shared storage datastores that may negatively impact VMs that need resources to function properly. How much resource strain you will encounter depends on the method used to back up the VMs. When backing up VMs using agents inside the guest OS, the resource usage on the host will be the greatest. If VMs are being backed up using image-level backups, the resource usage on the host will be lessened. When leveraging the new features in the vStorage APIs, the resource usage will be even further reduced.

Depending on the backup method you use you should take care to make sure you do not back up too many VMs simultaneously on the same host and shared datastores. When using agent based or over the network backups you can put too many resource constraints on the host. When using direct to datastore backup methods you can put too many resource constraints on shared datastores. This can cause your VMs to be resource starved while backups are occurring and VM performance can suffer greatly as a result. Doing some basic planning of backup schedules can help ensure that backups are staggered to not put too much burden on a single resource point at

the same time. Also be aware that snapshots are taken and deleted on all VMs being backed when doing image-level backups, so this can also cause additional resource constraints in the environment, particularly with storage.

Leveraging the vStorage API methods to streamline backups and backing up to disk targets mean everything moves at a very fast pace. Perhaps one of the biggest bottlenecks can be the backup server that is handling all the backup coordination tasks. When backup operations are running, there is more to it than just copying data from point A to point B. Backup operations also do a lot of CPU processing to determine what data to back up and what not to back up, deduplicate data, and also compress data that is written to the target. Having an undersized server especially in the CPU area can greatly reduce backup performance. Therefore, it is important to not skimp on resources for your backup server; running on a physical server or VM with at least 4 CPUs is necessary to get the best backup performance possible.

Virtualization advantages

Because of its unique and flexible architecture, virtualization provides many advantages over traditional physical servers when it comes to backup and recovery. These advantages can help companies save time and money as well as provide new capabilities to the data center. Listed below is a summary of the advantages that virtualization can provide related to backup and recovery:

- **Easier bare-metal restores**—Because VMs are all presented with the same virtual hardware, bare-metal recovery is much easier as VMs will always see the same virtual hardware regardless of the physical hardware that the host is using.
- **Easier backup verification**—Backed up VMs can easily be restored to any host or workstation to be verified without disrupting the original running VMs. SureBackup Recovery Verification simplifies and automates this and allows you to do this directly from the backup target datastore.
- **Backup usability**—With Veeam vPower you can actually put your backups to work instead of leaving them sit there taking up space until a restore is needed. Backed up VMs can be brought online for testing and troubleshooting purposes without affecting the original VMs.
- **Multiple backups methods**—Traditional servers typically need to do separate image-level for bare metal restore capabilities in addition to standard file-level backups. Virtualization only needs to back up once at the image level but can provide both image-level and file-level restores from that one backup.
- **Agentless**—Backup agents do not need to be installed on each server to be backed up as the backup is performed at the virtualization layer without going through the guest OS.
- **Resource-free backups**—The VM's virtual disk is accessed directly when backed up without involving the VM's guest OS, so no resources are tied up on the VM while the backup is running.
- **Network-free backups**—The backup server can connect directly to the VM's virtual disk using the storage network, so data being backed up does not need to be dragged across the network where it can impact other servers communicating on the network.
- **Easy snapshots**—VM snapshots provide an easy short-term backup solution for VMs and also enable easy image-level backups by freezing the VM's disk during backups.
- **Easy application-item recovery**—With Veeam vPower, application-level items such as database records and emails can easily be recovered from an isolated environment and copied as needed back to the original environment.

SUMMARY

In this chapter we compared traditional backup methods to those used in virtual environments. Implementing virtualization requires a different mindset and techniques as traditional methodologies are not as efficient and practical in a virtual environment. Virtualization architectures can be a big enabler for backup and recovery in the data center and open the door for new and more efficient ways of doing things. There are some people that continue doing backup and recovery the way they always have after implementing virtualization despite it being less efficient. It's important to leverage the strengths of virtualization and change your methodologies so you can take advantage of the great features that virtualization has to offer to improve your backup and recovery methods. After all, backup and recovery are never easy, so why make it harder than it needs to be.

CONTENTS

INTRODUCTION

Backing up your virtual environment is a fairly easy and straightforward process, but the whole point of backing it up is to have the ability to restore data when needed. All those 1s and 0s that are stored on hard disks throughout your datacenter are the most critical assets that your business has and you can't afford to lose any of them. In any environment, whether it's physical or virtual, you can't afford to have your restores fail. There is no worse feeling then finding out that after backing up your servers for many months, or years, that your restores do not work properly. In many cases we blindly trust that our restores will work well without actually testing them. After all, our backups seem to work just fine so shouldn't our restores do the same? Performing periodic test restores can help validate that your backups are working properly, but it can be a time-consuming and complicated process.

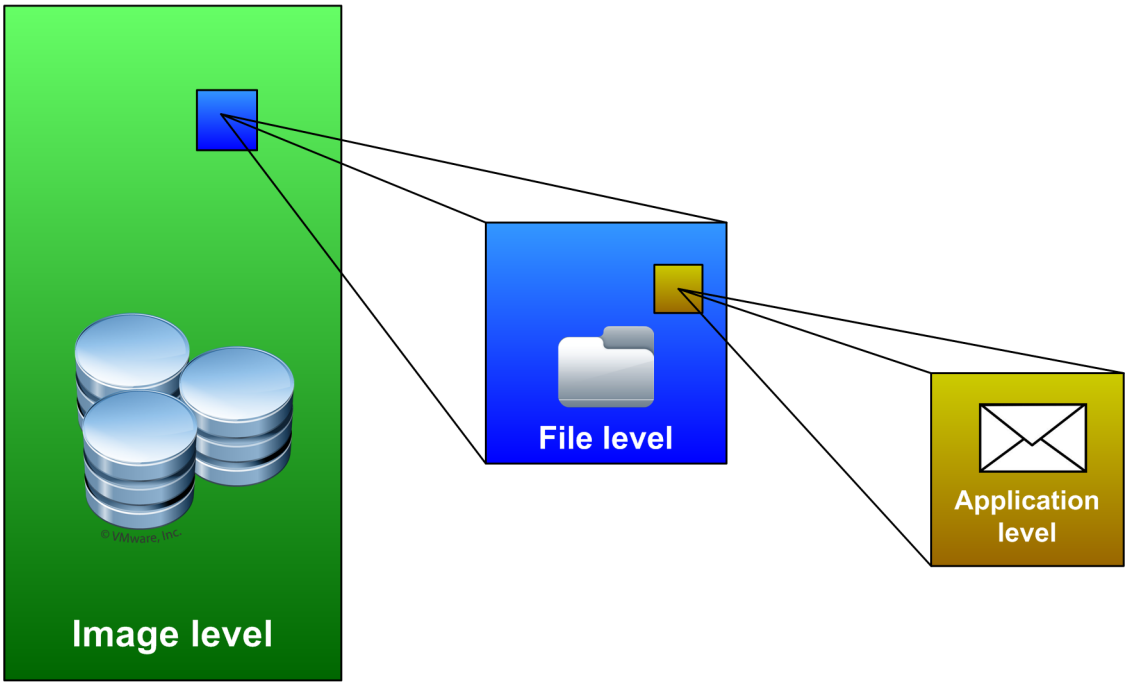
Restoring data in virtual environments can be a bit more complicated than traditional physical environments as there are more options and methods available to restore the data as well as more moving parts. But this is also a positive aspect of virtual environments as it provides you with greater flexibility and more innovative ways to restore your data and verify that your backups are working properly. In this chapter we will explore the methods and challenges for restoring and verifying data in virtual environments. We will also demonstrate how the innovative features of Veeam's Backup & Replication v5 with vPower can make restoring data easier and simpler, and also help ensure that your backups are 100% reliable.

RESTORING VIRTUAL MACHINE DATA

There are basically three levels at which you can back up a server:

- Application level—Where you typically back up individual objects inside of a file
- File level—Where you typically back up files located on a disk partition
- Image level—where you typically back up all the blocks of a disk partition

These levels have a hierarchy to them starting at the top with the application level and then down to the image level.



Each backup level has different advantages and it is not uncommon to back up at multiple levels based on different restore requirements. Whichever level you back up data at is where you also have to restore it to; however, if data is restored at one level, it can be used to restore data at a higher level as well. For example, if you back up at the file level you can restore a file and then access the application objects within to do an application level restore. However, you cannot restore data backed up at a higher level to a lower level. For example, if you back up at the application level you can only restore application objects, you can't do file level restores.

With traditional physical servers, restoring data from one level to a higher level can be complicated and also require extra resources. For example, if you back up a server at the image level so you can perform bare-metal restores, you typically need to find an unused, similar physical server to restore to so you can access the file system and the individual files. If you back up at the file level you typically need to set up another copy of an application so you can access the individual objects in the file and copy the ones you need back to the original application. With virtualization, its unique architecture provides greater flexibility which makes restores at any level much simpler and easier.

File level restores

As we covered in the previous chapter, most backup applications in virtual environments back up at the virtualization layer using image level backups that are more efficient than using traditional agents installed inside the guest operating system. One common misconception when doing image level backups is that since you are only backing up a VM's large virtual disk file from outside the guest operating system, you cannot do individual file restores from inside the guest operating system. Individual file restores are indeed possible with image level backups; the methods used are simply different than traditional restore methods. With traditional file level backups using an agent inside the guest operating system, you typically create a catalog of all the files as they are backed up which is then used as a reference so they can be restored later on. Image level backups have this capability as well, by simply mounting the virtual disk file that is backed up; you can look inside the guest operating system to see the file layout. As a result any individual file that resides on the virtual disk can easily be restored.

With file level backups, individual file restores are fairly simple. You pick the file to restore from the backup media and the backup server connects to the agent on the target server, locates the file on the backup media, and copies it back to the original source. With image level backups there typically is no agent on the target server, so it's slightly different. What happens instead is the virtual disk file from the backup repository is mounted. This allows for the file to be copied from it to either a local disk or back to the original server. Once the file is copied, the virtual disk is un-mounted and the mounting process simply attaches the image level copy of the original disk to the backup server where it is seen as an additional drive and the file system can be accessed. With virtualization, there is no need to actually restore the entire VM image just to restore a single file. The process for individual file restores in virtualization is different but the end result is the same.

Bare metal restores

A bare metal restore is commonly used in DR/BC situations where you are restoring whole servers directly onto physical server hardware without first installing any software or an operating system. With virtualization doing bare metal restores is a very simple and easy process because the whole VM is backed up at the image level as a single file. Since the VM is encapsulated into one big file all you have to do is copy that file back to a virtual host and you have a complete copy of the server from the point in time of the backup. VMs can also be restored to alternate locations and powered on with other virtualization tools that support the same VM format. For example, you could restore a whole VM from an ESX host to a workstation where you could easily start up the VM using a tool such as VMware Player or Workstation. From there, you could use it in an isolated environment without affecting the original VM. The options and possibilities are endless as backups of virtualization environments offers many more options than traditional physical backups. Instead of just using backups as backups, you can make use of backups for other things like testing and troubleshooting.

Another big advantage of virtualization with bare metal restores is that VMs all have the same type of virtual hardware regardless of the underlying physical host hardware. This eliminates any hardware incompatibilities that may occur when performing a bare metal restore to a different host. With traditional backups, if you do a bare metal restore to a different server there are many pre- and post-restore steps that need to be done to make sure hardware drivers, disk partitions and system configurations are all correct for the new hardware. The virtualization architecture greatly simplifies how bare metal restores are performed and eliminates the complexities and headaches that are typically encountered in physical environments.

Application level restores

With application level restores you are typically restoring one or more data objects from an application back to its original location within a file or container. An example of this is restoring a group of rows back to a SQL database or restoring user/group objects back to an Active Directory database. These types of objects are not stored as individual files but instead stored in one large file along with many other objects. The application that is storing the objects in the file is aware of the file structure and how to read/write to it. There are two ways you can backup and restore individual objects in the file. You can use a special backup agent that understands the application and can back up and restore individual application objects, or you can restore the file and use the native application to copy the objects from the restored file back to the original application file.

Application level restores can be complicated when you try to copy objects from the backed up file back to the original copy of the source file. If you were to try to restore certain rows back to a SQL database, you couldn't just restore the database file that they were located in. You need a copy of the application running as well so you can access the database, select the records that you need to restore, and finally write them back to the original database. With virtualization, this process becomes easier because you can restore, isolate and power on a backup copy of the original VM so data can be copied from it using the application without impacting the original VM.

TESTING & VERIFYING BACKUPS

While having good backups is very important, having good restores is even more important. Backups are worthless if you cannot properly restore files when necessary. You should never make the mistake of assuming that just because your backup software doesn't report any errors when the data was backed up that everything is okay and you won't have problems restoring the data. You must regularly test and verify that your backups are working properly by restoring data from your backup media and then accessing it to ensure that it can be read properly. This means more than just restoring a single file and making sure you can open it. For proper verification of your backups, you need to test your restore capabilities at multiple levels: the file level, application level and operating system level. This validates your backups and ensures that you are able to restore data for any type of situation whether it is a few records in a database or an entire VM.

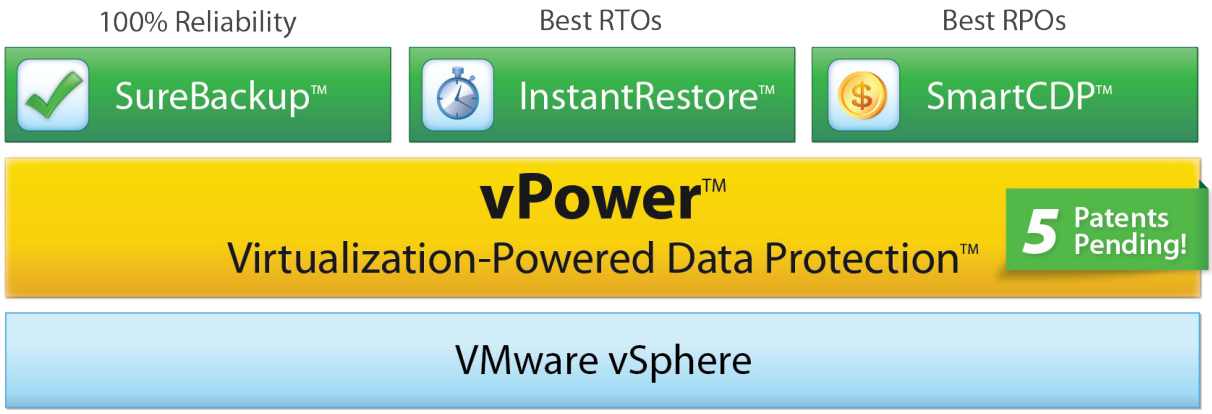
Backing up regular non-transactional types of files is pretty straightforward and you will usually not encounter any problems with them. But backing up transactional files can be tricky because if the application is not in a proper state when the backup occurs then you may not be able to restore it properly. To prevent this from happening, applications can be quiesced before they are backed up so the data is in an application-consistent state. This means it can be restored properly if needed.

Testing the restore capability of application level backups can be more difficult than testing simple file restores because you have to restore the application to a working state to see if you can read the data that is contained within it. Finally, when backing up an entire operating system you also need to verify that it will boot and function properly if it is ever restored. If any critical boot files are missing or modified the operating system will continue to function normally but your backup will be in a state that will not allow a restored copy of the VM to boot properly. Consequently, you need to verify that your image level backups are working properly so an entire VM can be restored bare-metal if needed.

Testing file level restores with traditional backups of physical servers is pretty straightforward since you can restore files just about anywhere in your environment and then open them to check their integrity. Testing application level and image level restores can be difficult and time-consuming and often extra server hardware is needed so you do not disrupt the original servers. Virtualization can make this a much simpler process because VMs can be restored and isolated on hosts without overwriting and affecting the original VMs. This makes testing a restores of individual files, applications or whole VMs an easy process. While verifying and testing backups with virtualization is easier from a process level, it still can be time-consuming as you must manually restore data and verify it; especially if you do it on a regular basis for many VMs.

VPOWER IS A GAME CHANGER

Veeam's revolutionary vPower, part of Veeam Backup & Replication v5, addresses some of the pain points that are typically associated with backup and recovery by introducing some innovative new features that fully leverage the strengths that the virtualization architecture provides. vPower is short for Virtualization-Powered Data Protection and consists of three feature silos: SureBackup, InstantRestore and SmartCDP.



Earlier in this chapter we covered how the virtualization architecture can help simplify and improve recovery operations compared to traditional physical architectures. vPower takes that a step further in more creative ways to really take full advantage of the flexibility and features that virtualization provides to take recovery in virtual environments to the next level.

The magic behind vPower is based on an innovative method of transforming the backup server into an NFS server which allows it to publish a compressed and deduplicated backup file as a regular VMDK file. The hosts can then access this file as they would any NFS datastore and, as a result, the VM can be powered on directly from the backup repository without first extracting it.

In the next section we will cover each component of vPower in detail so you can understand how the features work and understand the significant benefits that they provide.

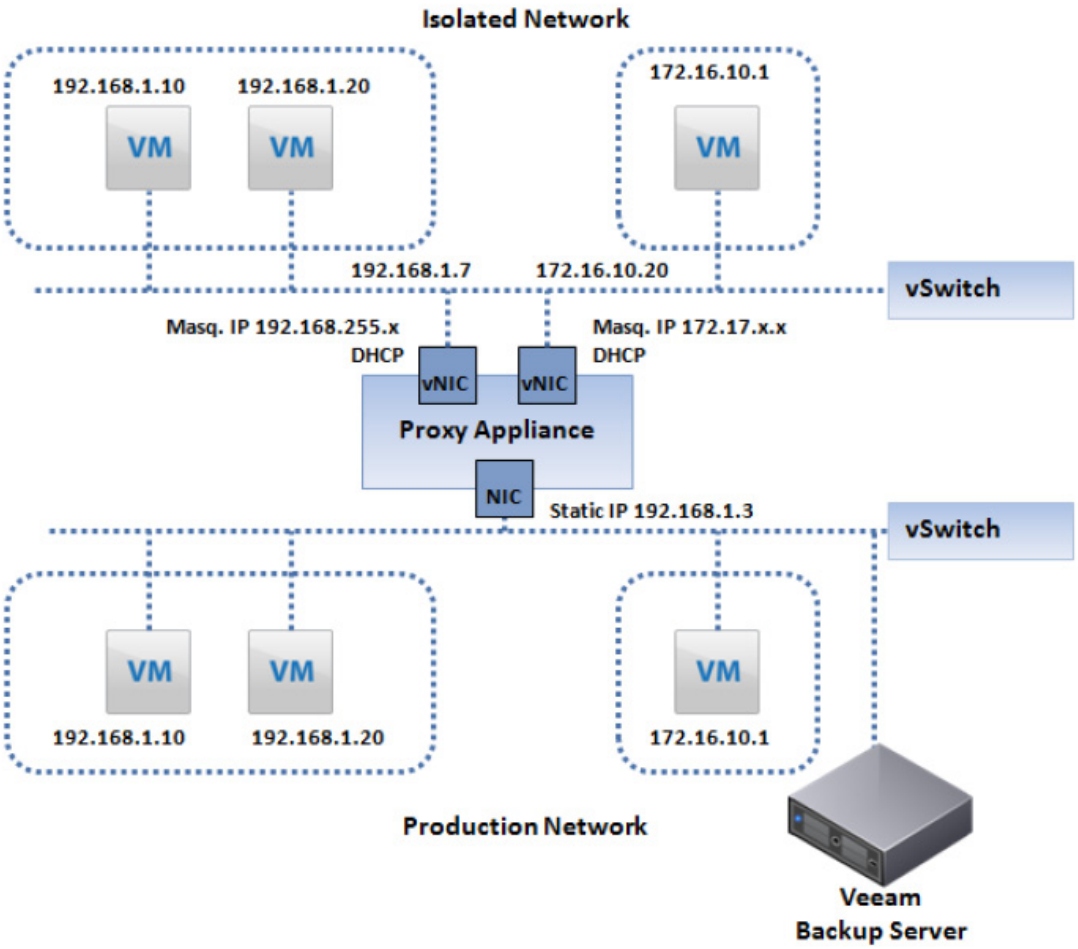
SureBackup

We talked about the challenges of testing and verification of backups and how virtualization can make this easier across all the different levels of backups. Veeam has made this even easier by introducing a new feature called SureBackup that automates the verification of VMs in a separate isolated environment on a host so you can verify that data and applications will function properly when restored. This allows you to verify your backups on multiple levels, at the image level by verifying that the VM will boot successfully, at the application level by verifying that applications function properly and data can be accessed, and at the file level by verifying that files can be opened properly.

While backup verification is easier with virtualization, it is still time-consuming. Normally, you have to copy the virtual disk files from the backup repository to a host, register the VM on the host, make sure it is on an isolated network so it doesn't impact the original VM, then power it on and verify that it boots and functions properly. If the VM is running a tiered application that relies on other VMs, you also need to restore those and bring them up so you can test applications. Once you've successfully verified that your backups work properly, you then need to clean up and remove everything that you restored. The whole process from start to finish can take quite some time, and trying to do this on a regular basis can be a full-time job.

The automation that SureBackup provides eliminates all the steps that you would normally have to do manually in order to verify backups. In addition, it also eliminates the need to restore the VM from the backup repository back to a host datastore to be able to power it on, which normally would require that extra storage be available. With SureBackup you can power on a VM directly from the target backup store without having to extract it to a datastore attached to a host and without interfering with the original running VM. Once the backup copy of the VM is powered on, it is kept isolated from the rest of the network so it will not affect the original production VM that the backup was generated from. This is made possible by the use of a virtual lab with an isolated vSwitch that contains no physical NICs that the backup copy of the VM is attached to. Connectivity to other networks is made possible by a proxy appliance that has multiple vNICs and acts as a gateway, routing requests from the production network to the isolated network. The backup VM's original IP address is preserved and, to avoid conflicts, the proxy appliance uses a combination of masqueraded IP addresses and special routing rules.

After the VM is powered on, the verification is performed. This consists of checking the VM's heartbeat that is generated from VMware Tools and also by pinging its IP address. This verifies that the operating system was able to successfully boot and it verifies bare-metal restore capabilities. You can additionally specify test scripts that can be run to make sure applications are functioning properly and data is accessible. Throughout the entire process, the original VM disk in the backup repository is kept as read-only so the backup image remains in its original point-in-time state. Any disk writes that occur as a result of the VM being powered on are deflected to a delta file, just like with a snapshot, that are discarded after the verification process completes.



In some cases, the application on a VM that you are verifying may be dependent on the availability of applications or services on other VMs for it to function properly. SureBackup has this covered as well since multiple VMs can be powered on as part of the verification process. Application groups can be defined that specify which VMs to power on, you can also specify boot priorities and delays to ensure that any application dependencies are met before VMs are powered on.

The whole verification process is completely automated and requires no interaction. Once the backup completes, SureBackup handles all the steps needed to verify it. If a problem occurs and the backup cannot be verified, SureBackup will alert you so you can investigate and take action to ensure you have proper backups. SureBackup can also provide you with the peace of mind of knowing that your data and applications are safely backed up and can be restored if needed.

InstantRestore

The InstantRestore feature allows you to quickly restore data at any level directly from the backup repository; this includes application objects, files and whole VM images. InstantRestore works in a similar manner as SureBackup in that a backup copy of the VM's disk is mounted from the backup repository to a host and powered on so data can be selected and quickly restored to the target destination. To restore files and application objects there is no need to restore the whole VM from the backup repository, which results in much quicker recovery times. InstantRestore can be done at the image, file and application levels as outlined in the following sections.

Instant VM Recovery

Instant VM Recovery allows you to instantly power on a point-in-time copy of any VM that resides in the backup repository that can be used to recover a failed or problematic VM. If a problem were to occur with a production VM that makes it non-functional, you typically have to spend time troubleshooting it and resolving the problem until it is working again. While this is happening, your VM and the services it provides are usually not available. With Instant VM Recovery you can quickly bring up a backup copy to stand in for it while you resolve any issues with the original VM. You can also permanently replace the original VM with the backup copy. Having backup copies of your VMs that are instantly available for use provides you with very quick recovery time objectives. Instant VM Recovery gets you back up and running quickly without having to use costly DR/BC solutions.

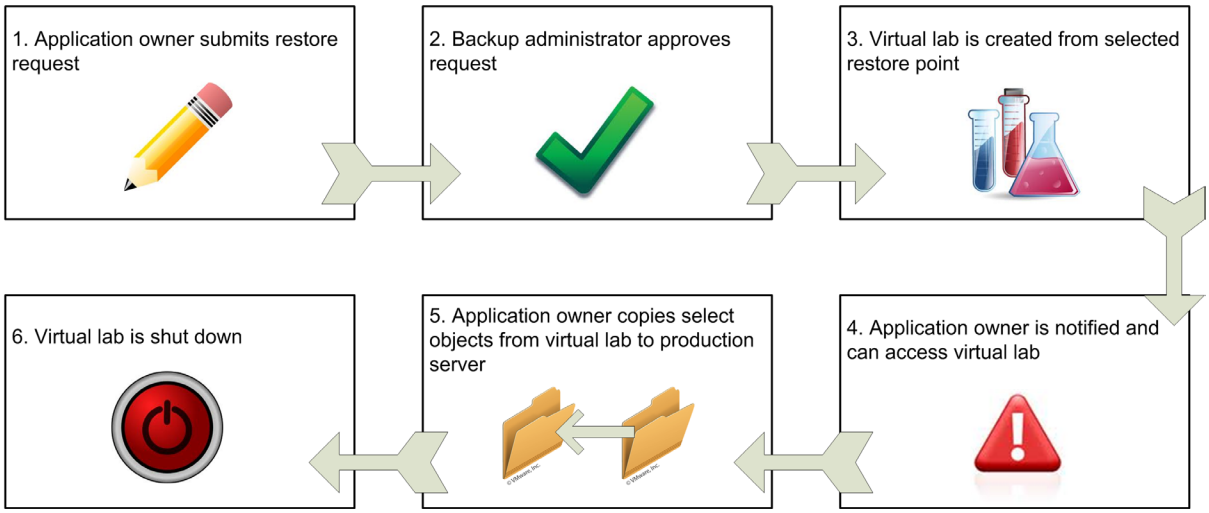
To quickly recover a VM from the backup repository, you launch a simple wizard and select a VM and then select one of the available restore points. You then select a host to run the VM and choose whether to connect it to the network. Once you've completed the wizard, the VM will be registered on the host that you selected and powered on. The backup copy of the VM runs from the NFS datastore that the backup server creates for the backup repository. If you want to use the backup VM for an extended period you can also migrate it from the backup repository to a host datastore while it is running using the Storage vMotion feature. You can also replicate, or "hot copy," the VM to another datastore with the features built into Veeam Backup & Recovery if Storage vMotion is not available.

Universal Application Item Recovery

Restoring individual application objects like database records or emails can be difficult and time-consuming. How many times have you had to restore a large file just to access a few small objects contained within it? Because application objects are contained inside a file, you usually need to restore the application as well so you can access them. Universal Application Item Recovery (U-AIR) is a feature that can greatly simplify this process for any virtualized application running on any guest operating system. U-AIR is universal in nature as it works directly with applications' native utilities, it does not require any additional agents or backup tools. The process is streamlined by using simple wizards to first configure a restore job and then by using a workflow process that allows backup administrators and application owners to work together to restore select data back to the original source.

U-AIR works by creating an isolated virtual lab environment that is used to bring up the application so that individual objects can be selected for restoration. It works in the same manner as SureBackup and all VMs necessary to restore the application data are powered on directly from the backup repository. To start the process, a virtual lab request is created that defines the lifetime of the lab, the source server/application and the restore point date. The lab requests can

be created by the application owners and after they are created, they are subject to approval by the backup administrators. Once the request is approved, a SureBackup job will be initiated which will create the virtual lab by starting the necessary VMs from the backup repository. Once the virtual lab is up and running the application owner that created the request will be notified that they can begin restoring the application objects from it. The isolated lab does not have direct access to the original production VM; it must be kept isolated so the two do not conflict. Instead, a special routing table is setup so the application owner has access to both environments simultaneously and can copy data from one to the other using the applications native utilities. Once the application objects have been successfully copied the lab is shutdown.



Because of the split workflow process, backup administrators do not have to get involved with the application, and application owners do not have to get involved with backup and recovery. Each group can stay within their specialty area which also results in better security because backup administrators do not need to be assigned access to application objects. U-AIR takes the complexity and burden out of restoring individual application objects and turns it into a simple and straightforward process.

Instant File-Level Recovery

Instant File-Level Recovery allows you to quickly retrieve individual files from image level backups on a VM with any guest operating system or file system. With image level backups that back up at the virtualization layer and do not access the guest operating system file system directly. It can make restoring individual files difficult. With traditional backups that use backup agents installed inside the guest operating system, the backup server only needs to communicate with the backup agent that runs natively on the guest operating system. Because the backup agent is running on the guest operating system it does not need any special file system drivers to read and write to the disk partitions. With image-level backups that run outside the guest operating system without any agents, to restore files back inside the guest operating system the backup server would have to have the appropriate file system drivers for every file system.

Instant File-Level Recovery makes this possible on any guest operating system by using several different methods depending on the file system that is used. Because the backup server is running Windows, it can natively write to any Windows operating system and the backup server can simply assign a drive letter to the destination VM and restore files to it. For VMs that are running any Linux variations, the backup server relies on a small virtual appliance that runs Linux

and can write to Linux file systems to use as a proxy to restore files to a VM running Linux. Finally, for the other supported guest operating systems like BSD, Solaris and Macintosh, the backup server relies on a helper VM running the same operating system as the server that files are being restored to that is able to read the file system of the image-level backup. The helper VM can be any VM running the same guest operating system (even the original source server), the image level backup file is mounted as an additional drive on the helper VM using the hot-add feature and then files can be copied from it and back to the source server.

Instant File-Level Recovery covers all the bases and solves the problem of restoring individual files from image-level backups from any operating system or file system. Leveraging vPower allows you to quickly mount an image-level backup directly from the backup repository and easily restore individual files without having to take any additional steps that you would typically have to take to copy files from image-level backups.

SmartCDP

Veeam Backup & Replication combines backup and replication into one package so you can not only backup and restore VMs but also replicate them to other hosts as well. This provides an affordable alternative for disaster recovery without having to rely on more expensive and complicated solutions. SmartCDP offers near-continuous data protection (near-CDP) so you can keep replicas of critical VMs continuously updated on other hosts and storage devices. Typically to obtain continuous data protection (CDP) you need to rely on a specialized product that replicates data at the storage layer between two storage devices. A CDP system uses real-time replication to allow for instantaneous recovery to achieve a recovery point objective (RPO) of zero data loss. Near-CDP typically relies on storage snapshots that are taken on a periodic basis which does result in a minimal amount of data loss since the last snapshot was taken. While it doesn't provide the protection of CDP, near-CDP provides a good enough solution for many where the minimal amount of data loss is acceptable.

SmartCDP leverages the replication feature in Veeam Backup & Replication to provide near-CDP for any VM. Instead of replicating at the storage layer it replicates at the virtualization layer so it requires no special storage hardware or software. SmartCDP takes advantage of the Changed Block Tracking (CBT) feature that is part of the vStorage APIs that allows for very quick incremental updates. CBT is a feature built into the vSphere hypervisor that keeps track of all disk blocks that have changed on a VM's virtual disk from a point in time. Applications can simply query the hypervisor and instantly obtain a list of disk blocks that have changed since the last backup or replication operation rather than spending time attempting to do it on their own. As a result, CBT can greatly decrease the amount of time it takes to perform incremental updates.

SmartCDP can be configured to replicate at specific time intervals or can be set to continuously replicate. When set to continuously replicate, SmartCDP will begin replicating immediately after the previous replication operation completes. The factors that will influence the RPO that you can achieve with SmartCDP are the amount of data that changes, the network bandwidth between the source and target hosts, and the available backup server and host resources. It is certainly possible to achieve an RPO using SmartCDP as low as a few minutes. SmartCDP delivers near-CDP and high availability for your VMs and, because it replicates at the virtualization layer, it works with any type of storage that is supported by vSphere.

vPower is more than just a backup & recovery feature

vPower is the engine that makes the innovative features in Veeam Backup & Replication possible, but it can be used for much more than simply backing up and recovering VMs. The virtual labs that can be created from the backup repository using vPower allow you to create on-demand virtual sandbox environments whenever needed. Having the ability to instantly bring up old point-in-time copies of any VM is a powerful feature that has many use cases. For example, if a problem occurs with your original production VM that makes it non-functional, you can quickly bring up a backup copy to stand in until the original is fixed. You can also use it for testing changes that are made to VMs such as patching, rather than impact a production VM when you make changes, you can use a backup copy instead to see the impact of the change. Once you have verified that there are no issues you can shut down the backup copy which will discard all the changes that you made and make the changes to the production VM.

vPower can be used for a variety of purposes and can help you troubleshoot issues, perform tests, recover files and much more without impacting production VMs. Backups are like insurance policies, you have to have them but they normally just sit around taking up valuable disk space until the occasional restore is needed. vPower allows you to actually make use of your backups for more than just recovery which gives you much more return on the investment you make in your backup infrastructure.

SUMMARY

In this chapter we covered the different recovery levels that data can be restored at and how the flexibility of the virtualization architecture provides benefits to VM recovery operations. We explained how even though VMs are backed up at the image level, data can still be restored at the file level, application level or image level. Virtualization can make recovery much easier as it provides you additional options that are typically not available in traditional physical server environments. Testing backup integrity is a critical task to ensure that your data can be restored when needed. Virtualization can help with this as well and reduce both the time and resource requirements for verifying backups.

We also covered how vPower leverages the virtual architecture and Veeam innovations to take data protection to the next level and make disaster recovery affordable. vPower eliminates some of the fundamental shortcomings of traditional approaches to reduce the cost, ensure the reliability and increase the value of data. Veeam Backup & Replication v5 with vPower empowers you to meet your recovery time and recovery point objectives for less time, effort, cost and risk.

CONTENTS

INTRODUCTION

Virtualization is an invisible layer that is inserted between physical server hardware and operating systems. When you implement virtualization in your data center, it affects almost everything in your data center that surrounds or connects to your virtual environment. This includes any type of hardware & software, such as storage, networking, monitoring and of course backup & recovery. With virtualization you're in a whole new world especially when it comes to your critical backups. These backups are your safety net for anything that happens in your data center, and you need make sure that you choose a backup solution that is made for that new world as well.

Backup solutions are seldom given enough consideration or attention when implementing virtualization. Often times, no thought at all is given to how virtualization will affect backup and recovery until after virtualization has been fully implemented in the data center. Ideally, when you are in the planning stages of virtualization, a new backup strategy should coincide with this plan that is based on a solution that is optimized for virtual environments. This ensures that you have the proper hardware in place to support your backup solution and the design of your virtual environment is compatible with the backup method you choose. In this chapter we will cover all the considerations and choices you need to make when choosing a backup solution for a VMware environment to ensure that you choose the right one, for the right reasons.

REQUIREMENTS AND SLA'S

Your requirements and SLA's are going to be one of the biggest influencers of choosing a backup solution. Most companies have limited budgets which can hinder the number of available product choices that are within a given budget. Additional critical requirements are how quickly data can be restored, also known as Recovery Time Objective (RTO) as well as how much data can a company afford to lose, also known as Recovery Point Objective (RPO). Backups are disruptive to any environment because they generate I/O workloads that can take resources away from applications that need them. In virtualization this is amplified as virtual machines all share common infrastructure components. Because of this, most companies want to backup data in the shortest window possible with the least disruption to production workloads. This makes choosing a backup solution very challenging since all of these factors must be kept in consideration when deciding on a solutions that meets all your needs.

Budgets

Budgets are always a big influencer on the backup solution that you choose; after all you can't buy what you can't afford. Most backup solutions consist of not just backup software, but also hardware components to support the backup target. Because price is often a concern, it's important to maximize the dollars that you spend on a backup solution so you get the most bang for your buck. Often times this means getting creative with the solution that you choose, instead of choosing a massive solution that meets your capacity requirements and fits in your backup window, a more affordable pairing of solutions may be a better choice. The unique architecture of virtualization opens the door for many creative solutions for backup & recovery that are not possible in traditional physical server environments. Using disk targets is becoming more popular in virtualized for a simple and affordable solution. You need to ensure you explore all your options when choosing a backup solution. There are more backup options available for virtualized environments then you might be aware of, so whatever solution you choose you need to also keep in mind your DR strategy since your backup solution plays such an integral part of a DR solution. Some backup applications like the one from Veeam Software allow you to not just backup VMs, but to replicate them as well which can play a big part in your DR strategy. So now, instead of deploying and maintaining two disparate solutions, you can have your backup and disaster recovery solution in one product. This provides you with not just a backup solution, but a DR solution as well.

Mixed virtual and physical environments

Virtualization takes time to implement, in most cases you are migrating your physical servers into virtual machines which can often take months or years to fully complete. During that time, you have to support both traditional physical servers and your virtual environment since most companies do not virtualize 100% of their servers right out of the gate, as a result they will have to support both virtual and physical environments. When it comes to backup applications, most companies have a solution already in place for their physical servers before they make the move to virtualization. These legacy backup solutions that were designed to for physical servers are less efficient in virtual environments, and as a result many companies look to implement a backup solution that is optimized for virtual environments.

Many of the backup applications like Veeam Backup & Replication that were built from the ground up to support virtual environments do not support backing up non-virtualized servers. This can result in what appear to be a potential increase in operational costs since you are forced to maintain two backup solutions, one for virtual environment and another for physical

environments. While you may be hesitant to invest in a new backup solution that can only backup your virtual environment, there are several advantages to doing this:

- Right tool for the job – It's really important to use the right tool for a job, tools that were created for physical environments are not as efficient in virtual environments. Even if tools are adapted to accommodate virtual environments they are typically not as effective as tools that are specifically created for virtual environments.
- Separation – Physical environments are very different from virtual environments. Having that separation between your backup systems can make things simpler as there is no mix between environments which can cause confusion and mistakes.
- Best-of-breed – This allows you to use the best possible solution in your virtual environment and not have to settle for one solution for both physical and virtual environments that may do one environment well but not the other.
- Virtualization percentage will only increase – You may not be 100% virtualized today but that percentage is only going to increase. Many industry experts predict that most companies will end up 80-100% virtualized. This makes having the best possible backup solution in your virtual environment very important.
- Physical backups can fit certain requirements– Most physical backup solutions write to tape targets (D2T), many virtual backup solutions write to disk targets (D2D). Most companies still rely on tape for offsite and long term retention of backups. Having two solutions allows you to back up your virtual environment to tape as well by having your physical solution backup the disk repository of the virtual solution (D2D2T).

So don't force yourself into the "one-size-fits-all" mindset, having separate backup solutions for your physical and virtual environments has a lot of advantages and is not as complicated to manage as you might think.

Backup Windows

Backup windows are the length of time you can perform backups within and are typically centered around avoiding any disruption to production environments. Backups are a disruptive operation by nature as they cause heavy I/O to occur as data is moved from a source to a target. In a virtual environment where resources are shared, this can be even more disruptive since the backup of one VM on a host takes away resources from all other VMs on the host. This can be partially offset by using technology such as the vStorage APIs that allow for direct access to the VM disk so it can be backed up without involving the guest OS. However, there is still resource overhead with backups as the source data must be read which generates disk I/O.

The challenge with backups becomes how ensure the shortest possible backup window and how to perform backups of all your VMs within that window. One way to accomplish this is to perform your backups in parallel. This allows multiple VM backups to occur simultaneously instead of serially where backups wait for one VM to be backed up before the next VM backup begins. When you start performing multiple backups at once, you put more stress on the backup server as well as more stress on the datastores that the VMs reside on. Being able to achieve shorter backup windows is dependent on having an architecture that can support the heavy resource demands that occur during backups. If a bottleneck occurs anywhere in journey from source to target, it will limit your backup speeds, the number of backups you can perform simultaneously and your backup window size.

As a result, it's important to make sure you architect your backup solution properly to meet your backup window requirements. One of the key factors is your backup server (physical or virtual)

which do more than just move data from a source to a target; there is also a lot of advanced functionality that occurs as the backup server determines which data to backup, what data is duplicate as well as data compression levels needed.. It is crucial that your backup server have enough resources to handle high network I/O, increased CPU & memory usage while also ensuring that data is de-duped and compressed during the backup. As you are determining your requirements for a backup solution, make sure the solution will scale to meet your backup window requirements and ensures that your backup server does not become a bottleneck.

Application & Operating System Compatibility

Having a backup system that will support all the applications and operating systems that run within your VMs is critical to guarantee your data is safe and can be properly restored when needed. Backups in VMware environments occur at the virtualization layer using image- level backups of virtual disks and do not involve the guest OS. Because of this factor, almost any guest OS that is running inside the VM can be backed up. The challenge with this approach only surfaces when c you need to restore individual files. To do this, your backup solution needs to understand the guest OS file system in order to restore files from the image-level backups. So whatever backup solution you choose needs to fully support the various operating systems that you plan to run on your VMs.

Veeam Backup & Replication has a unique approach for supporting many different guest OS types. Because the backup server is running Windows, it can natively write to any Windows OS. For VMs that are running any Linux variations, the backup server relies on a small virtual appliance. This appliance can write to a Linux file system used as a proxy in order to restore files to a VM running Linux. Finally, for the other supported guest OS's like BSD, Solaris and Macintosh, the backup server relies on a helper VM. This VM runs the same operating system as the server that files are being restored to and is able to read the file system of the image-level backup.

For applications running inside of the VM, you need to perform application level recovery of objects. Therefore it's important to quiesce applications before they are backed up to ensure that the backups are application consistent and the data can be properly restored. Most backup solutions in VMware environments can quiesce the OS and applications by leveraging VMware Tools that can interact with the operating system. If application consistent backups are important to you, make sure that your backup solution has this capability. Additionally, you should ensure that your backup solution can easily restore application level objects, which can often be a challenge with image-level backups. Veeam Backup & Replication contains a special feature called Universal Application-Item Recovery (U-AIR) that allows you to easily restore individual objects for popular applications like Microsoft Active Directory, Exchange and SQL Server.

RTO's & RPO's

How quickly you can restore your backups and how much data you can afford to lose are two critical drivers for determining the backup solution that you choose. In the backup world, these two key metrics are referred to as Recovery Time Objective (RTO) and Recovery Point Objective (RPO). RTO & RPO are defined as follows:

- **Recovery Time Objective** — This is the amount of time allowable that it takes to recover from a data loss event. For example: with an RTO of 60 minutes you are typically relying on a replicated system to recover from. With an RTO of 24-48 hours you are typically recovering from a disk or tape backup. An RTO of 0 means you need continuous availability which usually means you have a clustered solution or synchronous replication in place to avoid any downtime.
- **Recovery Point Objective** — This is the amount of data that you can afford to lose which is derived from the last backup time or replication cycle. For example: if you do a nightly backup

that completes at 3:00am that is your last recovery point. If a system experience data loss due t a failure at 5:00pm then you have lost 14 hours of data, the maximum recovery point in this scenario would be 24 hours. If you are using replication that cycles every 30 minutes then your recovery point would only be 30 minutes as you could fall back to the last replication cycle to recover from.

RTO's & RPO's dictate the acceptable data loss and amount of time you can be down in your environment. The smaller you try and make your RTO & RPO the more expensive it gets since you have to rely on larger, faster backup and replication products that can provide very quick or real time data protection.

Veeam Backup & Replication features both instant VM and file level recovery so you can quickly recover either a whole VM or individual files and achieve RTO's as low as 5 minutes. This provides an affordable way to get you back up and running quickly without having to use costly DR/BC solutions. Additionally, the replication built into Veeam Backup & Replication can help you achieve RPO's as low as 5 minutes by leveraging vSphere's Changed Block Tracking (CBT) feature which allows for very quick replication cycles and provides near-Continuous Data Protection (CDP) for your virtual machines.

BACKUP OPTIONS AND FEATURES

Every backup solution provides basic backup functionality. However, what many buyers are looking for though are the features and options that a backup solution includes that can provide additional value and more efficient and manageable backups. When looking for a backup solution it's important to ensure the base functionality works properly and performs well, but it's also important to compare the features from one solution to the next so you can maximize your investment. Just like shopping for a car, every car will get you from point A to point B, but the difference will be how efficiently and quickly it can get you there as well as how functional the car is. A similar comparison can be made for backup solutions. Let's look at some of the features and options you will see when comparing backup solutions.

DISK VS. TAPE TARGETS

Tape has traditionally been the backup target of choice for most enterprises, but using disk targets has become more and more popular due to some advantages that it offers over tape. Disk to disk backup has become especially popular for virtualization as the unique architecture of virtualization makes disk targets an attractive choice. Disk targets work by creating a large repository on a file system somewhere that holds the backup files that are created by the backup server. The target can be any type of disk such as local storage, shared storage or even a de-duplication appliance. The backup server accesses the target using a file sharing protocol such as CIFS, FTP or NFS and all backup data is simply copied to the target during the backup. This is similar to how the Virtual Tape Library (VTL) works, but instead of a VTL making disk storage look like a tape drive, data is written natively to the disk target.

Disk has some advantages over tape that has made their use as a backup storage target a popular choice for virtual environments.

- **Cost** — Tape has traditionally been cheaper than disk especially for long term retention periods, but with drops in storage costs using cheap SATA arrays, disk has become a more cost attractive option. Simple disk storage arrays are also typically cheaper than implementing tape libraries with multiple drives.
- **Replication** — Having a disk target means you can easily replicate your backup data to another storage system on-site or off-site giving you an additional copy of your available data.

Replication is typically done using storage array features so backup data is quickly and seamlessly replicated between storage arrays.

- **Off-site** — Disk targets make sending data off-site much easier as you can simply replicate or copy backup data from one storage array to another. With disk, the process of getting data off-site can be automated and data can be easily and quickly recalled if needed.
- **Recovery** — When it comes time to recover a whole VM or individual files, performing a restore from a disk target is much quicker and easier than tape. Disk is always accessible when you need it, yet with tape you have to find tapes (which are often off-site in many cases) and get them loaded resulting in delayed availability of the recovery data.
- **Simplicity** – Disk just works. It requires little or no maintenance, while with tape you have to constantly deal with maintaining tape drives and keeping track of hundreds or thousands of tapes.
- **RTO & RPO** – If you want the lowest RTO's & RPO's, disk is the way to go. Recovery is much quicker from disk than from tape and ensures low RTO's and RPO's.

Despite the many advantages that disk has over tape, there are still some drawbacks to using disk for a backup target. Perhaps the biggest is with long term retention. Many organizations are holding on to data for 7-10 years, and trying to do this on disk can require a massive amount of space on storage arrays. For this reason, many companies are implementing both disk and tape backup solutions that work together and provide benefits offered in both media types. Disk can be used for shorter term and on-site storage of backups and tape can be used for longer term off-site storage. With this method, the backup repositories that reside on disk storage are simply backed up to tape and then removed from disk. This ensures that you don't need to keep growing your disk storage and can instead move older backup data to tape.

vStorage API Support

The vStorage APIs were introduced in vSphere 4 to allow 3rd party applications to integrate easily with the storage related functions of vSphere. They were developed to replace VMware Consolidated Backup (VCB) and allow backup applications to directly interface with vSphere. They are grouped into categories with the vStorage APIs for Data Protection (VADP) being the most beneficial to backup and recovery applications. The most notable feature in the VADP is the Changed Block Tracking (CBT) feature which allows applications to quickly lookup which disk blocks of a VMs virtual disk have changed from a specific point in time. This is important because normally applications would have to figure this out on their own which can take some time. Being able to get this information instantly greatly speeds up incremental backup and replication operations. This results in shorter backup windows and the ability to achieve near-CDP without buying expensive storage hardware.

There are other features in the vStorage APIs that benefit backup operations such as the ability to hot add a disk from a target VM to a source VM running a backup application to allow a backup without going over the network in order to read the virtual disk. This allows for quicker backups and reduced network utilization on the host server. The vStorage APIs are a huge benefit to all storage related functions in vSphere and to use applications that don't leverage them is very inefficient. Veeam Backup & Replication was one of the first backup applications to embrace the vStorage APIs and make full use of their efficiencies. If you want the most efficient backups possible, make sure your backup solution leverages the vStorage APIs.

Deduplication and Compression

Over time your backup repositories can grow quite large. Before you know it you're running out of disk space to store your backups. When this happens, your only options are to purchase more storage or to limit the number of backups that you store in your repository. Sinceneither solution is desirable, you should use technologies that help reduce the size of the data that is stored on your backup repositories. Using data deduplication and compression can greatly reduce this storage need. Data deduplication eliminates duplicate blocks of data from being stored in the backup repository and data compression shrinks the data in your backup repository so it takes up less space. Another benefit of de-duplication is that empty disk blocks—that have been allocated to a VM, but have not yet been written to by the guest operating system—are ignored and not backed up.

There are different methods for doing data deduplication and not all backup products support it or include it natively. One of the most common methods is in-line deduplication it is done in real-time where hash calculations are done before blocks are stored in the backup repository. If they match a disk block already stored, they simply reference that one and are not stored again. While this is beneficial to reducing backup sizes, it can cause some extra overhead while backups are running as hash calculations have to be made on each disk block. Veeam Backup & Replication allows you to choose from multiple de-duplication options that use different block sizes when calculating hashes. With Veeam, you can choose whether you want maximum de-duplication at the expense of slower backups or minimum de-duplication for best backup job performance. Veeam also supports multiple compression levels that will vary the amount of compression that is done to meet the needs of the environment. Compression is very CPU intensive and can increase backup times so having at least 8 CPU cores on the backup server is recommended for maximum compression.

Replication

Traditionally, if you wanted to replicate data, you had to rely on expensive storage hardware and add-on replication software to accomplish it. This would handle the replication of the data at the storage layer from one storage system to another and is readily available for use as a backup if needed. The virtualization architecture allows for alternative methods to be used to achieve this. Because VM's are encapsulated into virtual disk files, the replication can also been done at the virtualization layer that sits in between the hardware layer and operating system layer. While doing replication at the virtualization layer may not be as efficient as letting the storage array handle it, it can still get the job done in an effective manner.

Replication of a VM is similar to backing it up. The primary difference between backup and replication is that backups are typically a daily event while replication is a continuous event. Replication is much like doing incremental backups on a very frequent basis, because of this replication is often bundled within backup products. Veeam Backup & Replication has had the ability to replicate VMs from a source host to a target host since its inception. And with the Changed Block Tracking feature in vSphere, incremental replication operations are much faster and as a result, achieving near CDP at the virtualization layer can now become a reality.

Backup Verification

There is nothing worse than backing up servers for months and months only to find out the data you have been backing up is no good once it is restored. The whole point of performing backups is for that one day that you actually have to rely on them to restore data when disaster strikes. If you are unable to restore data properly, why even back it up in the first place? The process of verifying backup data is more than just doing data verification on your backup media. This type of verification only confirms that disk blocks are properly written to the target device. If there is

something wrong with the source data, that will copy over to the target backup device as well. These types of issues can range from an operating system or applicationthat is not properly quiesced at backup and - to having missing or corrupt critical files on a server. Therefore, the only way to ensure your backups are working properly is to restore the data to a test server and make sure everything works.

Virtualization makes this verification process much easier. You no longer need spare physical hardware to restore a whole server since VM's can easily be restored hosts with spare capacity. This can still be a time-consuming process if done on a regular basis. Veeam recognized this challenge and developed a simple and automated part of the backup process that eases the burden of verifying the recoverability of backups. Veeam's SureBackup technology in Veeam Backup & Replication automatically verifies the recoverability of every backup. This capability is made possible by powering up the backed up VM directly from the backup repository, checking for a heartbeat from VMware Tools and pinging responses. Additionally, test scripts can be run to verify that applications are running properly and data is accessible. Backups are useless if you cannot recover from them, so having the peace of mind that you have proper backups is crucial when you get in a situation where critical data must be restored.

Empowering your Backups

Backups are much like an insurance policy; they continually cost you money and offer the security that you need but you don't get anything out of them unless you have an emergency. With virtualization, it is common to do disk to disk backups and optionally sweep them to tape as well. The backups of your VMs just sit around on your target disk repositories taking up valuable disk space and resources and are completely ignored. But since they reside on disk, in reality you have usable historical copies of your virtual machines available that could be used for various purposes. Imagine if you needed a quick sandbox or virtual lab to test an application upgrade or an isolated environment to do some testing or troubleshooting. Those VM backup copies are perfect candidates for this kind of operation especially if you isolated them on their own virtual network where you could do anything you wanted without disturbing the production environment.

Veeam has made this possible with their On-Demand Sandbox feature available with vPower technology. With vPower the backup server becomes an NFS server andthe backup repositories act as the storage devices. Any ESX/ESXi host can then connect to it using NFS and access the VM backups that are in the repository. The backup images are read only and any changes made to them while they are powered on are written to a separate delta file which is discarded afterwards. VMs powered on from the repository are kept isolated from the rest of the network using vSwitches and have no physical NICs assigned to them. In addition, a special routing appliance allows access to outside networks. Being able to actually make use of your backups for purposes other than the occasional restore allows you to make the most of your backup investment.

HOW TO CHOOSE THE RIGHT BACKUP SOLUTION

With all the options and choices you will face when shopping for a backup solution for your VMware environment how do you decide which product or products are best for you? You should start by mapping out your requirements, decide on which features are a must have and which are a nice to have. Budgets will typically determine what you can afford to purchase; you want to get the most for your money so be sure and shop around and compare features. Backups are not just about software, you will need to factor in hardware as well, andIf you choose to use disk targets, make sure you factor in enough storage to meet your retention needs. You need to choose a backup solution that is compatible with your virtual environment. If you are running a mixed environment that has mixed versions of VI3 & vSphere 4/5 hosts; you need a solution that supports that.

Evaluating products

Once you sort through the products and identify those that meet your needs, the next step is to put them to the test. This is probably the most important part of buying a backup solution; don't blindly buy a backup application without first trying it out in your environment. What the vendor tells you and how the product actually works in your environment may be completely different. Every environment is unique and things like backup speeds that are typically benchmarked by vendors are done in their controlled environments and may not reflect real life workloads and scenarios.

Almost all vendors will provide you with an evaluation copy of their software. By evaluating backup applications you can see firsthand how things perform in your environment and test the various features to see their actual benefits. You may not be able to evaluate backup hardware, but if you plan on using disk targets you should be able to use any disk in your environment to evaluate the backup software. When evaluating the backup software, make sure you do realistic testing, and don't just backup one VM to see if it works. You should also do a large scale backup so you can get an accurate depiction of the backup software performance. Doing so will help you when you need to size the storage to use for your disk target. Further, be sure to take notes when evaluating critical products such as backup software. You need a way to compare performance statistics and features between products. Also don't be afraid to ask a vendor for help if you get stuck or have questions, this will give you a feel for how good their support is and also make sure you are doing things properly so you get the best results. Remember: you're not just testing the software, but rather the entire vendor package. Many vendors have written guides that can help you evaluate their products. Veeam has detailed evaluation guide available for download on their website to help you evaluate Veeam Backup & Replication.

Comparison criteria

Below is a checklist that you can use that lists some of the most common features and support options that you will find in VMware backup applications. Be sure and compare each vendor's offerings to find out what they do and do not support. The checklist has been pre-filled out for the capabilities of Veeam Backup & Replication and you can use this to compare it with other products.

Feature	Veeam Backup & Replication	Product 2	Product 3	Notes
Support for ESX	■	☐	☐	
Support for ESXi	■	☐	☐	
Support for VI3 (which versions)	■	☐	☐	
Support for vSphere 4	■	☐	☐	
Support for vSphere 5	■	☐	☐	
Support for Microsoft Hyper-V	■	☐	☐	
Relies on Service Console Agents	☐	☐	☐	
Support for vStorage APIs for Data Protection	■	☐	☐	
Support for Changed Block Tracking	■	☐	☐	
Support for SCSI Hot-Add	■	☐	☐	
Support for Legacy Backup Modes (VCB & Network)	■	☐	☐	
Support for Backup Server on physical hardware	■	☐	☐	
Support for Backup Server as virtual machine	■	☐	☐	
Support for Multiple Simultaneous Backups	■	☐	☐	
Support for Virtual Applications (vApps)	■	☐	☐	
Support for Thin Provisioned Disks	■	☐	☐	
Support for Direct To Target	■	☐	☐	
Support for Disk Targets	■	☐	☐	
CIFS	■	☐	☐	
NFS (via Linux mount)	■	☐	☐	
FTP	■	☐	☐	
VMFS	■	☐	☐	
USB Hard Drives	■	☐	☐	
Support for Tape Targets	☐	☐	☐	
Support for Physical Server Backups	☐	☐	☐	
Support for Real-time Backup Reporting	■	☐	☐	
Support for Email Notifications	■	☐	☐	
Support for SNMP Notifications	■	☐	☐	
Support for pre-Backup tasks	■	☐	☐	

Feature	Veeam Backup & Replication	Product 2	Product 3	Notes
Support for post-Backup tasks	■	☐	☐	
Support for Incremental Backup Modes	■	☐	☐	
Support for Synthetic Backups	■	☐	☐	
Support for VM Replication	■	☐	☐	
Support for Replication Failover	■	☐	☐	
Support for VSS & Application Quiescing	■	☐	☐	
VMware Tools	■	☐	☐	
Proprietary Processing	■	☐	☐	
Support for Transaction Logs Truncation	■	☐	☐	
Support for Deduplication	■	☐	☐	
Inline	■	☐	☐	
Post-Process	☐	☐	☐	
Customizable Block Size	■	☐	☐	
Support for Compression	■	☐	☐	
Customizable Compression Level	■	☐	☐	
Support for vCenter Server Integration	☐	☐	☐	
Support for Scripting & PowerShell	■	☐	☐	
Support for File Level Restores	■	☐	☐	
Windows	■	☐	☐	
Linux	■	☐	☐	
Other OS's	■	☐	☐	
Support for Guest File System Indexing	■	☐	☐	
Support for Application Item Recovery	■	☐	☐	
Support for Full Image Restores	■	☐	☐	
Support for Instant VM Recovery	■	☐	☐	
Support for Instant File Recovery	■	☐	☐	
Support for 1-Click File Restore	■	☐	☐	
Support for Virtual Labs	■	☐	☐	
Support for Automatic Backup Verification	■	☐	☐	
Support for Near CDP	■	☐	☐	

Don't forget about DR

Any backup solution you implement is also going to be an integral part of your DR/BC solution and you should factor that in when deciding on a backup solution. Even if you do not have a DR solution today, you may in the future so be sure and plan ahead to ensure your backup solution has the required features and capabilities that can support your DR strategy. Replication products are often used to keep a DR site in sync with the main site, but storage replication can be very costly and complicated to implement. Virtualization offers a simpler and lower cost solution by replicating VM's at the virtualization layer. This provides you with the freedom of using dissimilar storage devices at the primary site and DR site. Even local storage can be used which can reduce costs by avoiding the additional purchase of an expensive shared storage device. Virtualization allows for creative solutions for implementing DR and your backup product can be the enabler for implementing any solution that you choose.

The Veeam Advantage

Veeam's revolutionary vPower addresses some of the pain points that are typically associated with backup and recovery by introducing some innovative new features that fully leverage the strengths that virtualization architecture provides. vPower extends the capabilities of VMware environments and takes full advantage of the flexibility and features that virtualization provides to take backup & recovery in virtual environments to the next level.

vPower enhances backup & recovery, but it can be also be used for a variety of other purposes as well such as troubleshooting issues, performing tests, creating development sandboxes and much more. Backups are like insurance policies, you have to have them but they normally just sit around taking up valuable disk space until the occasional restore is needed. vPower allows you to actually make use of your backups for more than just recovery which gives you much better return on the investment you make in your backup infrastructure.

SUMMARY

Backup software isn't glamorous but it is a critical part of your infrastructure. Don't rush into a decision when choosing a backup solution, take your time, ask a lot of questions, try products out and then make an informed decision. A product that may look good on paper may not be so good when you actually start using it. You wouldn't buy a car without test driving it and you shouldn't buy a backup solution without testing it first. Don't be afraid to ask for references and do your own research as well to see what existing customers think about a particular product. Remember: you are entrusting critical data to whatever backup solution that you choose so this is one decision that you can't afford to get wrong. By doing your homework you can help ensure that the product you choose is the right one for your VMware environment and guarantee that your data is properly protected.

ABOUT THE AUTHOR

Eric Siebert is an IT industry veteran, author and blogger with more than 25 years of experience, most recently specializing in server administration and virtualization. He is a very active member of the VMware VMTN support forums, where he's attained the elite Guru status by helping others with their virtualization-related challenges.

Siebert has published books including his most recent, "Maximum vSphere" from Pearson Publishing, and has authored training videos in the Train Signal series. He also maintains his own VMware VI3 information website, vSphere-land, and is a regular blogger and feature article contributor on TechTarget's SearchServerVirtualization and SearchVMware websites. Siebert has presented at VMworld in 2008 & 2010 and has been recognized as a vExpert by VMware in 2009 & 2010.

ABOUT VEEAM SOFTWARE

Veeam® Software develops innovative solutions for VMware backup, Hyper-V backup, and virtualization management. Veeam Backup & Replication™ is the #1 VM Backup solution. Veeam ONE™ is a single solution for real-time monitoring, resource optimization, documentation and management reporting for VMware and Hyper-V. Veeam extends deep VMware monitoring to Microsoft System Center with Veeam Management Pack™ (MP) and to HP Operations Manager with Veeam Smart Plug-In™ (SPI). Veeam also provides free virtualization tools. Learn more by visiting www.veeam.com.