

# VMware Insights

## Backup and Recovery Implications

*What they they don't tell you about VMware backup*

Presented by  
Mosaic Technology



*VMware Insights: Backup and Recovery Implications**Table of Contents*

<b>Introduction</b>	<b>3</b>
<b>ESX Server and Virtual Machine Backup Basics</b>	<b>4</b>
<b>VM Backup Methodology Alternatives</b>	<b>5</b>
1. <i>Backup like Physical Machines</i>	5
2. <i>Backing up VMs as Files on ESX Server Host</i>	6
3. <i>Backing up VMs as Files on Shared Storage</i>	6
<b>VMware Consolidated Backup (VCB) -- Overview</b>	<b>7</b>
<b>Taking Advantage of VCB</b>	<b>9</b>
<i>Vizioncore v Ranger Example</i>	10
<b>Summary</b>	<b>11</b>
<b>About Mosaic Technology</b>	<b>12</b>

---

## ***Introduction***

Virtualization gives companies new ways to deploy and manage application architectures. VMware lets you take management and virtualization to the next level with their latest release of VMware Infrastructure 3.

While most people initially look at VMware from a server consolidation perspective you soon discover that virtualization can alter the way you view your entire infrastructure – from apps to network to storage and disaster recovery.

The focus of this paper is what impact virtualization has on backup and recovery.

Virtualization does allow higher utilization of physical servers. One physical “box” can host up to hundreds of virtual machines and their associated applications and data. This means protecting your data and applications in a virtual world is as important – if not more so – than in the physical arena. A comprehensive backup and recovery strategy is a necessity for a virtualized environment.

VMware recognized this and made backups more manageable in VMware Infrastructure 3 with VMware Consolidated Backup (VCB). Interestingly VMware acknowledges the difficulty of backing up Virtual Machines – even when using VCB.

VCB is just the entry point to a sound VM backup and recovery strategy. You need to look beyond it and evaluate what other components you need to achieve real backup reliability.

---

## **ESX Server and Virtual Machine Backup Basics**

When you backup an ESX Server system, you need to think in terms of what needs to be addressed for recovery. With ESX Server the major components worth considering for backup are:

- Virtual disks
- Virtual machine configuration files
- The configuration of the ESX Server system itself

For virtual machines, all information normally backed up in the enterprise infrastructure, including operating system, applications, and data, is included in the virtual disks.

Virtual machines are encapsulated as a file on a host ESX Server and can be backed up as a single entity for restore purposes. Backup strategies designed for the physical world focus on backing up data and cannot take advantage of this aspect of virtual machines.

If a company emulates a “physical world” backup strategy and installs backup agents inside the guest virtual machine or on the ESX Server host, they would need to restore the host OS, applications, configuration settings, etc. first and then restore the backup data files themselves when in recovery mode.

Some additional challenges exist with backup methods that rely simply on using agents inside the guest or in the ESX Server host. They include:

- May not restore entire image of the virtual machine.
- Heavy overhead on the host ESX Server with either the Service Console or the virtual machines themselves
- Heavy network overhead
- Limited backup windows
- Need to heavily manage scheduling of backups
- High cost of disaster recovery

Many companies assume that VM backups can be performed by existing physical platform backup solutions — typically agent software.

On a virtual machine this agent is placed inside the guest OS where it consumes resources on the host ESX Server, guest OS, and network to run a backup job. Those backups also need to be scheduled appropriately. If backups run for

### **Deduplication and VMs**

VMs are by their nature very large files. VM backups can consume substantial disk space. Much of a VM is unchanged between backups. These two factors make VM backups ideally suited to deduplication solutions --which can give you 20x to 100x more logical capacity that turns 10TB of addressable capacity into 200TB to 1PB of logical. Data Domain and Exagrid are two examples of this technology.

extended periods they can generate large amounts of network overhead and problems can ensue. It's also not practical to run full backup images on a daily basis -- especially on more static virtual machines.

A fundamental value of virtual machines is that they are easily moved to different hosts with tools such as VMotion. Backup applications need to "follow" virtual machines through the infrastructure and continue to execute within a particular backup window. Maintenance updates of VM backup applications must be carefully planned to avoid inconsistent data, poor performance or catastrophic results on the ESX Server.

## ***VM Backup Methodology Alternatives***

Because of their nature you can manage Virtual Machine backup in a variety of ways. Here we consider three options: backup like a physical sever, backup as files on a virtual host, and backup VM as files on shared storage. Each method has plusses and minuses.

### ***1. Backup like Physical Machines***

A virtual machine is just like a physical machine. You can back it like a physical machine, using backup software running inside a virtual machine. This allows for traditional incremental and differential backups. There are some disadvantages to this approach.

Virtual machines provide complete guest operating systems on virtualized hardware. You can back up these operating system installations in the same ways as their physical counterparts.

- You can install a backup agent within each virtual machine and back up data over the network to other backup servers.
- You can copy data manually or with a script to another machine -- this is exactly like backing up a physical machine.
- You can -- in very specific cases -- attach SCSI tape to a virtual machine and run a media server within that virtual machine.

Using this approach you can use the same methodology to back up all servers in your data center -- physical and virtual. This approach also lets you to do file-level backups and restores and gives you more flexibility in your choice of backup software. This method also makes incremental and differential backups easier.

***Disadvantages:*** You cannot take advantage of the encapsulation of virtual machines into one or more discrete files. When you back up individual files within the guest operating system as files, you do not have the advantages of backing up

---

and restoring the virtual machine as a whole. Also, while backup loads of five percent may be acceptable on a physical host, when you multiply that by a number of virtual machines the backup load may be prohibitive.

## ***2. Backing up VMs as Files on ESX Server Host***

Using this approach lets you take advantage of the service console's ability to see each virtual machine's virtual disk as a file. ESX Server creates one .vmdk file per virtual disk. These files can be backed up -- which essentially protects an entire virtual hard drive in a single pass. To use this method while a virtual machine is running, you must use snapshots to back up the virtual machines, which use redo log files for writes while the backup is taking place. These redo log files are later committed and changes written to the .vmdk files.

A significant benefit of this approach is that it lets you back up or restore an entire virtualized server in one step. And with disk snapshots and redo log manipulation, you can do near-line backups.

***Disadvantages:*** Although this method is much simpler than traditional file-level backup, you need to restore the entire virtual machine even if you need to recover a single file. If you have extremely large virtual machines, you may need to restore tens of gigabytes of data to restore one needed file. And because this backup process treats the disk as a whole and is not application aware, the backups created through this process are only file-system-consistent. Additionally, the need to access large virtual disk files on VMFS that may be larger than 2GB — may limit your choice of backup software or require additional intermediate processing.

## ***3. Backing up VMs as Files on Shared Storage***

When virtual machine files reside on shared storage, you can use storage-based imaging on storage such as SAN, NAS, or iSCSI, or an independent backup server (a proxy backup server or NDMP) to back up virtual machine files without creating an additional load on the ESX Server host that runs the virtual machines.

---

## **VMware Consolidated Backup (VCB) -- Overview**

VCB is a platform that facilitates backups of virtual machines. It provides a way to get to the virtual machine's disks using a fibre adapter in a Windows Server 2003 machine. Because of that, VCB needs to be installed on a physical machine running Windows Server 2003.

VCB requires extra licensing (and cost) and is included in the Enterprise edition of VI3. You may also need some scripts (depending on backup software) that allow you to integrate VCB into your backup software. When you create a backup job, you run some pre and post-backup scripts to get the data onto the VCB server and to clean up afterwards.

The big advantage of VCB is that it allows you to take a LAN-free backup of your virtual machines with no load on the ESX servers. Of course, it is only LAN free between the VCB and ESX server.

There are basically two types of VCB-based backups:

- Full vm: each virtual machine file including all disks (vmdk files) are dumped to the VCB server so that your backup software can back them up
- File mount: the virtual machine's volumes are mounted on the VCB server so that your backup software can get to the files in each volume

For a full vm backup, you need enough space on the VCB server to dump the complete contents of the virtual machine. For a file mount, you do not need any additional space on the proxy because nothing is copied to it. The virtual machine's disks are merely mounted on a Windows folder so they are there for the taking by your backup software.

VCB does a variety of things well. It:

- takes the backup load off the ESX Server host
- eliminates the backup window
- removes backup traffic from the LAN
- eliminates the need to run backup agents inside the virtual machines to perform file-level backups of virtual machine data.

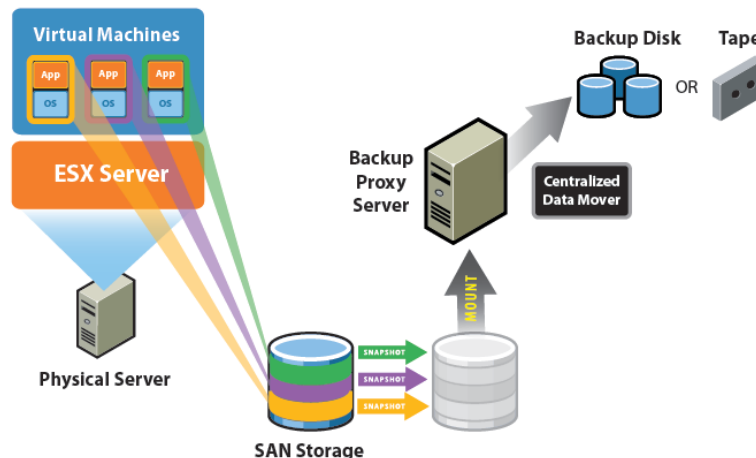
The way VCB manages backups is analogous to removing the disk from a running computer, attaching it to a dedicated backup server, and backing up appropriate files from the disk — while the original computer still sees the disk and continues to run normally.

---

In this process the virtual disk can be identified by the same drive letter when it is attached to the backup server — the backup proxy — as it is when it is attached to the virtual machine. This preserves the drive letter identification a backup agent needs if it becomes necessary to restore files to the virtual disk. Most advanced clients from backup vendors also incorporate the concept of off-host backups and can accommodate that configuration easily.

VCB uses functions in VMware Tools to quiesce file systems inside virtual machines. This ensures that when a snapshot is taken, all pending data changes have been written to disk so the snapshot contains consistent data. VCB also makes it easy to run scripts before and after a backup. This lets you freeze and quiesce applications and then restart them later.

Once a snapshot is taken, a separate physical machine — the backup proxy — mounts the base disk as if it were a locally attached file system. This allows a backup agent running on the proxy to read and back up files using the same features the agent uses when backing up physical drives, as shown in the figure below.



A central strength of VCB is that it removes the load of performing backups from the ESX Server system and places it on a dedicated computer -- the backup proxy. This effectively eliminates the need for a backup window.

With VCB you can use the same software and similar methodology to back up all the servers in your data center, whether those servers are physical or virtual. VCB also allows you to do file-level backups and restores, and it gives more flexibility in the choice of backup software. This method also supports incremental and differential backups. Backup software that has backup modules specific to your applications can be used to quiesce the applications when you prepare for the backup.

---

VCB is aware of the VMware Distributed Resource Scheduler and VMware High Availability features in VMware Infrastructure. It can use a SAN network for backup rather than the main data network. In short, for most use cases Consolidated Backup provides an efficient and consistent way to back up virtual machines.

Most VMware Infrastructure users will adopt Consolidated Backup. VMware is continuing further development, and increasing numbers of vendors readying their products to work with this technology.

**Disadvantages:** If you choose to back up individual files within the guest operating system as files, you do not have the advantages of backing up and restoring the virtual machine as a whole. Also note that to restore files backed up with Consolidated Backup, you must have a backup agent installed in the virtual machine and restore the files from within the running virtual machine.

While VCB brings innovations to protecting virtual machines, it may not be suitable in all situations.

- In VMware Infrastructure 3, VCB does not support file-level backup of guest operating systems other than Windows. Only image backups are supported. There is no incremental backup capability for systems other than Windows.
- If you need to perform frequent restores, VCB does not save you any cycles, because restores can be a two-step process, especially when you need to restore partial sets of files.
- Although few users of VMware Infrastructure face this problem, Consolidated Backup does not work with Windows snapshot tools such as VSS.

## ***Taking Advantage of VCB***

VCB is designed and implemented by VMware. It provides a much needed link to VMware products. It offers immediate out of the box value – but as the previous sections showed it is not a complete solution – you still need backup software.

The good news is that there are solutions now available that integrate with and take advantage of VCB.

As mentioned earlier some scripting may be involved with your backup software. However, you can also get advanced functionality and instant integration with products such as Vizioncore's vRanger backup and restore software.

---

### ***Vizioncore v Ranger Example***

Vizioncore's vRanger, is one of an emerging group of Backup and Restore solutions for VMware Virtual Infrastructure. vRanger was designed from the start for VMware. It offers agentless, image-level backups taken from outside the guest, without disrupting a running VM. vRanger provides VMware-specific functions such as:

- file-level restores from image-level backups;
- ability to skip VMDKs (virtual hard drives) for a given VM
- an advanced VCB integration with network failsafe option

Backups are most efficiently managed when there are no resources used on the host ESX Server, virtual machines or the network. VCB was designed to allow backup solutions such as vRanger Pro to provide a complete and cost effective backup and recovery solution.

VCB removes the backup process from the host ESX Server altogether by leveraging a proxy server to execute the backup functions performed by vRanger Pro. This combined solution addresses the challenges outlined previously.

1. The entire virtual machine/or individual files can be restored
    - a. A backup image is stored and accessible for restore needs
    - b. vRanger Pro will restore either individual files or the full virtual machine image.
  2. Overhead on the host ESX Server with either the Service Console or the virtual machines themselves is eliminated.
    - a. VCB snapshots the virtual machine and offers it to a secondary (proxy) server for backup purposes which offloads the overhead from the host.
    - b. vRanger Pro can then perform the backup of the .vmdk.
  3. Network overhead is eliminated
    - a. VCB leverages SAN architecture including fibre attached storage.
    - b. vRanger Pro will launch the VCB process and execute the backup using the fibre channel, not the LAN/WAN.
  4. Backups are less restricted to backup window constraints
    - a. Because VCB makes the snapshot available on a proxy server, fewer resources are consumed on the host or network, allowing backups to be done without concern for a specified resource window.
    - b. vRanger Pro compresses files to execute backups with increased speed. It also manages resources for backup windows, allowing backups to be done at any time.
  5. Scheduling of backups is managed through vRanger Pro
    - a. vRanger Pro integrates with VMware VirtualCenter, enabling hotbackups to be performed any time of day, by host, group, and individual virtual machine or by VirtualCenter attribute, to include new virtual machines added for scheduling purposes.
  6. The cost of implementing an effective disaster recovery strategy is reduced.
-

- a. With vRanger Pro and VCB, agents are no longer needed on every virtual machine for backup.
- b. Identical hardware infrastructure for recovery is not needed, as vRanger Pro includes the configuration files (.vmx) needed to rebuild the virtual machine's underlying hardware and can restore the virtual machines to dissimilar hardware.

vRanger also has the advantage of full integration with VCB with a simple check box during installation. And it integrates easily with your existing backup apps – so there's no need to rip and replace what you're already familiar with.

## **Summary**

VMware and virtualization changes the way we look at our infrastructure. Initially attractive due to significant ROI with server consolidation, VMware forces us to evaluate almost every aspect of our IT infrastructure.

One aspect that's frequently an afterthought is backup and recovery. VMware and Virtual Machines present an interesting backup/recovery dynamic. VMs are frequently very large files, the files contain everything from OS, to Apps, to data, they can be quite mobile (thanks to Vmotion) within the physical realm. From a recovery perspective you can either backup entire VMs or do a file level backup – each with its limitations.

You can backup virtual machines using your existing backup solution – but that may cause performance and/or storage problems.

VMware addresses some of these issues with VMware Consolidated Backup (VCB) – a platform that facilitates VM backups. However, it is a facilitator and not a complete solution.

In order to get a robust and reliable backup solution you need to design a system that lets you recover VMs either as whole entities or as individual files. This generally requires additional software (e.g. vRanger) and hardware (e.g. Data Domain, Exagrid).

---



## ***About Mosaic Technology***

Mosaic Technology provides IT infrastructure solutions for companies worldwide through three divisions: Value Add Solutions, IT Asset Management and Recovery, and Used and Refurbished IT Hardware.

Based in Salem, NH, Mosaic has sales and technical offices throughout the United States and associates worldwide. Mosaic provides virtualization, data storage, and management, and email solutions from companies such as VMware, EMC, Hitachi, Exagrid, Data Domain, CommVault, Dell, EqualLogic, Sun, Cisco, and Symantec.