Managing the information that drives the enterprise

STORAGE

ESSENTIAL GUIDE TO

Backup and Virtual Servers

Here's everything you need to know about protecting virtual servers.

INSIDE

TechTarget
The IT Media ROI Experts

Some creativity required for virtual server backup

VIRTUAL SERVER TECHNOLOGY has had an enormous impact in data centers by advancing server consolidation at a pace never seen before. Organizations of all sizes have slashed the number of physical servers they need, allowing them to trim costs, cut complexity, and reduce power and space requirements. But while they make a lot of IT chores easier, virtual servers make one key task a lot harder: backup.

Traditional backup apps tend to be oriented toward physical server environments and, as such, might not be appropriate or the most efficient means for backing up virtual machines. While stuffing many virtual servers into one physical server yields consolidation benefits, it also concentrates backup operations onto a single physical platform, leading to overtaxed processors and I/O systems.

Backing up virtual servers isn't exactly a dream job for anyone at this point, but it doesn't have to be a nightmare.

VMware and other virtual server vendors have made strides in addressing the backup issue but, as yet, none offers a complete solution. Similarly, backup app vendors have adapted their products to be better virtual world citizens, but gaps remain. And a handful of specialized apps—built from the ground up to protect data in virtual server environments—are gaining a lot of attention, but they, too, don't always offer the comprehensive data protection required in enterprise data centers.

In some cases, backing up virtual machines in a manner that allows bare-metal restores makes restoring individual files a lot harder. And the reverse may be true, too: The convenience of quick file recoveries may mean giving up full machine restores that can greatly facilitate disaster recovery.

With a perfect solution yet to arise, storage managers aren't without options—from using traditional backup apps to cobbling together combinations of all of the above. For a while, at least, it doesn't look like virtual server backup will be a set-it-and-forget-it affair, so some tinkering and adjustment of backup operations and processes will be required.

This Essential Guide offers a collection of tips, tools and techniques that you can use to back up your virtual servers more efficiently. Find out about the intricacies of using VMware's Consolidate Backup (VCB) application, get tips on how to fine-tune your current backup app for virtual servers, and read about the most common problems and solutions you're likely to encounter. ⊙

Rich Castagna (rcastagna@storagemagazine.com) is editorial director of the Storage Media Group.

THE SMARTEST STORAGE FOR VIRTUAL SERVERS

> “Compellent is the easiest storage to deploy and manage in a virtual data center.”

PETER FITCH
IT Strategic Planning and Infrastructure Manager
Rudolph Technologies

Virtually Eliminate Storage Management For Any Number of Virtual Machines

Only Compellent enables the fast and easy creation of storage for any number of virtual machines without wasting your time, money or disk space. Our SAN automates and optimizes data placement for all of your virtual applications, providing the perfect complement to virtual servers. Now you can take full advantage of the cost and energy-saving benefits of server virtualization.

www.compellent.com

Is your storage ready for server virtualization?

Read this white paper to find out.

White Paper: 3 Must Haves for the Virtual Data Center

Learn why easy provisioning, automated tiering and continuous data protection are must haves for a successful virtual data center.

compellent

Microsoft GOLD CERTIFIED Partner

2009 ADVANCED INFRASTRUCTURE SOLUTIONS
Storage Solutions
PARTNER OF THE YEAR
FINALIST

SearchServerVirtualization.com
BEST OF vmworld 2008
GOLD AWARD
HARDWARE FOR VIRTUALIZATION

InfoWorld 100

VIRTUAL BACKUP tips

There are three main ways to back up virtual servers. Here's how to determine which method is best for your storage requirements.

By Alan Radding

SERVER VIRTUALIZATION is supposed to simplify IT, and it does, except for one little caveat: It currently complicates backup and recovery. Companies quickly discover they can back up their virtual servers the same way they do their physical servers, but they may not get the same results.

"It's simple to back up virtual servers if you treat them like physical servers," said Scott Polly, director of technical publications at Vizioncore Inc. in Buffalo Grove, Ill. The firm provides vRanger Pro, a GUI-based application that automates much of the command line scripting typically involved in VMware backup. Simple, but "you don't get the benefits of server virtualization," Polly added.

The benefits Polly refers to revolve around efficient management. If you treat virtualized servers like physical servers, you have to manage each one individually, which undercuts any improvements in management efficiency. You also have to buy and install a backup agent on each virtual server as if it were a physical server, which will certainly require more work and may entail additional license fees depending on your backup product. In short, storage admins face more than the usual backup complications because there are more virtual servers and they have to run multiple concurrent backups.

In addition, the immaturity of server virtualization makes the backup challenge more difficult. VMware is still feeling its way when it comes to backup and recovery. "It's hard to get clear documentation," reported

virtual backup

BEST PRACTICES

- Provision sufficient physical computing resources (CPU, memory, storage, bandwidth)
- Plan for disk-based backup, copy to tape if required later
- Identify where you need file-level recovery and where machine-level recovery is desired
- Use snapshots and array-based replication to reduce server and network overhead
- Stagger backups to minimize possible resource contention
- Quiesce database applications before backup to avoid data consistency problems
- Take advantage of the latest tools from VMware (VMware Consolidated Backup, vCenter Site Recovery Manager)
- Adopt GUI-based tools to avoid writing extensive command line scripts

John Dolan, principal consultant at Viant Solutions in Suwanee, Ga. Dolan had been trying to pin down the backup of 24 virtualized servers spread across three hosts at Perimeter Church, a mega-church based in Duluth, Ga., but getting accurate information was frustrating. Early on, he feared he might need to write some software to make the church's version of Symantec Corp.'s Backup Exec (Version 10d) work with VMware ESX 3.0.2.

The solution turned out to be simple once he had the correct information. "Backup Exec [Version 10d] simply isn't supported by ESX 3.0.2," Dolan said. He finally discovered that after digging through VMware's latest compatibility guide. "We needed to be running Backup Exec 11d or later," he said. That meant Perimeter Church had to upgrade to the next release, which entailed an added cost. "But we'll only need one server license plus the Exchange and SQL license for our virtual servers," he added.

The VMware backup picture is changing almost weekly as VMware pushes out new documentation and products, as more vendors scramble to certify their backup tools with VMware and as consultants begin to identify some best practices (see "Virtual backup best practices," this page). VMware insists that backup isn't difficult: "It's not hard, but different," said Jon Bock, VMware's senior product marketing manager, adding, "It will change what you're probably doing now."

BACKUP AGENTS

The traditional way to back up servers is to load a backup agent on the server, set the parameters and let it run. “You can still install a backup agent on each virtual server, just as before,” VMware’s Bock said.

But backing up one server is different than backing up five, 10 or 15 virtual machines (VMs). The problem is resource contention. “Backup imposes overhead on the [physical] server,” said David Dale, chair of the SNIA Canada IP Storage Forum, a member of the SNIA board of directors and a NetApp executive.

> "In theory, it's possible that you could overload the physical machine by doing multiple backups, but we haven't hit that particular pain [point] in the real world."
>
> —Ashley D’Costa, enterprise solutions architect, Mainland Information Systems

Regardless of how many virtual servers are backed up, they’re guests of a single physical server and must share the CPU and network resources. One possible solution, Dale suggested, is to “delegate the backup overhead to the array.”

VMware’s solution is different: VMware Consolidated Backup (VCB). “VCB is a Windows-based proxy host,” said Scott Miller, president at Server Centric Consulting in St. Louis. Instead of installing individual backup agents on each VM, you do the backup on VCB, which offloads the backup process overhead to the physical proxy host, which is a dedicated server.

Some observers think the resource contention issue is a red herring. “In theory, it’s possible that you could overload the physical machine by doing multiple backups, but we haven’t hit that particular pain [point] in the real world,” noted Ashley D’Costa, enterprise solutions architect at systems integration and consulting firm Mainland Information Systems in Calgary, Alberta. The virtualization backup problems D’Costa encounters have more to do with performance and data consistency.

W.R. Berkley Corp., based in Greenwich, Conn., backs up almost 100 VMs running as LPARs on its IBM Corp. pSeries servers with Symantec’s Veritas NetBackup agent installed on each VM. “We haven’t hit any resource contention yet. I guess we could, but the boxes have a ton of ports,” said Tom Whelans, vice president of operations at the property and casualty insurer. The backup software is licensed by physical server so it doesn’t incur additional licensing charges.

BACKUP CHALLENGES

“People don’t put enough thought into the storage. They haven’t thought through the details of the number of VMs. Then there are one-off situations that may require raw data mapping,” which enables the storage admin to

specify a logical unit number (LUN) on the disk array for a given VM, said Ron Oglesby, director of virtualization and architecture services at GlassHouse Technologies Inc. in Framingham, Mass.

Also challenging are the different types of recovery: backing up and recovering the entire VM as a file, or backing up and recovering single files. To back up and recover the entire virtual machine, you "copy it block by block, disk to disk and capture the VM in a specific state," Oglesby said. "Here, you have to restore the entire VM as a whole."

To back up and recover individual files you "do a file-level backup within the VM using a normal agent or VCB, like a traditional backup. In this case, you can restore a single file easily but can't bring the entire VM server back online easily," he continued. You have to recover the VM piece by piece.

Virtualization is essentially file-oriented; the virtual machine consists of one or two encapsulated files. Restoring individual files is still a challenge. It's easy to back up and restore the entire VM. Restoring individual files isn't straightforward and, depending on the backup tools involved, may be quite cumbersome and involve restoring the entire VMware Disk Format (VMDK) file.

To back up and recover individual files you "do a file-level backup within the VM using a normal agent or VCB, like a traditional backup."

—Ron Oglesby, director of virtualization and architecture services, GlassHouse Technologies

"Everybody wants to rapidly back up at both the granular [file] and virtual machine level," Mainland Information Systems' D'Costa said. With VMware as it's configured today, that's not possible. "Right now you have to choose one or the other," he said. Granular backup means you can recover individual files without restoring and mounting the entire virtual server. With machine-level backup, you recover the entire virtual server as a single VMware encapsulated file; VMDK is fast and easy if you need to recover the whole thing.

VM BACKUP STRATEGIES

There are a number of ways to approach virtualized server backup. Your selection will depend on a variety of factors, including the backup tool vendor, the capabilities of the storage-area network (SAN) and the recommendations of the virtualization vendor, which in most cases is VMware.

D'Costa identified three main approaches to backing up virtualized servers: conventional server backup with the use of individual backup agents on each VM; use of the backup management tools provided by the virtualization vendor; and the use of a standalone backup proxy

server, such as VCB, or appliances from third-party vendors.

A storage manager might be tempted to simply back up the VMs on the physical host like regular files. In this case, the organization would back up each virtual machine's VMDK, a large disk file containing the VM configuration and data. Nice idea, but it might not work well. Unless the virtual machine is shut down, you might back up data in use, which is likely to result in inconsistent data.

The use of an individual backup agent on each VM can avoid data inconsistency by quiescing the application during backup. However, this approach can result in high backup software licensing fees if the organization must purchase an instance of the license for each VM. It also creates the potential for resource contention unless the organization staggers its virtual backups. Still, the advantages of putting a backup agent on each VM are simplicity and familiarity. The backup procedure runs no differently and admins can do all of the usual things, such as file-level recovery, and full or incremental backups.

However, the backup application, unaware of the encapsulated nature of the VMDK files, can't do things it otherwise might do to speed backup, Mainland Information Systems' D'Costa noted. This approach also undermines some of the efficiencies organizations hoped for from server virtualization in the first place, as each agent must be managed individually.

RAW DATA MAPPING

WHAT IT IS

- Enables a storage administrator to specify a logical unit number (LUN) on the disk array for a given virtual machine (VM)
- Prevents VMware from managing storage for that VM
- Used mainly for virtual machines running critical database applications

ADVANTAGES

- Improves performance of the virtual machine
- Gives the storage administrator direct control over storage and backup for that virtual machine

DISADVANTAGES

- Negates some of the value of virtualization by complicating management tasks
- Can't take advantage of VM mobility or Distributed Resource Scheduler (DRS)

SOURCE: Mark Teter, chief technology officer at Advanced Systems Group, Denver

D'Costa's second approach runs the backup software in the virtualization server itself, such as VMware ESX. This will likely be a Linux backup agent capable of backing up all of the VMs. However, this results in image-level backups that, although fast, don't allow for the easy recovery of individual files. It also requires scripting to automate the shutdown, snapshot and restart of the virtual machines.

The proxy server (particularly VMware Consolidated Backup), D'Costa's third approach, may become the most popular. The proxy server moves the backup from agents on the virtual machines or from the virtualization server to a dedicated server. In the case of VMware, "you put the backup agent on VCB and back up all the VMs," Server Centric Consulting's Miller said.

The proxy server is typically a Windows server connected to the same SAN volumes as VMware's ESX server. "With VCB you can now back up multiple VMs on the same physical host and VCB organizes the backups to keep them from overusing the resources," Miller added. VMware provides a load balancer called Distributed Resource Scheduler (DRS).

> "With VCB you can now back up multiple VMs on the same physical host and VCB organizes the backups to keep them from overusing the resources."
>
> —Scott Miller, president, Server Centric Consulting

But VMware Consolidated Back-up is far from a backup panacea. Mainland Information Systems' D'Costa points out that it requires a number of pieces to work right: a sync driver on the ESX server to flush the file systems and create a snapshot, a vLUN driver on the proxy server to present VMDKs to the proxy, and command line utilities to assist with scripting automation through the command line interface (CLI). It also will require an integration module provided by VMware or the backup application vendor.

Storage managers often mix different approaches. For example, Rockledge, Fla.-based Health First is a three-hospital healthcare provider network serving Brevard County. It uses VCB through a third-party tool, Vizioncore's vRanger Pro, to back up 1.5 TB of data from approximately 220 VMs every night. The data is backed up to disk. However, the hospital also uses IBM Corp.'s Tivoli Storage Manager (TSM) and installs a TSM backup agent on some of the virtual machines. The TSM backups go to disk and then to tape, which is shipped offsite.

VCB and vRanger Pro let the team "deal with the growth of VM sprawl," said Joel Otero, network engineer at Health First. Some days "we're building 10 to 15 VMs. Tivoli couldn't keep up with that. It's far too much to script," he explained.

W.R. Berkley combines VCB with Veritas NetBackup and FalconStor Software Inc.'s Network Storage Server for backup and recovery of its virtualized Windows and pSeries servers. The FalconStor product takes snapshots a few times a day and replicates them to a second data center. "So we risk, at worst, losing a couple of hours of data," W.R. Berkley's Whelans said. Using VMware Consolidated Backup, the company can restore individual files. Network Storage Server also spins off the snapshots to tape every night.

VIRTUALIZATION BACKUP TOOLS

The major array vendors are tuning their backup, recovery and replication tools to work with VMware, often as part of a two-step process in which companies use VCB to protect the VMs by creating crash-consistent copies and then deploying array-based replication technology to protect the data.

VMware recently introduced vCenter Site Recovery Manager to automate disaster recovery management.

Site Recovery Manager works with VMware's vCenter Server management console and with replication software from various storage partners, including 3PAR Inc., Dell Inc., EMC Corp., FalconStor, Hewlett-Packard Co., Hitachi Data Systems, IBM and NetApp.

> "Site Recovery Manager automated all this for us, all the scripting and changing of IP addresses. It also lets us test our DR plan non-intrusively."
>
> —Donald Wilkins, IT director, Navicure Inc.

Navicure Inc., a Duluth, Ga., provider of revenue-cycle management systems for physicians, relied on extensive manual scripting to back up as many as two dozen virtual servers using the replication capabilities built into its Dell EqualLogic SAN. "The scripting was taking us hours," said Donald Wilkins, Navicure's IT director. "We wanted to streamline the process." Just trying to recover a VM and promote it to production involved a cumbersome process of mounting files and changing IP addresses.

"Site Recovery Manager automated all this for us, all the scripting and changing of IP addresses. It also lets us test our DR plan non-intrusively," Wilkins said.

To validate this approach, Wilkins undertook the two-hour challenge: to bring up 10 VMs in two hours. "We started at 7 p.m. on a Friday night and began cloning VMs with EqualLogic. We changed IP addresses, replicated them and defined protection groups in less than two hours. We pressed a button and had all 10 VMs up and running 10 minutes later," he said. By 9 p.m. the team was heading home.

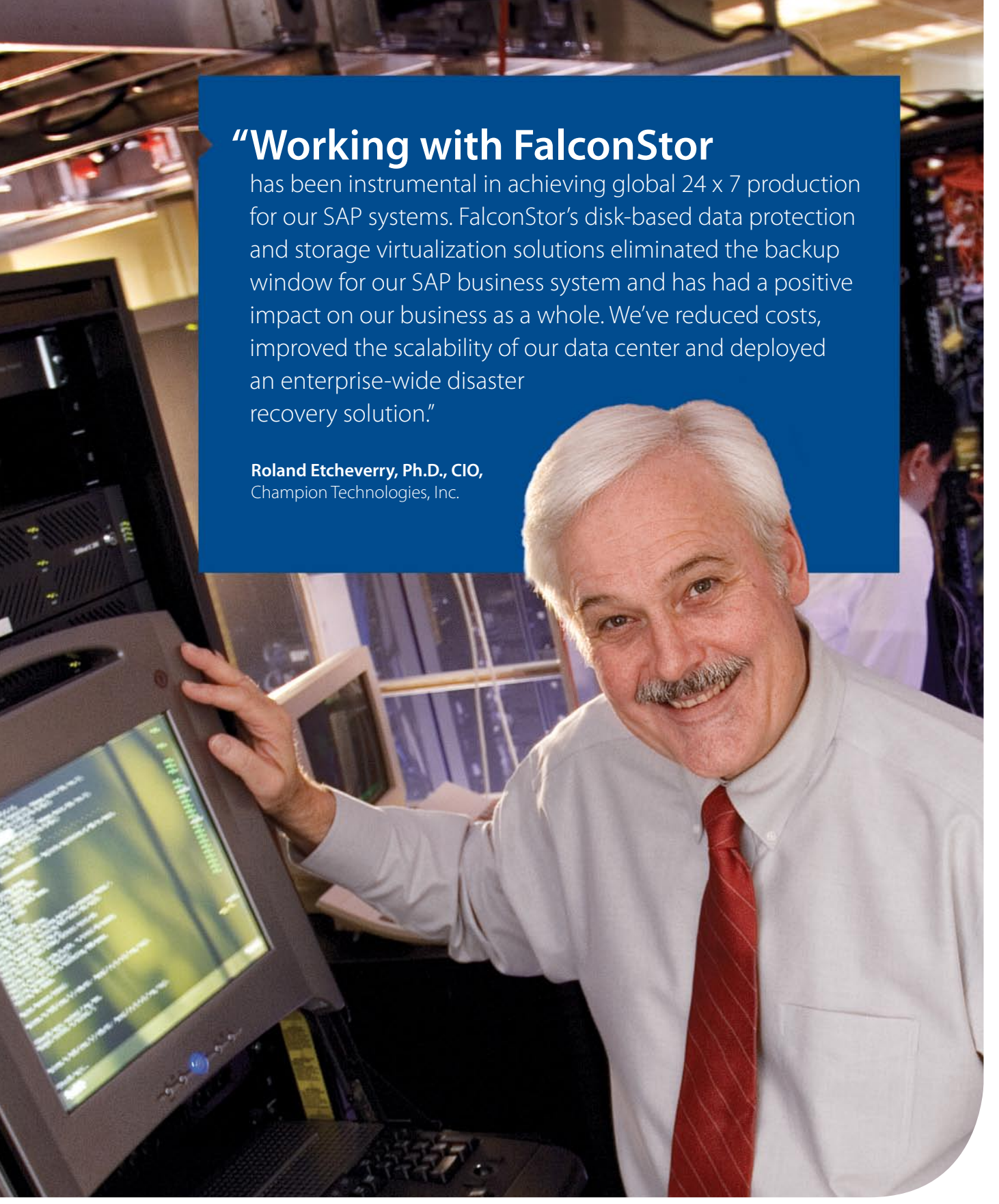
Newton, Mass.-based UGL Unicco turned to STORServer Inc.'s

STORServer Appliance for VMware Consolidated Backup. With approximately 100 virtual machines running on five VMware ESX servers and IBM TSM as its primary backup tool, UGL Unicco uses the STORServer Appliance to interface with Tivoli, VMware and VCB. For most of its virtual machines, it puts a TSM agent on the server and backs up in the conventional way to STORServer disk, which spins it off to LTO-4 tape. It uses VCB with approximately 30 critical VMs to allow for file-level recovery.

"We're an IBM shop and we liked what Tivoli does, but it's very complex, highly scripted and uses a CLI, so we went with STORServer as a GUI front end," said Darrell Stymiest, UGL Unicco's network services manager. VMware backup was similarly challenging, requiring extensive scripting through another CLI. With the STORServer Appliance for VMware Consolidated Backup, "we can take the VCB snapshot and write it to disk on STORServer and then send it to tape," Stymiest said. Recovering files with VCB remains a cumbersome multistep process, but it's much improved over previous VMware backups. "We want to use VCB with 90% of our VMs eventually," he added.

"Backing up virtual servers isn't simple and the tools aren't perfect," Mainland Information Systems' D'Costa said. But the virtualization backup process is still immature. "Backup tool vendors and array vendors are still trying to conform with VMware," which is still evolving its tools and APIs, D'Costa added. Until the virtualization industry matures in a few more years, backing up VMs will remain a challenge. ⊙

Alan Radding is a frequent contributor to *Storage* magazine online, SearchStorage.com and SearchSMBStorage.com.



"Working with FalconStor

has been instrumental in achieving global 24 x 7 production for our SAP systems. FalconStor's disk-based data protection and storage virtualization solutions eliminated the backup window for our SAP business system and has had a positive impact on our business as a whole. We've reduced costs, improved the scalability of our data center and deployed an enterprise-wide disaster recovery solution."

Roland Etcheverry, Ph.D., CIO,
Champion Technologies, Inc.

To learn how to deploy enterprise-wide data protection and storage virtualization with thin provisioning:

Contact FalconStor at 866-NOW-FALC (866-669-3252) or visit www.falconstor.com

FalconStor Software

Backup apps and VMware

Find out how well seven major backup applications work with VMware.

By Eric Siebert

SERVER VIRTUALIZATION has changed the way we back up and restore data. You can still use traditional backup methods for virtual machines (VMs) by using a backup agent on each virtual machine, but there are more efficient methods.

Most data backup vendors have adapted their software to better integrate with virtual server environments. And VMware Inc.'s VMware Consolidated Backup (VCB) technology allows backup vendors to easily integrate their applications into the VMware Infrastructure. While improving backup efficiency is important, performing whole VM and individual file restores easily and quickly is equally important. We'll look at the major backup applications to see how well they integrate with VMware.

CA ARCSERVE BACKUP R12.5

CA ARCserve Backup r12.5 integrates with VMware Consolidated Backup to perform off-host backups, and has minimal additional integration features or functionality. It backs up virtual machines using VCB to perform full virtual disk and individual file backups, and utilizes a VMware Configuration

Tool to populate the virtual machine information into its backup database. In addition, it has integrated antivirus protection to scan files while backing them up to eliminate the need to install antivirus on all of your virtual machines. To restore individual files to their original location on a virtual machine, you must install a backup client agent on it. To restore full disk backups, you must restore them to an alternate location.

When used in conjunction with CA's XOsoft High Availability and Replication products, you can replicate VM data to another server and then back up that server rather than the original server.

QUICK TIP: With Version 12.5, CA integrated VMware's vStorage API and Microsoft's Virtual Shadow Copy Service. Customers can back up servers at the hypervisor level without requiring a client on each guest OS and restore single files from snapshots.

COMMVAULT SIMPANA 8.0

With Simpana 8, CommVault introduced a single client for VMware ESX server and Microsoft's Hyper-V that can be deployed at the host level rather than requiring a client on each guest. Like other backup products for VMware, Simpana 8 integrates with the VMware Consolidated Backup API to allow virtual machine backups from a proxy server. Simpana can perform agentless full virtual disk backups and individual file backups utilizing VCB; it can also restore individual files directly to the guest virtual machine in a single step.

QUICK TIP: CommVault Simpana's host-based replication to and from virtual machines is a nice feature for disaster recovery, and Simpana includes an autodiscovery feature to automatically add new VMs in the environment to backup policies.

EMC CORP. AVAMAR

EMC Avamar integrates with VCB and includes an agent that can be installed on ESX hosts, as well as agents that can be installed inside the virtual machine guest OS. Its backup method relies on disk-to-disk data transfers from ESX hosts and VMs to an Avamar backup appliance. Avamar's strong point is its seamless data deduplication capabilities, which greatly reduce the amount of

data that's backed up. The data deduplication feature uses agents installed inside the virtual machine to perform deduplication at the VM so it's not sent over the network. While it does integrate with VMware Consolidated Backup, it currently doesn't integrate with vCenter Server or ESX hosts using the VMware Infrastructure SDK; however, this is planned for a future release. It also includes a virtual appliance that can be used as a single point of backup for remote locations.

QUICK TIP: Avamar has decent integration with VMware via VMware Consolidated Backup and its own agents, and while it's not as tightly integrated as some of the other products, its great data deduplication feature more than makes up for it.

HEWLETT-PACKARD (HP) CO. DATA PROTECTOR V6.1

HP Data Protector v6.1 also integrates with VCB to perform off-host backups. This version of Data Protector introduced a new VMware Integration Agent, which communicates with VMware vCenter Server and the ESX hosts using the VMware Infrastructure SDK. However, the VMware Integration component requires installing an agent on every ESX host you want to back up virtual machines on: the VMware vCenter Server server, the VMware Consolidated Backup server, and all physical or virtual systems that you plan to restore VM files to. It has some additional features such as Zero Downtime Backup, which performs application-level backups; and the Instant Recovery feature, which can be used to quickly restore files if you're using an HP Enterprise Virtual Array (EVA) system to make disk backups before writing to tape.

QUICK TIP: HP Data Protector has good VMware integration with VMware Consolidated Backup, but the additional overhead of installing and running agents on many systems can be cumbersome.

IBM CORP. TIVOLI STORAGE MANAGER 6.1

IBM Tivoli Storage Manager (TSM) 5.5 integrates with VCB to perform off-host backups by creating a snapshot of a virtual machine's virtual disk file and copying the virtual disk to the VMware Consolidated Backup proxy server so it can be mounted

and backed up. With TSM 6.1, IBM added a GUI for backing up virtual machines using VCB. Now it includes automated backup and recovery within the GUI at the image level. TSM is agentless on the guest VM, and the virtual machine's virtual disk to be backed up is mounted by the backup proxy server and not the virtual machine.

QUICK TIP: Tivoli Storage Manager 6.1 backup and recovery can be automated without manual commands or scripting.

SYMANTEC CORP. VERITAS NETBACKUP 6.5

Symantec Veritas NetBackup 6.5 has very good integration with VMware vCenter Server and ESX hosts when using NetBackup for VMware and NetBackup PureDisk for VMware options. NetBackup for VMware uses VMware Consolidated Backup to snapshot a virtual machine disk file and then make an off-host backup of it as a disk image. It then indexes every file within the virtual disk to allow a whole VM and individual file restores from within NetBackup so only one backup pass is needed to enable either type of restore. You can also perform VCB incremental backups to only back up changed disk blocks, which can greatly reduce backup times and storage requirements. NetBackup uses VMware's APIs to integrate directly with VMware vCenter Server and ESX hosts to automatically discover all virtual machines on the host servers. Additionally, it integrates with VMware vCenter Converter so you can quickly restore a whole OS with just a few mouse clicks. You don't need a backup agent installed on the VM unless you want to use the PureDisk data dedupe technology; PureDisk can greatly reduce backup I/O and storage requirements.

Symantec Veritas NetBackup 6.5 has very good integration with VMware vCenter Server and ESX hosts when using NetBackup for VMware and NetBackup PureDisk for VMware options.

QUICK TIP: Symantec Veritas NetBackup 6.5 integrates very tightly with VMware to improve efficiency. It also has advanced integration features to allow for easy restores and reduced storage requirements.

SYMANTEC BACKUP EXEC 12.5

Symantec Backup Exec 12.5 integrates with VMware Consolidated Backup just like Veritas NetBackup and has many of the same integration features. It also supports full disk and individual file restores without having to restore the whole virtual machine. Backup Exec integrates with VMware vCenter Server and ESX hosts for automated discovery of virtual machines on the host servers. In addition, it's agentless and doesn't require a special backup agent to be installed on VMs to be backed up. Because Backup Exec is aimed at small- and medium-sized businesses (SMBs), it doesn't have some of the enterprise features that Veritas NetBackup has, such as performing VCB incremental backups and using PureDisk deduplication technology.

QUICK TIP: Symantec Backup Exec 12.5 integrates very tightly with VMware, but lacks some of Veritas NetBackup's more advanced integration features.

All of the backup products discussed here integrate with VMware Consolidated Backup to provide off-host image and file-level backups of virtual machines. A few of the products provide additional integration by leveraging the VMware Infrastructure SDK to automatically read virtual machine information from VMware vCenter Server and ESX hosts. Of the products listed, Symantec's Veritas NetBackup seems to have the best integration with VMware and provides many advanced features. Virtualization is here to stay, so all of these backup products will continue to improve their integration with VMware to provide more efficient backup and recovery of VMs, as well as tighter integration with VMware's infrastructure and management applications. ⊙

Eric Siebert is a 25-year veteran of the IT world and has been specializing in virtualization for the last three years. He is a moderator in the VMware community VMTN forums and maintains VMware-land.com, a VI3 information website. He's also the author of *VMware VI3 Implementation and Administration*.

Double-Take Software

Microsoft

PRESENT

Ensuring the Availability of Hyper-V™ Virtualized Workloads

In this video webinar, hear from a Microsoft Evangelist and from Double-Take Software to discover:

- An in-depth look at Double-Take for Hyper-V, Windows Server 2008 Hyper-V, and Microsoft's virtualization strategy.
- A real-time protection and availability solution for workloads running on Windows Server 2008 Hyper-V
- A technical video demo of the new Double-Take for Hyper-V product.

ON-DEMAND VIDEO WEBINAR

Double-Take® Software is a leading provider of affordable products that are simple to use and enable you to move, protect, recover and more flexibly run critical IT workloads in physical and virtual environments, regardless of platform or location.

Keys to unlocking VCB

Learn these five principles that underpin VMware Consolidated Backup before you attempt to use it in your data center.

By Rich Brambley

ALTHOUGH VMware Inc. VMware vSphere 4 includes the VMware Data Recovery package, according to the company, it isn't a replacement for VMware Consolidated Backup (VCB). Customers using VCB will likely use it as a backup solution until a later generation of VMware vSphere comes along. But even before vSphere, VCB was often misunderstood; it requires preparation and understanding for backup administrators used to traditional physical enterprise backup. Here are five principles you should know about VCB before attempting to use it in your data center.

1. VCB is not the entire backup solution for virtual infrastructure. It's very rare that VCB allows administrators to completely remove all backup agents from virtualized servers. This is because VMware Consolidated Backup does *not*:

- Perform specialized application backups (like Microsoft Exchange Information Store or Windows Server System State)
- Perform file-level backups of non-Windows virtual machines (VMs)
- Provide management, cataloging or archiving of backup files
- Provide direct file restores to VMs

VMware Consolidated Backup is a framework of scripts that needs to be integrated with a third-party backup app to provide these features.

2. VCB should be installed on a dedicated Windows Server. It's recommended that VCB be installed on its own server. Also known as the VCB Proxy Server, this system has the following requirements:

- Microsoft Windows Server 2003 Service Pack 1 (32-bit or 64-bit) or higher
- Media repository managed by the third-party backup application's management server
- The same storage protocol access as the ESX hosts to the VMware vStorage VMFS logical unit numbers (LUNs) where the VMs are stored. (i.e., host bus adapters [HBAs] for access to Fibre Channel storage or initiator configuration for iSCSI storage). Depending on the version of Windows Server used, automatic partition mounting will have to be disabled before attaching the VCB server to the VMFS LUNs
- Dedicated disk storage for the VCB Holding Tank where backup and restore files are written
- Third-party backup agent

3. VCB needs a large disk volume for a Holding Tank. Along with the shared access to the ESX LUNs, VMware Consolidated Backup also needs a large disk volume formatted as NTFS, which will become the Holding Tank for backup images. This volume can be on the storage-area network (SAN) or the local VCB server's disks. The Holding Tank volume is where full VM images are placed during backups and restores.

Therefore, the size of the Holding Tank is critical in the design. For example, if a virtual infrastructure consists of virtual machines that take up 1 TB of disk space and the expectation is that a full VM backup is to be taken nightly, then the Holding Tank volume needs to be large enough to support 1 TB of backups. Another scenario would be to alternate groups of full VM backups to decrease the required size of the volume. In this case, administrators still need to make sure the Holding Tank is large enough to hold the virtual machine using the most disk space.

4. The role of the third-party backup agent. The third-party backup application does the actual backing up and management of the files. Once VCB copies a VM image to the Holding Tank, it's then up to the third-party backup application to move those files to whatever media repository is in use. It's also the function of the agent to clear out the Holding Tank so that the next scheduled job has available disk space to complete.

In the case of file-level backups, VMware Consolidated Backup also mounts the copied virtual machine image so that the backup agent can see the VM's file system. The backup agent can then perform full, incremental or differential file-level backups to the media repository. In some scenarios, the single agent on the VCB server can replace the multiple agents on the VMs.

VMware maintains a compatibility guide for supported third-party backup applications. Many of these supported applications have VCB integration modules that coordinate the scheduling of the VCB scripts and the agent backup from within the application's GUI.

5. Understanding VCB restore jobs. Restoring files leverages the third-party backup agent's ability to move files from the media repository back to the Holding Tank. Once the VM image is back, it can be copied in full to a VMFS volume or mounted like a thumb drive so that individual files can be restored. An administrator must manually copy files to the restore location in both scenarios.

VMware vCenter Converter, most often used to migrate physical servers to virtual machines, can also create VMs from VMware Consolidated Backup images. Therefore, VMware vCenter Converter can be a more effective full VM restore tool in some cases. Check out VMware's Virtual Machine Backup Guide for more detailed information on implementing VCB. ⊙

Rich Brambley is a senior infrastructure consultant/engineer working for Optimus Solutions in Norcross, Ga. He specializes in virtual infrastructure, and has been designing and implementing various virtualization technologies for the last five years.

Virtual server backup

Find out why backup capacities grow in a virtualized environment, how data deduplication can address the growing data store and what the restoration process looks like.

By Mark Bowker

HOW MUCH are virtualized server environments affecting backup processes? Mark Bowker, an analyst at Milford, Mass.-based Enterprise Strategy Group, answers the most common questions about virtual server backup from storage administrators.

How has virtualizing servers impacted the backup process?

With server virtualization, things are definitely different. There are a lot of advantages, but things are very different when you're backing up virtual servers. What we're finding is that the amount of data to be backed up has increased significantly.

However, in some cases server virtualization can reduce the number of backup licenses. This is a nice benefit. You don't have to run as many agents as in the past, and you can save some money.

In some environments, server virtualization has prompted the use of a secondary storage system, for example, a disk-to-disk backup system to maintain multiple copies of images.

What data is actually being backed up when a virtual machine (VM) image is backed up?

This is a big difference. Typically, with traditional backup, you put the agent on there, back up files and everyone's happy. If we need to restore a file or an email, it's fairly simple. We go back into the agent and recover it.

In a virtual server environment, the actual virtual machine image is stored as a single file. That's very different. The OS, applications and data itself is all stored within a single file.

In VMware, this is called a VMDK file. With Microsoft, it's called a VHD file.

What's the difference between traditional file-level backup and virtual machine backup?

Most traditional client/server architecture remains the same in the server virtualization environment where the client has an agent installed. However, this can become cumbersome. We still have to manage agents. There's processor overhead, which can be an issue, agent-based backup and VMs. This is especially true if multiple virtual machines are running on a single physical server, which is very much the case in a virtualized world. If those virtual machines all kick off the backup process at the same time, that CPU can be hit pretty hard and affect the applications.

An alternative is to use a different technology to perform the backup of the VM disk image directly. This requires the virtual machine to be suspended so that a consistent capture of the VM can be performed. Once the machine is suspended, the backup process can take place and then the machine can be restarted. This works well, but suspending the virtual machine is an issue because taking an app offline isn't acceptable in some environments. But there are other ways to integrate with the virtualization solution and perform a snapshot of the virtual machine. Then you can back up that snapshot.

Another thing you should be aware of is that you can recover the entire VM, but you can't recover a single file within that virtual machine. You have to restore the entire virtual machine, remount the virtual machine and then recover the file.

How has server virtualization affected the backup window and backup capacity?

The amount of capacity, just because of what's being backed up, has increased significantly. Enterprise Strategy Group just conducted a survey of people who recently adopted server virtualization, and 37% said the amount of data they back up increased after deploying server virtualization. It's significant—you're backing up the operating system, apps and data.

Then there's the proliferation of virtual machines, also known as VM sprawl. Once you get the infrastructure set up and in place, actually deploying the virtual machine is pretty simple. It's just a matter of clicking a couple of buttons, and you have a new machine set up. But after you do that, is the proper backup process in place?

So there's more data to be backed up and more VMs, but you still have to back up the data within the same window. It's important to be aware of that.

Can you perform a backup of a virtual machine when it's live?

There are a couple of different methods, depending on what your restore goals are. You can take a virtual machine image, and you can take a snapshot or backup of that while the VM is still writing to that image. But that's like walking up to a server and pulling the plug, and then plugging it back in, hitting the power button and hoping it turns back on. Oftentimes it works; the application may take care of itself and recover.

Another alternative to powering it down is to use the snapshot utility in the server virtualization software itself. This will quiesce the virtual machine image, freezing it temporarily while it performs the snapshot. Once the snapshot takes place, you can back up the snapshot. That gives you the complete system state of the machine. You can actually take that and bring it to a secondary data center. You could bring it to another environment for testing. It's a true point-in-time copy of the virtual machine image.

How does data deduplication affect the back-up process in a virtual server environment?

I'm hearing more and more questions about data deduplication and virtualization. And it's not just server virtualization; another area is desktop virtualization. With server virtualization, you have lots of virtual machine images out there and they have the same operating system. For example, Windows Server 2003 might be installed multiple times. I'm backing up multiple copies of something that I essentially only

need one copy of. Whether it's an operating system file, a patch, an application, a device driver, whatever it may be, I only really need one copy. So the benefit of data deduplication can be enormous.

In physical environments, next-generation backup has been somewhat slow and marginalized, maybe solving niche problems on the fringe of the data center. This is most likely due to cost. Implementing server virtualization requires a refresh of the data center, so it's an opportunity to do so.

Has the restore process changed at all?

It has. When you're looking at virtual machines, there are a few things you need to consider. The restore process can be just as it has been traditionally. If you have an agent-based backup system, you can restore a single file back into the VM just as we have in the past.

But if you're doing system-level backups where you're backing up the entire virtual machine image, you have to restore the entire VM image. This can be very time consuming. Once you recover it, you have to mount it somewhere in a virtualized environment. Often, this isn't the production environment because you can have conflicts with something else that's running there. Once you mount it, you can recover files from that point and transfer them back to where they need to be.

What types of innovations are being introduced in this space?

In a recent survey, we asked the question "Are you using the same backup tools for your virtual environment as you are for your non-virtualized environment?" Seventy-five percent of respondents are using the same tools and of the 25% implementing new tools, we're seeing that people want management tools similar to storage system management tools integrated in the server virtualization management platform.

Take something like VMware Inc.'s VMware Consolidated Backup (VCB). VCB is a kind of framework for the backup process that snapshots the network storage environment and allows the backup to take place without affecting the production environment. Using VCB, you can actually take virtual machine-level backups of those machines, but you can also provide file-level recovery. So something that had been a two-step process can be performed in one step. You're also removing the backup traffic from the production local-area network (LAN). And you don't need the backup agent installed in the virtual machine. ⊙

Mark Bowker is an analyst at Enterprise Strategy Group.

Channel Perspective

VMware backup FAQ

Find out about the options for VMware data protection and how to decide which approach is best for a particular customer.

By George Crump

While many customers are rushing to implement server virtualization projects with VMware Inc.'s ESX Server, very few are giving serious thought ahead of time to data protection of their new virtual servers. This oversight means that a lot of your customers have unanswered questions about the best backup method for their environment or may have taken a step in one direction only to learn it was the wrong one. To avoid some of these problems, find out what the options are for VMware backup and how to decide which is best for a particular customer.

My customer has implemented VMware. What are the options for data protection?

When a customer implements VMware, data protection is often an after-thought or, at best, the customer assumes data protection can be handled just like it was before virtualization. No matter which camp a customer falls into, there are three options for VMware data protection. The most common is guest OS backup. This method essentially treats each virtual machine (VM) as if it were a physical host, with the backup software's agent deployed within each virtual machine.

The second option is console backup. This method backs up at the physical host layer as if the virtualization host were a Linux server, and

a Linux agent is often deployed. The third option is VMware Consolidated Backup (VCB), which allows for off-host backup and recovery of the physical host and virtual machines.

What's important when determining whether guest OS backup makes sense for a customer?

Guest OS backups are the most popular approach with customers because the method most closely emulates what customers are familiar with, a (virtual) machine-by-machine backup.

Before you go down that road, you'll need to advise customers on the ramifications of this type of backup. First, it limits the number of virtual machines per host because if each backup job executes simultaneously, the performance on the host can be affected adversely and, in a worst-case scenario, the backup operation may cause a crash.

Guest OS backups are the most popular approach with customers because the method most closely emulates what customers are familiar with, a (virtual) machine-by-machine backup.

Second, even if the backup administrator carefully plans the backup schedule to limit the number of guest OS backups happening at the same time, it's extremely likely that new virtual servers won't be picked up by the backup administrator.

Third, guest OS backup doesn't protect the key VMware host files needed to recover the VMware host itself. An additional backup job will have to be created, monitored and maintained for full protection of the VMware environment.

Techniques like block-level incremental backups can aid significantly in guest OS backups by minimizing the client impact and the amount of data to be transmitted, making guest OS backup a viable option. But without block-level incremental, guest OS backup is too expensive administratively and too risky.

How do I know whether console backup is a good fit for a customer?

The Service Console is a special virtual machine based on Red Hat Linux that runs on each ESX server host. Backing up ESX servers via VMware's Service Console is the inverse to backing up using the guest OS backup method. The Service Console method backs up the ESX server as the operating system, and the virtual machines are just "files" under that operating system.

With this method, a Linux client is installed on the VMware Service

Console. The Service Console has visibility to the VMware File System (VMFS), which contains VMDKs, various VM and ESX configuration files, and redo log files.

The challenge with this backup method is that you can't restore individual files within the virtual machines. While console backup will work for customers with a Linux client and who can't afford to invest in one of the other protection strategies, it's not the best option.

What role should VCB play in the VMware data protection process?

First of all, it's important for both you and the customer to understand that VMware Consolidated Backup requires a shared storage-area network (SAN) environment; network-attached storage (NAS)-based VMware environments don't support VCB, nor does direct-attached storage (DAS). Second, the backup software a customer uses for all of their servers needs to support VCB. Third, they'll need enough excess capacity on the SAN to store the backup Virtual Machine Disk Format (VMDK) images. Finally, VCB backups require a Windows server to act as a proxy. For many customers, these prerequisites may be prohibitive, while other customers might be ready for VMware Consolidated Backup from Day 1. For the latter group of customers, implementing VCB isn't very difficult; it will enable them to recover both individual files and full VMDKs, and to protect the VMware host's primary files.

Does it make sense for customers to use deduplication with VMware backups?

Data deduplication counts on repetitive data to be effective, and VMware delivers. Probably more so than any other file system, VMware has plenty of redundant data that can be eliminated; deduplication is one of the best processes to introduce into the VMware backup process. Not only are there redundancies between files, but the files themselves are redundantly backed up. With a VMDK, the disk image of a virtual machine typically has to be backed up completely. There are limited ways to back it up granularly, and both of these factors make data deduplication a must-have addition to your customer's backup plans.

With deduplication, you can easily use some of the built-in VMware tools to protect the VMDKs and VMware host configuration files. Deduplication appliances can also be used as a target for backups generated by VMware Consolidated Backup.

This efficiency and flexibility make deduplication an ideal addition to any VMware protection process.

A customer is thinking of using NAS for their VMware environment. How does that change the backup options?

In two ways: First, you can leverage the NAS snapshot capabilities to provide granular restores of the files within the VMDK. Second, you can use Network Data Management Protocol (NDMP) to back up those hosts. This provides an ideal off-host backup option and is important because VMware's own VCB doesn't support NAS- or NFS-mounted VMDKs.

What's the best way to move VMware backups offsite for disaster recovery (DR) purposes?

There are several options. First, you can use the array-based replication provided by a shared storage solution. If your storage solution supports VMware's vCenter Site Recovery Manager (SRM), it makes sense to talk to your customers about that as well. SRM provides disaster recovery workflow automation.

Second, you could leverage one of the data deduplication or block-level backup solutions to use their replication feature to move data to a DR site. The VMDKs could then be recovered on a regular basis, and you could even automate this for the customer and provide an inexpensive near-real-time updated DR site for your customer.

VMware has several built-in backup utilities. Which ones should I consider for customers?

There are two primary tools provided for VMware: vmkfstools and vcbMounter. When vmkfstools is executed, a specific VMDK file can be copied from the VMware file system to a standard file system mounted by the Service Console. The VMDK files are separated into 2 GB files to maintain compatibility with standard file systems.

vmkfstools only works on VMDK files, and the VMware administrator must manage shutting down the virtual machine or creating a redo log for the virtual disk. It's not as effective as vcbMounter for backup and recovery. This tool is more appropriate as a way to move a virtual machine between VMware platforms.

vcbMounter, on the other hand, is used to export a backup copy of the virtual machine. The command is also invoked at the command line or can be part of a script. With vcbMounter, you can indicate the name of the virtual machine to back up and the destination directory for the target backup.

All of the files needed to re-create or restore the VM are also exported by vcbMounter to produce a complete file-system-consistent backup copy

of the virtual machine. As a result, vcbMounter is better suited for backup and disaster recovery than the vmkfstools utility because it ensures this consistency.

Working alongside vcbMounter is vcbRestorer, which imports a copy of a virtual machine created by vcbMounter and can be used to completely recover a virtual machine to its original state to the original ESX server host. It can also be invoked on a separate ESX server in a disaster recovery situation.

These tools are ideal adjuncts to a network-mounted deduplication device.

What effect does the backup method have on the ability to add virtual machines?

Anything that is "on-host" for backup is likely to cause performance issues. As the number of virtual machines increases, these performance issues multiply. The number of VMs also limits your ability to schedule around those performance bottlenecks. Block-level incremental backup and source-side deduplication limit the amount of time each virtual machine is impacted by the backup; block-level incremental backups, in particular, require very little in the way of compute resources.

Most off-host backups have no impact on the number of virtual machines, which explains why so many of your customers are considering the options. As virtual environments scale, you need to be prepared to offer either block-level incremental backup technologies or off-host backup technologies. ⊙

George Crump is president and founder of Storage Switzerland, an IT analyst firm focused on the storage and virtualization segments.

Check out the following resources from our sponsors:

compellent

White Paper: 7 Ways Compellent Optimizes VMware Server Virtualization

White Paper: Three Must Haves for the Virtual Data Center

Customer Conversation Video: Find Out How Customers Use Our Feature-rich SAN in the Real World

Double-Take Software

On-Demand Webinar: Ensuring the Availability of Hyper-V Virtualized Workloads

Workload Backup Solutions from Double-Take Software

Data Protection for Virtual Systems and Hyper-V

FalconStor® Software

Webcast: Understanding Backup and DR for VMware Environments

White Paper: Virtually Effortless Backup for VMware Environments

Case Study: Globalization Firm Creates Flexible, Low-Cost Storage Environment With Virtualization