SearchStorage

TechTarget

E-Guide

Mistakes Made When Backing Up VMs & How to Prevent Them

Backing up virtual machines is a complex *undertaking, and many IT professionals have no idea what they're getting themselves into. This expert E-Guide compares three virtual backup methods and offers tips to help you avoid making common mistakes as you perform this essential process. Discover why block-based VM image backup is quickly becoming the method of choice and explore the data protection strengths and weaknesses of three leading hypervisors.*

Contents

Top Five Mistakes Made When Backing Up VMs & How to Prevent Them

By: Eric Siebert

What you will learn in this tip: Creating backups in virtual environments *isn't as straightforward as it is in physical environments. While there are a variety of approaches for backing up virtual machines (VMs), there are just as many pitfalls you may encounter due to the unique nature of virtual environments. In this tip, learn how to efficiently create VM backups and avoid common mistakes.*

Don't back up through the guest OS

Backing up through the guest operating system (OS) is probably the most common mistake made when backing up VMs. You cannot use traditional backup methods that use agents installed on the guest OS to back up VMs. While this works, it is inefficient because the virtualization layer sits in between the guest OS layer and the physical hardware layer. The guest OS no longer has direct access to physical hardware where the data resides, so a backup agent inside the guest OS must go through the virtualization layer to get to the virtual machine data. This method also causes unnecessary resource usage on the host and if multiple backups are running simultaneously, it can cause performance bottlenecks.

Instead of using guest OS backup agents, backup servers should go directly to the virtualization layer and not involve the guest OS. By using this method, the guest OS is not aware of a backup process, nor is it wasting host resources. It is also much more efficient as the backup server can mount the VMs virtual disk directly from the host data store. This type of backup is known as an image-level backup because the VM's disk is backed up at the block level and not at the file level as traditional guest OS agents do. To properly do an image-level backup at the virtualization layer you need to use backup applications that are virtualization-aware and can leverage the APIs of the virtualization layer to access virtual disk files.

You should never try and back up virtual disk files directly at the physical storage device and bypass the virtualization layer. The guest OS and virtual disk need to be prepared so they are in a proper state to be backed up and if you bypass the hypervisor this does not happen.

Virtual machine snapshots are not backups

Virtual machine snapshots preserve the state of a VM from the point in time when the snapshot was taken. Additionally, multiple snapshots can be created to provide multiple restore points to choose from. While this can be useful in certain situations, it should never be used as a primary backup method for your VMs. One problem with VM snapshots is that once you revert back to a previous snapshot, you can't go back to the present. The current state of your VM is lost and you can only revert to previous snapshots. Snapshots are not useful for restoring individual files because they only bring a whole VM image back to a present state. Snapshots can also cause other problems because they grow in 16 MB increments: The entire LUN that a VM is on has to be locked when they grow in size, which prevents other hosts from writing to the LUN.

This process is known as SCSI reservations and too many of them occurring can decrease the performance of your VMs as they wait for LUNs to be unlocked. Each snapshot is an individual file that grows as data is written to it, and having a lot of snapshots running can cause your datastores to run out of disk space. Snapshots are useful as a secondary backup method for short-term or ad hoc backups if you need to permanently revert to a previous state, such as when applying patches or upgrading applications.

Make sure you are quiescing properly

Most virtualization backup applications back up at the image level and are not aware of what is going on inside the guest OS. Before you back up VMs, you need to ensure they are quiesced so they are in a consistent state to be backed up. If you don't quiesce them, you risk having data that is not in a state to be restored properly. The quiesce operation is handled inside the guest OS, and for Windows VMs, the Volume Shadow Copy Service (VSS) handles this. Since the backup server is backing up the VMs at the virtualization layer—and not inside the guest OS—it requires another application to tell the guest OS to quiesce the VM.

In vSphere, that application is VMware Tools, which tells the VSS service to quiesce the guest OS. The application installs on the guest OS to serve as a conduit between the guest OS and the hypervisor.

For VMs running Linux operating systems with no native services like VSS, VMware Tools also provides a special vmsync driver that can provide the same functionality as VSS. This makes it very important that VMware Tools be installed and kept up to date on all your VMs. There are also instances where VMware Tools may not support certain guest OS versions, so always check to see if your version is supported by the application.

Many backup vendors supply their own special agent that will handle the quiesce process if VMware Tools doesn't offer support.

Schedule backups carefully

VMs share the resources of a host and hosts share storage devices, and creating backups is a resource-intensive operation. In a virtual environment, creating a backup can cause resource starvation among your hosts and VMs. While backing up at the virtualization layer reduces resource usage on your VMs when backups occur, resource usage will still be high on your hosts and storage devices when backups are running.

To avoid too much concentrated I/O—which can affect the performance of your VM—you should schedule your backups to limit the number of concurrent VM backups on a host and shared datastores. Hosts typically share the same datastores in virtual environments, and bottlenecks caused by too many simultaneous VM backups on a single datastore will affect all hosts that have VMs running on that datastore.

Likewise, if too many VMs on the same host are being backed up at the same time, it will create bottlenecks for all the VMs on that host.

You should plan backup schedules carefully to ensure that backups occur in a balanced manner which do not cause resource problems for your VMs. And don’t rely on sluggish VMs to tell you that you have a problem while your backups are running. Instead, look at performance statistics taken at the virtualization layer to learn whether you have a problem. This allows you to monitor the I/O and make adjustments as needed to balance it out.

Don’t resource starve your backup server

Backup servers are basically like pumps: Data is read from a source and goes into the backup server, and then the data is sent from backup server to the target device. The volume that a backup server can handle is determined by the resources assigned to it, and the more resources available, the faster it can pump data. Backups can heavily tax network and storage resources, but there is more to backups then just moving data from point A to point B. Backup servers handle advanced functions like deduplication, compression and determining which disk blocks need to be backed up.

For your backup server to achieve maximum throughput, it needs to have sufficient resources to avoid creating a bottleneck in any one resource area.

You should monitor the resource usage of the backup server: In practice, it’s better for a backup server to have too many resources than too few. If those resources are maxed out, chances are the backup server will need more. By ensuring that your backup server has the resources that it needs, you can ensure that it pumps data at maximum speed and decrease the time of your backup windows.

The virtualization architecture introduces a lot of unique and creative ways to back up your VMs when compared to traditional physical environments.

Contents

Backup applications that integrate with virtualization can take advantage of these features and leverage them to increase the efficiency of backups. VMware has developed specific APIs that benefit backup applications, such as the vStorage APIs for Data Protection (VADP), which allows backup applications to interface directly with hosts and storage devices. VADP offers more efficient access to virtual disk files and contains features—such Changed Block Tracking (CBT)—which can greatly reduce the time it takes to perform incremental backups.

A big part of an incremental backup is figuring out what changed since the last backup. CBT queries a virtual machine's VMkernel, which keeps track of disk block changes, to quickly determine which disk blocks of a VM's virtual disk have changed since a specific point in time.

Backup applications normally figure this out on their own, so making this information instantly available can mean faster completion of the incremental backup process.

In order to achieve the most efficient backups possible, always make sure your backup application takes advantage of the many benefits provided by the virtualization architecture

Backing Up VMs

Traditional Apps vs. Virtual Machine Backup Software
By: Jacob Gsoedl, Contributor

Most organizations are still relying on traditional data backup and recovery software for virtual machine (VM) backup, but there are downsides to that approach. In Jacob Gsoedl's latest Storage magazine article, learn about the best way of backing up VMs, whether it's with traditional virtual machine backup options or new virtual machine backup software.

In most IT shops, virtual servers are backed up just like physical servers at first, but as the number of virtual servers increases traditional backup methods start breaking down. The fact that a single physical machine can host many VMs poses challenges that simply don't exist when backing up dedicated physical servers.

With multiple VMs competing for processing, storage and networking resources, contention for those resources is the No. 1 challenge of virtual server backup. Concurrent backup jobs on multiple VMs can seriously impact the performance of applications hosted on those VMs. And when traditional backup methods are used to protect virtual servers, some key capabilities are sacrificed, such as application-consistent data protection and the ability to restore sub-VM-level objects such as files without having to restore the whole virtual machine. As virtual servers proliferate in the data center, there's a clear call to storage managers to rethink backing up VMs and the applications they host.

VM backup options

Contemporary virtual server environments can be protected using one of the following backup methods:

- Backup agents on VMs
- Continuous data protection (CDP)
- VM image backup on the hypervisor using a backup proxy server

Backup agents on VMs. Backing up VMs by placing backup agents on each VM extends the most popular backup method of physical servers into the virtual server realm. Instead of having one backup agent per physical server, each VM gets its own agent and backup jobs run independently for each virtual machine. This approach is workable as long as the number of VMs is small; as the number of VMs per hypervisor grows, resource contention will create performance problems. Besides being able to leverage an existing backup product and approach, a backup agent can be the most straightforward way of ensuring application consistency. For many apps, especially non-Windows applications and applications that aren't integrated with Microsoft Volume Shadow Copy Service (VSS), backup agents may be the only way of ensuring application consistency of backup data.

Continuous data protection. A CDP product that runs on each VM is one step up from backup agents running traditional full and incremental backups. A CDP product captures changes on an ongoing basis and puts a much smaller burden on the host machine than traditional backups do. CDP products work either at the file or block level, and usually provide integration with mainstream applications to enable restoring to consistent points in time. In addition to large backup application vendors that have added continuous data protection to their suites, CDP products are offered by a list of smaller vendors such as FalconStor Software Inc., InMage Systems Inc., Vision Solutions (acquired Double-Take Software) and others.

VM image backup on the hypervisor using a backup proxy server. Backing up VM images on the hypervisor, rather than backing up virtual machines via agents within each VM, is appealing for many reasons: it enables efficient backups with little processing overhead; eliminates the need to install and manage backup agents on each VM; and by centralizing the backup of all VMs at the host, the backup of multiple VMs can be orchestrated to minimize performance problems and resource contention. To offload the backup task from the hypervisor machine, VM snapshots are usually replicated or mounted to a backup host or proxy server, minimizing the performance impact while backups are performed.

Host-side VM image backup, however, is usually only acceptable if VMs can be backed up in a consistent state; all major hypervisor vendors have added the ability to quiesce a VM while a snapshot of the VM image is taken. Another challenge with VM image backup is restoring granularity. While some backup products are only able to restore virtual machines, others are able to restore sub-VM objects such as files. Consistent data protection of apps within each VM is more challenging with VM image backup than it is with backup agents within virtual machines. Application-consistent data protection is usually limited to applications that are integrated with VSS. For apps that aren't integrated with VSS, crash-consistent backups may be the only option; but if application consistency is required, backup agents on VMs are the way to go.

Without question, the trend is toward VM image backup at the hypervisor level and offloading the backup task to a proxy backup server—and as the number of virtual servers grows it becomes even more relevant. Mechanisms to enable efficient VM image backup and capabilities vary significantly between Citrix XenServer, Microsoft Hyper-V and VMware vSphere.

VMware vStorage APIs for Data Protection

Data protection has been a sore spot for VMware and it took VMware until vSphere 4 to get it right. Prior to vSphere 4, VMware provided VMware Consolidated Backup (VCB) to offload backups from hypervisors to a proxy server, but it wasn't widely adopted due to some grave shortcomings. With VCB, snapshots of virtual machine disk (VMDK) images were taken and copied in full to a proxy server from which backups were run. Offloading the backup to a proxy server minimized the impact of backups on VMs, but it required additional storage for the snapshots.

The vSphere 4 vStorage APIs for Data Protection (VADP), the successor to VCB, addresses the shortcomings of VCB. To start with, VADP no longer requires copying data to a proxy server; instead, snapshots can now be mounted to a proxy server where they're backed up to disk or tape. While VCB only supported taking full snapshots of a VMDK, regardless of how much it changed, VADP supports efficient snapshots via its change block tracking (CBT) feature. CBT keeps track of changes within a VMDK at a block level and enables efficient snapshots of changes only.

vSphere 4 is fully integrated with VSS to enable application-consistent snapshots of VSS-enabled applications running on virtual machines. To be able to take advantage of VSS, however, VMware Tools needs to be installed on the virtual machine. vSphere communicates with VSS via VMware Tools.

To back up a VM via VADP, a "quiesce" command is sent to vSphere to instruct the VM to flush data in memory to disk and no longer accept writes. If VMware Tools is installed on the VM, VMware Tools can pass on the "quiesce" to VSS-enabled applications on the virtual machine to also "freeze" applications within the VM. A snapshot is then taken; on completion of the snapshot, the "freeze" is removed from the VM and VSS-enabled applications. Finally, the snapshot is mounted to the backup proxy from where it's backed up to disk or tape.

VADP is widely supported by third-party backup applications. In addition to major backup application vendors (Arkeia Software, CA, CommVault Systems Inc., EMC Corp., IBM, Quest Software Inc./BakBone Software Inc. and Symantec Corp.), smaller vendors such as PHD Virtual Technologies and Veeam Software offer virtual server backup applications with VADP support. Additionally, vSphere provides its own backup tool called VMware Data Recovery (VDR). VDR is delivered as a virtual appliance to perform snapshots and deduplication to a backup disk target. VMware has positioned VDR as a lower end backup product.

Microsoft Hyper-V and VSS

Thanks to Microsoft VSS, Microsoft got data protection for Hyper-V right from the get-go. In many ways, a VMware VADP backup cycle resembles backing up Hyper-V. A backup app dispatches a "quiesce" command to a Hyper-V VM via VSS to flush data in memory to disk; VSS then takes a snapshot and removes the freeze from the VM. Similar to VADP, the snapshot can then be replicated or mapped to a dedicated backup proxy server. The "quiesce" can be extended to VSS-enabled applications within VMs, but requires the so-called backup integration service installed on the VM, akin to VADP requiring VMware Tools.

VSS depends on several main components (click here for more information on Microsoft Volume Shadow Copy components) VSS-enabled applications need to implement a so-called VSS-writer that coordinates various components to create consistent shadow copies of one or more volumes. Applications, such as a backup application, need to implement a VSS-requestor to request a volume shadow copy. The key component, though, is the VSS-provider, which creates and maintains shadow copies (snapshots). While VSS-providers are included with the latest Windows OSes, software and storage hardware vendors can provide their own VSS-providers. Noticeably, hardware-based VSS-providers of arrays enable high performance and highly scalable data protection of Hyper-V environments. While it took VMware to implement change block tracking to get to efficient snapshots, in VSS it's a capability of the VSS-provider. For instance, the VSS-provider that's part of Microsoft's operating systems does incremental snapshots via a copy-on-write method; that is, when a change to the original volume occurs but before it's written to disk, the block to be modified is read and stored away.

The support for VM image-level backup of Hyper-V isn't as extensive among third-party backup applications as is support for VADP. For instance, IBM Tivoli Storage Manager (TSM) and PHD Virtual Backup don't have support for it at present. Similar to VMware, Microsoft provides its own backup solution for Hyper-V environments with System Center Data Protection Manager (DPM). DPM provides near-continuous data protection for virtual machines hosted on servers running Hyper-V. With advanced features like disk-to-disk and disk-to-tape support, the ability to recover sub-VM objects like files, protection of virtual machines while live migration is in progress, integration into Microsoft System Center, and support for clustered and standalone Hyper-V systems, DPM provides a state-of-the-art product rather than the entry-level product VMware provides with VDR.

Citrix XenServer backup

Third-party backup applications can initiate full or incremental disk image snapshots of Citrix XenServer VMs through XenAPI (XAPI). These snapshots are usually crash-consistent and depend on applications to regain a consistent state after a restore. Recovering after restoring a crash-consistent backup is analogous to powering up a virtual machine after a power failure.

With XenServer 5.6, Citrix added live memory snapshots to capture the state of a virtual machine when a snapshot is taken and allows reverting to a previous state on restore. Citrix memory snapshot leverages Microsoft VSS, so it's available for VSS-enabled VMs (Microsoft operating systems) but not for Linux virtual machines.

In addition to XenServer snapshots, XenServer supports shared storage snapshots for arrays supported by XenServer. Leveraging snapshot capabilities of arrays is the fastest and most scalable method to protect a

XenServer environment, but it's only an option if the storage infrastructure is supported by XenServer.

Like vSphere and Hyper-V, Citrix provides its own VM image backup application with VM Protection and Recovery (VMPR). A scaled-down version of VMPR that lacks features like scheduling is included with all versions of XenServer. A more advanced version that supports scheduling and automation is available as a paid option. Third-party backup application support for XenServer image-level backup is more tenuous than for vSphere and Hyper-V. Among the backup application vendors that support it are Arkeia Software, CommVault, PHD Virtual Technologies and Veeam Software.

Backing up VMs: The bottom line

Block-based VM image backup on the hypervisor host, ideally via a backup proxy server, is becoming the preferred way of backing up virtual servers. Maturing backup APIs in vSphere, Hyper-V and XenServer, as well as increasing support by backup applications for these APIs combined with performance and scalability merits, are among the main reasons for its adoption. Because most organizations run more than a single hypervisor (more than 70% of companies according to ESG's Whitehouse) and a mix of physical and virtual servers, multi-hypervisor support and the ability to support both physical and virtual server backups are important considerations when choosing virtual machine backup software.

TechTarget

Contents

Free resources for technology professionals

TechTarget publishes targeted technology media that address your need for information and resources for researching products, developing strategy and making cost-effective purchase decisions. Our network of technology-specific Web sites gives you access to industry experts, independent content and analysis and the Web's largest library of vendor-provided white papers, webcasts, podcasts, videos, virtual trade shows, research reports and more —drawing on the rich R&D resources of technology providers to address market trends, challenges and solutions. Our live events and virtual seminars give you access to vendor neutral, expert commentary and advice on the issues and challenges you face daily. Our social community IT Knowledge Exchange allows you to share real world information in real time with peers and experts.

What makes TechTarget unique?

TechTarget is squarely focused on the enterprise IT space. Our team of editors and network of industry experts provide the richest, most relevant content to IT professionals and management. We leverage the immediacy of the Web, the networking and face-to-face opportunities of events and virtual events, and the ability to interact with peers—all to create compelling and actionable information for enterprise IT professionals across all industries and markets.

Related TechTarget Websites

SearchDataBackup

SearchDisasterRecovery

SearchCloudStorage

SearchSMB

SearchVirtualStorage

SearchStorageChannel