Managing the information that drives the enterprise

STORAGE

SEARCHSTORAGE.CO.UK

essential guide to

VIRTUAL SERVER BACKUP

VM backup has come a long way in a few short years. Find out how the hypervisors and backup vendors are simplifying the job.

INSIDE

Brighter outlook for virtual server backup

CDP can simplify VM backup

Integrated backup

Hyper-V backup strategies

TechTarget

VM backup emerges from Dark Ages

After a difficult period of adjustment, hypervisors and the backup ecosystem that surrounds them have evolved to deliver tools that simplify the virtual server backup process and lighten the load on data centre resources.

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

VIRTUAL MACHINE BACKUP appears to have come of age, and we seem to have arrived at a situation in which users can now protect virtual machines and their data in a straightforward way using mainstream backup products. But now we can also protect virtual servers using replication-style methods such as CDP. In fact, so compelling a solution is the application of CDP to VM environments that we may be witnessing the beginning of the end of backup in its traditional form.

It has been a rocky road to get here. In the beginning it was as if the developers at VMware decided not to think very hard about how virtualised servers would be backed up. You could, of course, just do backup the traditional way, with an agent on every VM. But the I/O loads this brought about would soon show this method to be rather less than optimum, and you had to pay for and manage a lot of licenses.

VMware had its own on-board method of backing up in VMware Consolidated Backup, the consolidated part of which seems a hopelessly misnamed feature in retrospect. Using VCB, backups went through a two-stage process; first, the backup app signalled it wanted to do its work, and then a copy of the full VM or its data or both were snappshotted to a holding disk and the backup taken from there.

Then, late in 2009 VMware brought out the vStorage APIs for Data Protection (VADP), and the two-stage process of VCB was cast into the dustbin of history. With VADP, backup apps talk directly to the VM kernel, and backup products compatible with this include CA ARCserve, EMC Avamar, IBM Tivoli Storage Manager, and NetBackup and Backup Exec from Symantec. Specialist VMware backup products such as Veeam Backup and Quest Software's vRanger also interface directly with VADP.

The VADP revolution also saw the inclusion of vStorage APIs for multi-pathing to provide better I/O and storage path failover for storage arrays, plus vStorage

APIs for Site Recovery Manager to provide array-based replication for block- and file-based storage.

At the same time, Microsoft's Hyper-V also allows backup products to link directly to its hypervisor.

But despite these leaps forward in virtual machine backup, there is another way of protecting virtual machines and their data. Using CDP/snapshots, virtual machine users can now implement a regime of incremental block-level backups forever. By copying data at regular, frequent intervals, there need never be another full backup and the I/O, CPU and memory loads that implies.

So, just as backup as we know it has been made to work in the world of virtual servers, there really is no need for traditional backup after all. ⊙

Antony Adshead is the bureau chief for SearchStorage.co.UK.

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

Simplify virtual machine backup with image-level backups, CDP and near-CDP

Learn how to simplify virtual machine backup with image-level backup, continuous data protection (CDP) and near-CDP in this column.

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

IT'S TIME FOR users to pull their heads out of the sand regarding virtual machine (VM) backups. Server virtualisation software such as VMware or Microsoft Hyper-V may solve a lot of problems for system administrators, but they create a lot of problems for storage and backup administrators. The core problem is physics. While it may be possible for 10 to 20 Exchange or SQL Server VMs to successfully share a single physical server, two or three backup applications running at full speed will completely consume all the CPU, RAM and I/O resources of that same server. There are two key backup methodologies that attempt to address this issue: image-level backups, and continuous (or near-continuous) incremental backups.

IMAGE-LEVEL BACKUP

Image-level backups attempt to solve the virtual machine backup issue a couple of different ways. For example, they often use a server for the backup that is different to the virtual server being backed up. These other physical (proxy) servers are given access to the storage where the virtual machines are stored, allowing them to back up the "image" of the VM via a path that does not involve the virtual server. For this to work, the backup software and the virtual server software must cooperate so the backup software backs up a stable snapshot rather than an actively used volume.

This cooperation is also necessary in order for any applications running inside the VMs to know that they were properly backed up. For Windows VMs, it is typically provided via the Windows VSS service; other operating systems use a variety of methods. In addition to moving the physical I/O out of the virtual server, more recent image-level backup products are also able to

perform block-level incremental backups of the images, further reducing the I/O requirements for the storage. Readers interested in image-level backups should examine products that fully leverage the VMware vStorage Data Protection APIs or the Hyper-V VSS Writers.

BLOCK-LEVEL INCREMENTAL BACKUP TECHNOLOGIES

Dealing with virtual machine backup also provides a great opportunity to radically change data protection with the use of block-level incremental back-up such as continuous data protection (CDP) or near-CDP. CDP is replication with a change log that allows you to restore a system to any previous point in time (ie, sub-second granularity). Near-CDP is replicated snapshots that allow you to restore a system to any point in time when you took a snapshot, typically taken once an hour.

Depending on the product, these technologies are available as software that runs inside each virtual machine, in an appliance on the storage network, or as functionality provided by the storage array. The idea is to fundamentally change the way data backup and recovery is performed by never again doing a full backup, and by continually and incrementally copying changed bytes from the source system to the backup target.

By changing the backup process from a nightly, bulk transfer of data to a continuous, incremental transfer of changed blocks (that runs throughout the day), you fundamentally change the impact the backup system has on virtual servers. CDP and near-CDP technologies also fundamentally change the way restores are performed. Instead of requiring a bulk copy of data from the backup server (ie, a restore), both technologies support using the backup system as the primary system during an outage. The only downtime is the amount of time it takes to point the virtual server at the new storage location. Compare this to the hours required to do a typical restore, and you will understand the popularity of these technologies for this application. Readers interested in these technologies should ask backup software providers and storage providers about their support for CDP and near-CDP.

The way we back up our virtual machines has to change. Customers looking to make things incrementally better should examine image-level backup tools; customers looking to fundamentally change the way backups and recoveries work should examine CDP and near-CDP backup tools. ⊙

W. Curtis Preston (a.k.a. “Mr. Backup”) is an independent backup expert and has been singularly focused on data backup and recovery for more than 15 years.

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

Virtual server backup evolves from rudimentary process to integrated one

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

Virtual server backup is only now coming of age. VMware and Microsoft Hyper-V now allow backup products direct access to virtual machine system and file data. *By ANTONY ADSHEAD*

WHILE SERVER VIRTUALISATION has made life easier in many ways for IT departments, the virtual server backup process has until recently been a difficult and fragmented process.

In fact, while server virtualisation went mainstream a few years ago, the virtual server backup process lagged behind; only in the past year has it been integrated into the virtual server environment in a workable way.

What we've seen is an evolution that started by application of the tried and trusted methods of backing up physical servers to the virtual environment, via an awkward two-stage process, and finally reaching a good level of integration between backup and virtual servers with the release of vSphere's vStorage APIs for Data Protection (VADP) in 2009.

This overview will look at the various methods of backing up virtual servers available.

VIRTUAL SERVER BACKUP MEETS TRADITIONAL BACKUP

Many virtual server users have dealt—and still do deal—with the backup of virtual machine files by pretending they aren't virtual at all. They load the backup software agent onto the virtualised server just as they would have done with a physical server. It works—but with many agents on many virtual servers on one physical device, the potential for crawl-inducing I/O loads looms large. VMware and Microsoft Hyper-V users often adopted the same approach.

It's the backup equivalent of fighting today's war with the last war's tactics and technology; you may succeed, but it's an approach that could see you suffer losses to attrition or a catastrophic failure.

It's the backup equivalent of fighting today's war with the last war's tactics and technology; you may succeed, but it's an approach that could see you suffer losses to attrition or a catastrophic failure.

In answer to these difficulties VMware came up with VMware Consolidated Backup (VCB).

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

VIRTUAL SERVER BACKUP WITH VCB

VCB brought some relief from the potential I/O storms of traditional backup on virtual servers, but it was not an elegant solution.

VCB allowed backups to take place at the level of the virtualised server, but its inelegance lay in its requirement to go through a two-stage backup and restore with a discrete disk staging area between the device being backed up and the target.

The process goes something like this: First, the backup application tells VCB it wants to do a backup. This then triggers a Virtual Shadow Copy Service (VSS) snapshot (either of the data or the full virtual machine VMDK volume and data, depending on whether it is a Windows environment or not), and this is exported as a virtual drive image to the staging device. The backup application then takes its backup from that virtual drive image.

Restores of a virtual machine also require two steps. First, the backup software restores the data to the proxy server and then uses VMware vCenter Converter to restore that to the ESX server.

A number of point products were designed to address the shortcomings of VCB backup, such as Quest Software's vRanger Pro, PHD Virtual Technologies' PHD Virtual Backup for VMware (formerly esXpress) and Veeam Software's Veeam Backup & Replication. These products do VMDK-level full and incremental backups and file-level restores but work in very different ways, so you have to determine what works best for your environment.

VSPHERE AND HYPER-V SMOOTH THE VIRTUAL SERVER BACKUP PROCESS

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

In 2009, VMware introduced vSphere and the vStorage APIs for Data Protection (VADP), which have made virtual machine backup a much simpler process. Microsoft has meanwhile rolled out Hyper-V's backup architecture, to similar effect.

vStorage APIs for Data Protection allow users to carry out virtual machine backup without copying data to a proxy disk. Now a backup application can talk directly to the VMware VMkernel without scripts or agents mediating the process, and allow incremental or full backups directly from the virtual machine.

In 2009, VMware introduced vSphere and the vStorage APIs for Data Protection (VADP), which have made virtual machine backup a much simpler process.

The big four backup product vendors—CA, IBM/Tivoli, Symantec and EMC—now directly interface with vSphere as do Veeam and Quest Software, which have updated their products in the light of VADP.

Microsoft Hyper-V users only need to ensure the backup product knows it's talking to a Hyper-V server. This is similar to the operation of vStorage APIs, in that Hyper-V virtual machines can be backed up without guest-level backup inside the virtual machine.

The vStorage APIs for Data Protection also include vStorage APIs for multi-pathing, which provide for better storage I/O throughput and storage path failover for a specific storage array; and vStorage APIs for Site Recovery Manager, which enable array-based replication for block and network-attached storage (NAS).

vSphere 4 also introduced a number of other backup-related improvements over VMware Infrastructure 3. These include VMware Data Recovery (VDR), which provides snapshots and data deduplication; granular file-level

recovery without a two-stage process; changed block tracking (CBT), which allows the VMkernel to track the changed blocks on a virtual machine's virtual disk to allow only changed blocks to be copied; vCenter Server plug-ins, which provide monitoring and management views of a virtual machine's backup status; direct backup to shared storage; thin provisioning; and data deduplication. ⊙

Antony Adshead is the bureau chief for SearchStorage.co.UK.

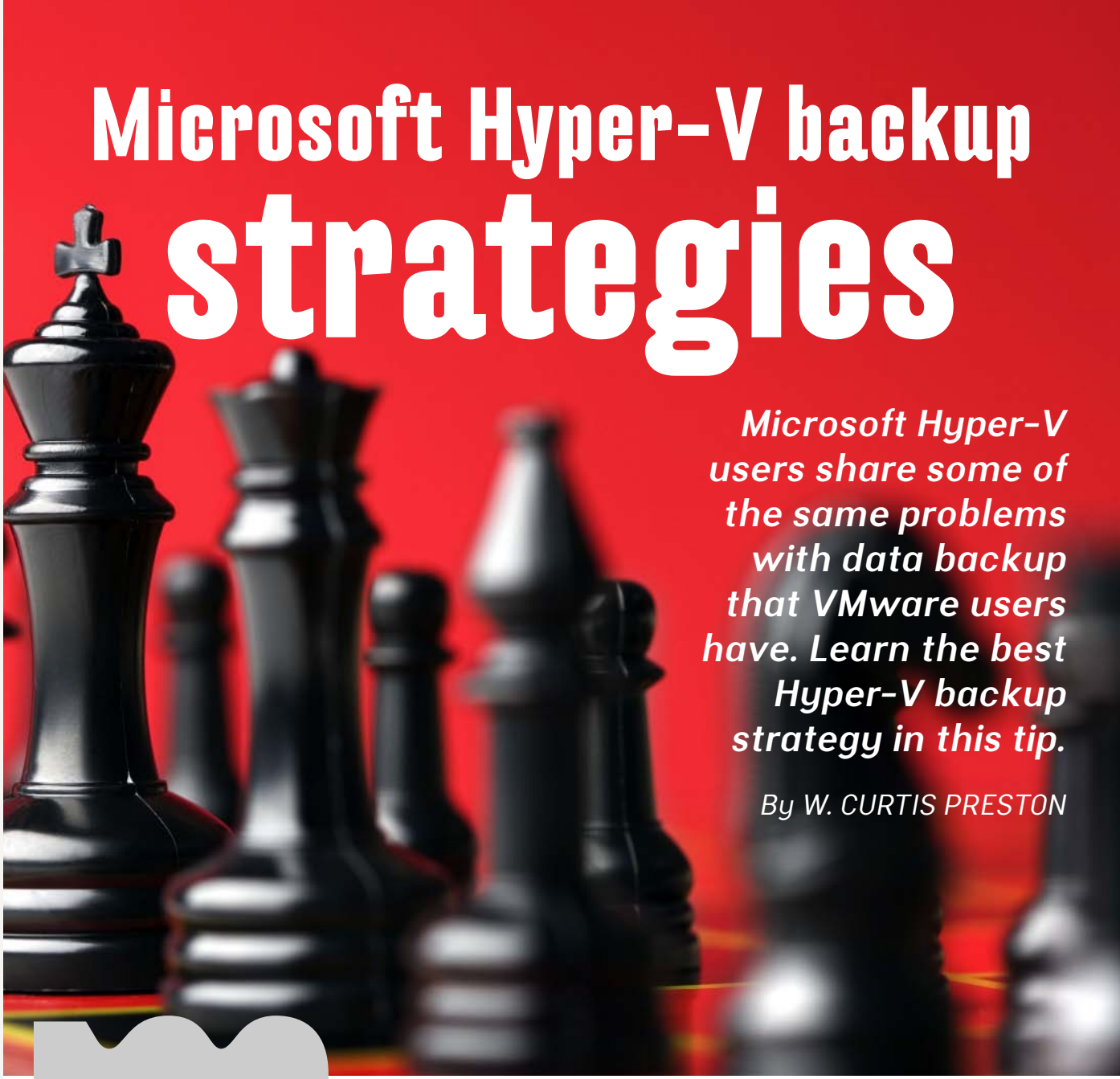
Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

Microsoft Hyper-V backup strategies



Microsoft Hyper-V users share some of the same problems with data backup that VMware users have. Learn the best Hyper-V backup strategy in this tip.

By W. CURTIS PRESTON

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

MICROSOFT HYPER-V USERS have the same core issue with backup and recovery that VMware users have—physics. When you move 20 or so physical servers into one physical server, many applications work just fine; however, one application does not—data backup. Backup is a very I/O-intensive operation that likes to take complete control over the CPU, memory and I/O resources of the server. That's why when you back up several virtual machines at one time, physics gets in your way.

One of the secrets to doing this right is to perform the backup at the Hyper-V level—not at the virtual machine level. A backup application that runs at the Hyper-V level can be cognisant of the underlying architecture of Hyper-V and behave accordingly.

HYPER-V BACKUP AND VSS

The challenge of backing up at the Hyper-V level is that the file systems on the virtual hard drives in the virtual machines are continually changing—and you need them to stand still while you back them up. The good news is that Microsoft already has a built-in framework for dealing with this issue, called Volume Shadow Copy Services (VSS). VSS can create a virtual snapshot of these drives and will give your backup system something stable and unchanging while it is backing the system up.

The challenge of backing up at the Hyper-V level is that the file systems on the virtual hard drives in the virtual machines are continually changing—and you need them to stand still while you back them up.

The VSS system consists of three parts: a requester, a provider and a writer. The requester is simply the application that requests that a snapshot be taken. The provider is what will create the snapshot. In a basic system, the provider is the Windows operating system itself, but in a larger system the provider could be a storage system that interfaces with VSS. Finally, each application wishing to be supported by VSS must create its own VSS writer that quiesces that application when it's time to create a snapshot.

For example, the VSS writer for SQL Server will put the database into a special mode prior to creating a snapshot. Once the snapshot has been backed up, the requester can tell the VSS writer that the backup was successful, and the SQL Server VSS writer will truncate SQL Server's transaction logs so they don't fill up.

This multi-tiered architecture of VSS allows the requester to control and back up applications without having to write a separate interface to those applications—it only needs to know how to talk to VSS. Hyper-V is really just another application running within Windows that can be controlled via VSS. Backup applications wishing to back up Hyper-V simply need to talk to the VSS infrastructure on the Windows server running Hyper-V. The VSS requester will then be informed that a Hyper-V VSS writer is present, and that it should communicate with it.

When requested to create a backup, the Hyper-V VSS writer becomes a VSS requester to the VSS systems in each underlying virtual machine

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

(VM). It finds out what writers are present in each VM, instructs them to do the appropriate thing, and creates a snapshot in each VM. Once all of this has been done, it can create the snapshot on the volume containing the Hyper-V virtual disk images. Once that snapshot has been created, it can tell the backup application to back up that snapshot.

OTHER HYPER-V BACKUP STRATEGIES: REPLICATION, CDP AND NEAR-CDP

Interfacing with VSS is not enough to solve the physics problem of backing up virtual machines. You also need to employ incremental-forever backup techniques. Full backups, however infrequent, can place quite an I/O load on Hyper-V and its underlying VMs. This is why you should strongly consider using a backup product that uses an incremental-forever backup technique. Replication, continuous data protection (CDP), and near-CDP are all examples of such technologies.

Another feature to consider is the ability to recover both an entire VM as well as individual files within that VM. Both recovery scenarios are very typical, so you want to make sure your backup product supports both image-level and file-level restores.

The most important piece of functionality you must look for in a Hyper-V backup product as part of your Hyper-V backup strategy is full integration with VSS. Once you have that, you should look for incremental-forever backup techniques and the ability to recover individual files as well as an entire VM. Once you have all that, all you need to do is argue over price. Good luck with that.

Brighter VM backup outlook

Simplification via CDP

Integrated backup

Hyper-V strategies

W. Curtis Preston (a.k.a. "Mr. Backup") is an independent backup expert and has been singularly focused on data backup and recovery for more than 15 years.