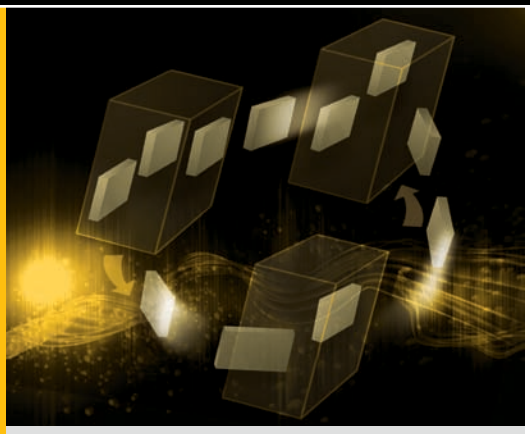

STREAMLINING DATA PROTECTION IN VMWARE ENVIRONMENTS USING COMMVAULT SIMPANA

By Brian Brockway
Zahid Ilkal

Using traditional backup methods in virtualized environments can quickly stretch processing, storage, and networking resources to their limits. CommVault® Simpana® software offers next-generation data protection designed specifically for virtualization, helping streamline backup processes, reduce storage requirements, and support rapid data recovery.

As virtualization continues to sweep through the enterprise—helping to consolidate server workloads and increase resource utilization throughout data centers, remote offices, and recovery sites—administrators are struggling to adjust legacy backup methods to this flexible, ever-changing infrastructure. Backup tools designed to protect traditional physical servers can introduce a variety of problems and limitations when deployed in virtualized environments, including taxing server processing resources, stretching networks and storage capacity to their limits, and slowing recovery times.

CommVault Simpana software, included in the Dell™ PowerVault™ DL2000 – Powered by CommVault, offers next-generation data protection designed specifically for VMware® and Microsoft® virtualization platforms to help overcome these challenges. This article outlines how following best practices for protecting VMware virtual machines (VMs) and taking advantage of key Simpana features such as incremental VM backup capabilities can help administrators streamline their backup processes, reduce storage requirements, and support rapid recovery of both individual files and entire VMs.

DESIGNING FOR VIRTUALIZED ENVIRONMENTS

Implementing data protection strategies in VMware environments presents a number of common challenges:

- **VM sprawl:** As VMs proliferate—often without the knowledge of data protection administrators—it becomes increasingly difficult to manage and apply consistent data protection services based on organizational needs.
- **Infrastructure overload from concurrent backup processing:** Not only can traditional backup agents be cumbersome to implement and maintain inside VMs, but the resulting concurrent backup operations can overwhelm host servers.
- **Balance between disaster recovery and granular recovery needs:** VMware Consolidated Backup (VCB) supports two distinct methods for backing up VMs: image level and file level. In large environments, however, having multiple policies for every VM can complicate management and increase backup media consumption.
- **Retention on and recovery from disk target:** Typical approaches to VM backup utilize the VCB image-level method to create full backups every day, resulting in an excessive volume of data transfer. The lack of incremental backup capability—a feature that is generally taken for granted for physical servers—can cause massive scalability problems. In backup-to-disk (B2D) configurations, the disk target can be exhausted rapidly, forcing backups to move to secondary and slower storage tiers. This in turn reduces the ability to perform fast

restores from disk and can negatively affect service-level agreements (SLAs).

Legacy backup software places a clear emphasis on the backup process, generally ignoring factors that contribute to scalability, resiliency, and operational efficiencies. Next-generation tools such as the CommVault Simpana Universal Virtual Server Agent (UVSA) provide a slew of advanced capabilities including automatic discovery, multi-streaming for concurrency, resiliency controls, incremental backup and deduplication for disk efficiency and rapid recovery, extended data retention and recovery options such as granular restore without restaging, and cross-platform restore. The UVSA is designed from the ground up to ease operational overhead for administrators; automate mundane, day-to-day tasks; and allow crucial resources to focus on business problems while providing extended disk-based retention and rapid restore capabilities to help meet aggressive SLAs.

CONFIGURING ROBUST BACKUP PROCESSES

In large environments, it is common for VMs to "float" across physical servers using advanced capabilities such as VMware vMotion™ technology, VMware High Availability (HA), and VMware Distributed Resource Scheduler (DRS). Integration with VMware vCenter™ Server (formerly VMware VirtualCenter) enables administrators to enumerate VMs as well as the VMware ESX servers and data stores hosting them.

Many backup solutions can integrate with vCenter Server during the initial backup policy configuration to locate and list existing VMs. In addition, however, the CommVault Simpana UVSA incorporates features such as automatic VM discovery to automatically assign new VMs to backup policies, backup process offloading, multiple backup options, and multi-streaming to help optimize backups in virtualized environments.

Automatic VM discovery

In large environments, new VMs can be added to the virtualized infrastructure almost daily, often outside the control of data protection administrators. Without automated tools, these administrators are left to manually hunt down new VMs and add them to the appropriate backup policy—a process that is both time-consuming and error-prone, potentially leaving VMs backed up multiple times or (worse) not at all.

The UVSA supports automatic VM discovery to automatically assign new VMs to backup policies based on predefined rules. Administrators create an initial set of backup policies based on desired protection levels, with discovery rules then applied that associate VMs with an appropriate backup policy. The discovery rules need to be defined only once, during the initial setup.

Backup process offloading

VMs are typically configured to balance and maximize physical server resource utilization for regular application workloads. A traditional backup model using an agent in every VM can quickly destabilize this balance. Scheduled simultaneous backups on VMs can consume all available ESX server resources, overwhelming the physical infrastructure and slowing production systems to a crawl.

While all VCB-based data protection tools benefit from some built-in VCB capabilities, VCB does have scalability constraints—and many ungoverned simultaneous VCB snapshot requests can cause failures that then lead to delayed or failed backups. The UVSA includes built-in controls designed to streamline simultaneous VCB snapshot requests and dramatically increase scalability and resiliency, even in very large environments.

Multiple backup options

VCB-based data protection tools typically require two backup policies for each VM: image-level backups for disaster recovery and file-level backups for granular recovery. However, this approach typically doubles the number of backups, increases storage requirements, and complicates recovery by requiring administrators to track two backup sets for each VM.

The UVSA offers three backup options to help balance disaster recovery with granular recovery and to help meet disk retention and recovery needs:

- **Disk-level backup:** This option backs up the full VM image, moving only the occupied portions of a VM image. (For example, for a VM image that is provisioned at 20 GB but consumes only 10 GB of space, the UVSA backs up only 10 GB.) Files and folders are indexed during backup, allowing granular recovery from an image backup. In addition, true block-level incremental (BLI) capability backs up only changed extents of a VM image, helping limit the amount of data that must be transferred and stored.
- **File-level backup:** This option enables backup of individual files within a VM. A separate file-level backup policy is useful when administrators need to separately retain both a disaster recovery copy and a subset of files in the VM. In addition, the ability to index files inside the VM also makes it available for search and e-discovery—a key capability as an increasing amount of critical enterprise data shifts to VMs, given that other e-discovery solutions typically require agents inside VMs that then complicate deployment, management, search, and discovery processes.
- **Volume-level backup:** This hybrid option takes advantage of the VCB file-level backup mode to mount the file system instead of copying the full image over, and then backs up blocks of the file systems for an image-like backup. Recovery options include the ability to restore individual partitions to either a physical system or a VM running on VMware or Microsoft Hyper-V™ platforms. This option also supports BLI backup.

Multi-streaming

VM proliferation can stretch the limits of already-shrinking backup operating windows—and sequential VM backup processing combined with full image processing of

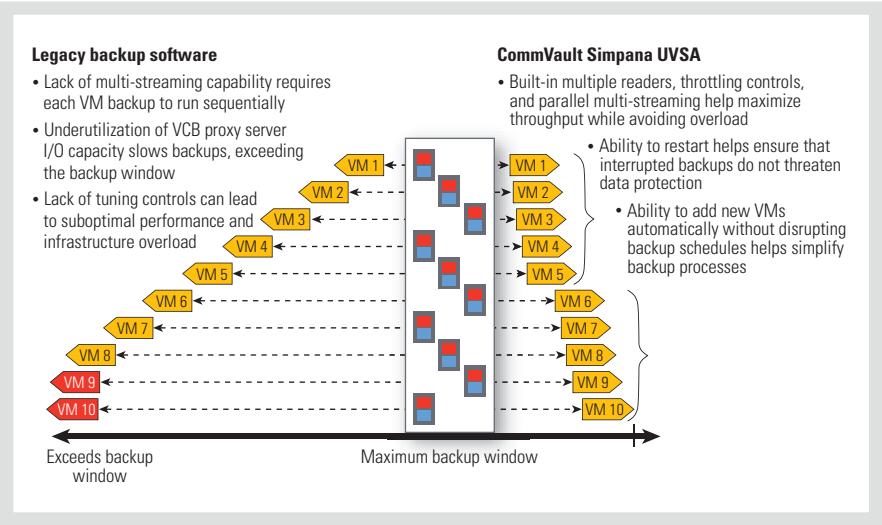

Figure 1. *Multi-streaming helps accelerate backups and avoid infrastructure overload*

every backup is almost certain to exceed operating windows. In these environments, multi-streaming is virtually a necessity for concurrent backups.

Uncontrolled concurrency, however, presents a different set of challenges. Simultaneous backup activity kicks off a period of intense processing, coordination, and movement within a confined operating window. In a virtualized environment, the shared operations and resources yield a large number of variables and conditions that can contribute to failure, rendering the diagnostic process long and tedious and making concurrency a critical unknown factor.

Figure 1 illustrates the results of running a set of backup jobs over 10 individual VMs using a single backup policy, highlighting the operational challenges of consolidated backup policies. The profile on the left, a traditional file system backup policy extended to support full VCB backups, performs sequential processing of each VM and therefore exhibits a much longer overall backup window than the profile on the right. Although the method on the left can help reduce the amount of concentrated load on the system, even underutilizing available resources on the physical server and backup targets, it also risks missing backup windows and exposing the environment to data loss.

The profile on the right uses multi-streaming to support the same set of full VCB backups in significantly less time. Because ungoverned parallel jobs can overload the infrastructure and cause system errors, the UVSA offers two controls to mitigate this problem and help ensure optimal performance. First, administrators can set the number of VMs that can be processed concurrently within a given policy. Second, the UVSA includes a built-in time delay between successive VCB snapshot calls, which helps limit the impact on the system. As each VM backup completes, resources are reapplied to the next VM in the policy. This approach imposes an operational governor across the overall process, helping minimize the impact of unknown conditions through the standard Simpana job management resiliency and job restart features.

OPTIMIZING DATA TRANSFER AND RETENTION

Environments using VCB typically utilize the image-level mode with granular recovery options. As more servers are converted to VMs, therefore, the total volume of nightly data transfer grows in full VM size increments, which can have a major impact on scalability. For example, under a typical schedule of weekly full backups and daily incremental backups, 50 traditional servers averaging 20 GB of data each with a 10 percent daily incremental change rate would require 1,600 GB of data transfer each week. If these servers were converted to VMs in an environment that supports only full VM image backups, the need to back up all 50 VM images on a daily basis would result in 7,000 GB of data transfer each week—an increase of 430 percent. In many environments, this massive inflation can quickly break the networks and storage budgets.

The BLI capabilities provided by the CommVault Simpana UVSA allow administrators to back up only the changed sections of a VM, enabling them to continue following weekly full backup cycles as physical servers are converted to VMs and helping to limit nightly data volumes to comparable pre-virtualization levels. Supporting continued use of existing infrastructure helps limit the need to invest in additional tier 2 disks or tape drives as part of the virtualization project.

Figure 2 illustrates the advantages of BLI backups and the trade-offs imposed when such features are lacking. As with the previous example, this environment has 50 VMs averaging 20 GB each, for a total of 1 TB of data. A typical B2D profile would include two to three times this production disk space for backups and a 30-day disk-based recovery point objective (RPO), or 30-day data retention; reflecting this typical profile, administrators have provisioned 2–3 TB of B2D disk space.

The workflow on the left reflects legacy backup software that does not support incremental VM backups, requiring daily full backups of each VM. In addition to stretching the limits of data transfer capacity each day, this example would fit only two days' worth of backups on the B2D disks—2 TB for two days of full backups, and 0.5 TB of room to recover granular file-level data from previous backups (a typical amount depending on VM size). Meeting the 30-day RPO would typically force administrators to store data older than two days on tape, add disk capacity, or invest in a deduplication appliance to help reduce storage requirements.

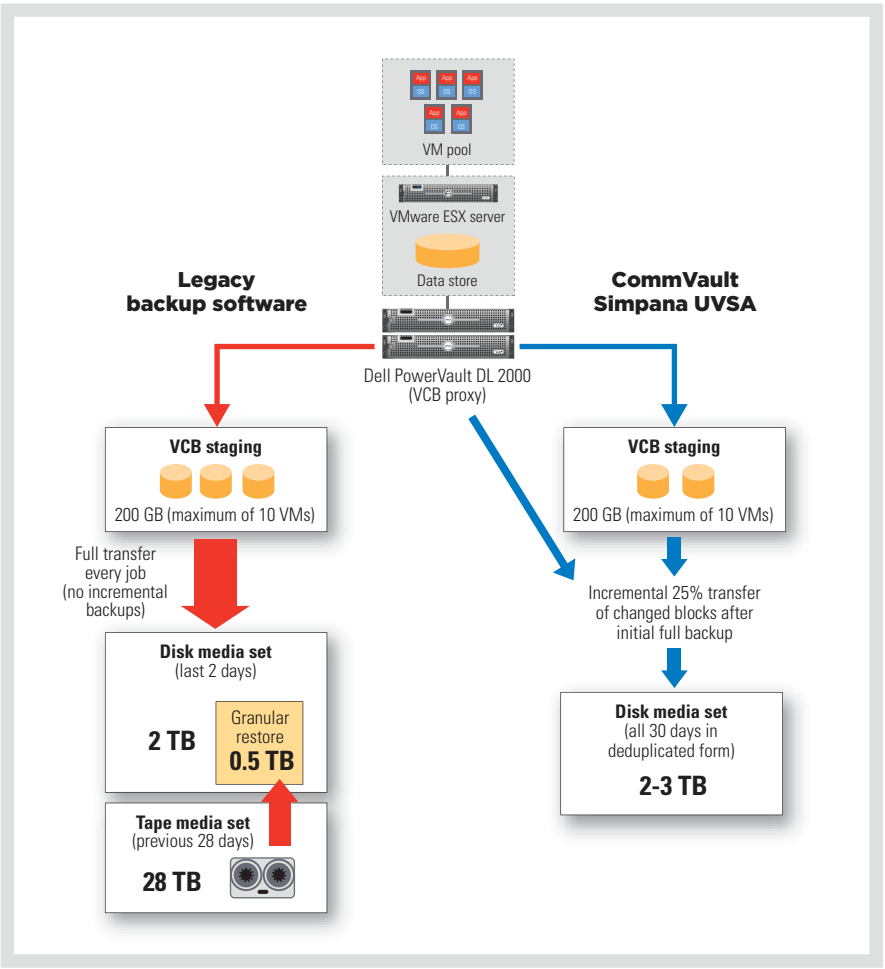

Figure 2. *Block-level incremental backup helps reduce storage requirements and accelerate recovery*

The workflow on the right illustrates the advantages of the BLI approach. In this case, the first full backup transfers the equivalent of 1 TB of data. Then, even assuming a high daily change rate of 25 percent, subsequent backup jobs incrementally transfer changed segments equivalent to 0.25 TB each day. Assuming a typical software compression of 50 percent, the 2–3 TB target can easily accommodate 10–14 days' worth of recovery copies. Adding built-in, in-stream, block-level deduplication would help reduce the volume of data transfer and increase disk utilization even further. With common workloads, and assuming a low 3:1 deduplication ratio, this solution can easily support 30 days' worth of backups on disk, helping meet the 30-day RPO for rapid recovery.

ENABLING RAPID DATA RECOVERY

Administrators often struggle to find the right balance between disaster recovery and granular recovery policies at acceptable performance levels. For example, as shown in Figure 2, a lack of incremental backup capability can quickly force data to tape. In this case, recovering files older than two or three days would require staging an entire VM image on disk—a process that could potentially take hours, defeating the purpose of granular recovery. Administrators face a difficult trade-off: investing in a large disk target to help maintain acceptable granular recovery performance, or using cost-effective tape media that can compromise recovery speed and SLAs.

The CommVault Simpana UVSA helps eliminate this decision. By enabling administrators to maintain a B2D capacity comparable to that required in a pre-virtualization environment while still meeting RPOs, it helps avoid the need to push data off to tape media and supports rapid granular recovery.

DEPLOYING FLEXIBLE DATA PROTECTION FOR VMs

Although virtualization can provide a range of benefits in enterprise data centers, continuing to use legacy backup tools not designed for these types of environments can lead to increased storage costs for data retention and slow, inflexible recovery. CommVault Simpana offers next-generation data protection designed for virtualization—helping organizations keep pace with the emerging challenges of disparate systems, rapid data growth, shrinking operational windows, and tightened recovery objectives.

Brian Brockway is vice president of product management at CommVault, and has over 10 years of experience in the software and storage industry. He has a B.S. in Aerospace Engineering from the University at Buffalo and an M.B.A. from New York University.

Zahid Ilkal is a senior product manager at CommVault, and has over 10 years of experience in the storage industry. He has a master's degree in Computer Science from the University of Pune and an M.B.A. from New York University.

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Dell and CommVault:
DELL.COM/CommVault

Dell PowerVault DL2000:
DELL.COM/DL2000