Overcoming data protection challenges in virtualized environments

By Scott Herold

Relying on traditional backup approaches in virtualized environments can be complex and time-consuming. By adopting image-based data protection technologies, IT staff can create a simplified, streamlined, reliable system for efficiently protecting and recovering virtual machines.

When virtualization was first moving into the mainstream, IT groups were largely concerned with "Step 1" problems—determining how to plan and deploy virtualized servers to consolidate physical infrastructure and reduce costs. These problems were generally straightforward to solve, and many organizations were able to achieve the types of consolidation benefits that they expected.

Now, however, these IT groups are deep into the challenges of "Step 2"—operating those environments efficiently. And the problems that have come up are not so simple to solve. Virtual machine (VM) sprawl, for example, took many IT teams by surprise. Setting up a new VM is significantly faster, easier, and more cost-effective than requisitioning an entire physical system, which is the analogous process in a traditional non-virtualized environment. Application owners can therefore obtain the necessary development, quality assurance, and production application servers relatively easily, which can quickly lead to an explosion of VMs. In a data center that has consolidated 50 physical servers to 4 virtualized servers, IT staff might expect a corresponding reduction in their administrative burden. But because the operating systems themselves don't go away, the workload actually continues to grow as these additional VMs are created. Six months in, this environment might have 100 VMs—or what amounts to double the workload.

Another Step 2 problem—and the one that is the focus of this article—is data protection. It's no surprise that many organizations still protect data in VMs as if they were working with physical systems, by deploying backup agents into the VMs and then using those agents to back up and recover files while often doing nothing to protect the VM image itself. But by instead taking advantage of image-based backup technologies, these organizations can overcome the numerous problems with these traditional backup approaches and create a simplified, streamlined, reliable approach to data protection in virtualized environments based on Dell™ systems.

Challenges of traditional backups

Traditional backup approaches are less than ideal in virtualized environments. Performing a complete system recovery using file agents in a disaster recovery scenario, for example, can be...
difficult and extremely time-consuming. In addition, because of the shared resource model of virtualization, scheduling backup jobs can be tedious or, in some cases, impossible. Backup administrators must identify the servers hosting each VM—which may change in real time in environments using live migration technologies—and ensure that only a certain number of backup jobs run at one time on a given server. When backup agents are individually deployed in every guest OS, the backup job is consuming underlying system resources the entire time it is running—slowing down not only the VM being protected, but also other VMs on the same server. Running simultaneous backup jobs on different VMs on the same server is generally out of the question.

Then there is the question of how backup data is moved from the virtualized server to storage. In a physical environment, a separate backup server is typically attached to each client to move that data. This approach can still work for virtualized servers—but all backup data must be sent over the LAN, which can disrupt use of that network for other users. To combat this, some organizations may use a separate backup network, which increases the cost and complexity of the infrastructure. Another option has been to use VMware® Consolidated Backup (VCB), but this feature requires a storage area network and has been superseded by the VMware vStorage application programming interfaces (APIs) introduced with the VMware vSphere™ release.

vStorage offers an enhanced way for backup software to capture VM data, by working with the VM image through a published VMware API (see Figure 1). Image-based backup software, such as the Vizioncore® vRanger™ Pro Data Protection Platform (DPP) solution for image-based backup and recovery, can offer major advantages in virtualized environments. It can accelerate the process of capturing data, helping reduce the impact on the system and the organization as a whole. It can fill a critical gap by helping protect the full VM image as well as the individual files within that image. It can help speed recovery and increase recovery reliability for individual files as well as for the entire image. And it can help increase data transmission speed—it’s generally far faster to transmit the whole image instead of the many individual files that comprise the image.

Advantages of image-based data protection
When designed well, image-based backup can provide a more reliable method for handling data than traditional methods. What IT teams worry about is the integrity of the backup data copies held in the archive, and creating those copies depends on a continuous, uninterrupted write of the backup data copy. Another concern is the integrity of the application data captured,

Putting image-based data protection to work
Image-based data protection can provide significant advantages over traditional backups in a wide range of typical IT scenarios.

- Rapid, efficient backup and recovery of large VM images, such as those that exceed 10 GB in size
- Support for local high availability of VMs by preserving them and then restarting them on alternate virtualized servers in the environment
- Disaster recovery through the transmission of VM images to an off-site recovery location, in which case recovering both the data and applications occurs in the single step of launching the replica VMs on different servers
- Rapid recovery of entire servers, especially those hosting business-critical applications such as customer relationship management, payroll, and accounting systems, in which case the ease of recovery—without requiring separate steps for operating systems and applications—can make image-based approaches invaluable
- Rollback of database systems to a point in time just before drive resizing and other optimization tasks—just in case
- Protection for end users and customers against disk problems, data corruption, and the troubleshooting steps required for repairs, replacements, and bringing storage environments back to operational readiness
- Rapid recovery from human errors such as storage provisioning and reallocation by mistakenly deleting logical units (LUNs) that contain critical data—a situation that can occur frequently when virtualized servers are underprovisioned and inexperienced administrators attempt to find space for backups, or when the steps of an infrastructure upgrade or cutover are performed out of sequence

Reprinted from Dell Power Solutions, 2010 Issue 2. Copyright © 2010 Dell Inc. All rights reserved.
which must be consistent at a given point in time to be recoverable and usable from that point.

Both problems can be alleviated with a well-designed image-based data protection system. First of all, this type of system can be extremely fast—incurring minimal overhead by bypassing the file system to directly read a disk, in addition to skipping empty blocks in the image. This approach also creates a smaller backup copy—in a single file—than would be created by the traditional method of creating a backup copy of the thousands or millions of individual files that make up the image. Faster means that there is less time for something to go wrong, while smaller means that the backup copy can be transmitted more quickly and stored using fewer blocks than it could otherwise—again helping reduce the risk of something going wrong.

In terms of application data consistency, because capturing an image that includes all of an application’s data from a single point in time is the definition of an image-based backup, this approach helps significantly reduce the odds of backing up data at inconsistent points. Even in the unlikely event that the data is not consistent, administrators could still recover and restart the application by rolling back to the nearest previous consistency point, which is captured in the backup image.

When asked to comment on image-based data protection, Lauren Whitehouse, a senior analyst from the Enterprise Strategy Group (ESG), agreed that this technology is the one most likely to help improve backup and recovery in virtualized environments. Citing ESG’s 2010 IT spending survey, in which over 500 respondents reported on their budgets for investing in virtualization technology this year, she said, “ESG research found that implementing data protection processes for server virtualization environments is a big pain point for many end-user respondents. These organizations have committed to fixing the issue this year with increased investments in solutions that improve backup and recovery of virtual machines, improving disaster recovery processes, and improving application backup and recovery.” She continued, “For x86 server virtualization environments, rapid image-level backup with flexible image- or item-level restore can address all of these challenges. It can provide nondisruptive and optimized backup, enable efficient disaster recovery strategies, and facilitate improvements in application-specific backup and recovery.” (For more examples of the benefits of image-based data protection in real-world scenarios, see the “Putting image-based data protection to work” sidebar in this article.)

Streamlined data protection for virtualized environments

Vizioncore has introduced its Backup 2.0 campaign to help organizations understand how image-based data protection can simplify their operations. The sooner that IT teams begin to adopt image-based methods for VM backups, the sooner they can overcome the Step 2 challenges of protecting their virtualized Dell servers—a critical step toward gaining the full benefits of virtualization.