Quantum®

When Traditional Data Protection Approaches for Virtual Environments Fall Short

Challenges Associated with Using Traditional Backup Methods to Protect Virtual Data

As server virtualization continues to make inroads into primary business applications, customers are experiencing the challenges that dynamic virtual environments can create. A recent study showed that nearly 1 in 2 IT managers are experiencing huge amounts of data growth associated with their virtual environment. This growth creates the need to re-evaluate existing data protection methodologies for their virtual environment, and in some cases causes customers to start thinking about how they are going to protect their virtual data infrastructure.

In looking at data protection for virtual environments, customers can be very quick to ask themselves, "Why can't I use my existing backup application to protect my virtual data?" In theory, they can, but there are some distinct challenges associated with using traditional backup applications to protect this data:

1) Increased Cost and Complexity

Most backup applications are sold on an agent basis, so licensing costs can rise very quickly, particularly as more hosts and virtual machines (VMs) are added to the environment. With this licensing model, some VMs might get missed altogether. If there is no existing agent on that host server where the VMs reside, and since VMs are constantly moving around, the backup application will not back up that host machine, so an agent on every host machine is a necessity. This methodology adds overall cost and complexity to managing the environment. Agentless tools provide much more flexibility and can ensure complete backups of all the VMs and the environment as a whole.

2) Performance Degradation

Performance issues can quickly surface when using traditional backup applications for virtual environments. Primary applications typically work better in virtual environments—and with other VMs—as their workload is variable by nature. This helps to eliminate resource contention with other VMs on the host machine and the entire infrastructure. But backup applications have scheduled times that run the jobs with limited flexibility, so multiple jobs running on the same host can place great strain and fight for resources on the host system causing overall performance issues, and subsequently extending the backup window. IT admins are quickly gravitating to VM specialty tools that work in the virtual infrastructure.

3) Lack of File-Level Access Provided by the Backup Application

Backups are about as sexy as an insurance policy, but restores are where the magic has to happen. Traditional backup apps in the virtual space inherently may have some difficulty providing file-level access during restores, causing customers to do complete restores of a backup set versus doing file- or object-level restores. For traditional backup apps it is easier to back up the whole image of the VM, particularly those VMs that are over-provisioned. This process limits restorability and can cause the user to restore the entire VM and then view the files inside the VM, making restores difficult and slow. Some tools back up guest file system information allowing individual file-level restores. This is great, but it is important to understand the complexities and processes required to restore those individual files, because it could be cumbersome and very time-consuming. A better approach is to use a VM backup tool that backs up both the whole VM and guest file system, providing maximum restore flexibility. The system administrator can either restore the specific files within the VM or restore the entire VM. Ideally, the backup application presents the virtual data in native VMDK format. This provides flexible backup and DR targets (disk, tape and cloud), and also limits vendor lock-in because there's no need for the backup application to restore from the backup medium DR location. This also allows for faster and easier restores of the data itself—something everyone can appreciate.

4) Handling Additional Data Generated by Virtual Machines

The final challenge in using a traditional backup application for virtual environments is handling the extra data normally associated with VMs. Virtual machines are typically over-provisioned for a number of beneficial business reasons. However, this creates a lot of empty space that does not need to be protected and more importantly, this empty data can take up precious resources. Dead data is always around with VMs: expired files, re-do logs, unassigned data, etc. Change Block Tracking (CBT) from VMware does a great job of eliminating old data; however, it keeps a lot of other data no longer needed by the file system. Many backup applications back up everything and constrain already taxed resources, consequently increasing the backup window—or worse, harm primary business application functions. Thus, it's better to have a backup app that provides a "clean" image for the backup target. This clean image provides massive data reduction, and relieves resource utilization across the network, host machine, and storage. This also minimizes the storage footprint regardless of the backup medium chosen—disk, tape or cloud.

Considerations When Shopping for Virtualized Data Protection

When shopping for a virtualization data protection solution, there are a number of considerations to think about: the importance of the backup format, how the solution integrates with other technologies, and manageability within a virtual environment.

Most backup applications generate files in proprietary format. This could be an issue as only this backup app can access the data going forward. This also slows restores and locks the customer into that specific backup application. But probably most important, this proprietary format reduces backup and DR choices. Customers are looking at how they can integrate disk and even cloud into their backup methodologies, so a backup application that uses a proprietary format will increase costs and the complexity of managing a DR solution—including backing up to the cloud—by requiring extra licenses at the DR site. Backup in original or "native" format is certainly more useful for today's media. This also allows for more flexibility, providing faster and more flexible restores, and giving more freedom for backup and DR targets. Imagine not needing to have the exact same hardware and software configuration at the DR site to restore data or even to run an environment from a DR site. Protecting data in native format allows administrators to do things previously unavailable with their data associated with disk, and even with cloud protection schemes.

Another consideration is the integration with other technologies that the VM backup software can provide. Customers looking to run two different backup apps can incidentally create more cost and complexity for themselves by deploying a specialty tool. It is best to have a product that provides unique integration with existing investments in current backup applications and associate processes.

It is important to know if a VM tool works only with its own deduplication solution or if it can use a secondary technology for this function. By combining software technologies with best-of-breed deduplication, customers can see dramatic improvements in their environment.

When deciding which VM data protection route to take, consider the manageability of the VM backup application and its integration with the VM environment. Taking advantage of virtual appliance deployment is something that customers can do with most VM backup tools. This creates easy delivery of the tool as a complete package in a simple VM. This can also reduce overall cost, as only the existing infrastructure is required, with no additional hardware for a host. As a virtual appliance, the tool can take advantage of VM functions such as vMotion and vStorage Motion to optimize performance. Seamless integration with VMware is essential. The VM backup app should work directly with vCenter, allowing the tool to auto-discover any new VMs and locate those that have moved to a different host since the previous backup. VMs are constantly moving around and constantly being created, so a tool that can go out to vCenter and auto-discover the new VMs and locate those that have moved to a new host gives comfort that all the data will be protected. The VM backup app should also provide native file system view of the VMs and any backup copies. This provides maximum flexibility in backup and DR choices and reduces vendor lock-in.

Questions to Ask a VM Specialty Vendor:

- Can the VM backup utility work directly with the traditional backup application?
- Can the VM backup tool allow the traditional backup app to see native file system structure inside the VM package, giving easy and fast restores?
- Can the VM backup application take advantage of data reduction techniques such as CBT and any additional cleanup?
- How well does the VM backup app integrate with existing processes, retention and management rules that are already in place with existing infrastructure?

The good news is there are VM backup specialty tools available today that meet these criteria and that can be adopted without disrupting existing infrastructures. As virtual environments gain momentum, backup administrators will be glad to know there's a suitable data protection solution for them.