Five Fundamentals of Modern Data Protection

David Davis

vExpert

Veeam Backup & Replication 6.5
More ways to WOW!

Veeam provides **powerful, easy-to-use** and **affordable** data protection that fully leverages the virtual environment and eliminates the need for agents.

Version 6.5 includes:

- Free e-discovery and item recovery for Microsoft Exchange
- Easy VM recovery from SAN snapshots
- New hypervisor support: vSphere 5.1 and Windows Server 2012 Hyper-V
- 50+ other new features and enhancements

» **Free Download**

For most company executives (including CFOs and many CIOs), the protection of company data has become a "given". Because backup software has been purchased and IT staff salaries are being paid, the assumption is that your critical data and the applications that manage it are protected. Sadly, in many cases this is far from the truth.

Protecting your company's applications and data is just as important today as it was 10 years ago. However, it's even more complex than in the past. With the introduction of virtualization, cloud computing, and many more applications, the aging data protection technology in use by most companies simply isn't adequate. This results in unnecessary risk for the company and unnecessary pain for IT administrators.

What you'll learn from this whitepaper is that data protection has changed and modern data protection software is now available that can much more reliably and efficiently protect your company's most critical applications and data.

Don't take data protection for granted. Continue reading to learn about the five fundamentals of modern data protection:

1. Use Data Protection Software Built for Virtualization

There are hundreds of data protection tools out there, but few of them are "virtualization-savvy". Legacy data protection tools tend to see every "server" the same way – as a physical server. By incorrectly assuming that all servers are the same, tremendous inefficiencies occur when you attempt to backup or recover applications and data. For example, lengthy file-based backups are performed when only small blocks of data have changed.

Data protection tools that are built for virtualization can talk directly to the virtual infrastructure. Through this communication, these data protection tools gain:

- Knowledge of virtual machines (VMs) and the hosts they run on
- Knowledge of virtual storage to know what needs to be backed up
- Ability to snapshot VMs and back them up with no downtime
- Ability to backup only those VM disk blocks that have changed to drastically reduce backup times and the amount of backup data (referred to as "changed block tracking")

This interaction also becomes a gateway to additional features that you might not expect from your virtual environment, much less from your backup software, including:

- Creating virtual lab environments where backups can be automatically tested or used to selectively recover application data
- Virtualizing the recovery process to make failed VMs available quickly

Figure 1. Numerous restore options are available for virtual machines.

Yes, some legacy data protection tools have been adapted, over time, to recognize the virtual infrastructure, but none was built specifically with virtualization in mind. With more and more of your servers being virtualized, the smart play is to select a data protection tool that is built for virtualization and, because of that, can offer you the most flexibility, functionality and efficiency.

2. Select an Agentless Data Protection Solution

Legacy data protection tools require you to install one or more agents on every VM that you need to protect. However, there are downsides to agents, including:

- They require a new piece of software installed on every VM. If a VM doesn't have the agents, it is unprotected and vulnerable to data loss.

- They can create conflicts with other applications.

- They are difficult to manage, and it's difficult to see which VMs do or do not have agents installed.

- They utilize CPU and memory on every VM.

The bottom line is that agents are inefficient, across the board.

Some data protection software vendors will say that they are "agentless" because they can do an agentless backup. However, many of these vendors require agents for file-level restore, proper application backup, or to restore application data. My advice is to make sure that your data protection tool is able to address all backup and recovery scenarios without the need for an agent.

Data protection tools that are built for virtualization can go directly to the hypervisor host or virtual infrastructure management system (vCenter Server or SCVMM), find out the names and locations of virtual disks, and then backup those VMs – all without agents.

That means that you won't have to install and maintain agents on each VM being protected, your VMs will run more efficiently, and you'll gain the maximum benefit possible from your virtual infrastructure.

When it comes time to restore a VM or files inside a VM, again, no agents are needed. As you see in Figure 1, there are numerous types of restores available – and none requires an agent.

3. Leverage a Layered Approach to Data Protection

Gone are the days of simply backing up your data to tape (and then hopefully storing that tape offsite). Modern datacenters use a layered approach to data protection.

This layered approach could include:

- Local backup to disk
- Storage-based snapshots
- VM replication – either onsite or offsite
- Archival to tape or cloud storage

The goal is to protect your applications and data in as many ways as are available, and to make restores as fast and easy as possible.

		Strength	Weakness
VM	Onsite backup	Streamlined recovery (access backups from disk)	Does not protect against site outages
	Storage-based snapshot	Frequent restore points	Does not protect against storage failures
	VM replica (onsite or offsite) VM	Fast recovery (failover to standby VM)	Cost of infrastructure to host
	Offsite backup	Protects against site outages, long retention	Slower recovery (takes time to retrieve)

Figure 2. A layered approach to data protection leverages new and evolving technologies to achieve the best RTOs and RPOs for all recovery scenarios.

Benefits of a layered approach to data protection include:

- Immediate access to backups for instant restore of entire VMs, individual files or application data
- Proven recovery by mounting local backups anytime for automated backup verification and disaster recovery testing
- Frequent restore points and ability to meet the most aggressive recovery point objectives (RPOs)
- Automated off-site backup and VM replication for disaster recovery (DR)
- Long-term archival to tape or the cloud for a final layer of data protection for peace of mind and audit requirements

4. Reduce Backup Data with Deduplication

It's a fact that the cost of data protection will vary greatly depending on the number and size of VMs you have to protect. To optimize your investment in data protection, you must take advantage of technology that reduces the size of your backups. Traditional data protection simply backed up raw data or, if you were lucky, compressed the data.

One of the most common ways to reduce backup size is to use deduplication. Modern data protection tools automatically perform deduplication as well as compression. Deduplication identifies identical data blocks in source VMs and stores each unique block only once. Since image-based backups used in virtualization capture the entire VM, including the guest operating system (OS), and since the OS is often the same between VMs, there tends to be a lot of duplication. By using deduplication, you can tremendously reduce the size of the backup repository, the time to backup VMs, the amount of backup data replicated offsite, and the amount of data sent to tape or cloud storage.

Don't attempt to protect your virtual infrastructure without using a tool that includes deduplication. Other space-saving features to be on the lookout for include:

- **Forever incremental backup** – performs an initial full backup but then performs incremental backups, forever. As incremental backups are performed, the full backup is updated with the changes to create a complete backup image of the VM that is always ready to be restored.

- **Support for hypervisor thin-provisioning** – your backup tool should understand that your hypervisor can create thin-provisioned virtual disks. It should support thin-provisioned disks and maintain thin-provisioning throughout backup, restore and replication operations.

- **Exclusion of unneeded data** – modern data protection tools should recognize that your VMs are made up of special files such as the configuration file, swap files, snapshot files, and the virtual disk. Not all of those files need to be backed up. You need to be able to specify which of these files you want to exclude from the backup to save time, network bandwidth, and space in the backup repository.

5. Select One Solution for Multiple Hypervisors

The release of Windows Server 2012 Hyper-V has enterprises of all sizes excited about using its advanced features for a very affordable price. However, most enterprises still want to run their tier-1 applications on VMware vSphere. By using both hypervisors in the datacenter, enterprises can significantly reduce the amount they spend for virtualization software. Whether you are using multiple hypervisors in the datacenter or not, it's smart to keep your options open.

By selecting a data protection tool that can protect both VMware vSphere and Microsoft Hyper-V, you are choosing the tool you need for a multi-hypervisor / tiered-hypervisor infrastructure.
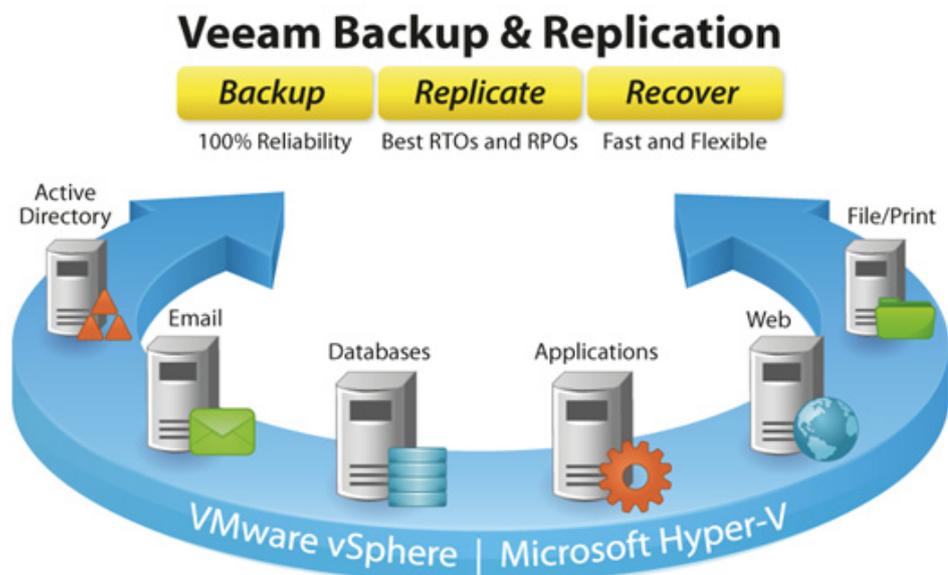
Figure 3. Veeam Backup & Replication provides modern data protection for both VMware and Hyper-V.

Summary

The job of managing a datacenter is all about "keeping all the balls in the air". With a constant flow of new applications, new technologies, and new problems, it's easy to take core datacenter responsibilities, such as data protection, for granted. In too many instances, datacenters stick with "the devil they know" and continue to use the same legacy backup application that they have used for the last 10 years. The problem with that route is that it leaves IT administrators inefficient, end users with more downtime and data loss than necessary, and the company with potential data protection holes. Modern data protection knows virtualization, is agentless, uses a layered approach, includes features to reduce the size of the backup repository, and protects VMs on the most popular virtualization platforms. I recommend evaluating modern data protection tools for use in your datacenter, today!

About the Author

David Davis is the author of the best-selling VMware vSphere video training library from TrainSignal.com. He has written hundreds of virtualization articles on the Web, is a VMware vExpert, VCP, VCAP-DCA, and CCIE #9369 with more than 18 years of enterprise IT experience. His personal website is VMwareVideos.com.

About Veeam Software

Veeam® Software develops innovative solutions for VMware backup, Hyper-V backup, and virtualization management. Veeam Backup & Replication™ is the #1 VM Backup solution. Veeam ONE™ is a single solution for real-time monitoring, resource optimization, documentation and management reporting for VMware and Hyper-V. Veeam extends deep VMware monitoring to Microsoft System Center with Veeam Management Pack™ (MP), and to HP Operations Manager with Veeam Smart Plug-In™ (SPI). Veeam also provides free virtualization tools. Learn more by visiting www.veeam.com.