

MORE ABOUT ERIC SIEBERT
Click here to read more about the author

MORE ABOUT VEEAM SOFTWARE
Click here to read more about the sponsor

written by Eric Siebert

Top 10 BEST PRACTICES for VMware Data Protection

VEEAM
www.veeam.com



© 2011 by Veeam Software and Eric Siebert

Copyright holder is licensing this under the Creative Commons License, Attribution 3.0.
<http://creativecommons.org/licenses/by/3.0/us/>

Please feel free to post this on your blog or email it to
whomever you believe would benefit from reading it.

THANK YOU!

Contents

<i>Don't back up VMs at the guest OS layer</i>	5
<i>Leverage the vStorage APIs</i>	6
<i>Virtual machine quiescing and VSS for ultimate protection</i>	7
<i>Don't under provision target backup resources</i>	8
<i>You don't need expensive storage hardware to achieve near-CDP</i>	9
<i>Getting the most out of your backups – putting your backup repositories to work for you</i>	10
<i>Use data deduplication and compression to reduce backup size</i>	11
<i>Verify your backups to make sure you can properly restore them</i>	12
<i>Granular recovery from an image level backup</i>	13
<i>Avoid the sticky situation of assigning permissions to all your sensitive data to backup administrators</i>	14



Summary

Backing up your virtual machines (VMs) may seem like a simple and easy process but there's a lot more to it than meets the eye. Backing up physical servers is fairly straightforward; you just install an agent on a server and add it to the backup schedule. But to efficiently backup VMs you need to use techniques and features that are designed specifically for virtual environments. If you treat VMs like physical servers when you backup and restore them, you waste resources and make backup windows longer than they need to be. Virtualization is a game changer in the data center, and once implemented, it requires that you change your procedures and methods to leverage its strengths and unique architecture.

The virtualization architecture offers many advantages to server backup and recovery operations. It changes the traditional techniques used to back up servers by leveraging virtualization features to streamline backup and recovery and make it more efficient. It also provides more flexibility and options for performing backups, restoring VMs and implementing disaster recovery. In this white paper we will offer 10 tips for assisting with the implementation of backup and recovery in a virtual environment. This includes using the proper methods, techniques and configuration as well as leveraging the features built into Veeam Backup and Replication™ version 5 to take your backups to the next level.

1

Don't back up VMs at the guest OS layer

“As a result more VMs can be backed up simultaneously and the host has more resources available for its VMs.”

When it comes to the method you use for backing up VMs, you shouldn't stick with the same method that you used to back up your physical servers. Physical servers are traditionally backed up using an agent installed in the guest operating system (OS) of the host, and the backup server connects to the agent in order to copy the data from it. This method will still work on a VM but ignoring the virtualization layer when you perform backups is inefficient and a waste of valuable host resources. The best way to back up VMs is at the virtualization layer, and for that you need a backup application that is built and optimized for virtualization.

By being virtualization-aware, a backup application doesn't have to involve the guest operating system of the VM in the backup. Instead, the application can connect

directly to the VM's disk file to back it up. Doing this means that there is no resource overhead on the VM while it is being backed up so workloads will not be affected while backups are running. Additionally, depending on the backup method used, it can reduce or eliminate resource usage on the host as well. As a result more VMs can be backed up simultaneously and the host has more resources available for its VMs. Veeam Backup and Replication was built from the ground up specifically to back up VMware environments and operates at the virtualization layer to achieve maximum backup efficiency.

So, make sure you change your backup methods when you virtualize and use a backup application that is able to work at the virtualization layer for maximum efficiency.

2

Leverage the vStorage APIs

“This allows for quicker backups as the data is read over the storage network as well as reduced network utilization on the host server.”

The vStorage APIs for Array Integration (VAAI) were introduced in vSphere to allow third-party applications to integrate easily with the storage-related functions of vSphere. They were developed to replace VMware Consolidated Backup (VCB) and allow backup applications to directly interface with vSphere. They are grouped into categories with the VAAI for Data Protection (VADP) being the most beneficial to backup and recovery applications. Perhaps the most notable feature in the VADP is the Changed Block Tracking (CBT) feature which allows applications to quickly look up which disk blocks of a VM's virtual disk have changed from a specific point in time. This is a big deal because normally applications would have to figure this out on their own which can take some time. Being able to get this information instantly greatly speeds up incremental backup and replication operations which results in shorter backup windows and being able to achieve near-Continuous

Data Protection (near-CDP) without buying expensive storage hardware.

There are other features in the VAAIs as well that benefit backup operations such as the ability to hot-add a disk from a target VM to a source VM running a backup application so it can be backed up without going over the network to read the virtual disk. This allows for quicker backups as the data is read over the storage network as well as reduced network utilization on the host server. The VAAIs are a huge benefit to all storage-related functions in vSphere, and to use applications that don't leverage them is very inefficient. Veeam Backup & Replication was one of the first backup applications to embrace the VAAIs and make full use of the efficiencies that they provide. If you want the most efficient backups possible always make sure your backup solution leverages the VAAIs.

3

Virtual machine quiescing and VSS for ultimate protection

“A file system consistent state is better because the operating system is in a proper state to be backed up.”

Quiescing is a critical function in backup and replication applications. It ensures that data and applications on a VM are in a state where they are able to be backed up so they can be restored properly. This is especially critical for transactional data that resides on email and database servers. Quiescing is basically a fancy word for temporarily pausing a VM so any outstanding writes and data held in memory can be written to disk before the backup begins. This is done by the operating system as well as applications that support being quiesced. Once a VM has been quiesced, a snapshot of the VM is taken to freeze its disk so it is read-only and then the backup begins. Without quiescing a VM before backing it up you may find that upon restore, some of it is corrupt or unusable because open files were not properly prepared to be backed up. On Windows VMs quiescing is handled by the Microsoft Volume Shadow-Copy Service (VSS). It is built in

to Windows and works with the OS and applications to prepare them for snapshots and backups

There are several consistency states associated with quiescing for restoring data: crash, file system and application consistent. A crash consistent state does not have any quiescing and is equivalent to a VM being powered off without being properly shutdown. A file system consistent state is better because the operating system is in a proper state to be backed up. Finally, the best state is application consistent where applications have also been properly prepared for backup. Veeam Backup & Replication makes full use of quiescing by leveraging VSS via the quiescing functions built into VMware Tools. In addition, Veeam also developed their own method to work directly with VSS for applications that VMware Tools does not support quiescing with.

4

Don't under provision target backup resources

“It's best to err on the side of too many resources when determining the size of your backup server hardware.”

Backup operations can be very resource-intensive on just about everything in the data center. As a result, most companies want the shortest possible backup windows for the least impact on production workloads. You may have fast storage, servers and network devices but the backup server can easily become a bottleneck if it does not have the resources it needs to keep up with all the data flowing in and out of it. As expected, backup operations can heavily tax network and storage resources as the backup server moves data from source servers to tape or disk targets. But there is more to backups than just moving data from point A to point B—the backup server is also responsible for advanced functions like deduplication, compression and determining which disk blocks need to be backed up and which don't. So in addition to heavy disk and network usage, the backup server will commonly have heavy CPU and memory usage as well.

You should make sure you size your backup server appropriately so you don't create a bottleneck in any one resource area. Make sure you give it enough CPUs and memory so it can stay on top of the data that it is transferring in order to achieve maximum throughput and shorter backup windows. It's best to err on the side of too many resources when determining the size of your backup server hardware. If your backup server is a VM, it makes it easier to fine tune your hardware resources to find the sweet spot. You can easily make changes to the virtual hardware until you find the configuration that gets you the best throughput without going overboard on resources that may be unnecessary. Veeam recommends at least 4 – 8 CPU cores for the backup server to achieve the best backup efficiency.

5

You don't need expensive storage hardware to achieve near-CDP

“As a result, achieving near-CDP at the virtualization layer is now a reality and is a more affordable alternative to storage replication.”

For many companies having near-CDP is good enough and provides a decent failover solution for critical applications. It used to be that to achieve near-CDP you had to rely on expensive storage hardware and add-on replication software. This would handle the replication of the VM data at the storage layer from one storage system to another so it is ready to be used as a backup if needed. The virtualization architecture allows for some other methods to be used to achieve this. Because VMs are encapsulated into virtual disk files the replication can also be done at the virtualization layer that is between the hardware layer and operating system layer. While doing replication at the virtualization layer may not be as efficient as letting the storage array handle it, it can still get the job done in an effective manner.

Replication of a VM is similar to backing it up with the main difference being that backups are typically a daily event and replication is a continuous event. Replication is basically like doing incremental backups on a very frequent basis. Veeam Backup & Replication has had the ability to replicate VMs from a source host to a target host since its inception. Near-CDP is defined as an RPO as close to zero as possible, typically less than 30 minutes. However, in VMware VI3, achieving near-CDP was not very practical because the amount of time and resources needed to calculate changed data between replication cycles and then copy it could be high. vSphere changed all that with the new CBT feature which made incremental backup and replication operations much faster. As a result, achieving near-CDP at the virtualization layer is now a reality and is a more affordable alternative to storage replication.

6

Getting the most out of your backups – putting your backup repositories to work for you

“How they achieve this is actually pretty clever, the backup server becomes an NFS server with the backup repositories as the storage devices.”

Backups are much like an insurance policy, they continually cost you money and you need the security they provide but you don't get anything out of them unless you have an emergency. With virtualization it is common to do disk to disk backups and optionally sweep them to tape as well. The backups of your VMs just sit around on your target disk repositories taking up valuable disk space and resources, and they are completely ignored. But since they reside on disk, you actually have usable historical copies of your VMs available that could be used for certain purposes. Imagine if you needed a quick sandbox to test an application upgrade or an isolated environment to do some testing or troubleshooting. Those backup copies of your VMs are perfect candidates for this, and if you isolated them on their own virtual network you could do anything you want without disturbing the production environment.

Veeam has made this possible with their new vPower™ technology that is part of Veeam Backup and Replication version 5. How they achieve this is actually pretty clever, the backup server becomes an NFS server with the backup repositories as the storage devices. Any ESX/ESXi host can then connect to it using the NFS client and access the VM backups that are in the repository. The backup images are read-only, and any changes made to them while they are powered on are written to a separate delta file which is discarded afterwards. VMs powered on from the repository are kept isolated from the rest of the network using vSwitches that have no physical NICs assigned to them, and a special routing appliance allows access to outside networks. Being able to actually make use of your backups for purposes other than the occasional restore allows you to make the most of your backup investment.

7

Use data deduplication and compression to reduce backup size

“Another benefit of deduplication is that empty disk blocks that have been allocated to a VM but have not been written to by the guest operating system are ignored and not backed up.”

Over time your backup repositories can grow quite large and before you know it you're running out of disk space to store your backups. When this happens your only options are to purchase more storage or to limit the number of backups that you store in your repository. Neither solution is desirable so to help avoid the situation you should use technologies that help reduce the size of the data that is stored on your backup repositories. These technologies are data deduplication (which eliminates duplicate blocks of data from being stored in the backup repository) and data compression (which compresses the data in your backup repository so it takes up less space). Another benefit of deduplication is that empty disk blocks that have been allocated to a VM but have not been written to by the guest operating system are ignored and not backed up.

There are different methods for doing data deduplication and not all backup products support it or include it natively. One of the most

common methods is inline where deduplication is done in real-time, hash calculations are done before blocks are stored in the backup repository and then, if they match a disk block already stored, they reference that one and are not stored again. While this is beneficial to reducing backup sizes it can cause some extra overhead while backups are running since hash calculations have to be made on each disk block. Veeam Backup & Replication allows you to choose from multiple deduplication options that use different block sizes when calculating hashes so you can choose whether you want maximum deduplication at the expense of slower backups or minimum deduplication for best backup job performance. Veeam Backup & Replication also supports multiple compression levels that will vary the amount of compression that is done to meet the needs of the environment. Compression is very CPU intensive and can increase backup times, so having at least 8 CPU cores on the backup server is recommended for maximum compression.

8

Verify your backups to make sure you can properly restore them

“Automatic verification is made possible by powering up the backed up VM directly from the backup repository and checking for a heartbeat from VMware Tools as well as ping responses.”

There is nothing worse than backing up servers for months and months only to find out when a critical restore is necessary that the data you have been backing up is no good once it is restored. The whole point of performing backups is that you may someday have to rely on them to restore data when disaster strikes or even when data is mistakenly deleted or the workload is otherwise useless. If you are unable to restore data properly why even back it up in the first place? The process of verifying backup data is more than just doing data verification on your backup media which only confirms that disk blocks are properly written to the target device. If there is something wrong with the source data to begin with it will copy over to the target backup device as well. This can range from a server that is not properly quiesced before being backed up to missing or corrupt critical files on a server. Therefore the only way to truly make sure your backups are working properly is to restore the data to a server and make sure everything works.

Virtualization can make this process much easier as you don't need spare physical hardware to restore a whole server to; VMs can easily be restored to hosts with spare capacity. But this still can be a time consuming process to do this on a regular basis. Veeam recognized the need for this and has made backup verification a simple and automated part of the backup process with their new SureBackup technology in Veeam Backup & Replication version 5. Automatic verification is made possible by powering up the backed up VM directly from the backup repository and checking for a heartbeat from VMware Tools as well as ping responses. Additionally, test scripts can be run to verify that applications are running properly and data is accessible. The only thing more important than backups is restores. Having the peace of mind that you have proper backups is a good insurance policy to have when critical data must be restored.

9

Granular recovery from an image level backup

“**The result is the ability to quickly and easily restore application items with minimal effort.**”

Backups in a virtual environment are done at the image level for maximum efficiency. So what happens when you need to recover individual files or application items? On most guest VM operating systems the backup server can mount the backup virtual disk file so individual file restores are possible. Once the virtual disk image is mounted you can browse it and select the files that you want restored which are then copied to the destination you choose. Veeam Backup & Replication supports instant file-level recovery so individual files can be quickly restored from any backup point in time. This allows for individual files to be restored, but what about individual application items that are contained inside a single large file for transactional applications such as Microsoft Exchange and SQL server? It is both time consuming and difficult to restore a huge database so a small amount of emails or records contained inside of it can be restored back to the original database.

Fortunately, methods exist for restoring application items without the need for restoring the whole file that they reside in. Veeam Backup & Replication contains a special feature called Universal Application-Item Recovery (U-AIR™) that allows you to restore individual objects for popular applications like Microsoft Active Directory, Exchange and SQL Server. How is this possible? In a nutshell, U-AIR leverages the ability to power up a VM directly from a backup repository that is kept isolated from the rest of the network. Once the VM is powered on the application is accessed so select records can be copied from it back to the original location. The result is the ability to quickly and easily restore application items with minimal effort.

As you can see, image level backups at the virtualization layer are very versatile and are able to provide image, file and application item restores from a single image backup.

10

Avoid the sticky situation of assigning permissions to all your sensitive data to backup administrators

“This makes your application data more secure and removes a potential vulnerability from inside your guest OS and applications.”

In order to backup the data on your servers, your backup server has to have access to read it. This means your backup administrators must have access to the local file system and applications on every server that you wish to backup. In addition, if you are running specific backup agents for applications like Microsoft Exchange & SQL Server to backup data at the application layer you also need to assign access to all your data inside the application as well. From a security standpoint this can be an issue as there is inherent trust that backup administrators will not maliciously access any of the sensitive data that they are exposed to.

Virtualization has changed the need for this as you no longer have to go inside the operating system or application to back up the data that resides there. Instead of installing a backup agent inside the guest OS and granting it administrator access, your backup application can access the data from outside the guest OS

by reading it at the virtualization layer. This is known as an image level backup since the entire virtual disk image is backed up at the disk block level instead of reading individual files like OS backup agents do. As a result, you no longer have to assign backup administrators permissions inside the guest OS file system or to applications running on the VM.

Veeam Backup & Replication with U-AIR still makes individual application-item recovery possible despite not using specific application agents, and application owners retain control of their data at all times. This makes your application data more secure and removes a potential vulnerability from inside your guest OS and applications. Veeam vPower is revolutionizing how backup and recovery is performed in virtualized environments by providing greater flexibility, better security and increased reliability of your valuable data.

About the author



Eric Siebert is an IT industry veteran, author and blogger with more than 25 years of experience, most recently specializing in server administration and virtualization. He is a very active member of the VMware VMTN support forums, where he's attained the elite Guru status by helping others with their virtualization-related challenges.

Siebert has published books including his most recent, "**Maximum vSphere**" from Pearson Publishing, and has authored training videos in the Train Signal series. He also maintains his own VMware VI3 information website, [vSphere-land](#), and is a regular blogger and feature article contributor on TechTarget's [SearchServerVirtualization](#) and [SearchVMware](#) websites. Siebert has presented at VMworld in 2008 & 2010 and has been recognized as a vExpert by VMware in 2009 & 2010.



#1 VMware Backup



100% Reliability



SureBackup™

Best RTOs



InstantRestore™

Best RPOs



SmartCDP™

vPower™

Virtualization-Powered Data Protection™

5 Patents Pending!

VMware vSphere

5 Patents Pending!

NEW Veeam Backup & Replication™ v5

vPower enables these game-changing capabilities in Veeam Backup & Replication v5:

- **Instant VM Recovery**—restore an entire virtual machine IN MINUTES by running it directly from a backup file
- **U-AIR™ (Universal Application-Item Recovery)**—recover individual objects from ANY application, on ANY OS
- **SureBackup™ Recovery Verification**—automatically verify the recoverability of EVERY backup, of EVERY virtual machine, EVERY time

To learn more, visit www.veeam.com/vPower

Veeam ONE

Solution for
VMware Management

VEEAM

vPower

Virtualization-Powered
Data Protection



Veeam Reporter 4.0 for VMware

Enterprise reporting, change management and capacity planning for VMware



Veeam Monitor 5.0 for VMware

The new standard for performance monitoring, capacity planning and troubleshooting for your entire VMware Infrastructure.



Veeam nworks Management Pack 5.5 for Microsoft Ops Mgr for VMware

VMware Monitoring with Microsoft System



Veeam Smart Plug-in 5.5 for HP Operations Mgr for VMware

VMware Monitoring with HP Operations Manager



Veeam Backup & Replication v5 for VMware

Virtualization-Powered Data Protection™

Veeam Backup & Replication provides 2-in-1 backup and replication for virtual machines running on VMware ESX(i). By leveraging the virtual environment and patent-pending Veeam vPower™ technology, Veeam Backup & Replication overcomes the limitations of traditional data protection and disaster recovery to provide fast, flexible and reliable recovery for all your virtualized applications, services and data.