VEEAM
IT JUST WORKS!™

Top 10 Best Practices for VMware Data Availability

Eric Siebert
VMware vExpert

AVAILABILITY™
for the Modern Data Center

Summary

Backing up your virtual machines (VMs) may seem like a simple 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. However, to efficiently backup VMs, you need to use techniques and features 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, you need to change your procedures and methods to leverage its strengths and unique architecture.

The virtualization architecture offers many advantages for server backup and recovery. 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 (DR). In this white paper, we will offer 10 tips for assisting with the implementation of backup and recovery in a virtual environment, including the proper methods, techniques and configuration as well as leveraging the features built into Veeam® Backup & Replication™ so you can take your backups to the next level.

1 – Do your backups from the VMware infrastructure level

When backing up VMs, you shouldn't stick with the same method you used to back up your physical servers. Physical servers are traditionally backed up using an agent installed in the guest OS of the host. 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. For that, you need a backup application that is built and optimized for virtualization.

A backup application that is virtualization-aware doesn't have to involve the VM's guest OS in the backup. Instead, the application can connect directly to the VM's disk file to back it up. This means that there is no resource overhead on the VM while you back it up, and workloads will not be affected while backups are running. This can reduce or eliminate resource usage on the host as well. As a result, you can back up more VMs simultaneously, and the host has more resources available for VM workloads. Additionally, your backup solution should leverage the VMware vSphere APIs for Data Protection, which features integration such as Change Block Tracking to allow the hypervisor to keep track of disk blocks that have changed between backup or replication cycles for faster operations.

As companies become increasingly virtualized, your backup solution should reflect that and work at the VMware infrastructure level. Veeam Backup & Replication was built from the ground up to back up VMware environments specifically. It is fully integrated with VMware, and it operates at the virtualization layer for maximum efficiency.

2 – Keep your VMs and critical data safe by using the 3-2-1 Rule

Your VMs and, more importantly, your data are critical to your business, and you cannot afford to lose anything. Backups are like an insurance policy for your data, and you hope you never have to use them. However, when you do have to use them, it's critical that they work properly and you can restore what you need. Failure is not an option when it comes to recovering data. If your primary recovery method somehow fails, you need a backup plan. Because many organizations do not routinely test the recoverability of their backups, you may find yourself in a situation where you need a Plan B or even a Plan C to get back the data you lost.

The 3-2-1 Rule will ensure that you have multiple options available for restoring your data so your backups do not have a single point of failure. Consider this a backup to your backup; if something happens to one backup, you have a Plan B. The 3-2-1 Rule works like this:

- **Have at least three copies of data** — this means you should also have at least two more backups in addition to your primary data. If something happens to one backup, you have another to fall back to.
- **Store the copies on two different media types** — this ensures that a failure of any one device will not affect the recoverability on another. For example, you can store one backup to tape and another to disk or any other target, such as a cloud provider, USB device, SAN/NAS, etc.
- **Keep one backup copy offsite** — this one is most important. Don't let a local event like a fire or flood take out both your primary data and all of your backup copies all at once. You can send tapes off site, replicate to another office location or even to the cloud. Whatever you do, make sure there is some type of long distance, physical separation between your backups.

Veeam Backup & Replication is equipped to ensure that you can follow the 3-2-1 Rule and safeguard your backups.

3 – How to protect your backup data and avoid losing it

Your backups essentially serve as a copy of your entire data center, all stored in one convenient location (or more if you're following the 3-2-1 Rule). While this is just the nature of a backup target, you need to ensure that your data remains secure wherever it resides. A lot of attention is given to securing hosts, networks, operating systems and applications, but you also need to pay attention to the security of your backups. What are you doing to secure your backups, which often reside outside of your traditional security points? If someone were to obtain your backups, they could easily restore VMs and access the applications and data within them. This threat could come from inside your network or from outside because files can be copied easily over networks or carried away on tiny USB devices. In addition, if your data goes off site, either to another location or into a cloud, you have to trust that someone else will protect it for you.

As a result, you need to ensure that you extend your security practices to include your backup repositories to reduce the risk of sensitive data being seen by someone that should not be seeing it. You can accomplish this in a number of ways, but perhaps the simplest way is to encrypt your backup repositories via your backup application. You could also consider encrypting the data via hardware using storage hardware that supports encryption. This can be more costly, though. You should also tightly limit access to your backup repositories to only the necessary administrators and audit the access to them. If you are backing up outside your data center to an offsite location or cloud provider, you need to work with your service provider to ensure you have sufficient security controls in place to keep your data safe.

Veeam Backup & Replication delivers built-in, end-to-end AES 256-bit encryption, giving you the ability to encrypt backup files and data at the source (during backup), in flight and at rest to help ensure that your company does not end up in the news headlines for a security breach.

4 – Leverage policy-based controls for smarter data protection

When it comes to doing anything in life, would you rather do it the hard way or the easy way? Doing almost anything the hard way and achieving the same results is generally a waste of time and resources and leaves you vulnerable to human error and forgetfulness. The virtual data center is full of hidden complexities that can increase management, reduce efficiency and lead to unwanted problems and downtime. A virtual environment practically begs for automation that can help ensure compliance while reducing management efforts and ensuring things run as smoothly as possible. A savvy vSphere admin is always looking for ways to work smarter instead of harder, and using any type of automation or policy-based controls is the easy way and ensures consistency.

Storage Policy-Based Management (SPBM) — introduced in vSphere 5.5 with VSAN and later for shared storage in vSphere 6.0 with Virtual Volumes (VVols) — allows you to define storage requirements for VMs based on storage array feature or hardware capabilities. SPBM is all about automation and ensuring VMs are compliant and aligned properly with storage resources. When it comes to protecting VMs, you can build policies based on specific RAID levels or other storage availability attributes that align with your SLA requirements. Another lesser-known feature of vSphere that can help simplify and organize the way you interact with VMs is the Tags feature. This feature lets you to customize the grouping of your VMs. Custom tags can be created and assigned to VMs and allow you to group them based on non-standard vSphere containers (e.g., by application, role, location, department, etc.). This can be used with vSphere features or third-party applications that might perform some type of action on the VMs with certain tags.

Veeam Backup & Replication fully supports the use of vSphere Tags that can be leveraged when configuring backup jobs that allow you to customize backup options more efficiently based on what tag is assigned to a VM.

5 – Know the impact that new vSphere features and architectures have on data protection

As VMware continues to evolve vSphere to conform to their software-defined data center vision, they introduce many new features and architectures that fundamentally change the way things are done in vSphere. No part of vSphere has been affected more than storage, with VMware introducing their new VSAN and VVols storage architectures that are intended to replace the traditional VMFS datastore. Networking has evolved as well with the new NSX networking architecture, and hyper-converged infrastructures — such as EVO:Rail — are the hottest trends that bring together servers, storage and networking into a single-appliance model. With all these changes, you might wonder what impact they have on how you implement data protection. Are you doing things incorrectly or inefficiently now? What changes should you be making to align with these changes in vSphere?

Introduced in vSphere 5.5, VSAN transforms server-side storage into a shared storage array that can be distributed amongst many ESXi hosts. With the SAN residing within a server, storage resources are much closer to the host, which is good for VM workloads because it shortens the I/O path. However, this can be a double-edged sword because backup operations can put much more resource pressure on a host, which can affect overall performance. In addition, it changes the backup logic necessary for ingesting data in the most efficient manner. As a result, you want to leverage QoS (Quality of Service) controls that can throttle backup operations and ensure that your backup application is VSAN-aware. With VVols in vSphere 6.0, this new storage architecture introduces many new components between a host and a storage array but it is mostly transparent to backup applications. There are some new APIs introduced with both VSAN and VVols, so you need to make sure your backup application can both support and utilize these features in the most efficient manner.

You can rest assured that Veeam Backup & Replication is highly optimized for both VSAN and VVols. It fully leverages the latest vSphere APIs to integrate with these new storage architectures, using advanced smart logic and optimal virtual proxy selection.

6 – How you can leverage the cloud as part of your data protection strategy

The cloud serves as an offsite, interconnected extension to your virtual data center that you can leverage to provide additional alternatives to doing everything in house. When it comes to data protection, the cloud can be leveraged as an offsite repository that can eliminate the need for you to deploy your own recovery data center, which can be very expensive to maintain. While the cloud is generally not a good target for primary backup storage, it does serve as a good complement to an existing primary backup solution when used in a layering model, making multiple copies of your backups available (3-2-1 Rule). Most backup and cloud vendors have made it very simple to integrate between onsite data centers and applications to cloud-based infrastructure and services, so you can easily get your data to the cloud and back.

When it comes to forming a cloud strategy, you might consider having short-term backups on site with long-term ones stored off site in the cloud. Additionally, instead of just using the cloud as cold storage for your VMs, you could leverage cloud as a replication target to serve as a warm site to run your VMs, if needed. There are many possibilities that the cloud provides, and it serves as a great complement to your data protection strategy. When looking at cloud providers, you should be aware that pricing is typically based on resource consumption, and you need to understand the costs that will be incurred because backup and restore operations can be resource-intensive. In particular, look at the ingress and egress rates; you may discover that while getting your data into the cloud is relatively inexpensive, bringing back large amounts of data can be much more expensive.

Whatever cloud strategy you choose, Veeam Cloud Connect provides a fully integrated, fast and secure way to backup and restore from the cloud so you can get your backups off site without the cost and complexity of managing an offsite infrastructure.

7 – Make sure you are meeting all the requirements for protecting your critical apps

When it comes to protecting whatever applications are critical to you, you cannot afford to have any mishaps, which can get very costly. You need to make sure that you are doing everything that is required for backing up applications to guarantee that data is 100% recoverable all the time. Database and email applications in particular can be the trickiest to back up and recover properly because they have special handling requirements that you must utilize. One of the most critical parts of the backup process is quiescing, which is a function that ensures that data and applications running inside a VM are in a proper state to be backed up so they can be restored properly. Quiescing temporarily pauses a VM, so any outstanding writes and data held in memory can be written to disk before 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.

Another special technique used during database backup is the truncation of the transaction logs that record all transactions and modifications to a database that you can used to recover a database if you need to. Transaction logs must be truncated on a regular basis to keep them from growing too large. Once a successful database backup completes, the transaction logs can be truncated because the backup can serve as a recovery point. Transaction logs after the backup can be used for recovery until the next backup completes.

Granular recovery is another a key requirement that provides the ability to restore only a subset of data, instead of an entire database or email object store. For example, if you only want to restore a few records from a database or one email from a mail file without overwriting the original, your backup application has to support granular recovery.

Veeam Backup & Replication fully supports application- and transaction-consistent backups to ensure that your critical data is backed up properly. It also supports log truncation with application-aware image processing. For recovery, the Veeam Explorer™ applications allow Veeam users to restore individual objects for popular applications like Microsoft Active Directory, Exchange and SQL Server with minimal effort.

8 – Know that backing up a VM at the disk level still provides you with plenty of restore options

In a virtual environment, backups are done at the virtual disk image level for maximum efficiency. While this is great for efficiency, real-life restoration requests often focus on restoring objects from within a VM, instead of the whole VM. So what happens when you need to recover individual files or application items? When you back up at the image level, you still have the ability to see inside the disk image because it can be mounted by the backup application and accessed via the OS file system within the image. This allows you to see the files within the image and restore individual files as needed. In addition, it allows you to access application-level data stores, such as a database or email file. However, it can be both time consuming and difficult to restore a huge database so that a small amount of emails or records inside it can be restored. As a result, your backup application needs to understand the application file format so it can peer inside and restore just the items that you need.

Aside from databases and emails, another common application-level restore is with Microsoft Active Directory, which is a critical part of any Windows infrastructure. With Active Directory (AD), restoring the entire AD structure and data is usually not desirable and special care must be taken to avoid disrupting anything. To accomplish this sensitive restore, a backup application must be able to access and browse AD natively so it can restore any deleted objects back to their original locations. You can see that image-level backups provide you with the ability to perform many different restoration types based on the requirements for a particular recovery scenario. Whether you need to restore a whole VM, a particular VM virtual disk, one or more files within a VM, or a single email or database record, your backup application should be equipped to handle these more granular restore scenarios.

Veeam Backup & Replication takes restorability to the extreme with 47 different restoration scenarios, ranging from a group of VMs in a vApp, to a full VM and all the way down to individual files and application-level items to be able to handle almost any restoration scenario.

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

Backups are very similar to insurance policies: they are an ongoing investment that costs you money and you must have the security they provide them, but you don't really get anything in return unless you have a mishap. With virtualization, it is common to do disk-to-disk backups and optionally sweep them to tape as well. Your VM backups just sit around on your target disk repositories, taking up valuable disk space and resources, and they are completely ignored. However, because 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 are perfect candidates, 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 in Veeam Backup & Replication by creating a Virtual Lab that leverages the backup server as an NFS server, with the backup repositories acting as the storage devices. Any ESXi host can connect to it 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 discarded afterwards. VMs powered on from the repository are kept isolated from the rest of the network, and a special routing appliance allows access to outside networks. This also allows you to automatically verify your backups so you can ensure they are recoverable. We covered the 3-2-1 Rule already, but when backing up with Veeam, it becomes the 3-2-1-0 Rule, where "0" means "0 errors" during the automatic recoverability verification of every backup with Veeam's SureBackup® and Sure Replica.

Being able to actually use your backups for purposes other than the occasional restore allows you to maximize the return on your backup investment and puts your insurance policy to work.

10 – Don't get caught short — do your backup math

Capacity planning for your backups is important to ensure that you can remain compliant with the retention schedule that you have set, either by choice or as the result of a compliance policy. Planning capacity for your backup repositories is not an easy task because virtual environments tend to expand at a quick pace due to VM sprawl. Do you have any idea how long your current capacity will last you? Do you know how adding five more VMs will affect it? If you end up short, either you have un-planned costs that must be incurred to expand it or your retention length must be cut. Neither is a desirable situation to be in. As a result, doing the backup math to find out when you will have to increase the storage that is dedicated to your holding backup repository can be a challenge. Backup math is not as straightforward as traditional storage capacity planning because you have many factors that can affect your calculations. This includes backup compression ratios, backup retention length, backup frequency and data change rates for incremental backups. What you need is an intelligent backup application that can do the math for you.

Veeam ONE™ provides deep analysis and monitoring of vSphere infrastructures and uncovers potential issues that can affect the performance of your backups and production applications. Veeam ONE can alert you when backup repository space reaches a certain level and includes a VM Change Rate Estimation report, which automatically analyzes your VMs' change rate and computes the potential amount of space required on your backup repository.

There's more to backups than just having available space to store them — you also have to have sufficient host resources available for them to complete in your desired backup window timeframe. The Datastore Performance Assessment report will help you examine your datastore's performance to identify potential issues that can occur during the backup process due to high latency or IOPs values. This information is helpful when defining latency thresholds so you can optimize your backup processing performance, increase resource usage efficiency and minimize the impact on production workloads. Veeam ONE is an excellent companion to Veeam Backup & Recovery and provides you with the visibility that is needed to maintain a healthy backup environment.

About the Author

Eric Siebert is an IT industry veteran, speaker, author and blogger with more than 25 years of experience who has been focused on virtualization since 2005. Siebert has published books including his most recent, "Maximum vSphere" from Pearson Publishing and has published hundreds of articles and white papers for Tech Target and VMware partners. He also runs and maintains his own VMware information website, vSphere-land.com. Siebert is a frequent speaker at industry conferences and events including VMworld and has been recognized as a vExpert by VMware each year since the programs inception in 2009.

About Veeam Software

Veeam® recognizes the new challenges companies across the globe face in enabling the Always-On Business™, a business that must operate 24/7/365. To address this, Veeam has pioneered a new market of *Availability for the Modern Data Center*™ by helping organizations meet recovery time and point objectives (RTPO™) of less than 15 minutes for all applications and data, through a fundamentally new kind of solution that delivers high-speed recovery, data loss avoidance, verified protection, leveraged data and complete visibility. Veeam Availability Suite™, which includes Veeam Backup & Replication™, leverages virtualization, storage, and cloud technologies that enable the modern data center to help organizations save time, mitigate risks, and dramatically reduce capital and operational costs.

Founded in 2006, Veeam currently has 30,500 ProPartners and more than 145,500 customers worldwide. Veeam's global headquarters are located in Baar, Switzerland, and the company has offices throughout the world. To learn more, visit http://www.veeam.com.

VEEAM
IT JUST WORKS!™

AVAILABILITY™
for the Modern Data Center

COMING SOON

NEW Veeam® Availability Suite™ v9

RTPO™ <15 minutes for ALL applications and data
Enabling the Always-On Business™
with *Availability for the Modern Data Center*™

To learn more, visit **www.veeam.com**