VEEAM
IT JUST WORKS!™

The VSS role in vSphere, Hyper-V and agent-assisted backups

Nick Cavalancia
Techvangelism

AVAILABILITY™
for the Modern Data Center

When it comes to Microsoft VSS (Volume Shadow Copy), there's more to backing up VMs (virtual machines) than meets the eye.

You've probably heard the somewhat contradictory statement: "Virtualization improves server backups, while at the same time it *complicates* server backups." This rings true, in part, due to virtualization's new approaches for capturing backups.

You've probably heard that once virtualized, a multitude of options present themselves for backing up a virtual server. On the one hand, you can continue backing up servers, just like in the physical world. For instance, you might install a backup agent into each VM and back up files one-by-one to disk or tape.

While this is an obvious tried-and-true method for backing up server data, it also provides no virtualization benefits. Restoring an entire VM isn't easy. This resource-heavy backup method impacts VM performance, and backups end up taking a long time to complete.

On the other hand, there's the host-based, backup approach. Using this approach, VMs are backed up with assistance from their virtual host. Entire VM disk files — **.VMDK** or **.VHD/.VHDX** — can be captured at once, which enables easy restores of whole VMs in case of failure. Smart vendors now integrate entire-VM backups with individual file restores, enabling files and folders to be restored with the same level of ease. Even smarter vendors go one step further by backing up server applications and data with the same granularity and performance as individual files.

Nearly every data center today places as much importance on application data as file data. Microsoft SQL Server, Exchange Server, Active Directory, Oracle and others are all applications labeled with Tier 1 data center priority. This prioritization means that that your backup solution must be able to handle application backups as richly as files, folders and entire VMs.

You might be surprised to learn that not *all application-aware backups are created equally*. While you may know that every backup solution is obviously different in the ways each performs its duties, you may not be aware that different backup solutions' on-the-server services used to gather application data can also be quite different.

As you may know, Microsoft's onboard VSS is a service used by Windows to create volume snapshots for backup and recovery purposes. But there are more pieces to VSS than you may realize. Understanding those pieces and how they impact the success of your backups and restores can have a big impact on your data protection.

What is VSS?

Microsoft's VSS is Windows' built-in infrastructure for application backups. A native Windows service, VSS facilitates creating a consistent view of application data during the course of a backup. VSS relies on coordination between VSS requestors, writers and providers for quiescence, or "to make quiet," a disk volume, so that a backup can be successfully obtained without data corruption.

At least, *that's the technical explanation*, the one with all the big words. In simpler language, VSS is a Windows service that interacts with installed applications to tell them when a backup is taking place. It also reports back to the server when the backup is complete, instructing the application and the server to perform important post-backup tasks such as truncating logs and other cleanup activities.

Why is VSS necessary? Specifically, it is for *coordination* between the applications, their data and the activities, which are being completed by your backup solution. This coordination is required to avoid situations like the one described in the following story:

Backing up Exchange without VSS = How a bad day starts

Here's an example of how a backup job could work if VSS wasn't around to coordinate activities. ***NOTE:*** *This is not a situation you would ever want to experience.*

One day you attempt to back up your Microsoft Exchange server named **\\exchange01**. At 10 p.m., your backup solution begins its backup job for this server and all its data. Being an Exchange server, **\\exchange01** is host to a set of files that contain its Exchange database. As the process begins, the backup server transfers files, including the Exchange database files, from **\\exchange01** to the backup storage device.

At 10:05 p.m., just a few minutes after the backup job starts, an Outlook user named Bob checks his email and sends and receives a set of mail for the day. Bob's process of sending and receiving email changes the data inside the Exchange database.

This presents a problem because the database has, at this point, already been partially backed up and its file on the disk is only partially transferred to the storage device. The data contained within the database files on **\\exchange01** is now slightly different than the data that was captured by the backup solution. *These two views of the database are no longer consistent.*

This is how a bad day starts, a day that will eventually result in a corrupted database upon restore.

VSS comes into play any time a transactional-based application is installed on a Windows server that requires backups. Those applications can be Microsoft Exchange, SQL Server, AD (Active Directory), Oracle or any of a number of applications which require open access to files on the disk. VMs are also transactional-based items in a data center, which require their own quiescence for proper backup. More importantly, VSS also comes into play with individual files on the disk to ensure that open files are correctly captured during the backup process.

You may recall that the primary job of the VSS is to "quiet" an application or file system just prior to a backup. This *quiescence* creates a point in time from which backups are then sourced. You'll often hear this point in time referred to as a snapshot, although snapshots in this sense are very different than the VM snapshots used by your favorite hypervisor.

Creating a single point in time ensures that a common starting point for backups is shared by both the server and the backup application, and it *guarantees that each maintains a consistent view of the data*.

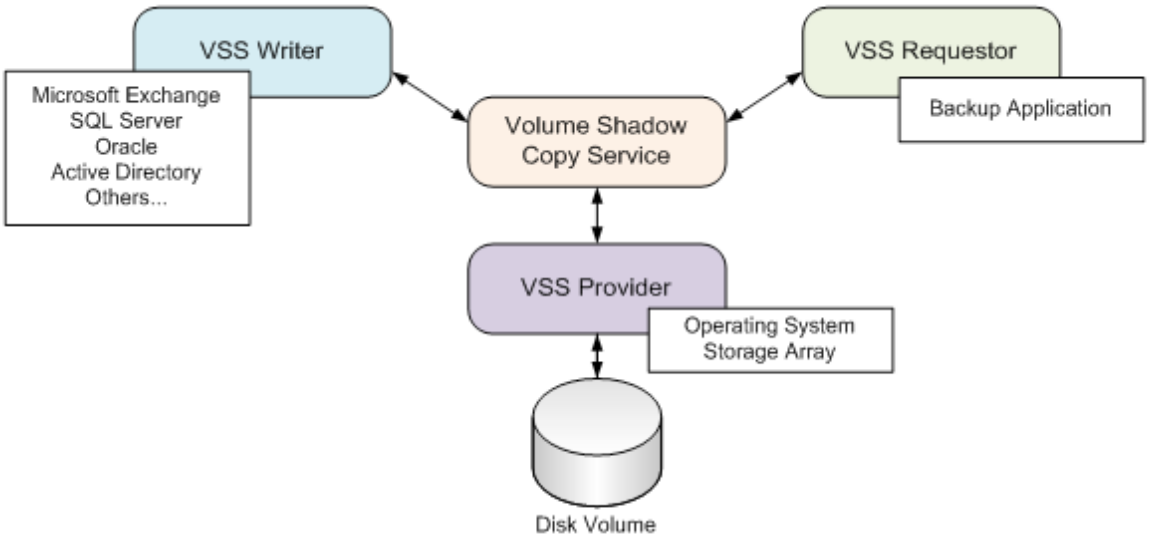

Figure 1: VSS components.

VSS relies on the coordination of three different components to maintain a consistent view. You can see those three components in **Figure 1**. In the upper-left is a set of VSS Writers. Each VSS-aware application installed onto a server also installs its own **VSS Writer**. The VSS Writer's job is to coordinate backup activities with the application and instruct the application to quiesce, or quiet, at the appropriate time.

VSS Requestors can be, among other things, the application you use for backups. The VSS Requestor's job is to coordinate VSS activities with those of the backup application. The VSS Requestor is also the component that actually requests that a volume shadow copy be taken. Once requested, the VSS Writer will instruct the application to perform the required actions to create that volume shadow copy.

The third component is the VSS Provider, which creates and manages the shadow copies. The VSS Provider can be either the OS (operating system) with its file system, or it can be a hardware provider on an external storage array.

NOTE: *You can use the command* ***vssadmin list writers*** *to list the VSS Writers that have been installed on a Windows computer.*

The role of VSS in virtualization

While VSS has long been used for backing up running applications, it has now become even more critical when paired with virtual environments so entire VMs can be backed up at once. Backing up an entire VM at once requires backing up that VM's disk file and either a **.VMDK** file for vSphere or a **.VHD/.VHDX** file for Hyper-V. By backing up that a VM's disk file as a point-in-time backup, it becomes possible to trivially and quickly restore that VM to a previous point in time.

Getting there, as you can imagine, requires the same sorts of quiescence that applications require. Since a VM file system is as interactive and always changes as an application's database, a mechanism to quiet the VM file system is needed if a host-based or externally-based backup solution is going to gather the disk file and maintain a consistent view.

That mechanism to quiet the VM file system isn't always the same, but depends on your backup solution and your selected hypervisor. That said, some architectures don't provide some functions, which are needed for true restores. Let's compare the approaches of three different native options and see how those compare with a true agentless, application-aware backup solution. The differences here will give you an idea about how very different the simple task of backups can become.

Option #1: Native Hyper-V data protection

A fully-native, Hyper-V environment automatically enjoys all the benefits of VSS components. This is because a Hyper-V environment runs completely atop Microsoft Windows. VMs in a Hyper-V environment are Windows (ignoring Hyper-V's Linux capabilities here). Windows Server is also the OS at the virtual host. Native Hyper-V uses the onboard Windows Server Backup as its backup application.

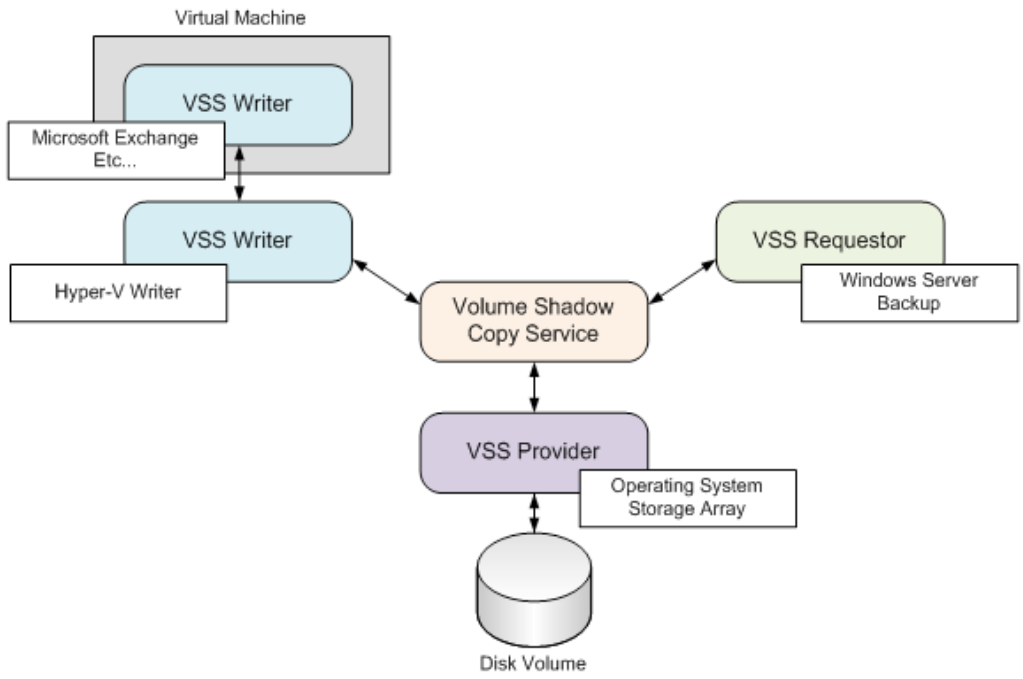

Figure 2: VSS in a native Hyper-V environment.

As you can see in **Figure 2**, these elements map directly to the original three VSS components described earlier. Windows installs a VSS writer named Microsoft Hyper-V VSS Writer with the installation of Hyper-V. The backup application Windows Server Backup serves as the VSS Requestor, with the OS and/or storage array filling the VSS Provider role.

In this configuration, the instance of Windows Server Backup on the virtual host requests the host's Hyper-V Writer to quiet the file systems of any running VMs so they can be backed up with a consistent view. But that's not all. Each VM has its own VSS components, as well as the host. Each VM also has its own installed applications that require quiescence. Quieting those applications requires coordination between the host's backup activities and those going on inside the VM.

That's why **Figure 2** also shows a VSS Writer inside the VM. Because a VSS snapshot is requested by Windows Server Backup, the Hyper-V VSS Writer on the virtual host integrates with any registered VSS Writers in the VMs (such as Microsoft Exchange in **Figure 2**) to ensure that the VM applications are also properly quieted. This integration is accomplished through the use of the Hyper-V **Integration Components**, which are a separate, yet required, installation to any Hyper-V VM.

In addition, beginning with Windows Server 2012, virtualized domain controllers can use snapshots without impacting the AD-DA (Active Directory Domain Services) environment. Before Windows Server 2012, the USN (Update Sequence Number) prevented a snapshot from being restored on a domain controller without creating problems for the AD DS environment. The primary problem was that the replication stopped functioning for the restored DC. The best way to fix that problem was to forcefully demote the DC, perform a metadata cleanup and then promote the server as a new DC. With Windows Server 2012 and later, a domain controller has an Invocation ID. When a snapshot is restored on a domain controller, the Invocation ID is reset which allows replication to function normally. Plus, the domain controller drops any local RID (relative identifier) pool, and non-authoritatively restore the SYSVOL (system volume) folder.

NOTE: *Windows PowerShell can also be used to configure VSS backups. Specifically, the* ***Get-WBVssBackupOptions*** *and* ***Set-WBVssBackupOptions*** *cmdlets are available to retrieve and set VSS options.*

Obviously, *there's an extra level of coordination* involved to maintain a consistent view of data across the host, VM and applications.

Option #2: Native vSphere

The situation gets slightly more complicated when VMs are run atop different hypervisors, such as VMware's vSphere. With VMware vSphere, there is no Microsoft Windows instance that operates as the virtual host. This means that there needs to be coordination between the backup software and VMware's framework for backups: the VADP (vSphere APIs for Data Protection).

vSphere added full support for VSS beginning with version 4.1 for all guests, including the Windows Server OS. This VSS support was introduced into vSphere-hosted Windows VMs through an update to the VMware Tools. Just like Hyper-V's Integration Components, the VMware Tools are a separate, but required, installation into any vSphere-hosted VM to perform application-aware backups.

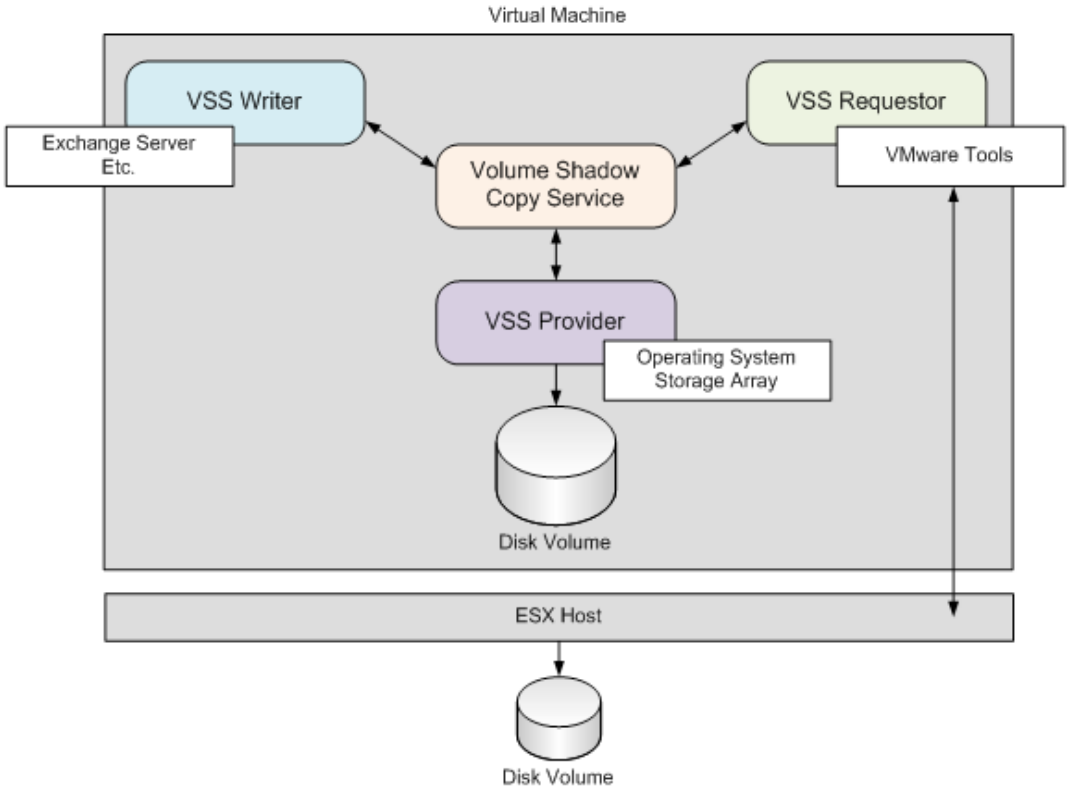

Figure 3: VSS in a native vSphere environment.

As you can see in Figure 3, a similar quiescence process occurs on a vSphere VM as the process experienced with Hyper-V. Here, however, the VMware Tools serve as the VSS Requestor, instructing registered VSS Writers to perform pre- and-post backup actions when whichever backup solution on the vSphere host begins the VM backup.

Option #3: Agent-Assisted Data Protection

While the architecture that makes up Solution Set #2 will work for backing up and restoring VMs, it does have limitations related to the very applications you are trying to protect with your backup infrastructure. However, the specific limitations have more to do with the recovery process than the actual backup process.

Backing up Microsoft Exchange, including DAGs (Database Availability Groups), is a simple process with Veeam®. Veeam uses VSS technology and application-aware Image Processing to capture transaction-consistent backups for applications such as Exchange and SQL. When using DAGs, a single backup job can be created for all nodes in a DAG. If you want to perform backups in more than one job, two best practices should be followed:

1. Ensure that backup schedules do not overlap
2. Ensure that all VMs are not being snapshotted at the same time

These best practices can be followed by modifying the settings of the Veeam backup jobs. In addition, for nodes that are geographically distributed, or have high-latency connections, it is recommended to modify the timeout values for the DAG. The methods to maximize the timeout values include:

```
cluster /prop SameSubnetDelay=2000:DWORD

cluster /prop CrossSubnetDelay=4000:DWORD

cluster /prop CrossSubnetThreshold=10:DWORD

cluster /prop SameSubnetThreshold=10:DWORD
```

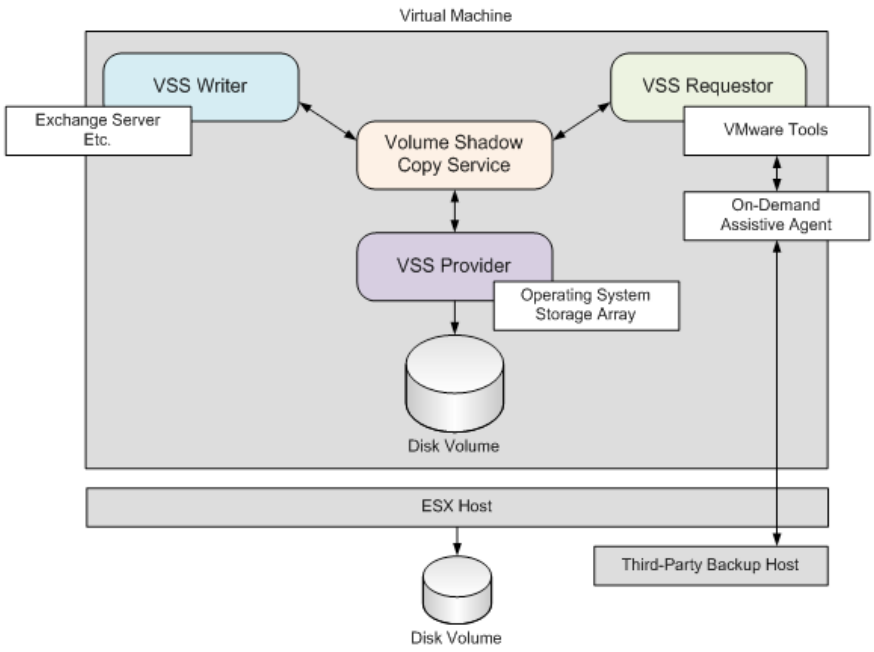

Figure 4: VSS in an agent-assisted vSphere environment.

One solution to prevent problematic situations is through the use of an on-demand agent installed to VMs during the backup process (see Figure 4). This agent is considered "on-demand" because it only resides on the VM during backups and later removed after the backup is complete. The presence of this agent facilitates the coordination between the vSphere VSS Requestor and the third-party backup host.

NOTE: *While not depicted here, the same on-demand assistive agent could be also used for similar results in a Hyper-V environment.*

It is important to recognize that an on-demand agent is one that is automatically available within the backed up VM. This means that the same agent will be available after the VM is later restored. The presence of this agent enables an immediate integration between the onboard agent and the third-party backup host and solution.

Being present on the host as it is restored allows the agent to control post-restore actions such as un-mounting Exchange databases. These actions ensure that restored servers and their data have a greater guarantee of successful restoration, with a minimum of accidental data destruction or corruption. This is key when servers are down, stress levels are high and the potential for mistakes is heightened.

Agent-assisted and transaction log handling

There's another important facet to agent assistance that benefits data protection. The agent-assisted approach offers greater support for handling application transaction logs both during and after a backup. You may recall that a VSS snapshot creates a point in time that enables the backup solution and the application to maintain a consistent view of data throughout the backup. Maintaining this view as data changes in the "real" database requires logging changes to a transaction log.

One significant limitation of some backup solutions is recognizing when the backup has completed successfully. Application transaction logs, such as those used by Microsoft Exchange and others, are an important source of data reconstruction if there is a failed backup. It is important that a backup solution instruct the server to flush those logs only after the backup has been deemed successful.

Some implementations, such as the VMware Tools implementation noted in Solution Set #2 above, are not equipped with the necessary instrumentation to know when a backup has completed successfully. This means that such implementations may either fail to prune transaction logs after a backup, or they may do so even if the backup has not completed successfully.

One benefit of using an assistive agent in a VM backup process is the assistive agent is more aware of the backup success. The assistive agent can then retry the backup if there is a failure, or it can prune the logs once the backup has been deemed successful. Both of these situations prevent the situation of needed transaction logs being inappropriately discarded, which can prevent the server from being restored is there's a failure.

Veeam application-aware, agentless backup and recovery

In addition to the native VSS provider for creating application-aware backups, Veeam has created powerful backup and recovery solutions for vSphere and Hyper-V through its close strategic alliance relationships with both VMware and Microsoft. These products rely on VSS, but expand the backup and recovery options greatly compared to the built-in backup and recovery tools. Veeam Availability Suite™ and Veeam Backup & Replication™ provide a wide array of backup, recovery and replication functionality. These solutions also offer specialized recovery for several Microsoft products, including:

- Veeam Explorer™ *for Microsoft SQL Server*
- Veeam Explorer *for Microsoft Active Directory*
- Veeam Explorer *for Microsoft Exchange*
- Veeam Explorer *for Microsoft SharePoint*

Veeam Explorer *for Microsoft SQL Server* provides agentless backup of SQL transaction logs. This provides transaction-level recovery of SQL databases and allows you to select the current restore point, a selected point in time or the state before a specific transaction.

Veeam Explorer *for Active Directory*, as well as other Veeam Explorer products, enable you to perform granular restores of backed-up objects. For the AD Explorer, it enables you to restore AD DS objects, containers, organizational units and user accounts, including individual or in-bulk passwords. This provides a method of restoring the objects without creating an additional VM or using a manual directory service restore method.

Veeam Explorer *for Microsoft Exchange* gives you a method of backing up and restoring emails, mailboxes and contacts directly from Veeam. It also lets you archive a mailbox to a PST file.

Veeam Explorer *for Microsoft SharePoint* gives you instant visibility into SharePoint backups with its quick search and recovery of individual SharePoint items or lists.

There's more to VSS than meets the eye

Virtualization can complicate backups, even though it also improves their usability. Once virtualized, you can absolutely enjoy the ability to restore whole servers just as easy as files, folders or application objects. You can only get there, however, if you implement solutions that really work. As you've learned here, Microsoft's VSS is one solution that does work if it is integrated with a well-designed backup solution.

About the Author

Nick Cavalancia has nearly 20 years of enterprise IT experience, is an accomplished consultant, speaker, trainer, writer, and columnist and has achieved industry certifications including MCSE, MCT, Master CNE and Master CNI. He has authored, co-authored and contributed to over a dozen books on Microsoft technologies. Nick regularly speaks, writes and blogs for some of the most recognized tech companies today on a variety of topics, such as cloud backup, virtual storage, Active Directory, virtualization and DevOps. Follow Nick on Twitter @nickcavalancia or @Techvangelism.

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**