

Realtime
publishers

Virtualization Backup and Recovery Solutions
for the SMB Market
The Essentials Series

How Traditional Physical Backup Imaging Technology Fits Into a Virtual Backup Solution

sponsored by



Jaime Halscott

Introduction to Realtime Publishers

by **Don Jones, Series Editor**

For several years now, Realtime has produced dozens and dozens of high-quality books that just happen to be delivered in electronic format—at no cost to you, the reader. We’ve made this unique publishing model work through the generous support and cooperation of our sponsors, who agree to bear each book’s production expenses for the benefit of our readers.

Although we’ve always offered our publications to you for free, don’t think for a moment that quality is anything less than our top priority. My job is to make sure that our books are as good as—and in most cases better than—any printed book that would cost you \$40 or more. Our electronic publishing model offers several advantages over printed books: You receive chapters literally as fast as our authors produce them (hence the “realtime” aspect of our model), and we can update chapters to reflect the latest changes in technology.

I want to point out that our books are by no means paid advertisements or white papers. We’re an independent publishing company, and an important aspect of my job is to make sure that our authors are free to voice their expertise and opinions without reservation or restriction. We maintain complete editorial control of our publications, and I’m proud that we’ve produced so many quality books over the past years.

I want to extend an invitation to visit us at <http://nexus.realtimepublishers.com>, especially if you’ve received this publication from a friend or colleague. We have a wide variety of additional books on a range of topics, and you’re sure to find something that’s of interest to you—and it won’t cost you a thing. We hope you’ll continue to come to Realtime for your educational needs far into the future.

Until then, enjoy.

Don Jones

Introduction to Realtime Publishers..... i

How Traditional Physical Backup Imaging Technology Fits Into a Virtual Backup Solution.. 1

 What Exactly Are Image-Based Backups?..... 1

 Challenges of Using Physical Server Backup Solutions with Virtual Servers..... 2

 Getting File-Level Backup as Part of Image-Based Backup 3

 Approaching Virtual Host Backups from a Physical Standpoint..... 3

 Benefits of Enhanced Protection Through Virtualization..... 4

 Traditional Approaches to RPO/RTO Aren't Good Enough..... 5

 What to Look For in a Host-Based Virtualization Backup Solution..... 6

 Coming Up Next..... 6

Copyright Statement

© 2011 Realtime Publishers. All rights reserved. This site contains materials that have been created, developed, or commissioned by, and published with the permission of, Realtime Publishers (the “Materials”) and this site and any such Materials are protected by international copyright and trademark laws.

THE MATERIALS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. The Materials are subject to change without notice and do not represent a commitment on the part of Realtime Publishers its web site sponsors. In no event shall Realtime Publishers or its web site sponsors be held liable for technical or editorial errors or omissions contained in the Materials, including without limitation, for any direct, indirect, incidental, special, exemplary or consequential damages whatsoever resulting from the use of any information contained in the Materials.

The Materials (including but not limited to the text, images, audio, and/or video) may not be copied, reproduced, republished, uploaded, posted, transmitted, or distributed in any way, in whole or in part, except that one copy may be downloaded for your personal, non-commercial use on a single computer. In connection with such use, you may not modify or obscure any copyright or other proprietary notice.

The Materials may contain trademarks, services marks and logos that are the property of third parties. You are not permitted to use these trademarks, services marks or logos without prior written consent of such third parties.

Realtime Publishers and the Realtime Publishers logo are registered in the US Patent & Trademark Office. All other product or service names are the property of their respective owners.

If you have any questions about these terms, or if you would like information about licensing materials from Realtime Publishers, please contact us via e-mail at info@realtimepublishers.com.

How Traditional Physical Backup Imaging Technology Fits Into a Virtual Backup Solution

Many small and midsize businesses are now familiar with image-based backup solutions. They have been used to protect physical machines with great success, and businesses wonder how or whether this technology can be applied to their growing population of virtual machines. The benefits of image-based backups for bare-metal restoration and fast, complete recovery are well known. As such, SMB customers naturally want to adapt these familiar backup solutions to virtual servers. What are the strengths to this approach and what challenges are likely to be encountered?

In order to understand exactly how image-based backup solutions fit into a virtual infrastructure, it is important to understand how various backup technologies for both physical and virtual servers perform, and then determine how virtual servers can benefit best from backup solutions designed exclusively for their unique attributes.

After understanding the benefits of virtual infrastructures for backup and recovery, the proper design of a virtual backup solution will be critical to realizing those benefits. Selecting the best option and path towards a virtual backup solution will drive down total cost of ownership (TCO) for the solution and maximize the return on investment (ROI) that comes from making smart choices up front.

What Exactly Are Image-Based Backups?

Many businesses still think of backups in terms of file-level recovery. The reason for this perspective isn't as surprising as one might think. The vast majority of restorations are still for the simple purpose of recovering a lost, deleted, or corrupted file. This might explain why the concept of imaged-based backups isn't the first thing that comes to mind when businesses think about backup and restoration.

This mindset is particularly true of the business where there is no dedicated IT staff in house. A non-technical employee might well be tasked with creating an appropriate plan for backup. In these small businesses, the concept of disaster recovery is no more than backup and restoration of files. There isn't a well-defined plan that goes beyond file backups and file restorations. The current concept of disaster recovery only addresses file recovery, but those files are of little value to business productivity if the systems on which the files run aren't also recovered.

Because file-level backups were the first type of backups that businesses used to protect their sensitive data, and because file-level backups existed as the standard for disaster recovery until recent years, it only makes sense that a different technology would take some time to catch on, no matter how much more advanced it is. Image-based backups function by taking a snapshot of the entire server's drives or volumes. Unlike file-level backups, with image-based backups there is no possibility of missing critical files. Restoration of the server in the event of a total failure can be done extremely quickly. There is no need to reinstall the operating system (OS) and then put back a patchwork of files in an effort to replicate the previous system. Recovery times from a complete failure even drop from hours to minutes in many cases.

Image-based backups, with some notable vendor exceptions, also allow the granular restoration of files from the image. All of the traditional benefits of file-level backups are now a subset of image-based backups. Image-based backups also have evolved to the point where incremental changes can supplement full backup images. This means that only portions of the server that changed are backed up, resulting in much smaller backup sets and thus more recovery points can be retained. These features have made image-based backups the standard for physical server disaster recovery.

So this begs the question, how do image-based backups fit with virtual servers? Are the solutions for physical servers the same ones that should be used for virtual host server and virtual guest servers?

Challenges of Using Physical Server Backup Solutions with Virtual Servers

When image-based backup solutions emerged as the gold standard of backup solutions as far back as 2005 to 2006, the concept of virtualization wasn't yet in the mainstream. Original image-based backup solutions were designed for the challenges of physical servers. In particular, image-based backup solutions ran inside of the OS, whether the OS was on a physical server or a virtual one. This wasn't too much of an issue because the only real difference from the standpoint of the software was the hardware variations.

However, because so many hardware variations were being used, some image-based backup software vendors realized the importance of the ability to support changes to hardware at the lowest level, thus enabling an image from one server to be restored "bare metal" to another server with wildly varying hardware. Software solutions that offer this functionality help save a tremendous amount of time. In addition, these solutions eliminate the need to purchase and maintain a set of identical spare hardware for each machine to enable quick recovery, which also leads to huge savings. The need to support hardware variations wouldn't seem like a problem when adapting physical backup solutions to virtual servers, but this is only true from a technical standpoint and only true with regard to running inside the guest OSs. However, very few image-based solutions, even to this day, can compensate for the hardware changes between physical servers.

There aren't any particular technical challenges with running physical backup solutions inside of virtual guest OSs, only problems with centralized management and costs associated with licensing and soft support costs. The technical challenges come with installation of the physical backup solution inside of the host OS and using the physical backup solution to back up the guest virtual OSs. The backup solution may not understand what it takes to ensure that the guest OSs are in a consistent state. Treating the guest OSs as just a series of files can lead to problems.

Getting File-Level Backup as Part of Image-Based Backup

Backup solutions are expected to perform a variety of roles. When only file-level recovery products existed, it was expected that a server would have to be rebuilt and everything that makes it do what it does would need to be put back on top in a file-by-file fashion. Once image-based backups became the norm, the ability to put everything back to the way it was replaced this lengthy rebuild and restore process. The challenges came when restoration of only a single item or a few small items meant that a file-level restore had to be performed from an image-based backup.

The ability to get file-level, granular restores from an image-based backup solution in a single backup pass has greatly helped to reduce the time it takes to recovery from accidental data loss, corruption, and a whole series of small problems that don't require restoring the entire server. This ability in both the physical server and virtual server environment has helped to move toward much better recovery time objective (RTO) and recovery point objective (RPO), but it hasn't taken these to the point that modern businesses expect.

Approaching Virtual Host Backups from a Physical Standpoint

The nature of virtual servers as nothing more than files on a host server means that backup should be nothing more than protecting these files, right? Well not exactly. Yes, the virtual machines themselves really are nothing more than a series of files that exist on a physical host server, but the nature of virtualization means that these files are in a constant state of change. Virtual servers aren't just static files that users edit like documents or spreadsheets. These virtual machine files encapsulate all of the necessary functions of live servers, including the OS, applications, settings, stored data, and even the contents of RAM.

The way that virtual servers are architected means that a solution designed exclusively for physical servers won't always work for a virtual environment, at least not in a fully vendor-supported way when you try to back up the virtual machines from the host server on which they reside. You can always install the backup solution directly inside each of the guest virtual servers using an agent, but doing so introduces several problems.

The licensing model for physical server backup solutions is naturally a per-server model. Just as a virtual machine OS doesn't care whether the hardware it sees is physical or virtual, a physical server backup solution doesn't care whether the hardware it sees is physical or virtual. A server is a server whether physical or virtual. This means that using a physical backup solution installed directly in guest OSs can get expensive quickly. If you have 10 virtual servers running on a physical host server, you are going to spend a small fortune by having to buy licenses for each of your growing population of virtual servers.

The other problem is management. One of the main objectives of virtualization is to reduce the amount of management needed to maintain the virtual infrastructure and thus reduce the costs overall. Using the same example as earlier, if you have to install, update, maintain, and manage 10 separate backup installations, it is going to get complex and costly quickly. Many organizations also want to avoid installing anything on servers unless it is absolutely necessary. Each piece of software installed on a server, no matter how reliable that software may be, increases the chance of unwanted interactions and stability problems.

Performance will also be degraded having so many guests with backup solutions installed directly on them, even if this is just an agent for a centralized solution. In fact, many organizations do not allow the installation of third-party agents on servers at all. If it has an agent, it won't be a viable solution.

Benefits of Enhanced Protection Through Virtualization

Virtualization in general provides a ton of cost savings and efficiency improvements over the traditional "one workload per physical server" model. After many years of virtual deployments, this has proven to be true time and again. So where do the benefits of virtualization fit with disaster recovery?

Simplification of the backup process is a key driver for businesses. Cobbling together a backup solution for physical servers of varying ages and from different vendors created a tremendous challenge for backup vendors.

What if instead of trying to balance the backup and restoration of servers between different hardware, there was a unified or greatly simplified hardware platform on which server OSs could run? This is exactly one of those benefits that virtualization delivers. The virtual hardware that a virtualization platform presents to the underlying OS provides all of the major I/O needs through a very small subset of device drivers. This means that migrating a virtual machine between physical servers is seamless. Most important, it means that a backup of a virtual server is going to work when restored to a different host. There is no extra layer of hardware to cause compatibility problems.

Another key benefit of virtualization that leads to fast recoveries is shared storage. Shared storage in the form of a storage area network (SAN) or network attached storage (NAS) allows for a pool of storage to be shared between all of the virtual hosts. In most cases, this is the single greatest expense for the deployment of a virtual infrastructure, but this cost comes with tremendous benefits.

This shared storage has to be designed to be highly available because it is the single critical point of failure for a virtual infrastructure. If all of the virtual machines are stored on the shared storage and it fails, then what would have previously been a failure of a single host now becomes a failure of all hosts.

The shared storage high availability combined with the need for redundancy and high performance is what contributes most to the cost. However, the ability to more fully utilize storage means that the traditional islands of storage on each host are eliminated, greatly increasing storage utilization. It also means that a failed host's virtual machines can be run from another host.

From a backup and recovery standpoint, the benefits are just as great. Virtual servers can be backed up to shared storage without the need to ever have backup traffic leave the storage. LAN-free backup means that virtual machines can be backed up directly to the shared storage without impacting the rest of the network. It also means the time required to take backups is much shorter than over the network.

Traditional Approaches to RPO/RTO Aren't Good Enough

Any features that can speed the time it takes to perform backups, allow more frequent backups, keep backups for a longer period of time, and do all of these more efficiently are going to be important factors when choosing a virtualization backup solution. The factors outlined earlier can be described appropriately as RPO and RTO. RPO represents the point in time from which you can recover and RTO represents how long it will take to do so.

RPO and RTO are quickly becoming two of the most important features in backup solutions. The reason for this is clear. Businesses have moved beyond the point of fretting over whether backups are going to work. It is just assumed these days that a backup is going to be consistent each time it is taken, particularly given advances in image-based backup solutions and disk-to-disk backups. Businesses expect that a backup solution is not going to fail.

Accordingly, attention turns to the problem of how much lost data is acceptable and how quickly the business can recover from a disaster. With physical server backups, the time it takes to recover can be several hours under ideal circumstances. For larger data sets, the recovery process can take much longer.

The backup windows for physical servers are decreasing while the time it takes to back them up always seems to be increasing as data growth compounds. Traditional physical backup methods mean that servers are only getting backed up daily in most cases, if even that frequently. It is not unheard of for critical servers to be backed up less frequently. This guarantees that there will be some data loss in the event of a disaster. This can be as much as a day or more of lost data in the event of a failure. This clearly isn't going to work in today's business climate. When businesses are reliant on hundreds or thousands of emails a day and mission-critical business databases, the loss of 24 hours might as well mean a catastrophe even if restoration is successful.

By embracing virtualization and the benefits that come with virtual servers, the RPO and RTO can be dramatically improved. Incremental image-based backup allows for many more recovery points to be kept while the time it takes to perform these backups in a virtual infrastructure can be reduced to allow for almost real-time backup. Likewise, restoration is quick, but more important, backups can be run directly from backup files without the need to restore, given the right solution. This means that RTO and RPO, no matter how aggressive, can be met.

What to Look For in a Host-Based Virtualization Backup Solution

As illustrated by the earlier discussions around physical versus virtual and host versus guest, there are many factors to consider when choosing a solution for virtual server backups. Of these factors, perhaps the most important is support from the virtualization vendor. Any chosen solution must support the means that the virtualization vendor has put in place for backups. Usually this is in the form of an application programming interface. An API is a supported means for a third-party piece of software to perform whatever function it does without adversely impacting the system. In the case of a virtualization backup solution, this is going to be critical given the fact that backing up virtual machines from the host, as we have discussed, is not as simple as just backing up files. Virtual machines are actively running servers. Just because these appear on the host's file system as physical files doesn't mean they can be treated that way. Doing so doesn't guarantee that applications and data will be in a consistent state and available for restoration.

Simplicity is perhaps the number two item to look for. Complexity is an inherent problem in IT. However, the more mission critical a piece of IT service delivery is, the simpler to operate it should be. I know that this sounds perhaps counterintuitive, but the best options are usually the simplest ones. Think of this like the IT equivalent of Occam's Razor, a principle that states the simplest solution is usually the correct one. When it comes to backup, it just has to work, all the time, every time. It doesn't make sense to spend hours configuring backups and options, most of which lead to confusion at best and failures at worst. Often this is only found out when it comes time to restore and it doesn't work.

If the learning curve for a product is measured in minutes instead of hours, days, or weeks, then it logically flows that costs to use it and support it will remain low. The more intuitive a product is, the more likely that a comfort level will quickly develop with the operators who have to use it to protect the virtual servers.

Coming Up Next

The next article will discuss ten of the must-have features for a virtual backup solution. This article will further explain the benefits that can be obtained with an appropriate plan for virtual server backup, and how to make sure that the solution you select will work best, provide the most ROI, and cost as little as possible to own.