WHITE PAPER

Virtual Data Protection – Which Way is the Right Way?

By Nick Cavalancia

UNITRENDS

TABLE OF CONTENTS

UNITRENDS

Virtualization was something that was supposed to simplify our lives. And, in some ways, it has, but in others it's complicated things. Virtualization has more than just allowed you to run multiple VMs on a single physical machine; it's ushered in an era where even the smallest company can host the most complicated multi-tiered applications. And while for some of you virtualization does just mean having a few guest OSes on a hypervisor, for others it's a more complex mix of clusters and server farms—all of which need to be protected in case of the need for recovery.

With that complexity comes the question of how to properly back up and recover each part of your virtual infrastructure. And the data protection possibilities don't exactly make it any more simple—with support for backing up and recovering at file, application, system, and hypervisor levels, as well as with the addition of using the cloud as a recovery target, it's no surprise there can be some confusion.

With support for backing up and recovering at file, application, system, and hypervisor levels, it's no surprise there can be some confusion.

Regardless of whether you only need to support the single hypervisor with a couple of guest VMs, or you have the environment so complicated you need a map to traverse it, you're each tasked with protecting every bit of data—whether at a hypervisor level, a VM, an application or even file level.

With so many levels possible, what's the right way to protect your virtual environment?

We all want the answer to be a simple "choose this" kind of answer, but the reality is selecting how to back up your virtual environment—whether at the guest or hypervisor level, as well as whether to incorporate the cloud as part of your recovery strategy - all depends on a number of factors involving what you want to protect, how quickly you want it recovered, how aligned your applications are to your virtual architecture, and what disasters you're trying to protect against.

Thoroughly confused? You should be. There's a lot to consider.

Then where should you start?

The answer may surprise you, as it's not the standard "take an inventory of your virtual environment, calculate space requirements, etc." type of

response. The right approach starts with you working from recovery backwards.

Start with Recovery

Despite how simple or complex your virtual environment may be, determining the right way to provide protection begins with the key differentiator of what needs to be recovered. Don't start with the various parts of your virtual infrastructure trying frantically to figure out which ones are critical and which aren't. By doing so, you're not focusing on the value that virtual environment brings to your organization.

What you really want to protect and recover is data, systems, and applications that are important to business continuity. Virtualization is simply the method by which the data, systems, and applications are delivered.

Determining the right way to provide protection begins with the key differentiator of what needs to be recovered.

Breaking it down

So begin with the specific data, applications, or machines you want to recover, defining what makes up the backup set. This could be anything from a simple group of files, to a databases and logs, to an intelligent selection (on the part of your backup solution) of everything that makes up an application—files, databases, logs, services, etc.—to an entire system.

Multi-tiered applications get a bit more complicated, as you need to identify each system's backup needs separately, but you do need to be thinking about recovery potentially as a single process.

You should next establish your recovery time objective (RTO) and recovery point objective (RPO) for each to define exactly how much time you have to recover and how far in the past you're willing to recover to. These will come in handy later as you decide whether to back up at a guest or hypervisor level.

Once you've identified the specifics of what needs to be recovered, then map it to your virtualized environment. It may be as simple as one or more guest OSes, a single host server, or can be far more complex, requiring specific Guest VMs from multiple hypervisors be included. This helps you understand what parts of your virtual environment are impacted by a given recovery set.

But, then, there's still the unanswered question of whether something needs to be backed up within the Guest OS, or whether the hypervisor-level backup is appropriate.

As, you'll see, the answer isn't perfectly straightforward.

Data Protection: Guest or Hypervisor?

If you're new to backing up a virtual environment, the simple rule of thumb is backups of the hypervisor get the entire machine and don't require any agent to be installed inside the VM, while backups at the guest level use agents and can provide more detail and granular insight inside the VM. So even if you know you need to protect an entire, single-role server, such as a domain controller, the answer may not be clear.

There are a few factors you should consider before choosing to back up at the guest or hypervisor level.

What determines when each backup type is the right choice?

In general, recovery of files, folders, and applications that are not mission-critical or performance sensitive are all suited-well for hypervisor level backups. The majority of your systems will fit into this category. But that small percentage of systems running applications or containing data that absolutely are a necessity for the business to operate that require more backup granularity and have tighter recovery objectives are prime for guest level backups.

To find the selection that's right for each recovery set, there are a few factors you should consider before choosing to back up at the guest or hypervisor level.

Deploying and Managing Agents

Deploying and managing individual backup agents inside virtual machines can be a tedious process. For VMs that can be backed up at the hypervisor-level, no agent is required so you can avoid this management challenge. However, for certain types of workloads that require more granular recovery and deeper insight, using backup agents in the VM can be extremely beneficial.

The RPO and RTO

Like RTOs, RPOs can also restrict your choice of backups approach.

Fast recovery points for a specific data set can't be met at a Hypervisor level because each backup includes the host machine and the entirety of all of the guest VMs. If every guest VM is part of the recovery set and the backup frequency of the hypervisor meets the recovery point objective, it's still a viable choice. Otherwise, you should backup at the guest level via a backup agent.

Backup Size

From a production standpoint, you only have a select window of time to backup, which impacts the potential size of the recovery set. Virtual machines these days are getting pretty large and may extend backups to a point where even with change block tracking to speed up the recovery set selection process, the size of the recovery set expands beyond the allotted backup window.

The choice of backing up at a guest or hypervisor level depends on many factors, each pushing the needle one way or the other.

From operational standpoint, you get better deduplication performance with guest level backups, as the deduplication process is able to compare the current backup data against past backups at a more granular level improving deduplication results, resulting in smaller backup data sets. Granular backups of specific data, which are possible with guest-level backup, can also shrink the dataset and allow for more frequent, smaller backups.

Application Requirements

Sometimes applications themselves will dictate the answer. Keeping with the Exchange server example, a hypervisor level backup of an entire server VM will certainly provide you the ability to recover the server, but the Exchange databases won't be in a consistent state to be mounted and accessible. So, in cases like this, you'd need to both backup the VM at a hypervisor level, as well as utilize an agent within the guest VM and backup Exchange. That way you can recover the entire server quickly and then restore the databases to bring servers like Exchange back into a functional state. Alternatively, you can simple stick with a guest-level backup of the entire server for recovery.

As you can see, the choice of backing up at a guest or hypervisor level depends on many factors, each pushing the needle one way or the other. Once you've identified the correct choice for each recovery set,

let's add another dimension into the conversation—the use of the cloud as part of your recovery.

Adding in the Cloud

The cloud adds a dynamic to both backup and recovery by providing new backup sources (with servers residing in the cloud), new recovery targets (systems can be recovered to the cloud), and new recovery options, such as continuous recovery (where a standby system is recovered in the cloud immediately after a backup, keeping it up-to-date) and certified disaster recovery (where backups are immediately tested and entire infrastructures are recovered in the cloud using automation).

No matter how much of it you take advantage of, the cloud addresses some level of protection against loss and provides continuity and availability for your business.

No matter how much of it you take advantage of, the cloud addresses some level of protection against loss and provides continuity and availability for your business. Whether it's loss of local backup data, you have an extra copy of them in the cloud. If a loss of operations, you can spin up one or more servers in the cloud. If a loss of location, you can spin up an entire business there if you need to.

There's obvious benefit the cloud brings to your backup and recovery efforts. But, like the guest vs. hypervisor topic, it's not as clear as to when it's the right choice to protect your virtual environment.

So, when should you consider use of the cloud?

It's important to first point out that this is not a case of either/or. That is, the cloud is not something that either your backup and recovery goes to or it doesn't. It's just not that simple. In reality, the cloud is an additional layer of protection that sits on top of your already existing on-premises efforts.

Using the cloud should always be an option. With the cost of cloud storage getting less and less expensive, every company should at least have a copy of company data offsite for long-term retention, and to protect against catastrophic failure. That means keeping copies of your guest level data as well as your VM images in the cloud just in case your on-premises storage fails you.

In situations where there's zero tolerance for downtime, while far more costly, the ability to spin up the compute side and bring a cold, warm, or hot virtual site live is valuable to maintain business continuity.

Picking the Right Cloud

You've got a few options around what kind of cloud environment to utilize. There's public cloud providers—like Unitrends, Amazon, Azure, and Rackspace—and private cloud providers that would more align with a service provider-hosted datacenter. And not all clouds are made alike.

You've got a few options around what kind of cloud environment to utilize.

In general, hyperscale public clouds today are only good for backups and archives. They simply don't have the recovery functionality to spin up the compute side necessary using the backup copies stored in the cloud. So use of these public clouds would fit with your desire to have a duplicate copy of all backup data offsite. Some backup providers are also providing public clouds that do allow for recovery and spin up of the computer resources in conjunction with storing backup copies.

Private clouds are generally used for both backup retention, as well as to act as that cold/warm/hot site (add heat as needed). There, you have the ability to failover systems, connect users via VPN, operate the business, and fail back data, applications, and systems to primary data center once the disaster has been rectified.

Mixing the Cloud and Virtualization

As you can see, the cloud definitely isn't a replacement for any of the backup and recovery approaches previous covered; it's an added option that provides a layer of recovery for disasters that cause your backup data to no longer be available, as well as up to and including when your entire virtual environment is no longer available.

Virtual Data Protection The Right Way

It's apparent by now that there isn't a "one size fits all" method to protecting your virtual environment. It's really more about what data and applications are important to your organization, and how those are represented in your virtual environment.

Bringing your recovery objectives into the mix establishes what's really important and determines recovery priority, which, in turn, provides direction on whether you should be backing up a specific recovery set at a guest or hypervisor level. And it's equally important to remember that some applications may require both to ensure the application is recovered to a functional state.

The cloud's role in your recovery will really reflect the maturity of your data protection strategy. If you're simply trying to protect against the loss of a specific server or simple application, the cloud doesn't need to be in the mix. However, if your strategy is mature enough to consider protection from multiple types of "disasters," which include the loss of data, application, system, location, and operations, the value of the cloud becomes more evident—both as a storage target, as well as a potential recovery target.

The cloud's role in your recovery will really reflect the maturity of your data protection strategy.

And with that growth in maturity comes the development of your recovery strategy, which shifts from the basic recovery of files, to recovering entire applications, to complete systems, to entire environments.

No matter how you choose to protect your virtual environment - whether based on files, applications, or systems, at the guest or hypervisor level, on-premises or in the cloud, by doing the work outlined in this whitepaper and taking steps to improve your recovery position, your virtual data protection is headed the right way.

With nearly 20 years of enterprise IT experience, Nick Cavalancia is an accomplished consultant, speaker, trainer, writer, and columnist and has achieved certifications including MCSE, MCT, MCNE and MCNI. He has authored, co-authored and contributed to over a dozen books on Windows, Active Directory, Exchange and other Microsoft technologies. He has spoken at conferences such as the Microsoft Exchange Conference, TechEd, Exchange Connections, and on countless webinars and at tradeshows around the world.

UNITRENDS