WHITE PAPER

Availability and the Always-on Enterprise: Why Backup is Dead

Backups certainly fit the bill at one time, but data center needs have evolved and expanded.

By Nick Cavalancia

VEEAM
IT JUST WORKS!™

Every business has experienced a significant increase in customer and partner expectations around application performance and business continuity. Response times once measured in terms of seconds are now milliseconds. Downtime measured in hours or days is now expected to be minutes (if even that).

To keep pace, many organizations are implementing modern data centers. They're deploying new technology to drive performance such as virtualized servers and networks; modern storage; and cloud-based services such as infrastructure as a service (IaaS), recovery as a service (RaaS) and disaster recovery as a service (DRaaS). In fact, 97 percent of organizations have or will have a modern data center within 2 years to address the 46 percent of their workloads they considered to be critical.

Until the last few years, the concept of availability was only a dream.

Implementing these new technologies—as well as the adoption of cloud services—demonstrates a shift in IT thinking. Instead of focusing on just servers and data and the technology needed to back them up, IT is now considering the concerns of the organization as a whole, including service delivery, operational efficiency, reputation, and most importantly availability.

The need for availability is a key driver for the modern data center in 68 percent of organizations.[1] Most organizations—65 percent—need real-time interaction with critical applications and systems[1] to meet the present demand. Essentially, they need to never be down.

Until the last few years, the concept of availability was only a dream. Backups were all that was available. Making a backup copy of some data set and storing it on tape or disk was enough to feel protected. Data sets and applications were much less complex, and recovery times in terms of hours were actually acceptable.

These days, with the expectations around applications and service availability and the pressure for organizations to be always-on, the stakes are higher. Organizations can no longer simply rely on age-old protection technologies and methods.

[1] Veeam, Data Center Availability Report (2015)

So, is it high time the always-on enterprise killed backup? To answer that question, look at where backups fall down. Define the gap that exists between availability needs and backup capabilities. And look at how to meet the need for availability within your organization.

You Can't Rely on Backups Alone

Backups were the answer when making a copy of a particular set of files or even a complete single server was considered adequate protection. Today's modern data center is far more complex. There are multi-tiered applications, virtual servers being moved from one host to another and the cloud handling both storage and compute functions.

In some ways, basic backup no longer has a place in the modern data center.

With organizations considering availability, it's no longer about simply needing to restore a single file or folder. It's about complex processes like recovering a multi-tiered application that spans multiple servers and bringing each server into a consistent state with the others. It's a far more involved than the restore jobs of yesteryear.

In some ways, basic backup no longer has a place in the modern data center. As new technologies have come into use—the foremost being virtualization—you can now easily move workloads from one location to another. You can even performing maintenance during the day. The options around availability are much greater and more flexible than what backups alone provide.

Even so, the idea of a backup—that is, having copies of your data—is still viable. Now data centers have moved to advanced concepts like replicating data at the disk level or entire virtual machines, both from one store to another or even one site to another. This provides both increased protection and faster recoverability.

Organizations today aren't just looking at availability on a per-application or per-server basis. The goal is to make everything available in the event of an outage. So it's not "we have our order processing back online, but e-mail is still down." Now it's essential to have the entire business back up and running, not just a few services.

Should you have an availability event, can you benefit from backup alone? Backups certainly still have a place. For example, if you're replicating changes to a VM and the source VM is somehow

corrupted, that corruption will simply get replicated. So having a backup of the critical data on that server can play a role in ensuring recoverability. However, backup as the only method is no longer an option for businesses focused on being available.

As newer technologies have emerged, the frequency of backups has also shortened. In previous years, your backup window simply couldn't be anything less than nightly. These days, backups occur much more frequently—even during production hours. And with technology like instant VM recovery in place, the concept of restoring a backup job is somewhat obsolete.

As newer technologies have emerged, the frequency of backups has also shortened.

You've modernized your data center, but have you also modernized your data protection strategy? Perhaps, it's time to do so. The process begins with understanding just how far away you are from availability before you can get there.

The Availability Gap

Once you've recognized the need for availability and acknowledged the inability for backups to meet the need, the first step in achieving availability for your data center is to define the gap in your organization between backup and availability.

With backups, you're thinking about the data that needs to be restored—a few files, a folder, a database or a server. The business doesn't define availability in terms of a specific data set, but in terms of resiliency in the face of an event that interrupts operations.

So if you're still relying on traditional legacy backup—with agents installed on every VM, nightly backups and recovery times that still take hours—but your organization has little or zero tolerance for downtime, the difference between the technologies in place versus what the business requires and your ability to meet those requirements is your availability gap.

Put Availability in Backup Terms

It's fine to discuss the gap as concept, but that probably doesn't help you without some tangible numbers to act as guidelines. So when you consider backups, take your recovery time objectives (RTOs) and recovery point objectives (RPOs). They likely differ from

system to system and application to application. You can think of availability as having an RTO and RPO of less than 15 minutes each. Consider that for a moment. That's 15 minutes to recover not just a single server or application, but potentially your entire environment—all applications and data—in that same short timeframe.

If you believe you have an availability gap, don't feel bad. According to the Veeam survey, 82 percent of CIOs said they couldn't currently meet the availability requirements of their business, which means they too have an availability gap[1].

The fact is, though, that you can make availability a reality in the near future.

In the past, availability was a difficult concept to comprehend. It simply wasn't cost effective. But today availability is attainable with the technologies in the modern data center, such as backup, replication, and storage snapshots, virtualization, advanced storage and cloud services.

But is availability truly necessary? That's up to your business. If it's expected, then yes. You can't deliver availability tomorrow. It may be far more of a journey in terms of what you can deliver, whether you have the budget, and so on than just enabling availability. The fact is, though, that you can make availability a reality in the near future. So how do you create availability?

Get to Availability

First, remember the goal of availability is to get your entire business operational within 15 minutes. That's a broad definition, so you'll need to start somewhere a bit more focused, like per application. For each application in your environment, talk to users, application owners, business managers—anyone who is affected by an application's availability. You need each one to define what the business requires. In doing so, you will shift your thinking from what IT normally can deliver to what the business needs IT to deliver.

Next, build a list of the types of interruptions that can affect your business. These can be anything from corrupted data or loss of a server to a disaster that makes your data center completely inoperable. You'll find the recovery tactics you use are very different

[1] Veeam, Data Center Availability Report (2015)

from one interruption to the next. Cross-referencing the availability requirements with the potential interruptions will help guide the conversation around what steps you'll need to take in each scenario.

The mix of business requirements and interruptions on a per-application basis will eventually need to be translated back to a definition of recovery execution, which should include:

Recovery Windows: While 15 minutes is one definition of availability, you need to define those RTO and RPO values that meet your business needs.

Recovery Methods: Include the use of replication, snapshots, disk synchronization, image and file-based backups and any other methods you currently employ as options.

In reality, though, your recovery plan is only as valuable as the testing you've done.

Recovery Targets: Consider the use of your data center's current physical location, as well as both the public and private cloud as potential choices.

Recovery Management: Determine if your IT team or another group will manage recovery. This could mean managing the recovery process, infrastructure, specific restores or anything else. Consider how your choice affects your ability to be available.

Regardless of the specific mix of recovery options with which you achieve availability, it's critically important you practice recovery.

Recovery: Practice over Concept

Almost 99 percent of the time, you're not trying to recover a specific item. You're in a steady state where everything's working. It's easy to plan for recovery and assume everything will work. In reality, though, your recovery plan is only as valuable as the testing you've done. Don't be lulled into a false sense of security thinking that because you have a backup of a given application that you're covered. You need to actually practice your recovery by testing all of the following:

- Test the recovery itself by attempting to bring up the application in an alternative location or in the cloud.
- Have it interact with the rest of the network to ensure it works with directory services and other relevant applications.

- Have users attempt to use it to ensure not only does the application work, but the data is viable.
- Ensure you can perform a failback to make sure once the availability event is no longer an issue, you can do so.

Having some form of automated recovery testing will be important here to ensure you can test every application that falls within your availability strategy, as well as ensuring you've performed every test. Remember, the goal here is to be able to say with confidence that when you boot up a recovered server or application, it will work.

Your organization must have standards for what is and is not acceptable downtime.

There are a few other use cases where availability can be extremely helpful. Before applying patches, internal development releases, and any other kinds of updates to critical systems and applications, you can take advantage of steady state backups in an alternate environment to test any needed updates—all without impacting production. Also, if you're having issues with a production application, you can spin up a replicated copy of that server in an isolated environment and troubleshoot it there instead of on the live server, thereby reducing the potential impact on production.

Achieve Availability

If your organization is like most, you've already begun or have made the investment in a modern data center. Despite your desire to simplify what you manage, it's a complex mix of virtual machines, servers, storage and networking. Because you've made the investment to meet the demand to maintain operations, traditional backups alone just won't scale to meet the availability needs of the organization in such an advanced data center environment.

Your organization must have standards for what is and is not acceptable downtime. Comparing the businesses' required levels of availability against what's currently attainable can help you create a service baseline from which to work towards availability. Begin with the business requirements around application and environment availability, instead of what your backups can do today. This will help IT look for ways to cost-effectively take advantage of current technologies or invest in new ones to make meeting availability requirements a reality.

As with many things, technology alone will only get you part way there. Put firm definitions around recovery, the interruptions against which you're protecting, and the process necessary to realize recovery times consistent with availability. You'll begin to craft a strategy you can implement, test and put to use.

Backup is little more than a means to an end. The modern data center has surpassed what backups can provide in terms of speed, capability, and performance when it comes to availability. While you may still keep your backup systems around for a few more years, with availability as your goal, you won't need them much longer.

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.