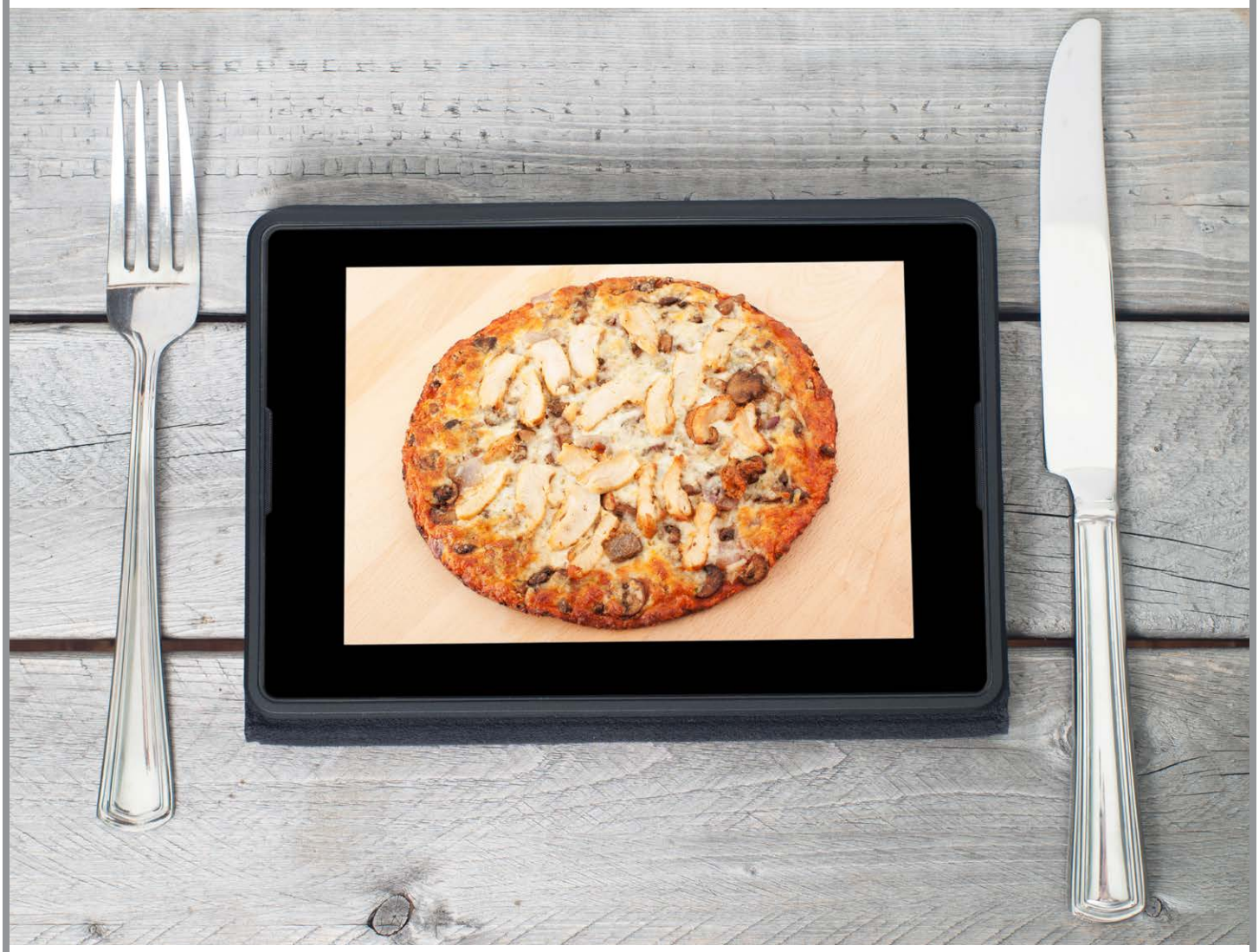
WHITE PAPER

Pizza, Slow Delivery, Data Protection, and You



By Nick Cavalancia

UNITRENDS

TABLE OF CONTENTS

UNITRENDS

Pizza delivery and data protection are more alike than you think.

We've all ordered pizza before. A seemingly endless choice of toppings, crusts, sides, and drinks - all promised to be delivered to your door in a reasonable amount of time. And, for the most part, we all get exactly what we ordered around the time we expected it. But, if you're part of the unfortunate few, you've seen it go wrong… really wrong. You know - 90 minutes late, completely cold and it wasn't even what you ordered.

So, what does this have to do with data protection?

Pizza delivery and data protection are more alike than you think. Pizza on a Friday night comes down to two simple concepts: the order and the delivery. That same Friday night can as likely be spent in a server room where you are frantically trying to recover some critical server and, again, it comes down to two simple concepts: the order (that is, what, and how, you chose to backup your critical data, applications and/or systems), and the delivery (that is, your expectation of the recovery in terms of recovery time objectives (RTO), recovery point objectives (RPO), and the usability of the system post-recovery).

For your data protection to be successful, like ordering pizza, there are a few parts of the process where things can go very wrong. So, in this rather unique whitepaper, we'll cover the three parts of the pizza delivery process you need to be focused on:

- What's on the menu?
- Slow delivery
- The order isn't right

We'll take these three parts, demonstrate how backup and recovery follows suit, and what you can do to ensure a successful and on-time delivery (of your data, applications, and systems, that is…).

What's on the menu?

Everyone's got their own preference of what kind of pizza they like (just ask a group of friends and you'll hear some strange choices), so, while everyone likes cheese pizza, there are times when no two orders will be identical. That's why pizza companies have so many choices of toppings, crusts, shapes, etc.

The same is true for data protection. There will be the simple "entire server" kind of backup, but when you consider the services, applications, data sets, and servers that need to be backed up, and apply the recovery time and recovery point objectives for each, you should find that, in general, data protection can be a bit more complex and, therefore, no two backup "orders" are the same either.

How can you be sure you have enough options at the time of ordering a backup to ensure a successful delivery at recovery time?

There are a number of options that should be on your data protection menu you need to consider to ensure a successful recovery:

Backups strategies usually align with assumptions around whether you'll be recovering to a physical or virtual machine.

Menu Choice #1: Backing up to Physical, Virtual or Both

Backups strategies usually align with assumptions around whether you'll be recovering to a physical or virtual machine. So if you are planning on recovering a physical machine, you'll backup system states, files, and folders. But if you are planning on recovering to a virtual machine, your thinking in terms of block-based image backups.

When should it be one, the other, or both?

The reality is you need to have the option of either, and base your choice of backup on whether or not you can meet the needed RPO and RTO with the current granularity afforded by each backup strategy.

Menu Choice #2: On-premises, to the cloud, or go hybrid

With each backup methodology, accessibility is the key. Both on-premises and cloud-based backups have situations where accessibility is a concern. On-premises can suffer a hardware failure, or simply not be viable if a disaster affecting the location occurs. And cloud obviously needs an Internet connection to even work.

Is the answer to just go hybrid?

Hybrid addresses the accessibility issue by giving you both on-premises and cloud sources, making it not only a viable option, but one that also

provides unlimited storage, high availability, and a shift from expensive capital purchases to a lower monthly expense.

Menu Choice #3: Using commodity hardware or an appliance

Off the shelf NAS or SAN solutions often times can be used as the local storage at a low cost entry point. However, many backup vendors offer hardware appliances as a way of integrating backups and simplifying administration. Some appliance also provide additional services, such as acting as a host for guest VMs for virtual recovery, deduplication, encryption, and WAN optimization when syncing to cloud storage.

So what's the right option?

Archives are one of those parts of a backup strategy that are usually only considered when compliance standards require it.

While commodity hardware can require less up-front costs, it does require you to first procure the right backup software first and then the right hardware to ensure compatibility. Also commodity hardware has hidden up-front costs around installation, implementation, testing and support.

Appliances, on the other hand, are integrated with backup solutions already, are purpose-built, and provide improved performance, greater scalability, and incorporate those additional backup-centric services previously mentioned.

Menu Choice #4: Backing up or Archiving

Archives are one of those parts of a backup strategy that are usually only considered when compliance standards require it. And if you do have archives, it's usually a case where your backups become an archive instead of making a backup for the purpose of it being an archive.

Do you really need to archive, or are backups enough?

In the event data becomes corrupted, maliciously deleted, or is needed as part of a legal action, it may be necessary to go back farther than the standard retention time to retrieve copies of data that either did not change within the retention window (and, therefore, weren't backed up), or were only viable (in the case of corruption or malware) prior to 30 days ago.

So, archives do play a role in data protection, which makes it even more important to have a strategy around what goes into your archive to ensure when you need something, it's in the archive.

Make sure what you want is on the menu

While the list covered isn't comprehensive, it does represent the need for you to ensure you have options before starting a single backup. Think about what you want to see upon delivery, and make the appropriate choices today. If you don't have the right options, you'll be stuck placing an order for a cheese pizza that you aren't going to be happy with.

Once you placed your backup order, the only thing that matters is how long it will take to be delivered. And, in the case of data protection, that can't take long.

You need the delivery of recovered data, applications, and systems to be comprehensive enough to be back up and operational, but you need that delivery to be as fast as possible.

Slow Delivery is Unacceptable

It's an obvious one. "30 minutes or less" is all about the expectation that a pizza must not only be hot upon delivery, but also needs to be delivered quickly. In business terms, you want it functional and you want it now.

Data protection is no different. You need the delivery of recovered data, applications, and systems to be comprehensive enough to be back up and operational, but you need that delivery to be as fast as possible.

Why, then, do some recovery jobs take longer than expected?

There are three aspects of data protection that impact delivery (as well as backup, in some cases): the level of backup & recovery used, how backups are processed, and the backup medium used. By focusing on these issues, you can determine what changes you need to make in order to make your recovery as fast as possible.

Slowdown #1: Backup/Recovery Levels

Backups, in general, are backed up (and restored) at one of three levels: block level, file level, or image level. There are clear circumstances when each is appropriate - for example, protecting an entire VM would

logically utilize image backups, or if you just need a project folder full of files protected, file-level is the obvious choice.

But is there a single correct choice?

No one backup level has it all together. All of them have pros and cons. Block level has the benefit of providing the smallest backup, but recoveries tend to be more complicated since they may require multiple restore operations.

File-level backup and recovery is ideal for enterprise applications like SQL Server or Exchange, where recovery is required at a deeper level, but may not include everything you need (e.g. OS, applications, data) in a single backup set.

File-level backup and recovery is ideal for enterprise applications like SQL Server or Exchange.

Image-level backup and recovery is the quickest thanks to technology like change block tracking, but is also the least efficient, as an image alone is an "all or nothing recovery..

The fastest route to choosing the right level of backups should be based on what you are wanting to recover. The mistake is choosing one level of backup as "the best" and having every backup utilize that same level of data backup and recovery.

Slowdown #2: Processing Backups/Restores

Think back to years ago when backups were little more than simply copying backups to tape, perhaps with some level of compression. If you were to attempt to use that same basic method of processing a backup today, especially when using cloud-based storage, you'd saturate your bandwidth pretty much all of the time (and probably never complete a backup).

What, then, needs to be a part of the process to speed up backups?

Specific technologies are in use today with many backup solutions that optimize what will be selected as part of a backup, ensuring what actually gets copied across the wire is unique, and how much space it takes up. The goal is to have as little as possible go across the wire,

while ensuring the most data possible can be recovered. Sounds a bit impossible, but with the right technologies in place as part of your backup solution, it's actually quite feasible. Data protection should include the following technologies:

Change Detection

It's most critical to only backup what's actually changed and track those changes. Backups solutions can detect changes at a block, file, and image level, so this is a fairly common, but important, feature.

Deduplication

This is the process of taking a set of data and looking for repeat patterns of information. For example, a pizza company only needs so many instances of the word "pizza" (and there's probably quite a few of them) in their backups. Deduplication shrinks down the backup data set by only sending the word "pizza" once and then a representative smaller data set in subsequent instances. The important part of deduplication is the data set scope – does deduplication occur within a single backup, across an entire server's data, globally across all backups, or something in-between? Knowing the answer for your current backup solution should influence how you group data sets in your backups. For example, if you backup your multiple file servers as separate backups, you may want to consider consolidating them to increase the influence deduplication has, thereby reducing your backups even farther. Deduplication should be application aware to also avoid a slowdown from trying to dedupe, say, your inventory database on SQL Server with your email on Exchange.

A pizza company only needs so many instances of the word "pizza" (and there's probably quite a few of them) in their backups.

Compression

Compression both speeds up WAN transmission, as well as optimizes the storage used. So every vendor has some type of compression technology, but not all are identical in methods or position in the process. For example, some backup vendors only compress before sending up to the cloud (but don't for local storage).

Slowdown #3: Your Backup Mediums

We've already talked a bit about mediums, but not from the perspective of speed. Whether you backup on-premises, in the cloud, physical or

virtual, the medium you choose will dictate how long the backup window is, as well as how long it will take to deliver your recovery. Let's look quickly at a few seemingly opposing pairs and figure out which is fastest.

- **Cloud vs. Hybrid** - When doing recovery, the cloud, as a general rule, is going to be slower than local backups as part of a hybrid implementation. Be sure your backup solution allows either the automatic selection of the fastest recovery source, or a manual selection.

- **Virtualization: Host vs. Guest** - You will likely protect at the hypervisor level, but for those application where the RPO is critical (an ecommerce system, for example), protecting the guest OS data (which will require an agent within the guest) makes sense to get the granularity and deduplication of that specific VM's data.

- **Local Storage** – When it comes to local storage, you have the choice of SAN vs. NAS, as well as Disk vs. Flash. SANs provide the better performance than a NAS, with their dedicated data paths and storage, as well as their ability to write to block directly (writing to a NAS device requires writing to the NAS OS first, and then to disk). Flash adds value with its' demonstrably faster speeds, but traditional disks are still much cheaper per GB.

Even if you are able to select the right data sets and make the right choices to ensure the quickest recovery, like your pizza order, it still may not be what you wanted to see.

Speeding up delivery

Once again, the answer isn't as simple as "choose this one for the fastest delivery." The answer lies in looking at what you are trying to protect, what disaster scenarios are you preparing for, and then choosing the right medium(s) that help get you there the quickest.

Even if you are able to select the right data sets and make the right choices to ensure the quickest recovery, like your pizza order, it still may not be what you wanted to see.

The Order Isn't Right

We all know in the case of pizza delivery, should the delivered order be wrong, it's probably the fault of the pizza place and not you, right? But in the world of data protection, restores can complete properly, but not

leave you in an operational state. And, unlike the pizza place's fault, it's more likely that it's either you, or the way you thought the backup would work during a recovery that caused you to still be scratching your head.

Take for example a bare metal restore of an Exchange server. You can setup your backup job to back up the System State and all the files on the Exchange server, but when you perform a Bare Metal Recovery of the server, you learn the BMR process simply gets the OS up and running (complete with Exchange services). The database files are either incomplete or, at best, in an inconsistent state, requiring you to properly restore the Exchange databases separately to get them back into a consistent state. So, the backup you thought would address your recovery, didn't. And it wasn't the backup's fault.

To have a successful recovery, you're going to need to be thinking about what's needed for each type of recovery.

So what should you consider to ensure the order is right?

Double-Checking Your Order

To have a successful recovery, you're going to need to be thinking about what's needed for each type of recovery and then, as mentioned at the beginning of this paper, work backwards to build a proper backup strategy.

Remember, when you back up at a file, application, OS, or image level, you're also dictating the recovery choices possible.

With the goal being to get a particular service or application up and running, you'll need to define what's necessary to restore operations – this could be a set of files, an entire server, a VM, or even multiples of each – and ensure those are all part of the backup set. Better yet, choose a backup solution that intelligently does this for you!

Lastly, you can check your order by putting it to the test before you ever need it. That's right – test your backups. Perform that BMR of the Exchange server and see what's missing! Don't wait for the disaster to come so you can see if your "pizza" will be delivered on time and as ordered.

Some backup partners even offer some type of Certified Recovery, where the partner orchestrates recovery testing in a sandbox where

virtual machines are tested to ensure they meet stated RTOs and RPOs. A great example is ecommerce using some kind of web service, DNS, etc. The certified testing ensures everything spins up properly and functions together. By either using a trusted partner, or simply doing the work yourself to make certain applications and servers that need to work together do more than just recover, but actually work together post-recovery, you'll ensure you're order will always delver as promised.

Turning Bad Delivery Into Good

You have a lot of choices when it comes to data protection – more so than when compared to pizza delivery. Unlike ordering pizza, the trick here is to simply think about what's needed at the end – in some cases on a per application-bases - and design an order (or, more likely – orders) that will meet your recovery objectives – including everything from where you choose to store your backups, to how fast they can recover, to what data sets to include for proper deduplication, to which servers and applications are needed to truly be considered operational.

You have a lot of choices when it comes to data protection – more so than when compared to pizza delivery.

By breaking down what data, applications, and systems are important, how they need to be recovered, selecting the right backup solution that meets your needs, and ensuring you are prepared for the disasters you deem likely, you'll find data protection to be nothing like bad pizza delivery, and when you need to place that recovery order, it will always come in as ordered, on time, as promised.

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