# Developing a Backup Strategy for Your Businesses

an INFOSTOR® Storage eBook

# Contents...

## Developing a Backup Strategy for Your Businesses





*This content was adapted from the Small Business Computing, IT Business Edge and Enterprise Networking Planet website. Contributors: Jennifer Schiff, Paul Mah and Peter Eicher.*

# A Guide to Backup and Recovery

By Jennifer Schiff

No matter what the size of your business, if you store your business and/or customer data on a computer (and who doesn't these days?) you need a safe, secure way to back up and store that data. Indeed, many experts argue that choosing a data backup system or service should be one of the first things you do as a business owner , right up there with finding an accountant, a lawyer and a bank.

When contemplating which backup method to use, businesses are faced with an overwhelming number of choices. Should they back up locally or use a cloud- or Internet-based backup service —or both? For many growing businesses, online backup options made a lot of sense when the business was starting out, but things are more complicated now.

If your business is now leaning toward backing up locally, it can back up to tape or disk — or to a virtual tape library (VTL) or removable disks — or some combination of the two (e.g., disk-to-disk-to-tape, or D2D2T). And if you decide to back up your data to a cloud provider, how do you determine which one is right for your business?

Patrick Corrigan, a senior analyst at Storage Strategies NOW, also believes that small and midsize businesses should use online providers as part of their backup and recovery strategy.

"Typically what I'll recommend to smaller businesses is that they take critical data and make sure it's backed up in some local fashion, whether it's to CD, DVD, tape or removable disk or some type of backup server," said Corrigan. "They [should also] take the same critical data and back it up to some online service like Carbonite."

He urges growing businesses to use both methods and not to rely on either option alone. He also recommends that businesses store copies of critical data in a safe place located offsite, in case of flood or fire or theft.

The leading advantage to backing up locally, said Corrigan: "It's generally quicker to [retrieve data] off a local system than it is online, especially if there is a lot of data to recover."

As to what type of local data backup system businesses should consider, Corrigan is a fan of disk-based systems — or systems that combine disk with tape. They've dramatically come down in price over the last few years (rivaling the cost of tape-based backup systems in some cases), and they're faster than tape.

Whichever method you choose, it's ultimately a matter of cost and (or versus) recovery time. Business owners should weigh all factors, including ease of use, support and especially recovery time, into their decision when choosing a local backup system, and they should not be afraid to ask vendors questions.

## Saving Data: What to Back Up

The general rule of data backup is to back up all information that you are legally required to keep or that is essential to running your business. Not sure what that is? One exercise is to think "If I walked into my office tomorrow and nothing was there, what are the most critical applications [and data] I need?" And don't back up files you haven't touched in years, unless you have a legal or regulatory requirement to do so.

Do you need to backup your operating system and software? If fast recovery is an issue, go ahead and image your computers, said Corrigan. If you have any proprietary applications, software the company built itself or had developed specifically for the business, absolutely back it up.

## How Often Should You Back Up Data?

Most data backup and recovery experts agree: You should back up your data daily. Set your system up to automatically back up once a night, even on days when the office is closed.

## Check That Your Data Is Being Properly Backed Up

Most (if not all) backup software and services provide review logs and reports, so you can check that your data was properly backed up (or if there was a media failure). And ideally, you or someone in your company should check the logs each morning to quickly make sure data was properly backed up the night before.

If that task is too time consuming, or just unrealistic, here's the bare minimum: schedule a time once a quarter (if not once a month) to review your backup logs or reports.

## If You Run an Online (Hosted) Business

If your business includes an online business, i.e., have an ecommerce site, chances are your hosting company backs up your site (including precious customer information) on a regular basis — but don't take it for granted.

That's why, before you sign up with a hosting provider you should ask about its data backup and recovery procedures, as well as what steps it is taking to keep your customer data safe and secure. And don't be afraid to ask the provider to do a test restore, even if the provider charges a modest fee.

As an added precaution, consider keeping a backup copy of any hosted sites or data site locally — or backing it up to another online service, so you have the data stored in at least two places, advised Corrigan. "No matter how reliable your ISP [or ecommerce host] seems, stuff happens," he said.

Stuff gets deleted; sites crash; breaches happen. "But if you have your data stored in two places, if it goes away in one place, at least you'll have it in another," said Corrigan.

The bottom line, said Corrigan: You can never have enough backup. And although it may seem expensive, backing up your business-critical data to a secure device and/or service, and making sure that the data is properly backed up and stored on a regular basis, is much cheaper than the alternative if something goes wrong. ■

# 3 Reasons Proper Data Backup Matters

By Paul Mah

We've all heard about the dangers of data loss, the importance of data backup and, in particular, ensuring that your key small business data is adequately protected. This is why most growing businesses these day make use of NAS appliances with at least RAID 1 configured for data mirroring protection.

While the most critical data is typically adequately protected, associated information such as source code, business documents and yes, even email messages are very often left out of regular backups due to cost considerations. In addition, proper backup procedures mandate that a separate copy of all pertinent data should be made, and preferably stored at a different location — stringent requirements that may not be adhered to in the first place.

Today, we take a look at three types of problems that can arise in the absence of proper data backups and disaster recovery planning.

## Malicious Intent

The disgruntled IT worker is hardly a new phenomenon. Indeed, the past year alone has seen some prominent cases where former employees took it upon themselves to erase entire banks of virtualized servers, or even wipe out all the corporate mailboxes in the company. Clearly, the use of a NAS (or SAN) would have offered no defense against such shenanigans by insiders. Only a comprehensive disaster recovery strategy where everything is backed up on a regular basis may have a chance of returning things to normalcy within an acceptable period of time.

## Damage to Storage Medium

Many workers make the erroneous assumption that storing data on a portable hard disk drive (HDD) or USB flash drive constitutes a backup. This is certainly not

the case, and such portable devices are in fact more susceptible to being misplaced, stolen, or damaged. In addition, it is important to remember that not all storage devices are designed for longevity in mind, or even for robust data preservation.

Do note that some mediums are particularly prone to damage, such as rewritable optical discs made with inferior dyes. Also, don't be surprised if that suspiciously cheap USB flash drive you bought at the flea market abruptly stops working after a year or two.

## Accidental Deletion

I recently had an experience with this exact scenario. When transferring thousands of email messages between two email servers, I carelessly opted to initiate the transfer moving key folders directly between the two locations. It was a bad decision because not all the data made it to the destination server before Outlook decided to call it quits (by crashing), even though the deletion from the origin took place immediately.

The loss of these emails was distressing because while I don't make money from writing emails, they play an enormously important role by helping me connect with my editors, PR folks, and various expert sources. Thankfully, some tips from an Exchange expert allowed me to restore the more than 35,000 emails.

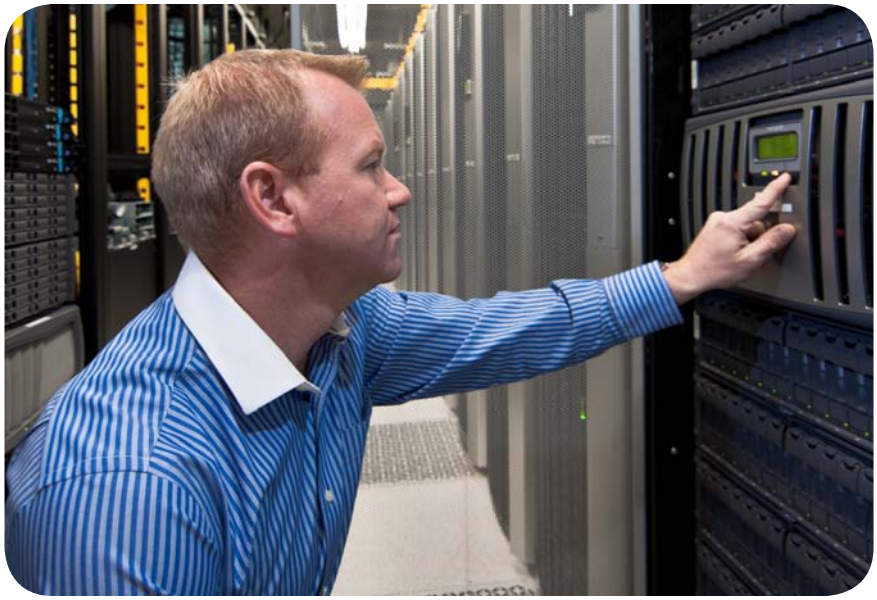
So what is my point here? Accidents can happen to anyone, even a seasoned computer user. The only real remedy would be to have separate copies of your data handy when mistakes happen. ■

# Re-examining Your Backup Strategy

By Paul Mah

Advancements in technology mean that new deployment options might become available or that previously impractical strategies become viable over time. As such, it only makes sense to review various facades of one's IT operations from time to time.

Let's take a closer look at the current state of storage technology so that you can better determine whether your organization's current backup strategy needs a refresh.



## Online Backup

Cloud computing is the latest catch phrase on practically everyone's lips, though it's not a new concept. Even as some businesses have had a somewhat knee-jerk reaction against storing their business-critical data on the Internet, others have already made cloud storage part of their business operations. Personally, I would urge growing businesses to consider online backup as they would any other storage option, and not adopt it blindly.

Some common issues include running afoul of compliance regulations, concerns over hacking or the storing of potentially sensitive data on servers located in foreign countries. As you can see, some of these deterrent factors might be insurmountable, though robust encryption of backup files will serve to protect against data leakage even in the event of a successful security breach.

On the plus side, businesses can stand to quickly tap into a ready, purpose-built backup infrastructure without having to spend a dime on capital cost.

## NAS/SAN Backup

The line between Network Attached Storage (NAS) and Storage Area Network (SAN) hardware is rapidly blurring where features and storage capacities are concerned. Today, it is not uncommon for mid-level NAS to sport dual Gigabit Ethernet ports, iSCSI support, or anything from five hard disk drives with the option for expansion to a single unit packing 24TB out of the box. In addition, most NAS is certified for virtualized environments these days, while others even offer replication between appliances.

On the other side of the spectrum, the prices for SAN have also come down to a price point where mid-sized businesses should have no problem affording it. Organizations that demand the highest performance or that want to architect a high-availability virtualized infrastructure will certainly want to consider this option. Of course, SAN deployments still rate on the high-end

of the cost scale, as they do generally require trained staffers to operate. In fact, improper ratification of a SAN error by an IBM engineer brought down a major bank's network for several hours in Singapore last year.

## Tape Backup

In some instances, tape might be mandated by compliance requirements for controlled, off-site backups, in addition to the inherent "off-line" portability afforded by tape cartridges. The latter offers protection against inadvertent corruption, and instills a certain level of traceability that can help guard against deliberate data modifications. In addition, alternative solutions such as cloud-based backups might be unsuitable for large volumes of data. Tape still offers the lowest storage cost today.

So while I would not advocate that every business deploy tape solutions, it might be an option worth examining for secondary or tertiary backup. One interesting tidbit of data that I've found out is that the largest consumer of tape is Google, which single-handedly consumes some 50,000 LTO (Linear Tape-Open) tape cartridges every quarter.

## In Conclusion

Which backup method should your business use? One possible strategy might entail the use of multiple tiers such as a NAS backed by a tertiary tape system for offline, offsite backup. Yet, the differences in organizations mean that there can be no standard reply on this, unfortunately. Complicating matters somewhat are hybrid abilities such as NAS that offers Amazon cloud-replication or Direct Attached Storage (DAS) cartridge systems that connect using USB.

In closing, my advice would be to carefully consider the merits of each technology to come up with a solution that achieves your RTO (Recovery Time Objective) and data backup needs, as opposed to being stuck with a specific appliance or category of hardware. ■

"So while I would not advocate that every business deploy tape solutions, it might be an option worth examining for secondary or tertiary backup."

# Five Imperatives for Extreme Data Protection in Virtualized Environments

By Peter Eicher

Transforming an organization through server virtualization requires a strategic and coordinated approach. Data protection — which includes not only backup, but also secondary storage and disaster recovery considerations — is an area that can easily complicate virtualized data centers if implemented hastily. It is essential that data protection efforts reduce hardware purchases, rather than require additional hardware to make it work. The following are five critical data protection imperatives that organizations must consider during virtual server planning.

## No. 1: Minimize Impact to Host Systems During Backups

In virtual environments, numerous virtual machines (VMs) share the resources of the single physical VM host. Backups — which are among the most resource intensive operations — negatively impact the performance and response time of applications running on other VMs on the same host. On a large virtual machine host with many VMs, competing backup jobs have been known to bring the host to a grinding halt, leaving critical data unprotected.

There are various approaches for minimizing the impact to host systems during backups, though each has drawbacks. The simplest approach is to limit the number of VMs on a given system, making sure you do not exceed the number you can effectively back up.

While effective, this goes counter to the purpose of virtualization, which is to consolidate applications to the fewest possible physical servers. It would also limit the financial benefits accrued from consolidation, perhaps significantly.

A second approach is to stagger the scheduling of VM backups. For example, if performance is impacted when four backups are running simultaneously, limit backups to three at a time. This can solve the performance issue, but it can create other challenges. For example, backup jobs cannot be scheduled without referencing all the other existing jobs. What if a particular job runs longer than expected and the next set of jobs start? Suddenly, the performance limit has been surpassed. As data grows over time, jobs may take longer to run, creating backup overlap. There is also no clear way to account for full backups and incrementals in such a scheme. Even if the scheduling is worked out, the total backup window has now been extended significantly by stretching backups over time.

An early technical attempt at solving the backup problem was the use of a proxy server. For VMware, this model is known as VMware Consolidated Backup, commonly called VCB. With VCB, a separate server is dedicated for running the backups directly off the storage. The virtual machines do not participate in backups. While this seemed good in theory, in practice there was still

significant performance impact due to the use of VMware snapshots. It also proved complex to configure. The result was that few users adopted this model and VMware has dropped support for it going forward.

A final approach is to install an efficient data protection agent on each virtual machine, and then run backup jobs just as they would be run in a physical environment. The efficient agent requires technology that deftly tracks, captures, and transfers data streams at a block level without the need to invoke VMware snapshots. By doing so, no strain is placed on the resident applications, open files are not an issue, and the file system, CPU, and other VMs are minimally impacted.

## No. 2: Reduce Network Traffic Impact During Backups to Maximize Backup Speed

Every aspect of the data center environment can stand a little improvement. But if your backup capabilities are like most, they are in dire need of an upgrade.

Reduction of network traffic is best achieved through very small backups, which dart across the network rapidly, eliminating network bottlenecks as the backup image travels from VM to LAN to SAN to backup target disk. Block-level incremental backups achieve this while full base backups, and even file-level incrementals, do not.

Minimal resource contention, low network traffic and small snapshots all lead to faster backups, which deliver improved reliability (less time in the transfer process means there is less time for network problems) and allowance for more frequent backups and recovery points. In a virtual environment, this also means more VMs can be backed up per server, increasing VM host density and amplifying the benefits of a virtualization investment. Technologies such as CBT and other block-level backup models are the best way to limit network impact.

## No. 3: Focus on Simplicity and Speed for Recovery

Numerous user implementations have revealed that server virtualization introduces new recovery challenges. Recovery complications arise when backups are performed at the physical VM host level (obscuring and prolonging granular restores) or through a proxy (necessitating multi-step recovery).

It is important to consider the availability of a searchable backup catalog when evaluating VM backup tools. Users of traditional, file-based backup often assume that the searchable catalog they are used to is available in any backup tool. But with VMs this is not always the case. Systems that do full VM image backups or use snapshot-based backups often are not able to catalog the data, meaning there is no easy way to find a file. Some provide partial insight, allowing users to manually browse a directory tree, but not allowing a search.

It is also important to understand how the tool handles file history. A common recovery use case is the need to retrieve a file that has been corrupted, but the exact time of corruption is not known. This requires the examination of several versions of a file. A well-designed recovery tool will allow input for both the file name and a date range to detect every instance of the file housed in the backup repository. While this may seem a minor point, it can make the difference between an easy five-minute recovery process and a frustrating hour or two hunting around for files.

"Every aspect of the data center environment can stand a little improvement. But if your backup capabilities are like most, they are in dire need of an upgrade."

Fast and simple recovery, at either a granular or virtual machine level, can be achieved if point-in-time sever backup images on the target disks are always fully "hydrated" and ready to be used for multiple purposes. In fact, with a data protection model that follows this practice, immediate recovery to a virtual machine, cloning to virtual machine, and even quick migrating from a physical to virtual machine are all done the same way — by simply transferring a server backup image onto a physical VM host server.

## No. 4: Minimize Secondary Storage Requirements

Traditional backup results in multiple copies of the entire IT environment on secondary storage. Explosive data growth has made those copies larger than ever, and the need for extreme backup performance to accommodate more data has necessitated the move from tape backup to more expensive disk backup. The result is that secondary disk data reduction has become an unwanted necessity.

Deduplication of redundant files can be achieved at the source or at the target. In isolation, each approach has drawbacks. Each new data stream needs to be compared with an ever-growing history of previously stored data. Source-side deduplication technology can impact performance on backup clients because of the need to scan the data for changes. They do, however, reduce the amount of data sent over the wire. Target-side deduplication does nothing to change the behavior of the backup client or limit sent data, though it does significantly reduce the amount of disk resources required.

A hybrid approach combining efficient data protection software with target-side deduplication can help organizations achieve the full benefits of enterprise deduplication without losing the other benefits.

No. 5: Strive for Administrative Ease of Use

Very few users have a 100 percent virtualized environment. Consequently, a data protection solution that behaves the same in virtual and physical environments is desirable.

A data protection solution in which a backup agent is installed on each VM can help ease the transition from physical to virtual. Concerns about backup agents needing to be added to every new virtual machine are overstated because each VM needs to be provisioned anyway — with an operating system and other commonly deployed applications and software. New virtual machines cloned from a base system will already include the data protection agent.

When evaluating solutions, it is vital to consider the entire backup lifecycle, from end to end. For example, if some data sets need to be archived to tape, a deduplication device may not allow easy transfer of data to archive media. This might then require an entire secondary set of backup jobs to pull data off the device and transfer it to tape, greatly increasing management overhead. This kind of "surprise" is not something organizations want to discover after they have paid for and deployed a solution. Ease of use can also be realized with features such as unified platform support, embedded archiving, and centralized scheduling, reporting, and maintenance – all from a single pane of glass.

## A Holistic View of Virtualization

To maximize the value of a virtualization investment, planning at all levels is required. Data protection is a key component of a comprehensive physical-to-virtual (P2V) or virtual-to-virtual (V2V) migration plan.

The five imperatives recommended here can help significantly improve organizations' long-term ROI around performance and hardware efficiencies and accelerate the benefits of virtualization. To complete this holistic vision, organizations must demand easy to use data protection solutions that rate highly on all five of the imperatives. Decision makers who follow these best practices may avoid the common data protection pitfalls that plague many server virtualization initiatives.