

# Data Storage Insights

for growing businesses



## Managing data growth

---

Save your IT  
budget with  
storage tiering

Expanding  
your storage  
capacity

Right-sizing  
your backup  
strategy

# Contents

CONTRIBUTORS: DREW ROBB AND PAUL RUBENS



3 Data storage advice for growing businesses



4 How storage tiering can help save your IT budget



6 Is it time to expand your storage capacity?



8 Has your business outgrown its backup strategy?

# Data storage advice for growing businesses

Research that shows how quickly the amount of electronic data that businesses generate and store is everywhere these days, but what's often missing from the conversation is information for businesses struggling to keep up in the storage race.

Growing businesses, in particular, can struggle to deal with today's storage demands. When they start out as a small business, a central file server, an external hard drive and an online backup service can be enough to get employees the files and data they need and protect them from disaster. But as the number of employees grows, and the demands in terms of data storage capacity, backup, archiving and regulatory compliance take hold, those solutions are quickly overwhelmed.

In this issue of *Data Storage Insights* we're going to focus

on some of the data storage techniques and strategies that growing business need to address to ensure they aren't suffering under the weight of their information storage and its costs.

Storage tiering takes the axiom "All data is not created equal" and puts it into practice. Storage media range from inexpensive media that are slower to access to more expensive media that are available almost instantly. By evaluating how valuable the different types of data in an organization are and then assigning an appropriate storage tier, businesses can breathe new life into their IT budget.

Capacity expansion is at the heart of the storage issues many companies are facing today. It's the natural byproduct of a world where more electronic data is created, needs to be stored and cannot easily be deleted for legal reasons. We'll look at how IT administrators know when it's time to expand their

What's often missing from the conversation is information for businesses struggling to keep up in the storage race.

data storage capacity, including 10 signs that it's time to upgrade your storage capacity.

Finally, we'll examine backup strategies, which need to evolve as businesses grow so they can protect their data and recover from incidents to quickly resume normal IT operations. We'll review the options that are available for businesses looking to right-size their backup strategy to ensure data protection.



# How storage tiering can help save your IT budget

BY DREW ROBB

A smart way for businesses to improve overall efficiency and save on data storage costs is by understanding the differences in the types of data they store. This can be achieved via storage tiering —the process by which different types of data are stored in the way that makes the most sense for that data.

Storage tiering, however, has long been thought of as the province of large enterprises due to the cost and complexity involved. This was perhaps true five years ago, but it is no longer the case. Tiering can now be accomplished economically even by small businesses. Storage tiering is also very important to midsize and growing businesses that are seeing the amount of data they store increase significantly and risk having their data storage infrastructure become overwhelmed.

According to Mike Karp, principal analyst at Ptak, Noel & Associates, the purpose behind tiering is to arrive at an appropriate balance between storage cost and performance. Those embarking upon tiering, he said, have to come to terms with a common misunderstanding—believing that all data is of equal value.

“Tiering involves the use of a number of different types of storage media (high performance/high cost at one extreme, and high-capacity/low-cost at the other end) and storage services (continuous backups for highvalue data, and weekly backups for lower value data, for example),” said Karp. “This gives an IT manager the ability to assign proper performance and protection levels to data, which can be a great money saver.”

Savings are realized by assigning storage hardware and software services according to the value of the data. If all data is regarded as having a monotone level of importance, it might all be running on a top-of-the-line system and be afforded a formidable amount of data protection. Particularly in rapidly expanding companies, this soon becomes an expensive proposition. When data is analyzed, though, it becomes obvious that not all files are created equal. Almost every analysis reveals that perhaps 10 percent of the total is frequently accessed. This is almost always data less than 30 days old. After that, access rates diminish and beyond a certain points, files are rarely looked at, if at all.

Karp points out, though, that there is more to it than time. Sales reports and accounts payable updates are examples of high-end data that might hold high value for a quarter or an entire year.

Based on such analyses, storage tiering puts high-value material on high-performance hardware and assigns it the best backup and recovery services. Lesser value data, on the other hand, deserves proportionately less investment, and belongs on less expensive devices/services.

When it comes to high-value data, “back them up frequently, and make sure they are on the most reliable, high-performance storage systems in the shop,” said Karp. “Lower-priority data might include the corporate phonebook or the accounting department’s football pool and should be assigned to cheaper storage that receives fewer services.”

## Elements of tiering

Traditional storage tiering consisted of speedy, but expensive Fibre Channel (FC) disks for Tier 1, cheaper disk such as SATA for Tier 2 and tape for Tier 3. But as technology has evolved and costs have plummeted, new tiering concepts have come on the market that can be implemented by small and midsize businesses. Solid State Drives (SSDs), for example, are frequently deployed as a Tier 0 to be used only by one or two mission-critical systems.

“There is Tier 0 using SSD where time is money or lives may be at stake, which is ultra fast but involves small amounts of space capacity,” said Greg Schulz, an analyst with Server and StorageIO Group.

However, it must be noted that SSD is many times more expensive than spinning disk. A clear value proposition has to exist if it is to be rolled out extensively. But inexpensive systems are coming on the market and they provide SSD capability at a reasonable cost.

Schulz suggested that smaller companies that use databases may want to consider using SSDs—but only for logs, tables and database metadata as this can provide a noticeable improvement in information retrieval rates even if the database file itself sits on SATA disk.

Tier 1 can either be FC disk or less-expensive SAS drives that can operate at the same speed of 15,500 revolution per minute (RPM). Placing all information on this platform, though, isn’t wise as this is the most costly form of disk.

Tier 2 is generally composed of inexpensive SATA disk running at 7,200 RPM and available in sizes as large as 1TB or 2TB. Tier 3 is



The purpose behind tiering is to arrive at an appropriate balance between storage cost and performance.

the lowest cost tier reserved for inactive data. Organizations can utilize tape, older SATA boxes or even the cloud. Note, too, that the above arrangements can vary from organization to organization.

### Implementing storage tiering

Karp counsels businesses new to tiering that it doesn't necessarily mean that they throw out old gear. As servers and storage arrays age, he said, incorporate them into the overall tiering strategy by using the old boxes in the lower tiers. He gives one caveat, however.

"Sometimes brand new gear may be cheaper in the long run than retaining an ancient machine that is stuttering along and is time consuming to maintain," said Karp.

For example, those that have tape drives could use them for Tier 3 and as an extra safeguard as part of their data protection strategy. It is well known, he said, that backing up to disk is far faster than tape. One possible solution is to backup regularly to disk and then once a week send a backup from a back-end disk system to tape.

Each environment, though, will be different. Discuss the best uses for existing equipment with the storage vendor or Value Added Reseller (VAR) in order to utilize it optimally in any tiering strategy.

When it comes to higher tiers, though, the most important action is to analyze data usage patterns and assign business priorities. A common mistake at this juncture is for management to want all data to operate on high-performance systems that supply top-grade business continuity services. The best way to overcome such demands is to ask the vendor or VAR for a price breakdown of various levels of service using a variety of tiering architectures. With full data in hand, it is relatively easy to determine a tiering plan that fits within budgeting realities.

"For primary storage, identify how much high-value data you really have and then put everything else on less expensive disks," said Karp. "Be aware, though, that data values can change over time."

Schulz adds that time should be spent isolating inactive data and moving it off to a tape archive or inexpensive disk. This might consist of old databases or

customer records from 10 years ago. Note, too, that tiering isn't only about hardware. Software is needed to track data patterns and move information automatically from one tier to another based on pre-set user parameters. Karp recommends, therefore, that organization investigate the software side before committing to a specific hardware platform.

### Keep it simple

Sabine Waterkamp, president of system integrator ACSLA Inc., advises IT organizations that are exploring storage tiering to keep it simple. Yes, three- or four-tiered systems may work for the largest of enterprises, but for those with less complex needs, two tiers might well be enough.

"For a typical SMB, I would suggest most often a 2-tier architecture using 15k RPM SAS for high transaction volumes and 7.2k RPM SATA for older data and backup storage," said Waterkamp. "Unless you deal with large volumes of data, anything else than some sound placement of data on a 2-tier storage architecture may well be overkill."

# Is it time to expand your storage capacity?

BY PAUL RUBENS

Ninety percent of businesses expect their data storage capacity to grow by at least 10 percent over the next two years, while about one-third expect it to double in that period, according to research carried out by the computing technology industry association CompTIA. Against this background of capacity expansion it is probably not unreasonable to expect that your own company will need to increase its storage capacity over time, possibly quite significantly.

There are two key questions that IT administrators need to be asking themselves: Is your current storage infrastructure sufficiently scalable to meet your future needs? If so, will it be able to offer the following important business benefits:

- Storage capacity that is inexpensive and easy to increase
- High reliability
- Simple and effective backup and recovery of all corporate data
- Fast database and other application performance
- Minimal administration requirements
- Ability to make use of lower cost second tier storage media
- Appropriate data security
- Ability to take advantage of technologies such as virtualization if required
- Reduced power and cooling requirements and associated costs.

If you can't answer "yes" to both of these questions then it may be time to consider adopting a new storage strategy. To help you

answer the questions, here are 10 warning signs that your storage infrastructure may need upgrading:

## **1. Your file servers are running out of physical space to add more hard drives.**

Many companies rely on internal hard drives or direct attached storage (DAS) devices, but this type of storage can't be expanded indefinitely. Buying additional servers is not a sensible option because it is not a cost-efficient way to provide extra storage. It also adds administrative complexity and complicates backup procedures. A better solution could be to invest in networked storage can be shared by all of your existing servers.

## **2. Some of your storage resources are running out, while other resources are underused**

If your company relies on internal storage or DAS devices that are only connected to one or a limited number of servers then you are probably using storage resources very inefficiently. That's because you may have to invest in increased storage capacity on one server when you have unused capacity in another. More sophisticated options such as storage area networks (SANs) can help solve this problem.

## **3. Backing up is complicated and time consuming, and involves multiple backup systems**

If you have expanded your existing storage capacity on an ad-hoc basis then it's likely that you have many different types of storage devices, and groups of storage resource that are isolated from each other. This makes backing up complicated, time consuming and ultimately costly. If it becomes too complicated to be easily manageable then there is also an

increased risk that your data is not all backed up properly, which can lead to catastrophic data loss. Centralizing your storage using a SAN array such as one from the Dell/EMC CX4 range combined with a Dell PowerVault tape-based backup system could make backup and recovery much more straightforward.

## **4. Data is difficult to find and share**

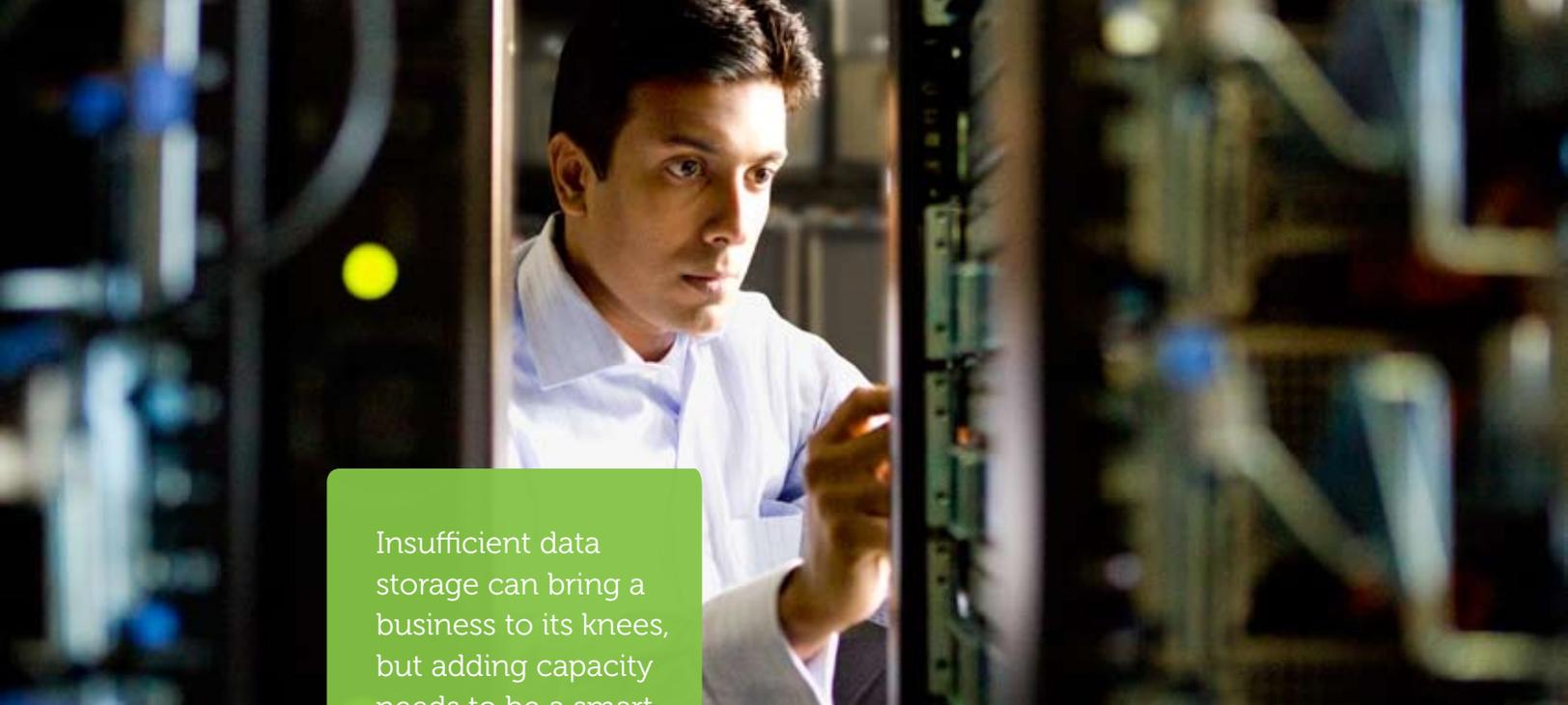
One of the most important reasons for incurring the cost of storing data is because there is a business benefit in having the data available for applications and employees to use. But if you are finding that much of your data is isolated or access is restricted to certain systems because of the way that it is stored then you are incurring storage costs without being able to extract the benefit from the data being stored. Network storage technologies can help employees leverage the value of your data by making it easily available to all who need it.

## **5. Data intensive applications are slow**

There are many reasons why applications may run slowly, but inappropriate data storage can be a major contributor to poor performance. That's because your application will have to spend much of its time idle while waiting for data to be made available to it from your storage systems. More appropriate storage systems such as a Dell / EMC CX4 SAN array can make your applications run faster and your staff more productive.

## **6. Your local area network grinds to a halt during backups and when certain applications are busy**

Applications that access large amounts of data using a legacy



Insufficient data storage can bring a business to its knees, but adding capacity needs to be a smart process.

storage infrastructure can easily clog up your company's local area network, slowing data traffic to a crawl and reducing the productivity of all of your staff. This can be avoided if certain types of data are stored in a SAN and moved around on a fast dedicated storage network using a technology such as Fibre Channel.

#### **7. Your company is not taking advantage of cheaper second tier storage**

Hard disks store and retrieve data very quickly but they are not the most cost effective storage media on a per gigabyte basis. Secondary storage media such as tape or optical discs offer a far lower cost per gigabyte, and are ideal for archived information that needs to be stored but which may rarely, if ever, need accessing. To take advantage of cheaper storage media you need to have in place some form of information lifecycle management system, which makes it easier to manage your data and archive it to secondary storage at the appropriate time.

#### **8. Your storage-related staff costs are increasing**

If you don't have a coherent storage strategy or you have been increasing your storage on an ad-hoc basis then it's likely that administrative activities such as provisioning, configuration and backing up are more labor-intensive—and thus less cost efficient—than they could be. Newer storage technologies often offer a high degree of automation that can reduce storage administration costs significantly.

#### **9. You have to buy more storage capacity than you need immediately**

Adding capacity to your current systems may involve buying a large amount of additional storage that you will only use over a period of years. This is an inefficient and costly way of buying storage, because the utilization rate may initially be very low and a large proportion of this unused capacity may also require power, making your energy bills unnecessarily high. Newer storage systems let you take advantage of thin provisioning and overallocation. These two storage technologies

enable you to prepare for storage capacity expansion while only buying storage media (such as hard drives) when they are required—instead of buying them up front and leaving them idle.

#### **10. You are paying for capacity you know you will rarely need**

If your company occasionally has sudden, large and unpredictable storage capacity expansion requirements for short periods then you may have storage resources that are mostly unused. Newer storage systems enable you to benefit from storage "cloud bursting"—making use of storage capacity in the cloud on a temporary basis and paid for per gigabyte per day.

Insufficient data storage can bring a business to its knees, but organizations that need to add capacity need to be smart about the process. A number of relatively recent data storage technologies such as thin provisioning and data deduplication, when combined with strategies like storage tiering, will help businesses get the most for their storage dollars going forward so they can smartly and efficiently meet their demand.

# Has your business outgrown its backup strategy?

BY PAUL RUBENS

The amount of data that companies use and store every day is growing at an explosive rate and the chances are that your company now handles significantly more data than at any time in the past.

There could be many reasons for this increase in data usage, including:

- Corporate growth
- Meeting requirements of regulations such as HIPAA, Sarbanes Oxley, PCI DSS and the EU Data Protection Directive
- More sophisticated IT systems
- Increased levels of online business activity

To accommodate all of this new data your company may well have increased its data storage capacity over the last year or two, perhaps by implementing a storage area network (SAN) to complement the storage resources directly attached to your servers.

As your business develops and your data needs change it is important to consider your strategy for back-

ing up this data to protect it from loss, and for recovering the data and resuming normal IT operations in the event of a disaster. If your data storage has increased significantly then full backups that in the past took several hours may now take an entire weekend, and if data growth continues the time is coming when they may take weeks rather than days.

The key question to ask is whether the current storage backup and recovery strategy you have in place is still able to provide the level of data and business continuity protection that your business requires and can scale to meet your likely future needs, or whether you have effectively outgrown it.

## Business objectives

To determine if your current backup strategy can meet your business needs it may be useful to take some time to consider your recovery point objective (RPO) for each application or system. An RPO of one hour means that only data stored up to one hour before a disaster may be lost, but any data

stored more than one hour before a disaster will have been safely backed up. Your RPO is therefore your possible data loss window.

Deciding on a suitable RPO for your business begins by establishing a baseline—the maximum data loss window that is acceptable.

The other factor that needs to be considered is your recovery time objective (RTO) for each application or system. This is the maximum amount of time that it can take before your application or system is up and running again with restored data after a disaster. The RTO for each application or system will be determined by the cost to your business of that application or system being unavailable.

## Assessing your current backup strategy

To assess if your current backup and disaster recovery strategy meets your business requirements or needs to be updated you need to examine whether each backup solution in place can meet your baseline RPO and RTO for the applications or systems they protect.

Even if your current solutions meet your baseline RPO it may still make sound business sense to upgrade to a system that offers a lower RPO if the benefit of doing so outweighs the cost. This will depend on the cost to your organization of losing data and the cost of a backup solution which could save it. In some businesses the cost of losing any data is so high that it makes sense to have an RPO of zero by implementing a high-end continuous data protection system.

Similarly, even if your current backup solutions meet your RTO it may make financial sense to invest in new backup systems that allow you to reduce it.



To help establish whether to upgrade your backup strategy—and what backup technologies you should be looking at—you will need to consider the following questions:

- What is the value of each type of data you are backing up?
- How much does it cost your business when a given system or application is down?

- How much time do you have to perform backups?
- What budget do you have available?
- Are you willing to completely change your existing backup strategy or do you want to stick with your existing backup processes as much as possible?

Technology options that you may want to consider are listed below.

A thorough evaluation of the current backup strategy will help businesses understand if the strategy needs to be updated because of new requirements, more data or business growth. There are a number of options available for IT administrators that need to right-size their backup strategy, and each options carries with it some advantages and disadvantages.

BACKUP OPTION	PROS	CONS
<p><b>Backup to Tape (B2T)</b></p> <p>B2T systems like the Dell PowerVault series are among the most common solutions currently in use although the RTO and RPO possible using tape is relatively long. The technology is proven and there is a wide range of backup software designed to work with automated tape backup systems.</p>	<ul style="list-style-type: none"> <li>• Low per gigabyte storage costs</li> <li>• Backup tapes can easily be moved offsite</li> <li>• Tapes consumes no power when not in use</li> <li>• Proven track record with many existing backup solutions</li> </ul>	<ul style="list-style-type: none"> <li>• Slow read speeds and sequential (not random) data access means low RTO is not practical for large amounts of data</li> <li>• Backup window can be long, making RPO relatively high</li> <li>• Tape media can be hard to physically manage</li> </ul>
<p><b>Backup to Disk (B2D)</b></p> <p>B2D systems typically stream backup data to a disk array made up of SATA, SAS or Fibre Channel disks. Storage costs per gigabyte are typically significantly higher than tape based systems due to the high cost of disks but RTO and RPO can be shorter.</p>	<ul style="list-style-type: none"> <li>• Reduced backup window</li> <li>• Shorter RPO and RTO than are possible with tape</li> <li>• Random data access means data can be restored faster</li> <li>• More reliable than tape because data can be protected using RAID</li> </ul>	<ul style="list-style-type: none"> <li>• Higher cost than B2T</li> <li>• May not be compatible with existing backup software</li> </ul>
<p><b>Backup to Virtual Tape Libraries (VTL)</b></p> <p>VTL systems are fast because they use hard drives instead of tape, but they present themselves to backup software as tape media so they can be used with existing tape backup software. This makes upgrading to VTL systems relatively easy. The higher media costs associated with hard drives can be mitigated to some extent using data compression and deduplication, and data can easily be copied to tape media for off-site disaster recovery purposes or archiving.</p>	<ul style="list-style-type: none"> <li>• Can be used with existing backup software and processes</li> <li>• Faster and more reliable than tape systems</li> <li>• Can take advantage of data compression and deduplication to reduce storage costs</li> </ul>	<ul style="list-style-type: none"> <li>• Higher cost than B2T and D2D</li> </ul>
<p><b>Continuous Data Protection (CDP)</b></p> <p>CDP systems record every primary storage disk write to backup storage as they happen, so the backup window is effectively zero. Data can be recovered from immediately before a disaster, making a very short or zero RPO possible.</p>	<ul style="list-style-type: none"> <li>• Zero RPO possible</li> <li>• Captures every change to data as it happens so recovery is possible to any point in time</li> <li>• Captured data can also be stored in secondary location for disaster recovery</li> <li>• Very low RTO possible using high performance disks</li> <li>• Systems can be recovered to any point in time</li> </ul>	<ul style="list-style-type: none"> <li>• Costly</li> <li>• Disk and network capacity intensive</li> </ul>