Why back up?

The importance of protecting your data

Introduction

Businesses of all sizes are witnessing an explosion in the volume of data they hold. Whether it is the result of the Internet, email, or increasingly heavy and media-rich application software, there is a massive growth in the volume of data all around. Conservative estimates from IDC in October, 2002, place data growth at approximately 80% per year. Data is increasingly being recognized as one of the real assets of a company, and losing this data would cause severe damage to any organization.

What is your data worth?

Data loss can be very costly, particularly for organizations in the small and medium business (SMB) market where the difference between survival and closure can rest on the ability to recover from a disaster. At the very least, critical data loss will have a financial impact on companies of all sizes:

"American businesses lost over $8 billion due to computer viruses in the first six months of 2001."

Source: *Computer Economics 2002*

The financial impact on a company is a combination of loss of business, low productivity, legal action, and the cost of re-creating data. In 2002, a study from Ontrack, a company that provides data recovery services, showed that the cost of re-creating just 20 MB of data can be extensive:

Data type	Time to re-create 20 MB	Cost
Sales and marketing	19 days	$17,000
Accounting	21 days	$19,000
Engineering	42 days	$98,000

At its worst, critical data loss can result in business collapse. Studies by the National Archives and Records Administration showed that 80% of companies without well-conceived data protection and recovery strategies go out of business within two years of a major disaster.

Your data under threat

A disruption of key primary or support operations can cause a crisis for any business. It need not be a catastrophic failure like those caused by a bomb; it could be an outage to a single circuit that carries patient records or payroll information. All computer data is at risk from threat or damage. Even with the most reliable equipment and the most secure operating environment, there is always the possibility of something going wrong.

How does the network going down affect business operation?

Obvious impact	Behind the scenes
No Internet	No payroll
No email	No accounting systems or finance information available
No printing	No access to client data (CRM)
No company website	No order entry
No e-commerce capability	No access to files—drawings, reports, and more
	No telephone (VOIP)

The Ontrack survey in 2002 shows the following to be the most common causes of data loss.

Natural
disasters
3%

Computer
virus
7%

Software
corruption
14%

Hardware
malfunction
44%

Human error
32%
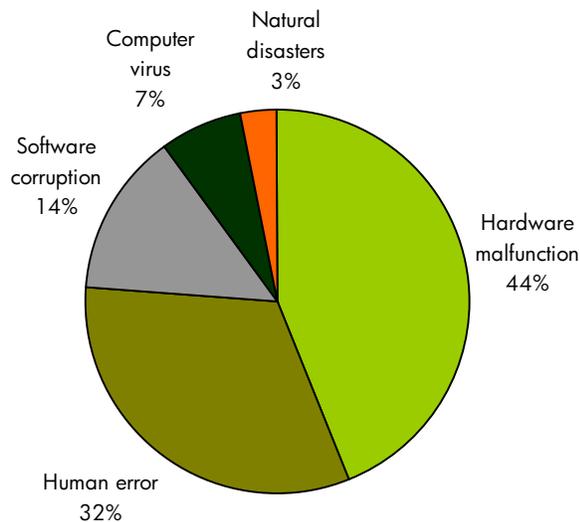
Compliance with government regulations

Recent world events including acts of terrorism, natural disasters, and large-scale company fraud have resulted in a new raft of legislation designed to protect company data from loss or corruption. This legislation includes:

- Sarbanes-Oxley Act (SOA) of 2002 intended to "deter and punish corporate and accounting fraud and corruption, ensure justice for wrongdoers, and protect the interests of workers and shareholders" (President Bush)
- European Union Data Protection Directive of 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector
- USA Patriot Act passed in September of 2001 introduced a plethora of legislative changes that significantly increased the surveillance and investigative powers of law enforcement agencies in the United States
- Gramm-Leach-Bliley Act (GLBA) of 1999 sets provisions to protect consumers' personal financial information held by financial institutions and companies
- Health Insurance Portability and Accountability Act (HIPAA) of 1996 involves legislation protecting patient privacy and records security

Compliance with any or all of this legislation means one thing: organizations are required to store, manage, and safeguard a lot more data than they might otherwise feel compelled to do on their own. This in turn has led to the need for more storage capacity and more efficient methods of backing up, retrieving, and archiving data.

The need for data protection strategies

With ever-increasing amounts of data to be stored, real threats of data loss, and increasing government legislation, it is little wonder that data protection has become a high priority. It is not just large and enterprise corporations that must protect their data, however. Predicted annual spending on data storage hardware, software, and services by SMBs is expected to more than quadruple between 2002 and 2006 according to a report by Access Markets International (AMI). The report concludes that worldwide storage spending will grow at a compound annual growth rate of 43% over this four-year period.

Five years ago, the terms "disaster recovery" and "data protection" were synonymous with large budgets and an army of IT personnel to manage the process. But today's business environment of low-cost disks, robust software, and high-performance tape drives enables companies of all sizes to proactively protect themselves from data loss and its threats to the business. Businesses, and especially SMBs, are looking for solutions that:

- Reliably use the best technology for the application. They cannot afford backup failure.
- Are simple to use and maintain—with ongoing low cost of ownership and low levels of complexity.
- Provide enough capacity and enough performance. The backup must happen within the available backup window.
- Offer a choice of suppliers and ease of data interchange.

Tape backup for cost-effective data protection

For the optimum balance between cost and performance, magnetic tape remains the clear first choice for backup today.

Increasing $/MB
Decreasing access time

Semi-conductors

Hard disk drives & RAID — Online storage

Optical, CD, DVD — Nearline storage

Tape drives — Offline storage

Paper
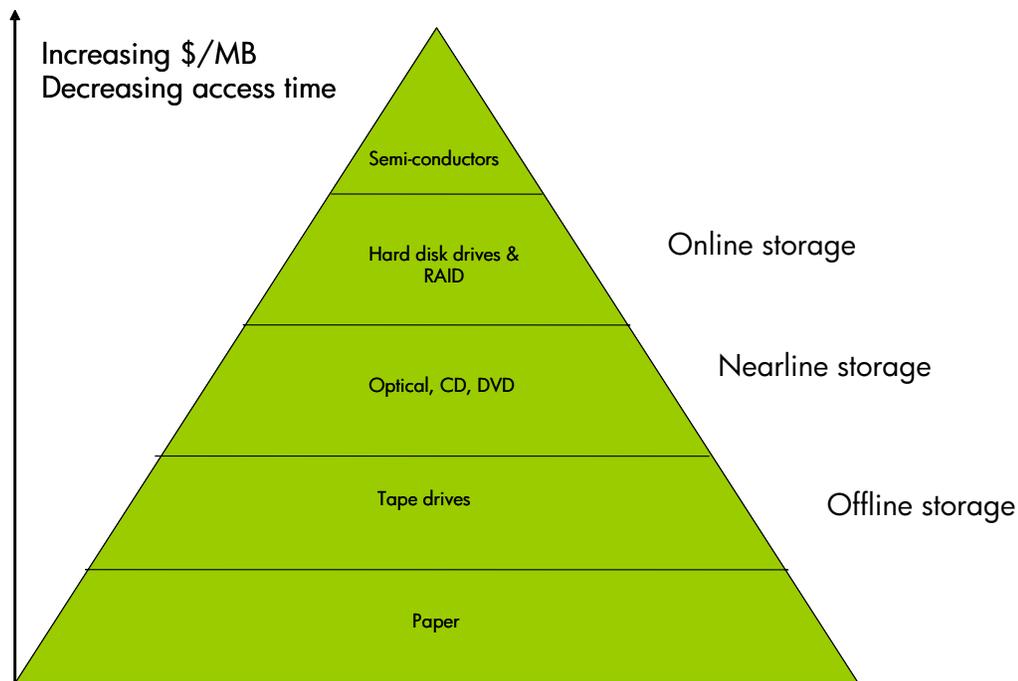
Semi-conductors and hard disk drives offer virtually immediate data accessibility, but there is a relatively high price to pay for each megabyte of storage. Consequently, they are more suited to online operation than backup storage.

Optical storage media, such as CD and DVD drives, offer moderate capacity with reasonably easy access to information. However, the cost per megabyte still makes them an expensive option for daily high-volume backups or data archiving. Also, their lower capacity is still a severe restriction where high-volume, regular backups are required.

Tape on the other hand offers long-term storage capabilities, ideal for medium-term archiving (of 10 to 15 years). A large volume of data can be stored on a single piece of media, making it easy to restore information as required. The fact that the media is small, removable, and easily transportable makes it possible to store the data separately from the system, which provides protection from everything from viruses to destruction of the site. In addition, tape still offers the lowest storage cost per megabyte, making it a viable option for even the most demanding backup schedules. Following are the relative costs.

Medium	Cost/GB	Notes
Disk drives	$1.00	Between 80 and 200 GB per disk
DVD+RW	$0.70 cents	Typically 4.7 GB on one piece of media
DAT/DDS	$0.50 cents	Typically 72 GB on one piece of media
LTO Ultrium	$0.35 cents	Typically 400 GB on one piece of media

Advantages of HP tape drives

HP is the market leader in tape drives and offers a range of tape drive technologies, including DAT/DDS, DLT, and LTO Ultrium. HP tape drives offer a number of advantages over competitor tape drives including:

- Broadest compatibility with servers, operating systems, and backup software
- Highly tuned support with HP servers and systems
- HP Library and Tape Tools, which comes with all HP tape drives, provides utilities for installation support, troubleshooting, and performance tuning
- One-Button Disaster Recovery that provides complete data protection from a single tape
- A choice of internal, external, rack-mounted, and hot-plug drives
- A complete solution with backup software in the box

HP also offers a range of autoloaders that can make establishing an effective backup routine even easier. Autoloaders contain a single tape drive and multiple data cartridges, plus a robotic mechanism that changes the cartridges automatically. Based on instructions from the backup software, the autoloader will perform the backups at the scheduled time, swap cartridges if there is a problem or the tape is full, and even run a regularly scheduled cleaning cycle. This type of unattended backup is ideal for sites where there are few or no IT staff to manage a daily backup. HP offers DAT autoloaders with six cartridge slots and Ultrium and SDLT autoloaders with eight or sixteen slots.

Open standard formats

In addition to the preceding advantages, the DDS/DAT drives and LTO Ultrium drives are built on open, industry-standard formats. An "open standard" means that multiple companies can manufacture the technology, adhering to a shared standard that ensures interchange. On the other hand, proprietary technologies are designed and developed by only one manufacturer who controls the development speed, roadmap, availability, and pricing. These include AIT, S-AIT, SDLT, DLT, DLT VS, VXA, and Mammoth.

Open technologies provide real benefits to customers due to the pressures of competition:

- Tape drives and media are available from multiple sources—multiple manufacturers, not just resellers—which means a greater protection for your investment and less risk to your supply of drives and media.
- Competition encourages the development of differentiating features that add value for the customer. Examples are the HP half-height Ultrium drive, One-Button Disaster Recovery, and DAT hot-plug drives.
- Competition tends to keep prices down.
- Competition encourages manufacturers to stick to their published timelines.
- Format agreements among competitors make for more credible roadmaps that cannot be changed unilaterally.

Implementing a backup strategy

Having decided to implement a backup system, thought must be given to choosing the most appropriate strategy. Disk drives must be backed up comprehensively on a regular basis, or the backup gives nothing more than a false sense of security. There are some important considerations.

Determine your system requirements

Every system that holds valuable information must be protected. In addition to databases that are clearly mission critical, it is important to remember the more mundane and routine systems. Decide which parts of the system are critical and which parts you may choose not to back up, for example an individual user's C: drive.

Selecting the right tape drive solution

HP has a broad range of tape technologies to suit different environments and needs. Choosing the right tape drive will be linked to business priorities where, generally speaking, SMBs will be more concerned to find a cost-effective, easy to use solution, while larger businesses might prioritize performance and ongoing cost of ownership. Consequently, the choice of HP tape technology is likely to depend on the scale of business needs. When you refer to the HP product portfolio included earlier in this white paper, it is clear that:

- DAT/DDS provides the optimum price:performance ratio for most SMB server environments.
- LTO Ultrium delivers ultimate performance for high-end and enterprise servers and networks.
- SDLT and DLT VS80 provide backward compatibility to customers who have already invested heavily in these technologies.

HP also provides these technologies as autoloaders. An autoloader enables a backup routine to be automated with each data cartridge used for a specific day's backup over the period of a week.

Selecting the right backup software

Most HP tape drives are provided complete with backup software in the box to enable customers to immediately install and begin backing up their system. This software provides all the functionality required for general backup applications. However, if more sophisticated features are required, a package like HP OpenView Data Protector allows sharing tape drives and libraries among multiple systems in heterogeneous environments and provides central management and control. There are several major features to be considered when evaluating backup/recovery software packages.

- Centralized administration: the software package should allow for the administration of multiple backup servers and devices from a single interface if your environment requires this.
- Usability: the application interface should be intuitive and easy to use, while simultaneously providing the power and flexibility to meet your needs.
- Scalability: the backup software should support any additional hardware required in the future. For example, if you are choosing a software package to back up local servers, be sure that it supports autoloaders if you might decide to upgrade in the future.
- Options and agents: ensure that you have full access to files even when they are open to ensure that even files in use get backed up.
- Compatibility and support: finally check that your tape drive and operating system is fully supported, that the pricing model includes full licenses for the whole network rather than a single server, and that technical support is fully available to you.

Planning the backup

There are three key ways in which backup can be performed—full backup, differential backup, and incremental backup as illustrated in the following diagram.

| DATA | Original system | Day 1 | Day 2 | Day 3 | Day 4 |
| Full backup | Backup of Original |

Differential backup:
- Day 1
- Day 1 | Day 2
- Day 1 | Day 2 | Day 3
- Day 1 | Day 2 | Day 3 | Day 4

Incremental backup:
- Day 1
- Day 2
- Day 3
- Day 4
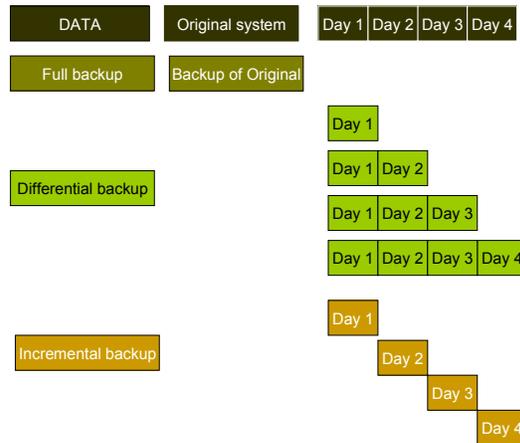
Full backups

All data on a system is backed up. Full backups are done periodically to ensure that all data resides on a single piece or a single set of media so that, when restoring information, only the full backup media is required. A full backup is the basis for all other types of backup.

Advantages

- All files are easily found when needed—all the information can be found on the last backup tape.

Disadvantages

- Since most system files rarely change, a good deal of the backup is redundant. Full backups take time to perform and so backing up static information is actually wasting time.
- If full backups are not performed frequently, then data restored may be well out of date.

Differential backups

A differential backup stores all files that have changed since the last full backup. When restoring differential backups, the full backup media and the latest differential media are needed.

Advantages

- The method requires less time than full backup.
- Undertaken daily, it means that there is a more current set of data to restore.
- Restores are more efficient than incremental backups, as only two sets of tapes are required—the full backup and the latest differential backup.

Disadvantages

- Because each differential tape contains the data that has changed since the last full backup, a good deal of redundancy occurs and this requires more capacity. However, if your differential backup still fits on one piece of media, this should not matter.
- Increasing amounts of time are required as you get further from the full backup.

Incremental backups

An incremental backup stores all files that have changes since the last backup. When restoring from incremental backup schemes, the last full backup plus all incremental backup sets are required.

Advantages

- It uses less media: only files that were created or changed since the last backup are included.
- Backups require less time to conduct.
- It reduces the backup "duty cycle" of the device (the number of hours per week it runs).

Disadvantages

- Multiple tapes might be needed when restoring data, making the process more complex.
- It takes more time to carry out the restore process.

Good housekeeping

Good housekeeping is essential to the validity of the backup. The following steps should be taken.

Manage the media rotation

A tape rotation scheme and schedule is vital to the reliability of the backup process. The following schemes are the most common in use; however, depending on the value of the media, the quantity of data, and the backup device in use, a schedule may be devised specifically for your needs.

1. Grandfather-father-son (GFS) media rotation strategy

Four pieces of media are used Monday through Thursday (the sons) for incremental or differential backups with a different tape for each day. These are recycled each week. A full backup is done each Friday. This tape is stored offsite and recycled every four weeks (the fathers, three pieces of media). In addition, a monthly full backup is done and kept offsite (the grandfathers). This tape can be recycled every third month or kept indefinitely. The advantages of this media rotation scheme are that it is both comprehensive and simple to use and manage. However, the downside is that the wear to the tape is uneven, with excessive wear on the tapes for Monday to Thursday.
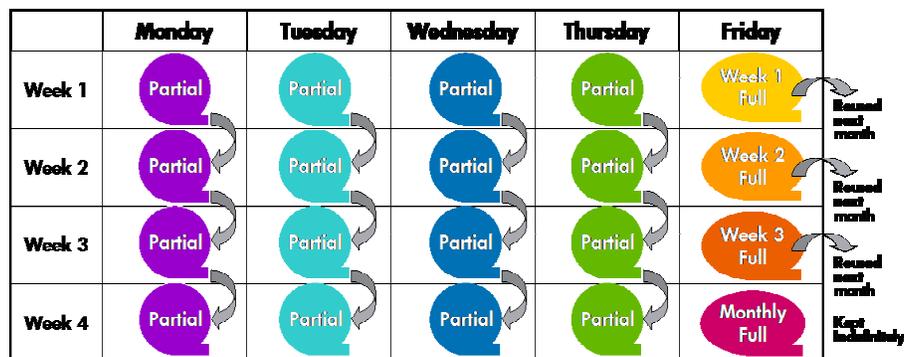
The tapes for the daily partial backups are reused each week. A full weekly backup is performed on Fridays and the tape is reused on the same week in following months. The monthly full backup tapes can be kept indefinitely or recycled every three, four, or more months.

2. Tower of Hanoi

This strategy is named after a game in which a number of different sized rings must be moved between three poles. In the game you start out with all the rings on one pole and must move them to another pole. The idea is that you must move them in a certain order to accomplish the task. In the case of tape backup we use this order to rotate the tapes. The pattern shows that tape set "A" starts on day one and repeats every other day. Set "B" starts on day 2 and repeats every fourth day, set "C" starts on day 4 and repeats every eighth day, set "D" starts on day 8 and repeats every sixteenth day, and so on.

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	A		A		A		A		A		A		A		A	
		B				B				B				B		
Media set				C								C				
								D								
																E

Return to day 1 at this point

This schedule can be used in either daily or weekly rotation. This is another scheme that is easy to manage because it maintains a good file history and covers each day with a different set of media. However, it is also prone to uneven tape wear as, for example, set "A" is used for backup four times as often as set C.

3. Media X scheme

Media X scheme eliminates the problem of uneven tape wear by rotating the media so that each piece is used the same number of times. However, it is complex and requires a well-documented schedule that is stringently adhered to.

For example, use a 10 X media scheme in a 10-week rotation, where Monday to Thursday are incremental backups and Friday is a full backup stored offsite. The numbered tapes would be rotated as follows and, after 10 weeks, new tapes would be introduced.

	Monday	Tuesday	Wednesday	Thursday	Friday
Month 1	1	2	3	4	5,6,7,8
Month 2	2	3	4	5	6,7,8,9
Month 3	3	4	5	6	7,8,9,10
Month 4	4	5	6	7	8,9,10,1
Etc…					
Month 7	7	8	9	10	1,2,3,4
Month 10	10	1	2	3	4,5,6,7
Repeat from month 1					

Document all procedures

To be of any value, a backup regime must be followed regularly and accurately. To ensure that there are no variations to the schedule, it is important that all procedures are clearly documented. Logs should be kept and held where they are accessible to anyone following the backup procedures.

Purchase high-quality media

It is important to choose reputable, high-quality media for your tape drive. HP supplies and media program invests heavily in the ongoing testing and qualifying of media for its tape drives.

Regularly clean and maintain the tape drive

Many drives must be cleaned with a high-quality cleaning cartridge as part of the backup schedule or when the cleaning light is on. Debris on the tape or on the tape head can cause backup errors.

Store media securely onsite and offsite

Secure offsite storage provides additional protection against flood, fire, and theft. Onsite storage should be carefully thought through. Fireproof safes provide additional protection. Archived data should be periodically inspected to ensure that no obvious damage or corruption has occurred due to environmental or other damage.

Rehearse data recovery and disaster recovery procedures

If disaster strikes and data is lost, you must be sure that you can effect a swift and complete recovery from your backup tapes. To ensure that there are no hiccups or unforeseen glitches, it is important that recovery procedures are rehearsed regularly. Doing a test restore on a regular basis both ensures that you know how to follow the procedure and confirms that your backups are doing their job.

Summary

An effective backup strategy is essential if your business is to be safe from the effects of data loss. Tape drives still represent the most versatile and economical backup solution. The broad range of tape products from HP provides a complete solution to meet the needs of everything from a small business to a large enterprise. With careful consideration and planning you can create a backup schedule that provides full protection of your data with minimum disruption to your business.

For more information

www.hp.com/go/tape

hp ®

i n v e n t