HP

A Short Guide To Successful Data Backup
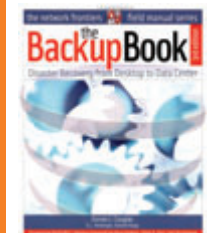
The essentials of protecting your data

the network frontiers field manual series

the BackupBook

Disaster Recovery from Desktop to Data Center

COMPUTER VIRUSES; CUSTOMER PRIVACY; REGULATIONS; AND FIRES, FLOODS AND EARTHQUAKES

It seems as if a day doesn't go by without some new threat or issue that organizations have to address to protect and manage their critical information.

Small and medium sized organizations, as well as remote offices and branch locations of larger organizations must address the same data protection issues as large organizations. Compounding the problem is the fact that many organizations do not have sufficient resources—either in personnel, time, or financial—to fully plan for the direct and indirect impacts that these issues can have on their activities. This lack of advance planning often translates into heightened risk of unprotected data—arguably the most important asset to in your organization.

That's why Hewlett-Packard is pleased to bring you this guide on how to implement a "quick start" backup and recovery plan that may help to protect your organization's data.

As much of the content is taken from Dorian J. Cougias' award winning book "The Backup Book," readers should find that the information is both easy to read and understand and should be considered a "must read" for anyone responsible for data protection activities in your organization.
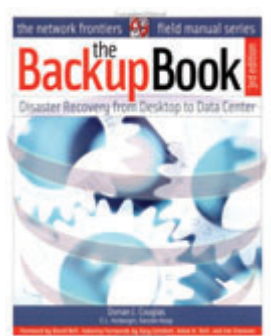
As much of the content is taken from Dorian J. Cougias' award winning book "The Backup Book," readers should find that the information is both easy to read and understand and should be considered a "must read" for anyone responsible for performing the backup function in your organization.

Topics reviewed in this guide include:

- Planning for the "What If(s)"
  – The Three Basic Types of Protection

- Data Backup & Archival Technologies
  – What to Store Your Backup Data On?

- Implementing a Basic Data Backup, Recovery, & Archival Plan

- Know What to Backup

- Why Use HP Backup Media?

PLANNING FOR THE "WHAT IF(S)"

What if a fire, flood, or computer virus destroyed all of your data…how quickly, if at all, would you be able to recover?

What if your biggest customer asked you to immediately share with them your data protection policies that safeguard their data?

What if an auditor asked you to quickly locate a critical customer file dating back to last year… could you? How about one that was made two years ago? Five years ago? Seven?

What would happen if an unhappy member of your staff decided to destroy all your electronic records, could you recover them before the organization ceased to operate?

Unlikely scenarios? Not in today's world.

These are real world scenarios that affect businesses each and every day. The bottom line is that well run organizations recognize the importance of planning for these events and have put into place comprehensive backup, disaster recovery, and archival plans that protect their data and allow them to maintain access to the information when needed.

The Three Basic Types of Protection

A comprehensive backup and disaster recovery plan is a blend of protecting both data and hardware. Data and hardware protection falls into three basic categories:

(1) Data backup (2) Data duplication or mirroring (3) Hardware and software fault tolerance

These three classifications are illustrated below. Data mirroring and fault tolerance are very important areas to consider as part of a comprehensive data protection plan especially as the network infrastructure increases in size and complexity. More information on these technologies is available from **ww.hp.com/go/storage** as well from Dorian Cougias' "The Backup Book". The scope of this guide however, focuses on the first category—data backup—to ensure that if all else fails, your organization has at least one copy of your data in case you need it.

Protection Type	Description	Benefit
Data backup	• A second copy of data is made to backup media as protection against loss of the first copy. The copy can be made to many different types of storage media including disks and/or tape media. Tape backup media can be used for offsite storage and archival.	**You can recover your data to a previous point in time after data or hardware loss**
Data duplication or mirroring	• A second copy of data is maintained at a different location with changes made to the primary data copied to the data at the different location.	**Quick recovery from failure of hardware storing primary data**
Hardware and software fault tolerance	• A duplicate set of hardware (and sometimes software) in is used case of component failure. In case of failure, the duplicated hardware component can take over while the failure is fixed or replaced.	**Continuous or near continuous operation if a hardware component fails**

Data Backup vs. Data Archiving

Many people confuse "Data Backup" with "Data Archiving" because both processes are often used together as part of a comprehensive data protection plan. It is important to note that these are two different data protection processes and both need to be planned for separately.

See page 8 for details on the differences.

DATA BACKUP & ARCHIVAL TECHNOLOGIES

ULTRIUM LTO

Six Generation Roadmap

Source:
The LTO Program

	Generation 1	Generation 2	Generation 3	Generation 4	Generation 5	Generation 6
				Encryption	Encryption	Encryption
			WORM/RW	WORM/RW	WORM/RW	WORM/RW
Compressed Capacity	200GB	400GB	800GB	1.6TB	3.2TB	6.4TB
Compressed Transfer Rate	up to 40MB/s	up to 80MB/s	up to 160MB/s	up to 240MB/s	up to 360MB/s	up to 540MB/s

DAT Roadmap

		DDS-1	DDS-2	DDS-3	DDS-4	DAT 72	DAT 160	DAT 320	DAT 8th Gen
Read/Write Capability									
Media Type		90m Tape	120m Tape	125m Tape	150m Tape	170m Tape	150m Wide Tape	Wide Tape	Wide Tape
Native Capacity		2 GB	4 GB	12 GB	20 GB	36 GB	80 GB	160 GB	~300 GB
Native Transfer Rate		183 KB/s	≤720 KB/s	≤1.5 MB/s	1-3 MB/s	≥3 MB/s	≥5 MB/s	12 MB/s	≥16 MB/s
		1990	1993	1996	1999	2003	2007	2009	~2 Yrs

Tape remains the most used backup media. As disk media has become more affordable it can be used with tape to combine the low cost, portability and durability of tape with the speed of disk for single file restores and frequent incremental backups.

There are several different backup and archival technologies that organizations may use to protect and archive their information, the most widely used fall into two camps—disk drive based products and tape based products. The two technologies can be used together to provide comprehensive backup and archival protection.

	Backup to Tape	Backup to Disk
Cost	Cost effectiveness – backing up to tape is the least expensive way to backup data per gigabyte (GB). No power costs for media unless accessed	The cost of disk technologies varies depending on the type of disk used (generic disk or virtual tape libraries) Disk prices continue to decrease. Ongoing power cost for spinning disks.
Performance	Excellent for full backup and restores. • Speed and Capacity – storage manufacturers are continuing to invest heavily in improving the speed and capacity of each new generation of tape backup products. Ultrium 4 tape drives and media can now store up to 1.6TB of data per cartridge at 240MB/sec! • High-performance tape drives provide very fast back up for larger files and sequential backup and restores (but slower than disk for single file backup and restores)	Good for frequent incremental backups. Requires large disk system to match tape streaming throughput. Faster than tape for single file backup and restores.
Reliability	• Shelf Life – media that is properly used and stored can last up to 30 years or more • Longevity – tape has a long, defined future incorporating major R&D and technology enhancements Mobility & Storage – data that is backed up to tape can be easily moved to an offsite location for safe data storage	Reliability is increased through technologies that add redundancy such as RAID and mirroring.
Technologies and Products	Two tape technologies dominate the backup market today: LTO Ultrium tape drives: highest performing drives with high capacity data cartridges. DAT tape drives: Economical and easy to connect tape drives with very affordable data cartridges. HP provides both single and automated tape drive solutions for both LTO and DAT.	• Virtual Tape Libraries: backup appliances that present the disk capacity as virtual tape cartridges to allow easy integration with existing backup processes • SATA disk systems: low cost disk suitable for backup • SAS, SCSI, FC disk systems: more expensive and higher performance disk • Removable disk systems (RDX): economical and simple protection to disk with advantages of portable media.

Backup & Recovery Software—Whatever backup devices are chosen, backup and recovery software is required to create and schedule backup jobs and manage backup devices and media. Choice of backup and recovery software depends upon:
• The number of servers and desktops to be protected
• Operating system used
• Applications in use and whether they need to be backed-up online
• Other

HP offers HP Data Protector Express for smaller organizations and HP Data Protector for larger organizations with more complex backup requirements.

Implementing A Basic Data Backup, Recovery, & Archival Plan

So far this booklet has covered why it is important to protect your data and the choices of backup devices, media, and backup software. The following section shows you how you can define a basic backup and recovery plan for your organization.

The steps covered include:
1. Understand Common Backup & Recovery Terms
2. Knowing When To Use Full, Differential, or Incremental Backups
3. Know What To Backup
4. Develop & Implement A Backup Rotation Strategy
5. Confirming Your Backup Strategy Works

Step 1: Understand Common Backup & Recovery Terms
Having a good understanding of some basic storage concepts and terms provides a good background as you begin to develop your own backup strategy for your organization.

Archive
The purpose of data archiving is to store the primary copy of data for possible use in the future. It is not backup. Often archiving involves moving (copying then deleting) data from disk to tape and storing that tape in a safe long term location. In planning, you need to be clear about when data needs to be backed up (for protection) and when it is to be archived (stored for possible future use).

Backup
The purpose of data backup is to make a second (and maybe third) copy of data so that data can be recovered if the source (primary copy) of that data is damaged or lost. Data backup keeps copies of data for the short and medium term. Backup data is only used if the source data is damaged or lost.

Catalog
Backup software is used to manage the backup and recovery process. The catalog is a database used by the backup software to record which source files have been backed up and the destination(s) of the backed-up files. The catalog also stores information about backup devices and media. It will hold all the backup jobs that have been created as well as information about when they will run and the outcome when they have run. Different backup software calls the catalog by different names: catalog file, the storage management database, the share resource library, or just library.

Media Set
The backup software copies files from the source and writes them at the destination. The destination has one or more pieces of media to store the backed up data as well as information in the catalog about what has been backed up. All the media and the metadata related to an instance of a backup job is called the media set. A backup job scheduled to run multiple times can have multiple media sets.

"Source" vs. "Destination"
The backup "source" is the site of the information that you want to backup. The backup "destination" is the site to where you want the copy the source data.

Tape format
Backup software writes data from the source to the destination using a special format. This format is normally proprietary to the backup software which means that a backup written by one backup software will be unreadable by different backup software. The layout of the backup data is defined in the tape format. A tape format is used to efficiently write the backup data and metadata as well as allow quick recovery (reading) of the backup data. The term "tape format" is used because the way backup data is written was developed before disk media and optical media became viable for backup. When backing up to these media the same "tape format" is used.

HP Data Protector Express

- Backup and recovery software designed for smaller organizations
- Backup to tape, disk and CD/DVD
- Windows and Linux support
- Simple installation, operation and licensing
- Wizards for frequent tasks and pre-configured job schedules

www.hp.com/go/dataprotectorexpress

HP Data Protector

- For medium and large organizations
- Automates high performance backup and recovery from disk or tape over unlimited distances
- Enables 24x7 business continuity
- Acquisition and deployment costs 30% to 70% less than similar software
- Integrated with HP StorageWorks disk and tape products and HP OpenView management solutions

www.hp.com/go/dataprotector

Combining Full Backups with incremental or differential backups can reduce backup times and allow more frequent backups.

Full Backups

Whatever backup types you choose, you will need to start with a full backup. A full backup copies all source data to the destination. Restoring data from a full backup is the fastest method, because the files that are being restored to any one computer are stored on one media set.

Incremental Backups

Incremental backups only copy the source data that has changed since the most recent backup (of any type). This can significantly reduce the time it takes to run the backup, particularly if the source data changes or grows slowly. The reduced backup time can allow more frequent backups to be run for added protection. Incremental backups are typically used between full backups. Restoring data from an incremental backup may require multiple media sets if data is changing regularly. The recovery and identification of media needed will be managed by the backup software but is more time consuming than restoring from a full backup. If the recovery device is a tape drive this may involve a lot of manual tape handling.

Differential Backups

A differential backup only copies the source data that has changed since the most recent full backup. This can reduce the time it takes to run the backup, particularly if the source data changes or grows slowly and the latest full backup was run recently. Restoring data from a differential backup will require a maximum of two media sets—one differential backup media set (depending upon the recovery point required) and the most recent full backup media set. Using differential backups, is a way to get many of the advantages of incremental backups while maintaining a reasonably simple recovery process.

Many media set rotation schemes (described later) use a mix of these backup types.

KNOW WHAT TO BACKUP

Step 3: Know What To Backup

Think of all the information and systems that are part of your organization. You will find that some of the information you have is more important than other information.

In some instances you might be required to keep customer data backed up on a daily basis…other pieces of data may require periodic backup at first but then after a period of time the data becomes much less important therefore reducing the need for frequent backups. Some data needs to be easily accessible via online disk systems whereas other data can be stored either offline (archived) or off site.

The fact of the matter is that every organization is different. What is consistent, however, is the fact that all organizations need to determine what information is required to operate, determine the data's overall value, determine the type of storage medium best utilized to store the information on, determine what information they are required to keep, and finally, document the policies and procedures to get rid of data that is either no longer valuable or needed.

If you are struggling with this step, the following matrix may help. The matrix describes a simple model that many companies have used to help identify, categorize, and prioritize their data. Once categorized appropriately, Storage Administrators can more easily assign "values" to their data, making it easier for them to determine the best means to store, backup, archive, and eventually delete their data (if needed).
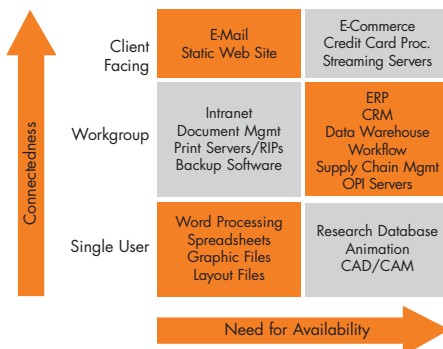
For example, think through the organization(s) you work with. How much of what you need to backup is in the single-user, low-need-for-availability category, and how much is in the client-facing, high need-for-availability category? What you have, and how connected your group to be backed up is, will play a major part in determining the nature of your backup plan.

As a guide, the greater the number of people who would be affected by the loss of a piece of data, or the greater the cost of having to recreate it, the higher up the backup priority list it should be.

13

Preconfigured backup media rotation schemes available in Data Protector Express

Name	Backup Scheme	Max data loss	Oldest version available for recovery[1]	Media sets required
GFS – 62	each 15 mins: incremental each hour: differential each weekday: full backup	15 mins	2 years	62
GFS – 54	each 1 weekday: incremental each week: full backup	1 hour	2 years	54
GFS – 30	each 1 weekday: incremental each week: full backup	1 day	2 years	30
GFS – 25	each 1 weekday: incremental each week: full backup	1 day	2 years	24
GFS – 20	each 1 weekday: incremental each week: full backup	1 day	2 years	20
Simple - 12	each 1 weekday: incremental each week: full backup	1 day	4 years	12
Simple-11	each 1 weekday: incremental each week: full backup	1 day	3 months	11
Simple-10	each 1 weekday: full backup each week: full backup	1 day	3 months	10
Simple-6	each 1 weekday: full backup each week: full backup	1 day	2 months	10
Simple–4	each 1 weekday: full backup	1 day	1 month	4
Daily Append	each 1 weekday: incremental each week: full backup	1 day	1 week/per media set used	1 or more

[1]As well as considering the oldest version that can be recovered, you should also look at the number of recovery points available. The more media sets used, the more recovery points (and file versions) available for recovery.

Step 4: Develop and Implement A Backup Rotation Strategy

There are dozens of backup strategies that you can employ at your organization and picking the right one will depend on your requirements.

Hewlett-Packard suggests you consider a "Grandfather, Father, Son" tape rotation strategy as depicted below. This strategy allows you to protect data while keeping things simple as it enables backup on a daily, weekly, monthly, quarterly and annual basis.

Different backup and recovery software has different ways to implement backup rotation schemes. HP Data Protector Express allows you to choose from 11 pre-configured "Grandfather, Father, Son" and other backup schemas or create you own backup schema depending on your needs. These pre-configured backup rotation schemes are described on page 14.

"Grandfather, Father, Son" Tape Rotation Strategy (Month 1)				
Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Monday Tape[1]	Tuesday Tape[1]	Wednesday Tape[1]	Thursday Tape[1]	Week 1 Tape [2]
Monday Tape[1]	Tuesday Tape[1]	Wednesday Tape[1]	Thursday Tape[1]	Week 2 Tape [2]
Monday Tape[1]	Tuesday Tape[1]	Wednesday Tape[1]	Thursday Tape[1]	Week 3 Tape [2]
Monday Tape[1]	Tuesday Tape[1]	Wednesday Tape[1]	Thursday Tape[1]	Week 4 Tape [2]
Monday Tape[1]	Tuesday Tape[1]	Month Tape[3]		
				Quarter Tape [4]
				Annual Tape [5]

Strategy:

1. Back up data on a different tape set every working day. The same Monday tape is used every Monday and Tuesday tape every Tuesday, etc.

2. On the fifth day, use a weekly tape set. Weekly tape sets are then stored for the entire month, until a monthly backup is performed in the last week of the month.

3. The monthly backups are stored until the next year.

4. Quarterly backups are performed at the end of every 3 months and are kept until the next year.

5. One annual backup is performed at the end of every year and is kept as long as needed.

Step 5: Make Sure It Works—Confirming Your Backup Strategy Works

Periodically you should test that you can easily and reliably recover your data as needed. Doing so will also ensure you are familiar with your restore process. HP also provides HP Library and Tape Tools (HP L&TT), a free, downloadable diagnostic tool that makes it easy to troubleshoot backup issues and get statistics on tape drive and media.

Visit **www.hp.com/support/tapetools.**

WHY USE HP BACKUP MEDIA?

The quality of the backup media determines the reliability of your restore

Whatever backup technology you use, it is very important that you choose the highest quality media to maximize your chances of successful restore when needed.

Backup to Tape Media:
Although today's tape media products can last up to thirty years or more if used and stored properly, large duty cycles can accelerate tape degradation much more quickly. And unfortunately, there are no hard or fast rules as to how long it takes to wear out a tape. Some media vendors quote a figure of 1 million passes, while others stick to the number of full back ups. Both metrics are hard to judge. If you are following a well known backup routine like "Grandfather, Father, Son", the system will instruct you to replace those tapes long before wear and tear becomes an issue. And using the same backup tapes over and over again will probably cause you more issues than simply buying more backup media on a regular basis.

Backup to Disk Media
If you are using disk as a backup media, you should consider a dedicated Disk to Disk (D2D) backup appliance as they employ redundant disks and do not normally allow general purpose access to the disks. If you are backing up to a general purpose disk, it is important that the disk is configured for reliability and that access is controlled to prevent accidental or malicious deletion of the backup data. If you are backing up to a RDX removable disk, you should choose HP ruggedized, durable and shockproof (to 1 meter drop) RDX data cartridges. These are tested to similar high quality control standards as HP's tape media.

Why Use HP Media?
The price of storage media is small in comparison to your need to save and store critical information reliably. This is the ultimate measure of cost, not cost per tape or cost per gigabyte.

One of the ironies of backup is that while the media may be low cost today, there is no way of quantifying the potential value of the data in the future. In other words, an inexpensive tape could one day be the home for some of your most precious company information. When a critical database has to be restored, a 25% or even a 50% recovery is not what you need. That tape might be worth its weight in gold in the future.

For over 20 years HP has been has been a leader in data storage products, technologies and services. With this experience, HP has developed a quality control program that tests HP media to extremes. HP tape media cartridges must pass an exhaustive battery of tests that mimic real-use situations. This testing program, incorporating some 170,000 tests in excess of 400,000 test hours in 2005, is designed to ensure that HP media will always offer maximum reliability even in the most extreme conditions.

By purchasing HP data cartridges, you can be certain that you are receiving the most exhaustively tested product in the industry. HP tests its media to extremes so that you don't have to. And while no one can predict the future, or offer a certain guarantee, the scale and breadth of our testing gives us more confidence that your data will be safe on HP storage media tapes.

Here's some more information designed to go in one ear and actually stay there.

Dorian J. Cougias is the founder and CEO of Network Frontiers, a company that focuses on disaster recovery, security, compliance and IT infrastructure consulting, training, and books. Over the last twelve years, Dorian has overseen the launch of Network Frontiers, has served as CIO of two of the leading advertising agencies in the world, and has written extensively. He is an Adjunct Professor of Technology teaching courses in security and disaster recovery at the University of Delaware, College of Human Services, Education, and Public Policy.

UNIFIED COMPLIANCE FRAMEWORK

Network Frontiers has reviewed hundreds of international and domestic standards and regulations to simplify and centralize your compliance efforts. The UCF displays the minimum number of controls required to comply with all the authority documents you select, including PCI-DSS, Sarbanes-Oxley, HIPAA, and more. Visit **www.unifiedcompliance.com** for more details.

THE BACKUP BOOK

The Backup Book combines what can go wrong and how that applies to the various backup methods, couples that information with recovery and business continuity tactics, and then works through multiple real-world disaster recovery scenarios.