WHITE PAPER

February 1996

Prepared by Systems Division

Compaq Computer Corporation

CONTENTS

Backup Basics

EXECUTIVE SUMMARY

Developing a successful company-wide backup strategy requires an understanding of the system's network architecture and the demands placed on the system by its users. Compaq, a leading-edge provider of technology for backup and storage management, is developing a series of white papers that explain the basics of backup, as well as how to:

- conduct a network or enterprise backup needs analysis, and
- use that information to create and implement an effective, company-wide backup strategy.

This introductory white paper is the first in that series and focuses on backups of small to medium networks. It provides a foundation of basic information by answering these questions:

- Why are backups essential?
- What information should be backed up?
- How much information should be backed up?
- How much time is available for the backup?
- How can backup performance and storage capacity be optimized without exceeding the company's cost restraints?

Other white papers in this series will discuss data compression and other performance-tuning options, the many variables that affect overall backup performance, the various backup options available from Compaq, etc. Key terms used in this document are defined in the Glossary.

Please direct comments regarding this communication via internal bmail to this address: Tech Com@HW Tech@Sys Hou
Compaq field personnel may also send comments via the internet to this address: tech_com@bangate.compaq.com

COMPAQ

NOTICE

The information in this publication is subject to change without notice.

COMPAQ COMPUTER CORPORATION SHALL NOT BE LIABLE FOR TECHNICAL OR EDITORIAL ERRORS OR OMISSIONS CONTAINED HEREIN, NOR FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES RESULTING FROM THE FURNISHING, PERFORMANCE, OR USE OF THIS MATERIAL. THIS INFORMATION IS PROVIDED "AS IS." COMPAQ MAKES NO REPRESENTATION OR WARRANTIES OF ANY KIND, WHETHER EXPRESS, IMPLIED, OR STATUTORY AND EXPRESSLY DISCLAIMS THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR PARTICULAR PURPOSE, GOOD TITLE AND NONINFRINGEMENT.

This publication does not constitute an endorsement of the product or products that were tested. The configuration or configurations tested or described may or may not be the only available solution. This test is not a determination of product quality or correctness, nor does it ensure compliance with any federal, state or local requirements. Compaq does not warrant products other than its own strictly as stated in Compaq product warranties.

Backup Basics
First Edition (February 1996)
Document Number 371A/0196

WHY ARE BACKUPS ESSENTIAL?

The backup process copies important information (in many companies, this is *vitally* important information) onto magnetic tape or other disks. This enables the restoration of anything from one file to the entire system, should the need arise. Backups have helped companies recover from data losses caused by power surges and outages, static electricity, lightning strikes, other natural disasters, terrorist bombings, simple errors made by users, simple accidents (such as a spilled cup of coffee), sabotage, equipment malfunctions, viruses, and so on. Data recovery tools and services[1] exist, but they are limited and can be expensive. While users might be able to recreate some lost data, retrieving *all* of the lost information is unlikely. Complex application and network configurations, customized setups, even passwords and IDs will be difficult and expensive—perhaps even impossible—to recreate.

The sudden loss of a mission-critical server that stores and maintains corporate records and data (one of a company's most valuable assets) can be financially disastrous. In most companies, just the downtime before recovery can be much too costly. A well-designed backup system safeguards crucial information, providing the most efficient and cost-effective insurance against a potentially disastrous loss of data, time, and money.

HOW TO DETERMINE BACKUP NEEDS

Developing a successful backup strategy begins with a carefully planned backup needs analysis. The administrator first identifies the company's *total* backup needs and then matches those needs to the appropriate backup hardware and software. The backup needs analysis begins with the following questions:

- What information (programs and data) should be backed up?
- How much is there of this information?
- What types of backups are needed?
- How much time is available to do each backup?
- How often should backups take place?

What should be backed up?

In many companies as much as 40% of the company's data is changed every month. Ultimately, *all* company data and programs should be backed up, so that the entire system can be restored in the event of a catastrophic disaster. In addition, certain groups of users may have special backup needs. For example, a group of key users responsible for developing mission-critical data may require special backups of their data every Monday through Saturday evening. This would be over and above the regularly scheduled full-system backups. To develop an effective backup strategy, *all* corporate and user needs must be identified and spelled out.

[1] These tools and services are a poor alternative to a good backup. The term "data recovery" refers to restoring information that has been physically damaged or corrupted on a disk or tape. This can be caused by viruses, bad software, hardware failures, power failures that occur while the magnetic media is being written, etc. "Data recovery" always takes place *after* the damage has been done.

How much information must be backed up?

This involves the following factors:

- How much corporate information should be backed up now? How large is a complete system backup?
- How much additional capacity is needed because of backup redundancies resulting from special, user-defined backups?
- How much will corporate information increase in size? For many companies, the amount of data stored on the typical server doubles every year. Future storage demands will only increase, especially as imaging and multimedia applications become more common. A successful backup strategy will accommodate this growth.

Knowing the total amount of information to be backed up helps to determine the appropriate backup technology in two ways:

- ***Storage Capacity***. The total amount of information to be backed up indicates the capacity required of the drive and the media. If planned backups will be unattended, then the selected backup device must have enough capacity to hold the full amount of information to be backed up. If the amount to be backed up requires more than one tape, then a backup tape drive with an autoloader may be appropriate. (The autoloading feature uses a robotic mechanism to change tapes, thereby reducing administrative costs.)
- ***Performance.*** The selected drive's *typical backup rate* together with the *appropriate backup software* (and several other factors[2]) must make the backup system capable of accomplishing the entire backup during the time that the network is available and the server can be taken off line (the backup window). Otherwise, off-line backups will not be possible.

What are the types of backup?

Backups are classified by the status of the network server or servers (*off-line* or *on-line*) when the backup takes place, and by the amount of information that is backed up (*complete*[3] or *partial* backups).

Off-line and On-line

For an off-line backup, the system administrator's first step is to take the server off line, making it unavailable to users for the duration of the backup operation. The typical off-line backup takes place when user demand is at its lowest.

An on-line backup takes place with the server on-line and available to users. Depending on the network architecture, users may see a degradation in network performance while an on-line backup is taking place. In addition, there can be a danger to data integrity caused by file contention. This danger can be minimized by selecting the appropriate backup software.

File contention occurs during a backup when the backup system attempts to open and copy a file that has already been opened and is in use by one or more other users.

[2] The overall performance of any backup system results from the combination of *many* factors. A future white paper will discuss this in detail.
[3] *Complete* backups are often referred to as *full* backups. The meaning is the same.

When a file is initially opened, it can be opened privately (no sharing allowed) or it can be opened to allow sharing. If sharing is not allowed for an open file, then many backup programs cannot back up this file on line. These backup programs will log this as an error ("Unable to back up file XYZ"), and will move on to the next file.[4]

A file that has been opened with sharing allowed has two additional options. The open file can be opened by others for reading only, or the file can be opened for writing by more than one user. This latter condition is especially common with the multi-user databases that are included in many LAN applications.

Read-only sharing is most often allowed for EXE and COM files. The typical backup program that opens a file for read-only sharing will attempt to open the file with the "Deny Write" option in effect. If this attempt is successful, no other user will be allowed to write to the file as long as the backup program has it open. As long as other users are not writing to the open file, the file can be backed up with no risk to data integrity.

However, if the backup program attempts to open a file that has already been opened by other users who have write access, the backup program's "Deny Write" request will be denied. As long as other users have write access to the file, there is no way of knowing whether changes are being made to the file while the backup program is copying it. Under these conditions, file contention can be a problem, depending on how the backup software proceeds from this point.

If the backup program attempts to open the file without denying write access to other users (without the "Deny Write" option), the attempt will succeed. If the file in question is a database file in the process of being edited, it is possible that the file could be copied with an incomplete transaction, and no backup errors would be logged. If this file is later restored, serious problems could result. The incomplete transaction will probably contain bad data. It is even possible that the bad data could spread silently to other databases, depending on how they are linked. By the time the errors become evident, weeks may have passed and recovery may be impossible.

For some networks, not backing up the file at all may be the best approach. In other cases, having a "suspect" backup copy may be better than having none at all. It depends on the user's needs. One backup program, *ARCserve (from* Cheyenne Software, Inc.), backs up these files whenever possible and enters an error in the log to notify the administrator that the file's protection is suspect. The administrator can then take the appropriate action (which may be to restore an earlier, non-suspect version of the file).

Other products that protect open files during backup include *ARCserve* database agents, which are available from Compaq.[5]

Complete and Partial

These terms describe the amount of information that is copied. A *complete* backup is a full backup of the entire server or PC client hard drive. For a server, this includes all volumes, directories, and files. For a PC client, this includes all drives, directories, and files. A *partial* backup can be any of the following. All backups, whether complete or partial, can be done *on line* or *off line*.

[4] At the time this white paper was being developed, a new software package, *Open File Manager,* was able to back up all open files. ARCserve agents also enable the backup of open database files. These agents are special programs that enable *ARCserve* to gain access to the data on a workstation's or server's hard disk.

[5] An *ARCserve* agent is a special program that enables *ARCserve* to gain access to the data on a workstation's or server's hard disk. For an ARCserve backup to be successful, an *ARCserve* agent must be loaded on the workstation or server. Special ARCserve agents for *Oracle, Sybase, MS-SQL,* and *Gupta* are available from Compaq.

- *Differential*. Copies all files that were changed since the last *complete* backup. Differential backups are useful when it is important to have the latest version of each file. If the same tapes are used for consecutive differential backups, the newer versions of backed up files are often allowed to overwrite older versions of the same file on the tape. Typically, backup programs do not reset the file's archive bit after a differential backup; the archive bit remains turned on until the next *complete* backup.
- *Incremental*. Copies all files that were changed since the last backup, *regardless of what kind of backup it was*. This type of backup is used when each revision of a file must be maintained. If the same tapes are used for consecutive incremental backups, the newer versions of backed-up files are not allowed to overwrite earlier versions. Rather, the newer files are usually appended to the backup medium. Typically, backup programs reset the archive bit following each incremental backup.
- *User-defined*. Copies a user-defined set of files. Often this is a special backup requested by a group of employees on a mission-critical project.

The information that is backed up can be:

- *all applications.* This type of backup saves all files in the area defined by the user, including settings, customizations, passwords, etc. Application backups are particularly useful after a major change or upgrade in software.
- *applications and data.* This type of backup creates a standalone copy of the user's information base. Application and data backups allow easy restoration of the user organization's records. These backups can also be used to migrate information to another server.
- *data only*, which may be segregated by project or department, or which may include all information created within a certain time frame, or both.

Clearly, the amount of backed-up information varies with the type of backup selected. This, in turn, directly affects the overall strategy in terms of capacity and transfer rate.

How big is the backup window?

Administrators typically perform backups when user demands on the server are at their lowest. Ideally, this time period, the *backup window*, is when user access can be restricted or the server shut down. As more and more companies move to 24-hours-per-day, 7-days-per-week operation, backup windows are shrinking. For many companies with worldwide operations accessing their servers, no clear backup window exists. The system administrator must determine how to get the backup done without impacting the productivity of users or seriously degrading network performance.

How often should backups take place?

Backups must be performed regularly. The actual frequency of backups will be determined by considerations such as:

- the acceptable amount of work that could be lost, if any, in the event of a catastrophic failure;
- the allowable down-time for recovery from this failure; and
- the volume of update transactions that normally take place.

An effective backup strategy should also incorporate *redundancy*.

The *Grandfather-Father-Son (GFS) tape-rotation scheme* is the most commonly used and requires a weekly-backup capacity of at least double the server's storage capacity. It uses three levels of backup to provide redundancy and security. Among other things, this scheme allows for different levels of data retention. The system administrator can choose which generation of tapes to store temporarily and which to archive.

The name GFS refers to these three levels of backup: the monthly "*grandfathers*," weekly "*fathers*," and the daily "*sons*." Typically, the system administrator performs a full backup every Monday (*father*) and does incremental backups on Tuesdays, Wednesdays, and Thursdays (*sons*). The administrator performs another full backup at the end of the week (*father*) and yet another at the end of the month (*grandfather*).[6] The media containing the weekly and monthly backups are usually stored in a location away from the site of the server. To help reduce media costs, many companies reuse older weekly backup tapes.[7]

Of course, the implementation of this GFS backup strategy varies from company to company. One engineering and construction firm has developed the plan shown in the following generic calendar (Figure 1). In addition to GFS backups, this plan incorporates special, user-defined backups.

Sun.	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.
	FULL store offsite (*Father*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	**INCR**. (*Son*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	**FULL** store offsite (*Father*) **SPECIAL** store offsite	
	FULL store offsite (*Father*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	**INCR**. (*Son*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	**FULL** store offsite (*Father*) **SPECIAL** store offsite	
	FULL store offsite (*Father*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	**INCR**. (*Son*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	**FULL** store offsite (*Father*) **SPECIAL** store offsite	
	FULL store offsite (*Father*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	**INCR**. (*Son*) **SPECIAL** store offsite	**INCR**. (*Son*) **SPECIAL**	***End-of-Month*** **FULL** store offsite (*Grandfather*) **SPECIAL** store offsite	

Figure 1. Typical GFS Backup Schedule

[6] The GFS tape-rotation scheme is intended to ensure that a company can always restore lost data to within a day of a disaster.
[7] Never overwrite a *recently used* backup tape for this reason: if a hard-disk crash should occur during the backup, not only will all of the data on the disk have been lost, but the tape will no longer be useful for a complete restoration.

The company that developed this plan has an engineering staff that works on several design-and-development projects simultaneously. One of these projects is mission critical. The management team of that project has requested that the network administrators perform a full, special backup of their applications and data every day. (These are the daily special backups shown in the calendar.) These managers have also requested that their backup tapes be stored offsite every Monday, Wednesday, and Friday. This special backup requirement is in addition to the routine, company-wide GFS backups.

No weekend activity is shown in Figure 1. However, if any mission-critical development is done over a weekend, the network administrator is notified to schedule one or two additional special backups. Also, if the last day in the month is a Saturday or Sunday, the *grandfather* backup takes place on that day. If any of the other engineering projects work on Saturday or Sunday, the normal Monday *father* backup will cover that work.

Notice that the applications and data of the mission-critical project are backed up twice every Monday and Friday, the same days when *father* backups (which are complete backups) take place. The company's management does not see this as an issue. To them, this redundancy is additional insurance against a catastrophic loss of a very important project's information.

SELECTING APPROPRIATE HARDWARE

After determining the company's backup needs, the system administrator determines the specifications for an appropriate, cost-effective backup solution that will best meet those needs in terms of:

- performance,
- capacity,
- cost,
- life expectancy, and
- compatibility with the operating system.

Performance

Needed performance is determined by dividing the amount of information (in gigabytes) that must be backed up by the size of the backup window (in hours). This simple calculation yields the required performance as an overall transfer rate expressed in gigabytes per hour (GB/hour).

Table 1 lists capacities and transfer rates of drives available from Compaq. It also includes typical system transfer rates when doing a local backup using those drives in a system running *ARCserve* from Cheyenne Software, Inc.[8] Table 1 also includes the typical time for that system to do a
10-GB backup using each of the drives.

By comparing the results of the required-performance calculation with the typical backup performance rates listed in Table 1, the administrator can determine if the required backup performance is achievable. If one of the listed drive types will do the job, then local, off-line backups are an appropriate choice.

[8] Other tape-backup programs deliver similar results. However, one new backup program, *JETserve* from Cheyenne Software, Inc., is capable of delivering significantly higher backup performance. This program is available from Compaq and will be discussed in a future white paper.

TABLE 1

	Compaq Tape Drives					
	525 MB QIC	1.2 GB QIC	2/8 GB DAT	4/16 GB DAT	10/20 GB DLT	15/30 GB DLT
Drive's native (uncompressed) capacity	<525 MB	<1.2 GB	<2 GB	<4 GB	<10 GB	<15 GB
Drive transfer rate without data compression (GB/hr)	0.720	1.03	0.658	1.44	4.5	4.5
Typical transfer rate in a system (GB/hr)	0.2 - 0.5	0.4 - 0.8	0.3 - 1.0	1 - 1.75	3 - 6	3 - 6
Typical time for a 10-GB backup	13.9	9.7	14.8	7+ hr	2+ hr	2.2 hr

Capacity

A 1993 study[9] by Strategic Research Corporation estimated that the annual cost for personnel to manage information storage may be as high as $7 per megabyte. Compared to the initial hardware costs, this administrative cost is much too high. Unattended backups can significantly reduce administrative costs and obviously should be used whenever they are appropriate.

For unattended backups, the tape drive and media must meet the backup requirements in terms of overall storage capacity as well as backup performance.

Because of variables such as the different types of files to be backed up, only the drive's native capacity should be evaluated during the drive-selection process. (In other words, zero compression should always be assumed.) The compressed capacity of the typical drive is calculated for ideal conditions, which are very hard to predict.[10]

As Table 1 indicates, both performance and capacity vary with the type of drive technology selected. The older QIC technology generally has the lowest capacities and slowest transfer rates. The newest DLT technology has the highest capacities and fastest transfer rates.

- **QIC** (**Q**uarter **I**nch **C**artridge) drives meet the half-height form factor of desktop computers. QIC tapes are virtually industry-standard for standalone machines. However, with capacities limited to 1.2 GB, they are not generally suitable for backing up servers with 2 GB or more storage capacity.
- **DAT** (**D**igital **A**udio **T**ape) drives are usually the appropriate choice for servers with 2- to 8-GB capacity. Standard DDS1 DAT tape drives (2/8 GB DAT) can store approximately 2 GB without compression. DDS-2 DAT (4/16 GB Turbo-DAT) tapes can store 4 GB without compression. An Autoloader for use with DAT tapes is available from Compaq. Autoloaders reduce administrative costs by using a robotic mechanism to load and unload tapes.

[9] *Cost of Managing Storage*, Sept., 1993. © 1993 by Strategic Research Corporation.
[10] Once a given system has been backed up, then the actual, average compression ratio for that system will become clear. Beginning at that point, the actual compression ratio can be used as a factor in future calculations.

- **DLT** (**D**igital **L**inear **T**ape) drives use simultaneous, multichannel/multihead read/write technology to achieve capacities up to 10 GB without compression. A DLT drive is the appropriate high-end backup solution for systems with 10 to 30 GB.

Another white paper in this series will describe the three drive technologies in greater detail.

Hardware and media costs

The needed drive performance and drive capacity must be evaluated in view of current drive and media costs and the company's budget. A realistic budget for the purchase of the appropriate drives and media is important. If too little is budgeted for drives, the company will probably incur increased labor costs. An all-too-common example of this involves a system administrator who must work overtime every evening just to change tapes on a drive without an autoloader. Or, the company may end up with a drive that is simply too slow to complete the backup during the backup window. This could result in a degradation in server performance outside the backup window. It will certainly result in premature wearout of the tape device. On the other hand, a drive with ten times the required capacity may have the advantage of low cost per gigabyte, but the initial purchase price would be difficult to justify.

Because the costs of backup hardware and media constantly change, Figures 2 and 3 illustrate the *relative* prices of different drives and media types. Note that the QIC technology (drives and media) is the most expensive per GB. The TurboDAT drive with the 12-cartridge magazine is the least expensive per GB (Figure 2), but the media costs are essentially the same for the 4/16 GB TurboDAT Drive, the TurboDAT with the 12-cartridge Autoloader, and the TurboDAT with the 4-cartridge Autoloader (Figure 3). This is because the same cartridge is used for all three drives.

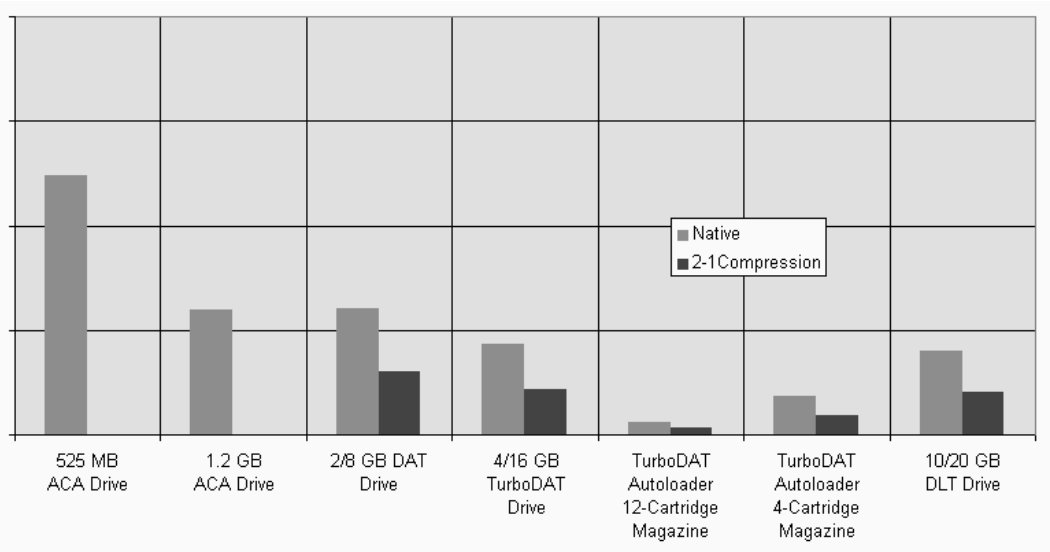

Figure 2. Relative Drive Cost per Gigabyte

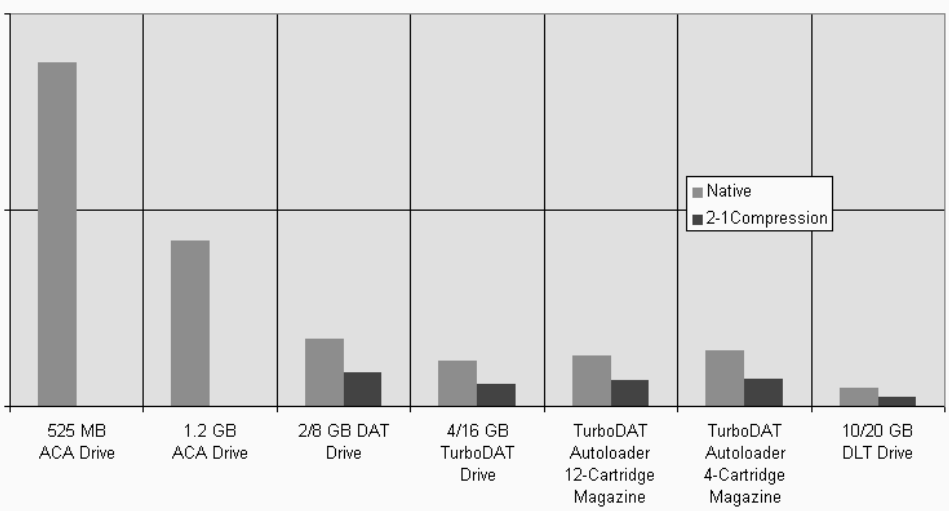

Figure 3. Relative Media Cost per Gigabyte

Hardware Reliability

The reliability of any backup device is directly related to its duty cycle (the number of hours per day that the device is in use). For example, if a tape drive designed for 1-GB backups is being used to back up a 10-GB server, the results on that drive will be

- a need for accelerated preventive maintenance—especially head cleaning,[11]
- premature aging, and
- reliability problems.

The best method for building hardware reliability into a backup strategy is to ensure that the backup hardware is matched to the server(s). Table 2 relates various servers with the appropriate tape drives. If the company needs special, partial backups in addition to the routine backups, it might be appropriate to select the next larger drive size.

TABLE 2

	Compaq Tape Drives					
	525 MB QIC	1.2 GB QIC	2/8 GB DAT	4/16 GB DAT	10/20 GB DLT	15/30 GB DLT
Corresponding Server Capacity	<1 GB	<2 GB	<2-4 GB	<4-8 GB	<5-15 GB	<10-25 GB

[11] This applies primarily to DAT drives. DLT read/write heads make very little actual contact with the tape and require much less cleaning. In DAT drives, the tape is wrapped around the read/write heads; hence the need for additional head-cleaning in the situation described above. DAT heads *must* be frequently cleaned for successful backups.

If a DAT drive appears to be the appropriate choice for a given system, the projected duty cycle of the drive becomes an important consideration. Consider the example of an organization that plans to do an unattended, 8-GB complete backup every night, using one 4/16 GB DAT drive. Eight gigabytes are well within the capacity of the drive. However, at a backup rate between 1 GB/hour and 1.75 GB/hour, that tape drive will be in operation for 6 to 8 hours every night. With this amount of usage, the tape drive's read/write heads should be cleaned *every other night*. Assuming that a person can be found to do this, these backups could hardly be called "unattended." If the drive heads are not cleaned when they need cleaning, soft errors will increase (as will overall backup time) to the point at which something will fail. For this particular customer, even though a 4/16-GB DAT drive would seem to be the appropriate choice, a DLT drive would be a better choice. This is because of the difficulty most customers would have finding someone to clean the DAT heads every other night.

Operating Systems and Backup Software

The first step in selecting the best backup software application program is to evaluate what is available for the operating system being used. Most backup applications on the market today offer appropriate capabilities for most environments. Table 3 (on the following page) pairs currently available backup software with common operating systems and indicates which Compaq tape drives are supported.

TABLE 3

		Compaq Tape Drives				
		QIC	DAT		DLT	
BACKUP SOFTWARE (See Keys)	**OPERATING SYSTEM**	**525 MB and 1.2 GB**	**2/8 GB and 4/16 GB**	**Internal TurboDAT Autoloader**	**10/20 GB**	**15/30 GB**
AS	Novell Netware (3.1x and 4.1x)	●	●	●	●	●
JS	Novell Netware (3.1x and 4.1x)		●		●	●
AS 2.01	Windows NT	●	●	●	●	●
BE	Novell Netware (3.1x and 4.1x)	●	●	●	●	●
	Windows NT	●	●	●	●	●
NT	Windows NT	●	●	●	● (See Note)	● (See Note)
NV	OS/2 and LAN Server	●	●	●	●	
PL	Novell Netware (3.1x and 4.1x)	●	●	●	●	●
ST	OS/2 and LAN Server	●	●			
UX	SCO UNIX	●	●	●	●	●
VN	Banyan Vines	●	●		●	●

Keys:
- AS: Cheyenne *ARCserve* Ver. 5.01 and Ver. 6.0 (from Compaq).
- AS 2.01 Cheyenne *ARCserve* Ver. 2.01 (from Compaq).
- JS Cheyenne JETserve (Ver. 1.0) (from Compaq).
- BE: Arcada *Backup Exec* Software.
- NT: Microsoft Windows NT support for Compaq tape drives is included in the operating system.
- NV: Novastor *Novaback* for OS/2.
- PL: *Palindrome* Software.
- ST: *Sytos Plus* Software.
- UX Device drivers for Compaq tape drives are available from SCO in the SCO Compaq Supplement.
- VN: Banyan Vines support for Compaq tape drives is included in the operating system.

NOTE: Support for Windows NT versions 3.5 and 3.51 is included with the operating system. Other support is available through *Backup Exec* from Arcada Software.

CONCLUSION

As long as companies continue to entrust their mission-critical data to their computer systems, networks, and enterprises, and as long as the possibility of a catastrophic failure exists, a well-thought-out backup strategy is crucial. Developing the optimum backup strategy for a particular corporate situation requires an understanding of the performance, capacity, life, and costs of the various backup solutions available today and in the near future.

In general, the system administrator will weigh performance, capacity, projected duty cycles, and cost factors against the volume of backup information, choosing the backup technology that is both appropriate and cost-effective. Table 4 summarizes the technologies in use today.

TABLE 4

	Backup Volume		
Factor	**Low**	**Medium**	**High**
Performance	QIC	DAT	DLT
Capacity	QIC	DAT	DLT
Usage	QIC	DAT	DLT
Cost	DAT	DAT	DLT

GLOSSARY

Applications (*or* application software)

The general term for software programs that perform specific tasks such as accounting, word processing, and database management.

Archive bit

Each file has a bit which is used by backup software to determine whether the file has been backed up or accessed since the last *complete* backup. After each complete backup, the backup software resets the archive bit (turns it off) for each file. After a file has been opened (or opened and modified), the operating system sets that file's archive bit (turns it on), indicating that the file has been opened or (opened and modified) since the last *complete* backup.

ARCserve

An enterprise-wide data-management and backup program that enables backing up and restoring all information on Netware servers and on all workstations attached to those servers. A product of Cheyenne Software, Inc., ARCserve runs on file servers running NetWare 3.11 and above. ARCserve can back up and restore information from the following environments: Windows, DOS, OS/2, UNIX (using Cheyenne's *UAgent* software) and Macintosh (using Cheyenne's *MacAgent* software).

Backup

Used as a verb, the process of creating a duplicate of all or part of the data and/or programs on a computer system. The data is copied to a removable storage medium, such as a tape. Depending on the type of backup, the tapes may be stored remotely (away from the main system).

Used as a noun, the result of such a process.

Backup application

The specific software used to control the backup process (for example, ARCserve). Using this application, the system administrator can define which information to back up, where to send the backed-up information, the amount of file compression, etc.

Backup path

The route the backed-up information travels in moving from the source to the destination.

Backup source

The server or workstation being backed up.

Backup window

The time period within which backups are scheduled. Typically, a backup is performed at night or on weekends, when user demands on the server are at their lowest.

Client/Server Network

A communications network that uses dedicated network servers for all client workstations in the network. Compare this with the definition of peer-to-peer network, which allows any client to be a server also.

Complete (*or* full) backup

A copy of the entire system on one tape or set of tapes that can be used to restore the entire system in the event of a catastrophic failure.

Complete restore

A restoration of all files copied from the backup media and copied to the source media. Depending on the backup application software, a complete restore might be destructive, replacing all files on the destination drive and thereby destroying whatever information was previously stored in those files.

Database

A file or set or files containing information generated by a database management system (DBMS). Many companies store large amounts of mission-critical data in company-wide databases. Since the information is being continuously accessed, these databases pose particular difficulties for backup. Many backup applications also generate a database to keep track of information such as the files included in the backup.

Differential Backup

A type of partial backup that copies only those files that were changed since the last *full* backup. This type of backup is useful when it is important to have the latest version of each file. If the same tapes are used for consecutive differential backups, the newer versions of backed up files are often allowed to overwrite older versions of the same file on the tape. See Incremental Backup.

Driver

A relatively small program (typically, only one or a few files) that controls the specific interactions between applications (such as backup applications) and a particular piece of hardware (such as a tape drive).

Duty Cycle

A manufacturer's rating that indicates the percentage of time that a given hardware product (for example, a tape drive) can safely be in use. A tape drive with a 10% duty cycle is expected to be in operation 10% of a day, or 2.4 hours per day. See MTBF.

Full Backup

See Complete Backup, above.

GFS

Refers to the **G**randfather-**F**ather-**S**on tape-rotation scheme. This scheme has become the standard in the mainframe and minicomputer environments and is one of the most popular in the PC-networking environment. The GFS scheme is intended to ensure that a company can always restore lost data to within a day of a disaster. It is a strategy for maintaining backups on a daily, weekly, and monthly basis. A complete backup is performed at least once per week. (This is the Father tape or tapes.) Daily backups are performed on the other days of the week. (These are the Son tapes; they are often incremental or differential backups.) At the end of the month, another complete backup is run. (This is the Grandfather tape.) The popular backup software programs incorporate the GFS strategy into their tape-management routines.

Incremental Backup

A backup of only those files that have changed since the last backup, regardless of what type of backup that was. In doing incremental backups, most backup application programs will copy all specified files that have been opened for any reason other than read-only access. See Differential Backup.

Mission Critical

Vital to the operation of an organization. In the past, mission critical information systems were implemented only on mainframes and minicomputers. Increasingly, they are being designed for and installed on personal computer networks. Many firms consider their network-backup functions to be mission critical.

MTBF

Mean **T**ime **B**etween **F**ailure, the average time that a hardware product (for example, a tape drive) will run before a failure occurs, assuming that the drive's duty cycle is not exceeded. A drive with a 20,000-hour MTBF rating at a 10% duty cycle should last for 20,000 hours of operation if the drive is in operation less than 2½ hours per day. See Duty Cycle.

NetWare

A family of network operating systems from Novell, Inc., that support DOS, OS/2, Mac and UNIX clients and various LAN access methods including Ethernet, Token Ring and ARCNET. NetWare is the most widely-used LAN control program. *Personal NetWare* is a peer-to-peer network operating system, which allows any client workstation to be a server. It supersedes earlier peer-to-peer versions known as *NetWare Lite* and *NetWare ELS* (Entry Level System). *Personal NetWare* is also included with Novell's DOS operating system. See Peer-to-Peer Network.

NetWare 2.x (originally *Advanced NetWare 286*) runs in a dedicated file server (286 and up) and supports up to 100 concurrent users per server. This version is no longer being updated.

NetWare 3.x (originally *NetWare 386*), which supports up to 250 concurrent users, runs on 386 servers and up and takes advantage of 32-bit architecture.

NetWare 4.x, introduced in 1993, is backward compatible with *NetWare 2.*x and *3.x* and includes the *NetWare Directory Service* (NDS), which provides X.500 compatibility. A *NetWare 4.x* server supports up to 1,000 concurrent users.

Except for *Personal NetWare*, NetWare is a stand-alone operating system that runs in the server. It does not use DOS or any other operating system. The hard disks in a NetWare server are formatted with a Novell format, not a DOS format.

SFT (System Fault Tolerant) NetWare provides automatic recovery from network malfunctions.

Portable NetWare provides NetWare source code for conversion to other platforms.

Off-line

A server that is off-line is unavailable to users. Typically, an off-line backup (a backup of an off-line server) takes place during off-hours, at a time when user demands are at their lowest.

Off-site storage

A location at some distance from the site of the server that has been backed up. Some subset of the backup tapes is typically stored at this distant location to guard against theft and/or damage to backups stored in the same location as the server(s).

On-line

An on-line server is available to users. On-line backups take place while users are using network services and probably updating files that are to be backed up.

Peer-to-peer Network

A communications network that allows all workstations and computers in the network to act as servers to all other users on the network. Dedicated file servers may be used, but they are not required as in a client/server network. *Netware Lite* uses this method of networking to share resources across the network.

Restore

The process of copying files from an earlier backup back to the server or workstation. A restore is performed to get the system back to a known state after information has been lost or files have been corrupted. Restores can be complete or partial.

Verification

The process of comparing backed-up files with the original files on the server.