white paper

How to optimize the performance of hp ultrium tape drives

Abstract

This white paper explains how to get optimum performance from the new generation of high-performance HP Surestore Ultrium drives. It describes the hardware, software, network and system configuration requirements that enable the drive to deliver its optimum performance.
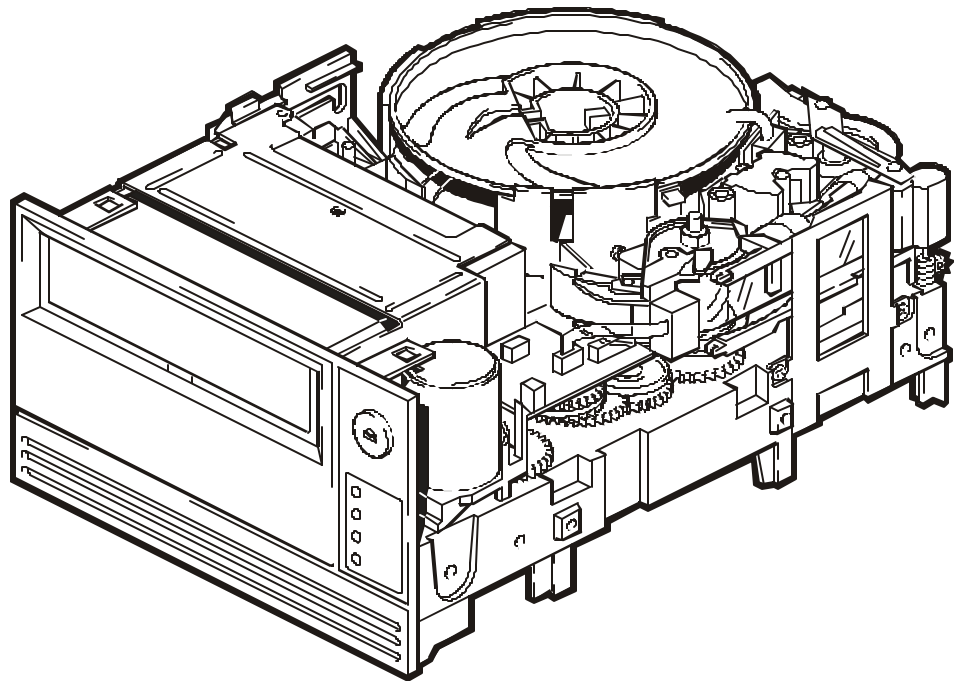

Figure 1: HP Surestore Ultrium tape drive

How to use this white paper

This white paper describes a number of different scenarios in some detail and is, therefore, quite lengthy. It also includes a glossary. For ease of navigation, select an area that interest you from the hyperlinks on the next page or use the bookmarks (enabled from Acrobat's Window menu) to jump to a particular topic.
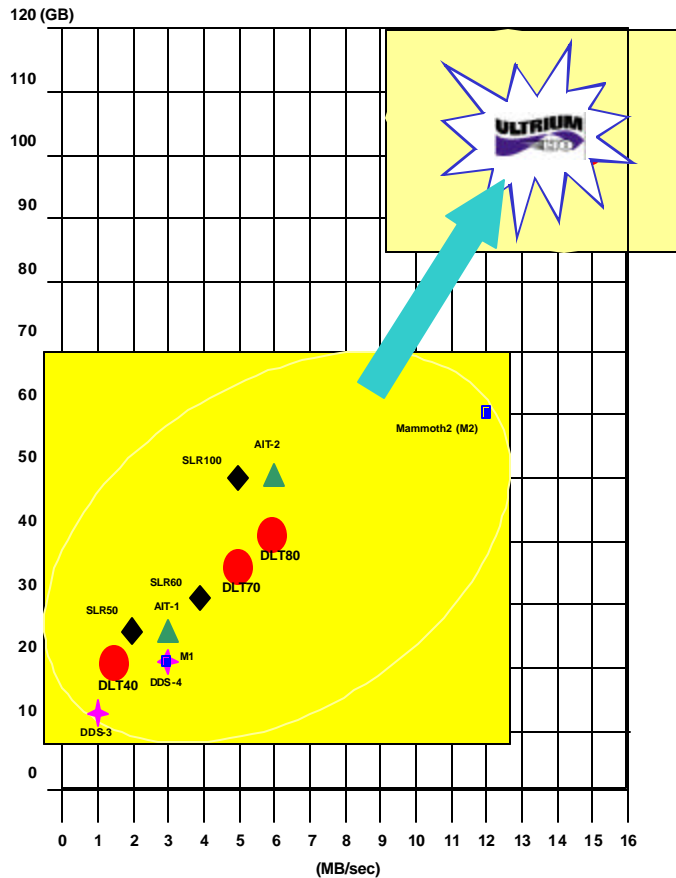
Contents

The main topics in this white paper are:

Executive summary

Figure 2: Ultrium a new paradigm in tape backup

The new, high-performance drives require systems that can supply data at up to 30 MB/sec

With the explosive growth of data in today's computing environment and the relentless requirement for high availability systems, a new generation of tape drives has been developed that can back up higher capacities of data in a shorter amount of time than ever before. This new tape technology is Ultrium and it represents a new paradigm in tape backup.

Assuming a compression ratio of 2:1, these next generations of tape drives are capable of streaming data at up to 30 MB/sec (108 GB/Hr). The issue lies in whether the systems to which the tape drives are attached can supply data at this rate. Failure to do so can result in:

- Misled customer expectations. Longer than expected backup windows lead to less than desirable system availability at full performance levels.

- Potential reliability implications. Once a drive cannot sustain a transfer rate it goes into "re-positioning mode" which will cause multiple passes of the head over the media, therefore reducing media life and increasing mechanical wear.

Tape can no longer be an "add-on" accessory that can be expected to perform well on any system.

To use today's high-performance tape drives to best advantage, backup performance must now be considered and planned for at the system design stage. It is no longer simply a question of adding tape drives to a system as and when they are required and expecting them to perform well.

We will discuss system scenarios in more detail in the rest of this white paper, but to summarize:

- If tapes are attached directly to the server, the fundamental enabler in being able to stream Ultrium technology is a high performance Disk RAID subsystem

- In centralized (network) backup, the fundamental enabler in being able to stream Ultrium technology is Gigabit Ethernet, now becoming widespread and cost-effective

- In Storage Area Networks, providing the source data is available at more than 30 MB/sec, there should be no problems streaming Ultrium technology.

- Using the Virtual Device Interface (VDI) to SQL server, (with an appropriate performance RAID subsystem) Backup Application Agents can achieve high performance.

The ability to stream Ultrium technology with lower performance RAID subsystems could be achieved if Independent Software Vendors (ISVs) would consider introducing "multi-threading" into the backup application software to drive the RAID hardware more efficiently

Table 1 on the next page illustrates the Ultrium 1 technology specification.

Table 1: Ultrium technology specification

Technology	Native Capacity	Native Transfer Rate	Speed Enabler	Comments
HP Ultrium Generation 1 A new standard in high end tape technology, also being developed by Seagate Technology, IBM.	100 GB	15 MB/sec Entry Level Product also available with 7.5 MB/sec native transfer rate ALDC Compression delivering typical compression ratio of 2.55:1[+]	Linear Tape Technology 8 tracks simultaneous write/read, High tape speed 4.1 m/sec	Due Fall 2000 Roadmap for 4 generations up to 800 GB (native) and 160 MB/secs announced

+ = Source = Avax International test results using Calgary/Canterbury compression corpus.

http://www.avax.com/compress.html

Key to Figure 3

Figure 3 on the next page provides an overview of the factors affecting tape performance. The numbers on figure 3 (on the next page) relate to how these factors may be adjusted.

(1) only configurable at system design or install time

(2) system upgrade

(3) maintenance procedure

(4) adjustable online (4*) Available on HP NetRAID 4M only

(5) automatic feature

(6) requires further technology enhancements

*Before continuing, you may like to familiarize yourself with the glossary of terminology at the end of this document. An * will indicate where a reader should read the glossary for a detailed explanation of the concept involved.*

Figure 3: Overview of factors affecting tape performance

RAID

RAID Controller (Board)

uProcessor

SCSI Protocol Chips

Cache

SCSI or FC Bus(s)

PCI Bus

Memory

PCI Bus

Network Interface Card (NIC)

Network Topology and Transport

OTHER DATA SOURCES FOR TAPE

Fibre Channel HBA (FC)

Storage Area Network (SAN)

Host Bus Adapter (HBA)

Tape Drive

	Disks	Raid	File System	Application	Tape	
PARAMETERS	• No. of Disks • Speed of Disks • Raid Levels • Stripe Size	• Processor Speed • Cache Settings • No. of SCSI Bus • Bandwidth of Bus • RAM	• Processor Speed • Memory • File Size • Access Pattern • File System Type (*6*)	• How Files are Called Random vs Sequential • No's of Threads Enabled • Buffering Efficiency • Block Size	• SCSI Bus Speed • Compression Algorithms • Block Size • Buffer Size	**PARAMETERS**
INCREASE PERFORMANCE	• More Disks (*1,2*) • Higher Speed Disk (*1,2*) • Raid Level (*1,4**) • Larger Stripe (*1,4**) • Queue Depth (*6*)	• Faster Processor (*2*) • Faster Bus (*2*) • Faster Cache (and Algorithms) (*2*) • Cache Parameters (*4*) • More Cache (*2*)	• Sufficient CPU and Memory (*2*) • Fragmentation Effect (*3*) • Cluster Size (*3*)	• Block Size/ Buffer Size (*4*) • Multi-Threading (*6*) • Image vs File by File (*2*)	• Data Rate Matching (*5*) • Dedicated HBA per Tape Device (*2*)	**INCREASE PERFORMANCE**

Scenario 1: optimizing performance with direct tape attach

This section describes the main contributing factors which affect direct tape drive attach performance, looks at future enhancements and ways to assess file system performance. For a quick summary of the information in this section, see page 16.

There are five major components that can affect performance:

- the disk subsystem

- file sizes and fragmentation

- disk file access - random or sequential

- tape drive performance

- operating systems and backup applications

The disk subsystem

The performance of the disk subsystem is by far the most critical factor in being able to stream Ultrium tape technology directly attached to the server. With the latest (Ultra2) 18 GB 10,000 rpm drives you will be able to achieve transfer rates through typical third party applications of around 10-14 MB/secs from a single drive with typical user data. This speed is barely enough to stream the tape drive and repositioning is likely to occur. To realize the full potential of the tape drives (with speeds above 15 MB/sec) and improve your application performance, a disk RAID subsystem is an absolute necessity

Types of RAID controller

CAUTION: Not all RAID controllers are created equal. The Gartner Group segments RAID controllers into three types:

- RAID on a chip (ROC)
- RAID on the motherboard (ROMB).
- Host based add-in RAID controller (ADD-IN)

ROMB based implementations are around 50% less expensive than single channel add-in RAID controller implementations. RAID controllers can operate over more than one channel to optimize performance of disk drives over the SCSI bandwidth. 1,2,3,4 channel RAID cards are currently available.

In HP's Tape Performance Lab we found the ROC & ROMB implementations on today's servers can provide reasonable (around 17 MB/sec) tape performance with typical user data on 8 disks with a single channel Ultra 160 or dual channel Ultra 2 RAID controller (specified at 100 MHz).

For the best backup and application performance (30 MB/sec and above) with typical user data on an 8-disk RAID subsystem HP recommends the use of an add-in RAID controller such as:

- HP NetRAID 4M 4 channel Ultra160, 233 MHz, 128 MB cache, 64 bit PCI,

- DELL PERC 2 4 channel Ultra 160, 233 MHz 128 MB cache 64 bit PCI

- Mylex eXtreme 1100 3 channel Ultra 2, 233 MHz 64 MB cache, 64 bit PCI

- Compaq 4200 SMART Array controller

Below are two characterization graphs of an HP NetRAID 4M RAID controller performing sequential reads from disk. Figure 4 shows perfomance characterisation with high queue depth and Figure 5 shows the performance characterisation with a single queue depth. Please note the following important points:

1. The performance increases as the number of disks connected to the RAID controller increases.

2. The performance increases as the transfer size increases.

3. The performance increases as the queue depth increases. Queue depth is the number of "stacked" read requests the application can input to the RAID controller and which the RAID controller can then optimize, locating the nearest data to its current position. The ability to utilize this feature is application-dependent. Databases, for example, generate queue depths because they are transaction based. Current backup applications however do not generate queue depths, so the backup application performance is limited really by the single queue depth, as shown in figure 5 below.

4. Remember this is sequential read from disk and so represents best case performance. Some applications, such as SQL, do have a sequential read characteristic from disk. File-by-file backup, however, tends to have a more random disk access pattern.
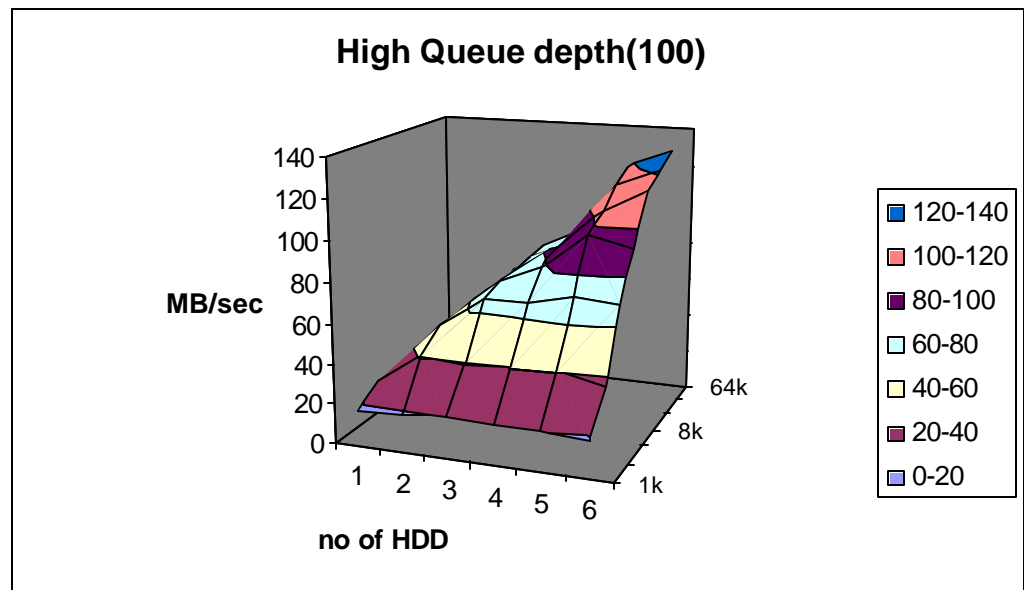

Figure 4: HP NetRAID4M performance characterisation with high queue depth
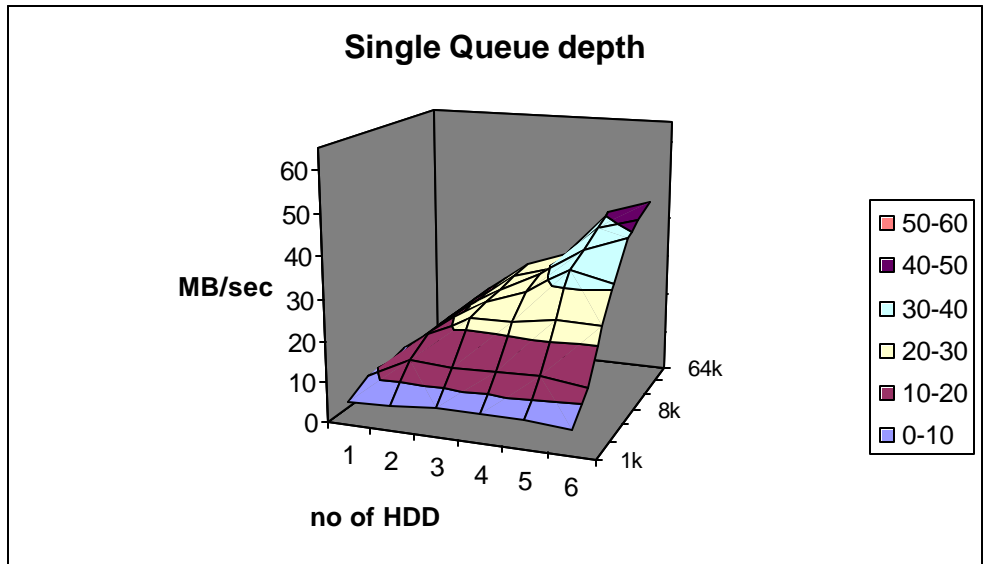
Single Queue depth

Legend:
- 50-60
- 40-50
- 30-40
- 20-30
- 10-20
- 0-10

Y-axis: MB/sec (0, 10, 20, 30, 40, 50, 60)
X-axis: no of HDD (1, 2, 3, 4, 5, 6)
Z-axis: 1k, 8k, 64k

Figure 5: HP NetRAID4M performance characterisation with single queue depth

Improving backup performance from a RAID controller

- Ensure the disk drives are allocated equally across the multiple channels on the RAID controller. This ensures each channel has maximum SCSI throughput for each disk attached.

- The number of disks in the RAID 0 or RAID 5 subsystem can be increased. This will also increase application performance, but of course this comes at a price. With Windows2000 and its integrated Logical Disk Manager (LDM) it is possible to create, break and rebuild RAID levels 0,1,5 ONLINE, making it easier to increase performance by adding more disk drives. For NT 4, users must purchase Veritas Volume Manager to accomplish the same effect. HP NetRAID 4M also allows dynamic online reconfiguration.

- Ensure the stripe size is set correctly at install time to best match the application the server will be running. For example, stripe sizes can be set to match the database record sizes running in the application. HP NetRAID 4M allows this to happen dynamically online.

- Adjust RAID cache settings*. Each RAID controller manufacturer can have a different caching algorithm so it is difficult to give a generic solution here. The cache settings available are explained in the glossary, but a certain amount of experimentation will be required. Optimal RAID cache settings for the application may not necessarily be best for backup. In general, it is the cache read parameters that will impact backup performance the most.

- Connect the tape drives to a separate SCSI bus from the disk to avoid conflict on the same SCSI bus and improve throughput (this will be automatic if RAID controllers are being used). For tape libraries with more than one drive installed, allocating a separate Host Bus Adapter (HBA) to each drive will ensure maximum performance.

For reasonable tape performance, you need a single channel RAID subsystem with a minimum of four Ultra 2 drives and a minimum 100 MHz, 64 MB Cache RAID controller plugged into a 64 bit/66 MHz PCI bus. Ensure Read cache is enabled.

For best tape and application performance, add more disks and a higher performance 2 or 4 channel RAID controller (233 MHz or better) again into a 64 bit, 66 MHz PCI slot. Ensure Read cache is enabled.

- There are normally application sizing tools, which determine what transaction performance can be maintained. Additional allowance (in terms of transaction capability) for the backup application to run effectively whilst maintaining an acceptable application performance should always be taken into account at system design time.

- On NT, use the performance monitor to assess processor loading and memory usage once an application is installed, and to see how much margin the system has when the backup is running. If CPU loading is greater than 70% and memory usage is more than 80% you should consider adding more processing power and memory.

File sizes and fragmentation

Luckily adding more disks to a disk array, more processors to a system and more memory is a relatively easy way to improve performance.

The distribution of file sizes on a system is key to performance. Large numbers of small files cause a high overhead to the file system, therefore, the amount of time spent in the file system is high compared to the actual amount of data transferred. Large files on the other hand have a lower overhead for the amount of data transferred. The file size distribution is determined largely by the application. For example, File Print servers tend to have a higher proportion of small file sizes, which will cause performance issues in file-by-file backup mode (see below).

- Microsoft Exchange servers tend to have transaction logs that are fixed at 5 MB and an information store (total mailboxes on system), which is a very large file indeed. These large file sizes make it easier to stream today's higher performance tape drives, especially if the access to these files is sequential from a disk perspective.

- Microsoft SQL server also utilizes transaction logs together with database pages (8K) that occupy contiguous disk space. As long as the backup application addresses the database pages sequentially, reasonable performance should result.

File size is everything in performance terms. The bigger the file size, the better the transfer rate.

Ensure you implement a regular defragmentation strategy.

File Fragmentation* whereby a file, instead of being contiguous, is spread out over various random parts of the disk can also severely impact performance (especially on NT). NT uses a cluster size* of around 4 KB, which means that 4 KB is the minimum amount of data to be written or read in one transaction and leads to defragmentation. Windows 2000 now includes a defragmentation tool as standard.

Unix, in general, can write read in individual disk sectors (512 bytes) leading to less overall fragmentation.

NetWare NDFS has the ability to allocate sub-blocks within a data block as an option at volume creation time, which can also lead to fragmentation.

Disk File Access - random or sequential?

The way the backup application accesses the data is also critical to performance. Three approaches are in general use today:

- File-by-File*. This follows the normal file system access paths, and translates into random access on disk.

If your system has many small files, consider using the image backup option to improve performance.

- Image Backup*. Here the normal file system access is bypassed and the disk is read sequentially sector by sector, resulting in less file system overhead and no random motion on disk. The down side is that a single file restore can take some time to complete because the disk image on tape must be searched along its whole length to reconstruct the file. Image Backup is also sometimes referred to as raw disk backup in the Unix world.

- Snapshot* & Storage checkpoint backup technology*. With the increasing size of databases, it is becoming more and more impracticable to back up the whole database every night and still meet the high availability and performance requirements required by users. Snapshot/storage checkpoint technology just being introduced will allow the creation of a clone of a file system without making a physical separate copy of the file system. Storage checkpoints keep track of changed file system blocks. This saves disk space and significantly reduces I/O overhead by maintaining the changed file system blocks in the same underlying disk space instead of a separate storage pool. Because checkpoint technology can track changed blocks (not changed files), backup applications can be enhanced to provide incremental **block** level backup. Although this may result in random disk accesses to gather the changes, the block level changes can be gathered into a contiguous file that can then be streamed sequentially with good performance to tape.

The table below compares worst case performance in file-by-file mode with image backup of the same data. It is based on a single Ultra2 disk drive with many 1K files on NT and a high degree of fragmentation.

Drive Type Wide Ultra2 SCSI	Sequential Disk Access (Image Backup)	Random Disk access (File by File Backup)
10,000 RPM	12 MB/sec	> 1 MB/sec

Table 2 - file-by-file versus image backup on a single drive with small fragmented files

Tape drive performance

ONLY connect to a separate LVDS bus.

Use the recommended tape block size.

Use a tape drive that supports data-rate matching.

For maximum performance always ensure the tape drive is on a separate SCSI interface to the disk subsystem (this maximizes SCSI bandwidth).

This new generation of high performance tape drives have an LVDS Ultra2 SCSI interface that offers up to 80 MB/sec burst transfer rate. These new tape drives should only be connected to an LVD bus. They must **NOT** be connected to a single-ended SCSI bus or have a single-ended drive connected to the same bus, because the SCSI bus will default to 40 MB/sec and severely affect tape drive performance.

The bigger the tape block size the less the SCSI overhead for transfer of data. 32K and 64K are typical values. The larger the tape block size, however, the more wastage there may be in the application "padding the tape blocks". HP Ultrium Technology recommends a minimum 32K block size. Use the backup application software to ensure the tape drive supports the block size you set up.
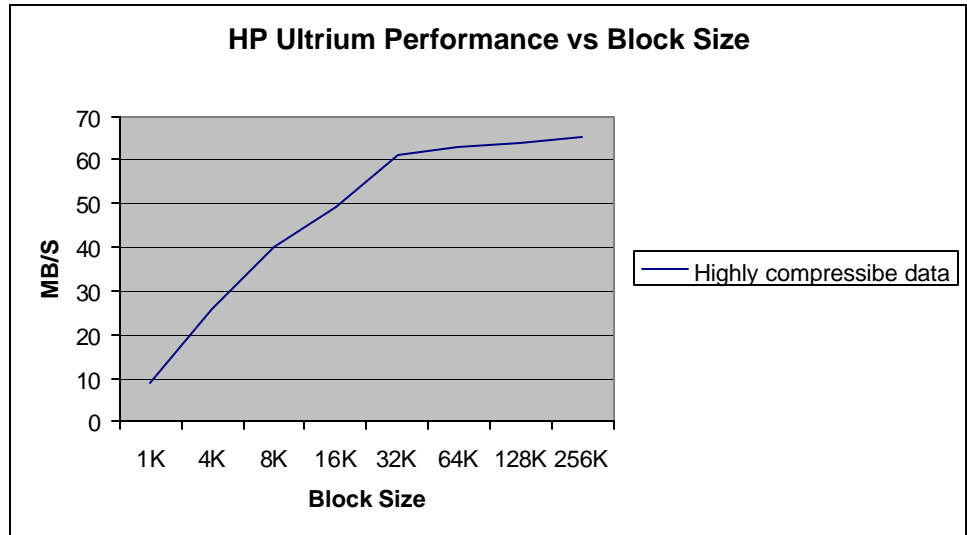
Figure 6: HP Ultrium drive performance compared with block size and highly compressible data

Some of the new higher performance tape drives such as HP Surestore Ultrium offer data rate matching* whereby they adapt the tape speed (within a certain range) to match the incoming data rate from the host. This prevents repositioning with subsequent performance gains. The HP Surestore Ultrium tape drive, for example, can data rate match incoming native data rates from 6-15 MB/second.

Operating Systems and Backup Applications

The HP Tape Performance Lab performed a benchmark using identical user data (3.2 GB) mapped into NT4, Windows 2000 and Red Hat Linux 6.1 operating system file structures on an 8 disk RAID system with NetRAID 4M controller. Data was backed up to an HP SureStore Ultrium 230 tape drive and the same hardware was used throughout. Performance results were as follows using the "native", built-in backup applications:

Operating System	Native Backup Application	File by File Performance on same data set
Windows NT 4.0	NTBackup	2 MB/sec
Windows 2000	NTBackup	12 MB/sec
Red Hat Linux 7.0	tar	12 MB/sec

Table 3 - Comparison of different OS performance in file-by-file mode

NT4 with its poor performance tape driver architecture could not get above 2 MB/sec. The Windows 2000 tape driver was better but still below the streaming rate of Ultrium (luckily, data rate matching will help here). Disk performance also limited Windows 2000 performance - adding more drives it could have been improved further. Linux was limited by the small file sizes (10k) used by tar.

Netware recently changed its tape driver to be multi-threaded, which is why its native backup utility should now perform well, but the Netware 5.1 driver for Ultrium was unavailable for testing at the time this paper was written.

Using additional third party ISV backup application packages can significantly improve on the native backup utilities because of better memory management and other features built into the backup applications. Within the backup applications there are generally three areas of tuning

- Increase tape block size. This is generally preset to a default but can be increased if the tape device supports the higher block sizes.

- Increase the memory available for backup application usage and "watermark" buffers to maintain streaming.

- Increase SCSI transfer size through the operating system (typically, the default is 64 K). Some ISVs may use "pass through" drivers to accomplish this.

These changes are generally referenced by backup ISVs but they may involve changing registry entries (NT) or driver editing (Linux) and should be done with caution by people who know what they are doing. The glossary contains examples of how to do this on NT.

In our in-house test at HP we concluded that—except for slightly improved performance with fragmented and compressible data—Windows 2000 is virtually identical to Windows NT when specific third-party backup applications are run. The trends of the two operating systems are the same. In terms of the backup application, changing to Windows 2000 from Windows NT would be of no real advantage/disadvantage to the customer. See figure 6.

Figure 6: comparing Windows 2000 performance with Windows NT using ArcServeIT 6.61

With the current (July 2000) ISV shipping products used for normal file-by-file backup, the HP Performance Test Lab found that no single vendor offered major performance advantages over another, taking into account the whole spread of file sizes on a system. Figure 7 below shows two major backup applications tested on a 550 MHz Pentium III machine with a NetRAID 4M high-performance RAID controller and an 8 disk array under NT4. The data sets were at predefined sizes from 1 K up to 100 Mbytes and consisted of random data (1:1 compression) and compressible data (around 1.5:1). The files were contiguous on disk and the tape drive was an HP SureStore Ultrium 230 drive. Veritas and Legato drivers were unavailable at testing time.

Backup Application File by File Backup Rates

Legend:
- ArcServeIT Compressed
- ArcServeIt Random
- OmniBackII Compressed
- OmniBackII Random

Y-axis: MB/s (0, 5, 10, 15, 20, 25, 30)
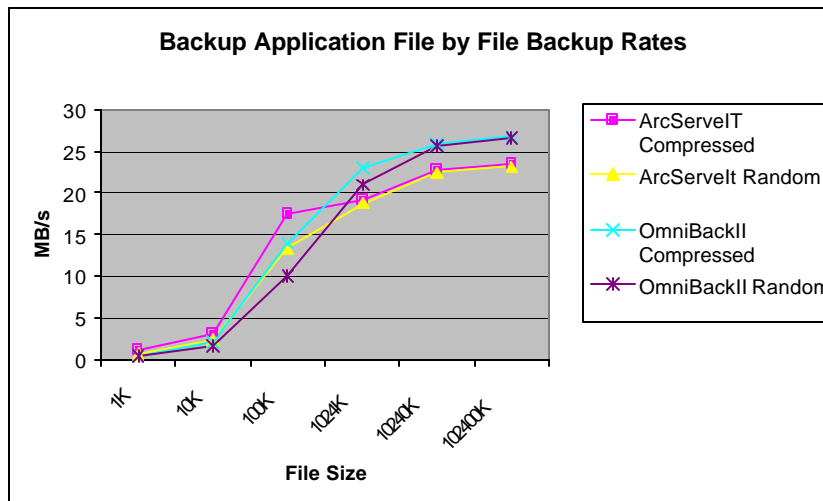X-axis: File Size (1K, 10K, 100K, 1024K, 10240K, 102400K)

Figure 7: file-by-file comparisons for two major ISVs

The options and "feature enhancements" offered by ISVs vary. These options and feature enhancements are the areas that can be made to benefit the performance offered by today's high-performance tape drives. The major options from the major ISVs that can affect direct tape attach backup are shown in table 4.

Table 4: ISV features to enhance direct attach tape performance

ISV	Feature	Comment
Computer Associates	Image Backup Option ArcServe 2000 has dynamic memory allocation, the performance scales with available memory	Allows a sequential Backup of the Disk, thus improving performance. Single file restore can however take up to several hours. HP Performance Test Lab found dynamic memory allocation could give 10-15% performance improvements.
HP OmniBack II	Supports Concurrency* Supports Raw Disk backup on Unix file systems	Multiple data streams from separate disks or file systems can be run concurrently onto a single tape. Up to 32 streams. Use raw disk backup with large numbers of small files.
Veritas Backup Exec	Veritas VxFS Filesystem is available now for NT (June 2000) and offers the possibility of improved file system performance because the file system can be tuned to RAIDed volumes. In addition, VxFS resolves the defragmentation problem by consolidating free blocks and making small and large files more contiguous. VxFS also offers Storage Checkpoint facilities for improved Backup and Restore and online data backup. This file system enhancement is compatible with all ISV software.	Tunable File system performance and reduced defragmentation should improve backup times with today's high performance tape drives. Backup Exec and NetBackup can take advantage of Storage Checkpoints to enable "snapshot" images of file systems and databases to be made in seconds. By backing up storage checkpoints, less data needs backing up (see glossary).
Veritas NetBackup	FlashBackup Option An image backup technology Supports Concurrency (Multiplexing) of data streams	Aimed at enterprise servers with large numbers of files, such as NFS, email and file servers. Currently available on Solaris and HP-UX servers with UFS,Online JFS. or VxFS file systems. W2K support planned for 2001. Allows direct-attach or over-the-network data streams to be multiplexed onto a single high speed tape device.
Legato	Power Edition (option)	Contains built-in support for 64 parallel data streams (concurrency*) and, when source and target are on the same machine, it optimizes the data path by removing the TCP stack overhead (since all devices are local). Can support throughput on the local server of up to 100 GB per Hour.
Dantz	Architecture improvements for high performance tape drives in Restrospect for Windows 5.2 include:- Tuning internal buffer & OS pipe sizes, and available real memory Dynamically adapting to utilize changing virtual memory Using only the faster synchronization primitives under Windows NT (a pre-cursor for implementing multi-threading) Developing more optimized SCSI transport layers (ASPI) specifically for high performance sequential tape devices.	Dantz seem to be one of the few ISVs who are proactive in understanding the fundamental changes required to stream today's high-performance tape drives.

Performance Tuning on NetWare

The HP Tape Performance Laboratory, discovered during testing of Ultrium Tape drives on NetWare that the Storage Management Services (SMS) architecture used on NetWare 4.x/5.x has performance limitations because of the API interface used by the Backup Application to access the file system. The backup engines have to open files, request namespace, close files etc all of which gives reduced performance. The look ahead algorithms of the various Backup applications in securing access to files all worked slightly differently so the Target service Agents (which the Backup applications generally access the filesystem through) could not optimize access to the file system in any one way. The net result is that on any NetWare system there is a limit of around 6mB/sec today iusin large file transfers.

When HP spoke to Novell on this subject, Novell indicated that for NetWare 6, a completely new API for Backup Applications to access the file systems was in development and would be supported by all major ISVs by Fall 2001.

In the meantime HP recommends the following activities to improve performance on NetWare.

- If you are using CA ArcServe consult the following urls

  http://support.cai.com/techbases/as61/10015.html

  http://support.cai.com/techbases/as61/NAS61519.html

Of particular relevance is the effect of increasing the buffer size available to the Backup Application.

The introduction of ArcServe 7.0 for Netware in early 2001 has improvements to Backup over the Network using improved Netware client agents and also supports concurrency, backing up multiple targets to a single tape.   CA inform us that NetWare & will also have an image Backup utility for NetWare (bypassing the files system) developed by Mid 2001

- If you are using Veritas Backup Exec look at

  http://www.veritas.com/us/products/whitepapers.html

and chose "How to Optimize Backup Exec 8.5 for NetWare"

- Some Applications such as Yosemite Technologies TapeWare

  http://www.tapeware.com

 avoid using the TSA drivers and have their own file access drivers, interfacing into the name spaces and doing direct file I/O using regular file I/O and File Engine Libraries. The HP Tape Performance Lab has proven this to work and have achieved over 13MB/sec on NetWare 5 with TapeWare onto Ultrium 230 using 100MB file sizes.

- Consider using "Image " Backup for NetWare available from Columbia data Products as their SnapBack Product, for more details see..

  http://www.cdp.com/

Future enhancements to improve direct attach tape performance

- Continuous improvements in processor speed together with disk and RAID controller technology will allow higher overall throughput and benefit backup performance.

- The implementation of multi-threading*(see glossary) by major backup ISVs will significantly improve overall throughput by taking advantage of SCSI command queuing within RAID controllers. Dantz and Yosemite Technologies plan to do this in the near term.

- The possibility of building "snapshot" backup technology into the RAID controllers themselves will save disk space and significantly reduce I/O overhead. It will also help to have the tape attached to the RAID controller and able to perform a direct attach tape backup from the RAID without using host resources.

Ways to assess file system performance

It is important to understand the data throughput that your disk subsystem can deliver in order to determine if it is capable of supporting one of today's high performance tape drives. To measure the maximum/typical sustained read performance of a disk and file system subsystem, the following tools are available. Ideally the sample data needs to represent your typical backup data sets.

1. Formal Performance Assessment Tools such as Intel's I/O Meter (see "further information section for url) can analyze RAID performance. This utility needs additional space on the partitions to write test patterns. Typically set the transfer rate at 64 K, a queue depth of 10 and 100% sequential reads to establish the absolute maximum throughput of your RAID subsystem. Use a queue depth of 1 to establish what the backup application performance will be like.

2. Most ISV software will allow you to create a "null" device. This is a bit bucket to sink data into. For example HP OmniBack under NT will allow you to set up a :NUL device. You can select directories to pass to this NUL device and the size/time this takes will give you a measure of the file system performance. Similarly, on Unix you can create a /dev/null device file to accomplish the same thing.

3. Use the performance monitor in NT to assess disk performance. Using the Command prompt in NT you can:-

   Enable Disk Performance monitor by typing: `Diskperf -ye`

   Access performance monitor and use "Physical Disk" to measure Read Mbytes per second by typing: `C: copy <File name>  NUL:`
   (The file should be as representative of your backup data as possible. Unfortunately :NUL is not supported on xcopy, which would allow nested subdirectories to be copied.)

   Look at performance monitor trace to see what disk performance was achieved.

4. Many disk array management tools allow access to performance graphs whilst the RAID is running. Use these to assess read performance and consider copying sample backup files to NULL whilst running the monitor to gauge the throughput versus the file size. Examples are HP NetRAID Utility and Adaptec FAST tool.

5. Use the HP Performance Assessment Tool (see Appendix C). Windows NT and Windows 2000 support is only available at present.

6. Also available for purchase from Intel is the "Ipeak" storage assessment tool. This tool has the advantage of being able to sit in the background and monitor disk activity.

If ISVs do not implement multi-threading in the near future, we will be unable to stream the second-generation Ultrium products.

Direct attach high performance tape - performance tuning summary

DO

- Attach your high performance tape drive to a separate Ultra 2 channel or better, either on the motherboard of the server or by adding extra host bus adapters. HP recommends the Adaptec 29160 range of HBAs for this purpose.

- Use a RAIDED disk subsystem with four or more disks (Ultra 2 10k or better).

- Use the best performance RAID controller available for your system, 100 MHz minimum

- Ensure read cache is enabled on your RAID controller

- Add in extra capability for your system at the design stage to ensure higher performance backup capability
    - Multi processors (two minimum) in line with your application requirements, determined by sizing.
    - System memory 512 MB or above in line with your application requirements, determined by sizing.

- Ensure the data is unfragmented (NT mainly) using inbuilt defragmentation (W2K) or 3rd party solutions (NT4).

- Check with your ISVs if your application can
    - Increase tape block size
    - Increase memory allocation
    - Increase SCSI transfer size above 64K

- Use image backup mode (ISV options) if the system consists of a large number of small files. (< 5MB)

- Use a tape drive with proven data rate matching wherever possible. The drive should have as wide a range as possible and be continuous in its matching capability (as opposed to stepped increments in data rate matching).

- Invest some time in understanding your system's disk performance using tools such as I/O meter or the HP performance assessment tool. If your current system cannot support transfer rates of 30MB/sec you can still use Ultrium technology by choosing the lower performance models indicated in Table 1.

- Do watch out for ISV enhancements in the future such as multi-threading and serverless backup. These will further improve backup performance and reduce loading on system resources.

DON'T

- Connect the tape drive to the same SCSI channel as the disks

- Connect LVDS tape drives to a single-ended SCSI channel or connect a single-ended device to the same SCSI bus as the tape drive. This will severely reduce performance.

- Use a single disk drive

- Use Native Backup applications

- Expect good performance on NetWare until after the release of NetWare 6.x

- Back up lots of small files in file-by-file mode to the high performance tape drive and expect good performance.

- Expect good performance if using agent backup of individual mailboxes or specific database records. It is NOT good for matching the performance of the tape drives. (See Scenario 5.)

Scenario 2: optimizing performance on a network

This section describes the main contributing factors which affect network backup performance, looks at future enhancements and ways to assess file system performance. For a quick summary of the information in this section, see page 23.

The main contributing factors that affect network backup performance to a dedicated backup server are:

- Network Transport, as illustrated in table 5.

Transport	Typical User Data Transfer Rate - Dedicated Backup LAN GB/Hr
10 BaseT	3.2
100 BaseT	32
FDDI	32
100VG Any LAN	32
GigaBit EtherNet (1000BaseT)*	320
Fibre Channel (for comparison)	360

Table 5: different network transports

- Expected throughput from the remote disk subsystems, are they single disks or Raided ?

- Performance capability of the dedicated backup server - operating system, processing power and memory.

- Backup software - can it support concurrency* and does it have intelligent compression software on the clients, for example "agent accelerators" (see below)

- Type of drive - use a tape drive that supports data rate matching to best optimize performance, since it can accommodate fluctuations of speed of the data stream.

"Push" Agents & "Pull" Agents

A "Pull" Agent is an agent that sits on a client machine to service requests from the main backup application running on the dedicated backup server. It responds to *single file* requests.

A "Push*" Agent* on a client machine is one where the Agent can receive multiple file requests and the Agent then supplies all of them. Most agents today are "push agents"

We'll now consider two typical implementations, illustrated in figures 8 and 9. Because of the potential higher volumes of data involved in centralization, dedicated backup servers tend to use libraries where tapes can be automatically changed. In the first example we assume the dedicated backup network is 100 BaseT and already in place. In our second example we assume customers have decided to upgrade their dedicated backup LAN to Gigabit Ethernet technology. The remote servers are of a higher specification in terms of disk performance, which means RAIDed disk subsystems on these servers will be capable of sustaining at least 30 MB/sec.

The new generation of high performance tape drives increase the appeal of moving to a centralized backup scenario to reduce cost and improve manageability.

For a centralized backup environment HP's "rule of thumb" is to always use a dedicated LAN for the backup data traffic. This ensures optimal backup performance.

Example 1: using existing, dedicated 100 BaseT backup LAN

The dedicated backup network is 100 BaseT and already in place. The network performance is adequate in terms of the data rates likely to be coming from non-RAID disk installations on remote servers as long as the backup software supports concurrency.
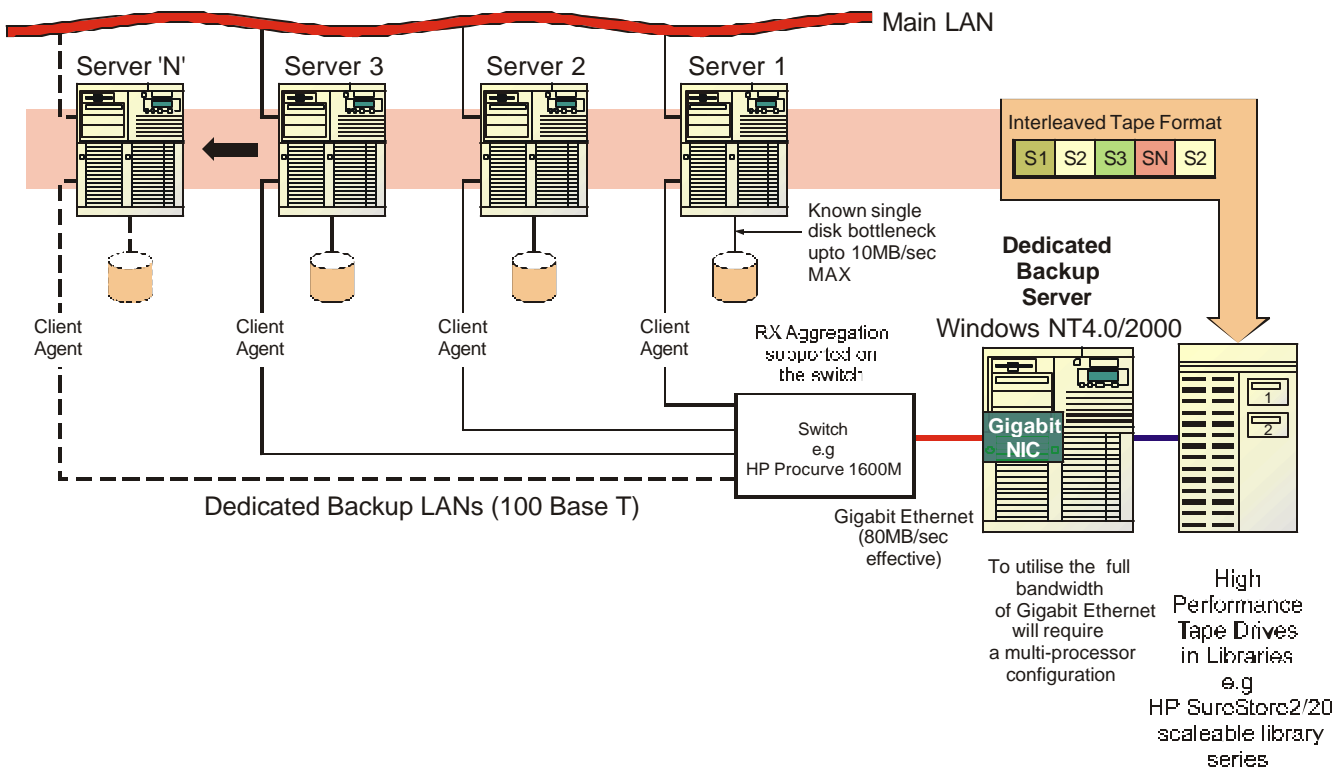

Figure 8: centralized backup using existing dedicated backup LANs, aggregating switches and high performance tape drives

The issue is how to concatenate these "slower" streams into a data stream that can make use of the high performance tape drive. This can be done by using an Rx (receiver) aggregation switch with a Gigabit Ethernet * output, such as the HP Procurve 1600M switch with Gigabit port output module (Fibre or copper).

There are two reasons for not using teamed ethernet into multiple NIC cards in the dedicated backup server

• the additional cost of multiple NIC cards

• the fact that each card would need processing power to maintain the TCP/IP stack, which would reduce processing power to handle the actual backup.

To interleave this data stream onto tape you must use backup software that supports concurrency* (HP), multiplexing (NetBackup) or "parallel streams" (Legato) (see table 6). Such software enables efficient backup and provides the ability to restore a file from an interleaved data set on tape. Improved concentration of data from the client side will also be required, such as Agent Accelerator technology* (Veritas) or "client-side parallelism" (Legato). See the glossary for an explanation of Agent Accelerator technology. The HP Tape Performance Lab found that in an NT/W2K environment the operating system was perfectly capable of sustaining an 80 MB/sec data stream from a Gigabit NIC.

HP Data Points

Table 6 illustrates the results, using two servers "Back to Back" and transferring files over a Gigabit Ethernet link (no discernable difference between W2K and NT4):

	Single Processor (550Mhz)	Dual Processors
Gigabit throughput (file transfers)	56 MB/sec	Full 80 MB/sec
CPU loading	100%	65%

Table 6: comparing gigabit throughput and CPU loading using single and dual processors

The server had 256 MB of memory. A single processor was incapable of utilizing the full gigabit Ethernet capability, however, by using dual processors it was possible to handle the full bandwidth without having high CPU loading.

In the HP Tape Performance Lab a test was set up as shown above using four remote servers with single drive spindles and using concurrency with HP OmniBack II. A dual 667 MHz backup server was able to sustain a system backup rate of over 25 MB/sec (around 90 GB per hour).
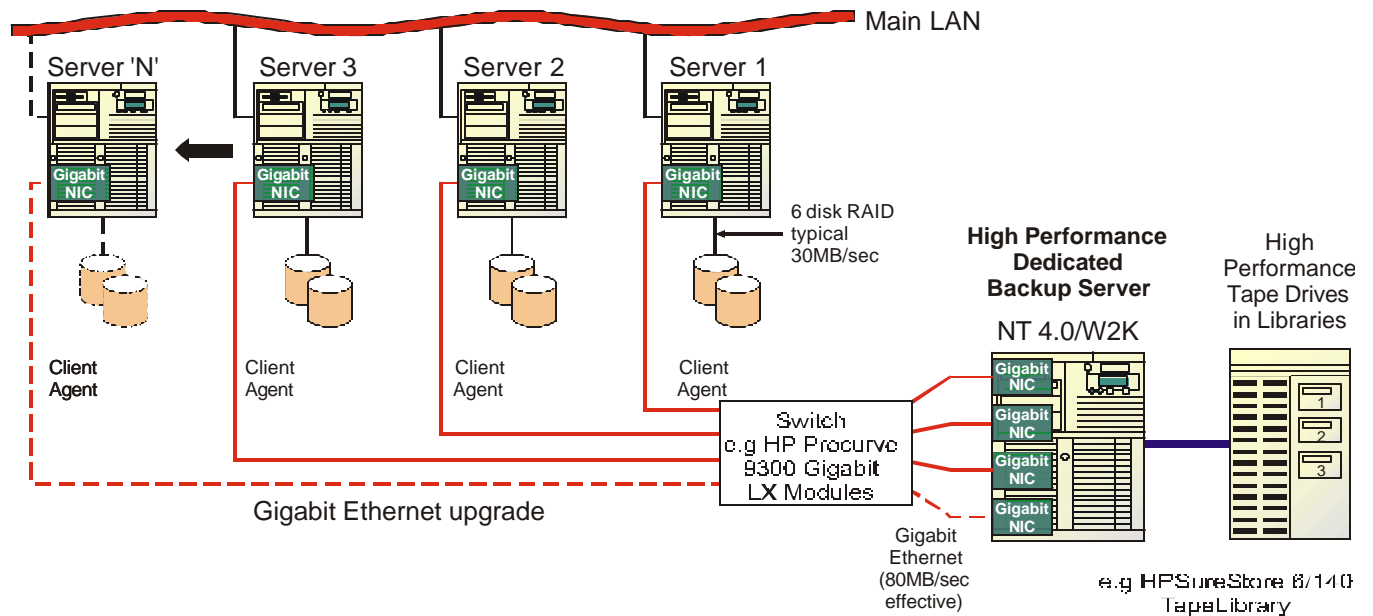
Gigabit Ethernet connections will be required to enable a sustainable data rate to high performance tape drives over a 100 BaseT dedicated backup LAN

Software supporting concurrency will also be required.

Example 2: upgrading to gigabit ethernet technology

We assume that customers have decided to upgrade their dedicated backup LAN to Gigabit Ethernet technology and the remote servers are of a higher specification in terms of disk performance. This means RAIDed disk subsystems on these servers will be capable of sustaining at least 30 MB/sec.

The most likely applications running on these servers will be email or databases. The most popular type of backup for a 24*7 SQL or Exchange server is an online backup to a "disk dump" data file, which is then backed up as part of the daily backup process over the network.

Main LAN

Server 'N'   Server 3   Server 2   Server 1

Gigabit NIC   Gigabit NIC   Gigabit NIC   Gigabit NIC

6 disk RAID
typical
30MB/sec

**High Performance
Dedicated
Backup Server**

NT 4.0/W2K

High
Performance
Tape Drives
in Libraries

Client Agent   Client Agent   Client Agent   Client Agent

Gigabit NIC
Gigabit NIC
Gigabit NIC
Gigabit NIC

Gigabit Ethernet upgrade

Switch
e.g HP Procurve
9300 Gigabit
LX Modules

1
2
3

Gigabit
Ethernet
(80MB/sec
effective)

e.g HP SureStore 6/140
Tape Library

To sustain all the gigabit ethernet streams simultaneously would require a very high performance server
(probably 8-way with large amounts of memory)
Alternatively each remote server could be backed up in sequence to a separate piece of media in the Library.

Figure 9: centralized backup using gigabit ethernet and high performance tape drives

To be able to sustain multiple gigabit input streams simultaneously would require a very high performance server (8 way) with a large amount of memory. Even then the theoretical maximum of the PCI bus backplace (64 bit/66 MHz) is around 528 MB/sec, so around five gigabit NIC cards would be the limit. A more practical solution is to use gigabit Ethernet and back up the remote servers in sequence to a high performance library. Concurrency is not really required here since the remote servers can supply data at a speed capable of streaming today's high performance tape drives.

Table 7: ISV features to enhance network backup performance with high performance tape drives

ISV	Feature	Comment
Computer Associates	"Push Agent Technology" for remote servers.	Allows remote servers to be backed up to a dedicated backup server with reasonable performance. The edition of ArcServe you purchase determines which agents you receive. No concurrency support.
HP OmniBack II	Supports Concurrency over the Network as standard Supports Pull agents on clients. Allows client to pre-compress data before transmitting over slow networks.	Multiple data streams from separate servers can be run concurrently onto a single tape. Recommended interleave is 3 to optimize the restore time.
Veritas Backup Exec	Agent Accelerator Technology for NT, W2K & NW. (an option to enable remote server protection)	Delivers higher throughput over the network than normal push agents by more efficient packaging and compression over the network. (see glossary) No concurrency support.
Veritas NetBackup	Supports Multiplexing Backups and Restores over the network and automatic multi-streaming client sessions from different mount points.	"Multiplexing" is Veritas terminology for concurrency, but it is designed specifically to optimize the usage of high-speed tape devices. Automatic multi-streaming also dramatically increases performance by initiating parallel backup sessions in response to a single backup request from the NetBackup server.
Legato Networker	Standard product supports parallel streams (concurrency) Client Agents support "client side parallelism" & data compression" for improved transfer performance over the network.	Capable of supporting up to 64 parallel streams from clients. Recommended optimal setting for restores is 5.
Dantz	Incremental PLUS™ feature	Delivers faster and more efficient network backups by only backing up new, changed or unique files from each computer.

Ways to Assess Network Performance

A quick and easy way to test your Network performance is to "ftp" some large files from your remote servers onto your dedicated Backup server. "Ftp" has very low overhead and should give an indication of what is possible.

Tools that can be accessed from the World Wide Web include:

- Intel I/O meter - Dynamo utility (downloadable) can be used to assess network performance. It can be accessed from:
  http://developer.intel.com/design/servers/devtools/iometer/News.htm

- Perform3 - Netware utility can be accessed from:
  http://developer.novell.com/research/appnotes/1994/may/04/03.htm

- For Unix try Netperf at http://www.netperf.org

Future Enhancements

- The NDIS5 Microsoft specificaton allows more of the TCP/IP stack checksums to be offloaded into the hardware, (rather than being done in the operating system layers). Providing the gigabit Ethernet cards are able to support NDIS5 in their drivers, this should free up more processing power to enable simultaneous processing of gigabit NIC cards.

An example of this can be found at…

http://www.alacritech.com/home.html

By offloading the TCP/IP stack activities into hardware using specially designed NIC cards, they are claiming an 800% increase in Network performance, and a 70% reduction in Tape Backup time. These products would be an ideal fit with Ultrium Technology for over the network backup.

- ISVs, such as Dantz, have plans to introduce concurrency but WITHOUT the restore overhead in the near future. It will be interesting to see how this is achieved.

Backup over the network with high performance tape drives - performance tuning summary

DO

- Always use a dedicated backup LAN

- Implement a Gigabit Ethernet upgrade to your network, especially on your dedicated backup LAN

- If a total gigabit ethernet upgrade is not possible, consider using an RX aggregation switch for 100baseT with a gigabit ethernet output such as the HP Procurve 1600M and family. Gigabit NIC must be installed on the dedicated backup server.

- For configurations using the slower 100BaseT input streams the backup software must support concurrency in order to stream the drives.

- The dedicated backup server requires enough processing power (processors and memory) to service an incoming data stream of up to 100 MB/sec—typically, you will require multiprocessing and larger (1 GB) of memory. However, the overall cost of the dedicated backup server can be much lower than the cost of an application server with direct attach tape, because the backup server does not need the high performance RAID subsystem. The dedicated backup server MUST have 64bit/66MHz PC slots to accommodate the Gigabit NIC cards and sustain the throughput.

- Agent Accelerators* or "client side parallelism" agents must be loaded on the remote clients to improve the throughput over the network by better utilizing the network bandwidth and providing pre-compressed data if required.

- Use a tape drive with proven data rate matching wherever possible. Th drive should have as wide a range as possible and be continuous in its matching capability as opposed to stepped increments.

DON'T

- Use the same LAN for normal network traffic and backup traffic

- Expect to be able to stream the new high performance tape drives using single, existing Ethernet NIC cards when backing up over the network.

Scenario 3: using high performance tape drives in Storage Area Networks (SAN)

This section describes the benefits of Fibre Channel technology and looks at future enhancements. For a quick summary of the information in this section, see page 27.

The advantages of Fibre Channel technology

The advent of Fibre channel technology with its high bandwidth and scaleable architecture brings a new dimension to server/storage data communications. Fibre channel was invented to provide the highly reliable and high performance network environment required for storage. Fibre channel combines the best of both SCSI and networking, implementing the SCSI common command set over the Fibre channel network protocol layer. Unlike the SCSI architecture, Fibre channel is configurable via switches as a true network (called a Fabric), which enables a robust, reliable, scaleable network for storage. Fibre channel offers the real prospect of application servers being available at maximum performance for 24X7 for two main reasons. The backup process can now take place without using server resources and fibre channel offers the ability to share tape storage in libraries across multiple servers, thereby providing the IT manager with easier manageability and cost effective automation.
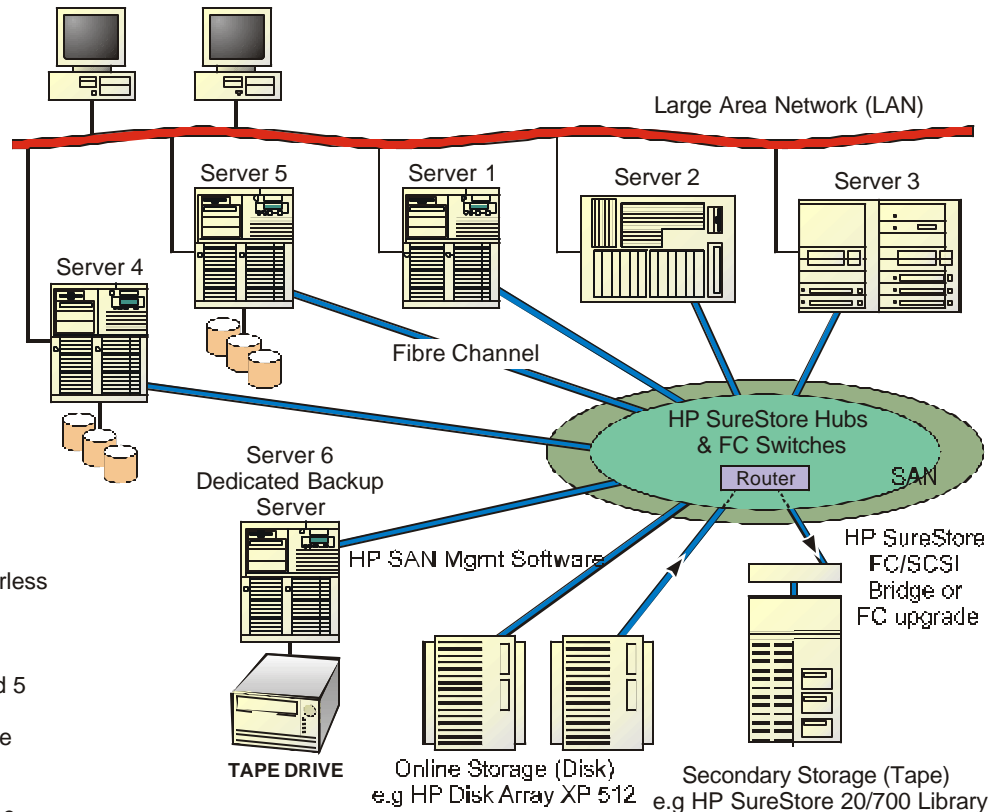
SAN Key Features:

- Shared Storage
- Higher Bandwidth
- Centralized Storage Management
- Scalable Architecture
- Longer Distances (Up to 10 KM)
- Heterogeneous Enviroments

Servers 1,2&3 all share online storage and tape library. Backup could be done in serverless backup mode.

Sever 6 is dedicated a Backup Server to backup Servers 4 and 5 connected to SAN could easily stream todays high performance tape drives.

Both drives and Libraries can be shared amongst servers on the SAN using ISV 'Shared Storage Options'

Large Area Network (LAN)

Server 5   Server 1   Server 2   Server 3

Server 4

Fibre Channel

HP SureStore Hubs & FC Switches

Router   SAN

Server 6 Dedicated Backup Server

HP SAN Mgmt Software

HP SureStore FC/SCSI Bridge or FC upgrade

TAPE DRIVE

Online Storage (Disk) e.g HP Disk Array XP 512

Secondary Storage (Tape) e.g HP SureStore 20/700 Library

Server-less Backup, data does not pass through server but uses 3rd party copy command implemented in switches or tape drives. Transfers upto 80MB/sec could stream multiple high performance drives.

Figure 9: using high performance tape drives in a switched fabric SAN environment

Theoretically Fibre channel* is capable of up to 4000 Mbs (500 MB/sec) although current hardware limitations mean 100 MB/sec is more realistic. 200 MB/sec in full duplex mode is possible if the FC-HBA supports full duplex.

In tests at the HP Tape Performance Lab, lightly loaded servers with FC HBAs installed have achieved up to 95 MB/sec over the SAN infrastructure using Fabric Switch* loop topology and communicating with tape drives using a FC-SCSI bridge.

Fibre channel* offers four connection topologies (see glossary):

- Point to Point generally used for direct attach of disk subsystems to servers (maximum bandwidth = 100 MB/sec

- Arbitrated loop - up to 126 devices, servers, tape, disks, etc connected in a loop. Point to point bandwidth is still up to 100 MB/sec but all other devices on the loop go into repeater mode whilst the two devices are communicating. This is cost effective but not ideal for tape backup, since there will be periods of inactivity as other devices arbitrate for the loop. If disk and tape are on separate "loops", the implementation can support application traffic to the disk subsystem and still stream a high performance tape drive (see glossary).

- Switched Fabric - using fabric switch multiple "paths", all with maximum bandwidth (up to 100 MB/sec), can be established - this is what we mean by scaleable performance of Fibre channel. Fabric switches are expensive. (8 ports = $20K)

- Fabric Loop Attach - an arbitrated loop connected to a fabric switch. This will probably become the most popular means of connecting tape drives, libraries and disk storage. It allows multiple devices to be connected whilst only occupying a single "f-port" on the switched fabric. As with a normal arbitrated loop, the bandwidth inside the loop is limited to 100 MB/sec. Attaching a three or four-drive library via Fabric Loop attach will allow best use of the bandwidth available from a backup perspective.

Implementing a SAN

Using Fibre Channel connect almost guarantees the ability to stream today's high performance tape drives, providing the disk subsystem can deliver data at > 30 MB/sec.

To take advantage of the performance offered by today's high performance tape drives in a SAN environment you should:

1. Ensure all Fibre channel components support class 3 Fibre Channel implementation as the lowest common denominator. Some support class 2 for "confirmed delivery" and some niche applications such as Video transmission support Class1. Class 3 is the preferred implementation for data rates to streaming tape drives since it uses the SCSI protocol itself to confirm delivery.

2. Consider an arbitrated loop implementation if there are only a small number of devices in the SAN loop, to ensure bandwidth is available. For more complex (large numbers of devices many to be shared) implementations, the switched fabric implementation will yield the better performance. Tape devices can be attached via a Fabric Loop to the Fabric switch without impeding performance.

3. Ensure all switches support "non-blocking" switched fabric. This ensures no two input or output ports share any common loops and so limit the throughput.

4. Use the highest performance FC Host Bus adapter that you can afford. This ensures the HBAs can support full duplex mode (reading from FC disk array whilst simultaneously writing to tape device).

5. When backup is instigated, ensure that the server being backed up is at a low usage period on the SAN and NOT trying to service two separate loops on the SAN at the same time.

6. Because SAN technology is not yet "plug and play", we recommend the use of consultants to assist in prototyping the production environment in which the high performance tape drive will be used. This is especially recommended, if you are considering using the "extended copy" feature of switches to allow backup of SAN on-line storage direct to tape without using any server resources. This is known as "serverless backup".

7. The role of ISV software in this environment is focussing on two areas:

   - Allowing multiple servers to "share" online or offline storage to maximize manageability and cost effectiveness. Even if a server goes down the data will still be accessible via another route.

   - Allowing server processing power to focus on the applications they are running, rather than performing the backup - by implementing serverless backup.

A summary of ISV software is shown below in table 8.

A SAN implementation is the most guaranteed way of being able to utilize the performance of today's high performance tape drives (providing the disk throughput is in excess of 30 MB/sec), but unfortunately the most expensive in terms of additional hardware and software and integration of components. As an example, certain serverless backup applications have been shown to reduce CPU loading by 97%, increase I/O capability by 97%, free up 50% of the memory and easily stream a tape drive at over 30 Mbytes per second (108 GB/Hr).

Table 8: ISV SAN implementations for high performance tape drives

ISV	Feature	Comment
Computer Associates	Serverless Backup (option)	Uses CA's Image Backup to take a snapshot of the disk subsystem by using extended copy command in Crosroads/Gadzoox and Pathlight SAN switches. Capable of volume level and file level restores.
HP OmniBack II	Supports a wide range of HP SAN components	Does not yet support serverless backup or shared storage access.
Veritas Backup Exec	Backup Exec Shared Storage Option (SSO) for NT/W2K & NetWare (option)	Allows Libraries to be shared amongst multiple mixed Windows NT, W2K and Netware servers. Marketed as "LAN-Free" Backup.
Veritas NetBackup	NetBackup Shared Storage Option for NT and Unix	Targeted at more mixed enterprise environments. Serverless Backup product planned for Fall 2000
Legato Networker	Serverless backup in the Solaris and HP-UX environments. Needs two additional options to Legato Networker 5.6 - Celestra Power 1.1.2 - Celestra Data Mover	Only copies used files in a file system. Block level increments possible where only changed blocks are copied.

Likely Future Developments

- Over the next few years the high performance tape drives will evolve to have direct Fibre Channel connect interfaces on the drives themselves. This will enable tape drives/libraries to connect more cost effectively to Fibre Channel SANs without the need for costly Fibre Channel to SCSI converters.

- Support of extended copy by Fibre Channel tape drives.

- Increased ISV support for Serverless backup applications.

Read this page for summary recommendations for scenario 3 - high-performance tape drives in a SAN

Optimizing the performance of high performance tape drives in a Storage Area Network (SAN)

DO

- Ensure the source of the data to be backed up can supply the data at a minimum of 30 MB/sec, which probably means a 4-5 disk raid subsystem and 233 MHz or above RAID controller with a minimum 64K Cache.

- Evaluate your current and future likely growth plan. Fibre channel technology is a key enabler to being able to utilize the performance offered by today's high performance tape drives. Consider:

  - FC-Arbitrated Loop implementation* (126 devices max) with full duplex FC HBAs if the SAN is likely only to ever support a shared disk system and maybe two to three servers and a single high performance tape drive. The maximum bandwidth of this system is 200 MB/secs and it can be very cost effective

  - The use of a full switched fabric* implementation. As long as non-blocking switches are used, this can offer very high levels of performance (100 GB/Hr) on multiple simultaneous paths but the switches are expensive. This type of system can support multiple high performance tape drives, all streaming in parallel.

- Use a tape drive with proven data rate matching wherever possible. The drive should have as wide a range as possible and be continuous in its matching capability (as opposed to stepped increments).

- Consider shared storage options, particularly in the switched fabric implementation, to enable cost effective sharing of libraries whilst backing up at maximum performance.

- Use a total SAN solution supplier such as Hewlett Packard to ensure all the various SAN components "plug and play" together.

DON'T

- Place more than three high performance tape drives on a single arbitrated loop or fabric loop. The topology will not be able to stream them all simultaneously.

- Consider implementing a SAN tape backup strategy without using professional consultants, the technology is not yet "plug and play".

Scenario 4: tuning Unix environments

This section describes how the Unix environment differs from the NT environment and describes ways of testing throughput and performance. For a quick summary of the information in this section, see page 30.

The Unix environment

The Unix environment differs from the NT environment in terms of tape backup performance in the following ways.

- In-built unix backup applications will give very poor backup performance because the coupling of the data source to tape drive is not optimal. Unix in-built backup applications tend to use small block sizes (e.g. Tar 10k, cpio 5k) which limit performance. Third party applications offer much better data flow control.

- All third party backup applications use the Unix inbuilt SCSI tape driver but wrap around it much more efficient data flow. The driver must be set to fixed or variable mode to suit the application that is using it. The major Unix backup applications (OmniBack, NetBackup, Legato) all support concurrency* allowing efficient parallel paths from different file systems to be run simultaneously and multiplexed onto tape. Using 3rd Party ISV software is the only real practical way to get good tape performance from Unix systems.

- Unix offers raw disk backup, which will be high performance but suffers from an "all or nothing" restore capability.

- Unix file systems suffer much less fragmentation than NTFS filesystems, 2 to 3% typically, meaning files tend to be much more contiguous on disk and therefore can be read more sequentially - which in turn is better for backup performance.

- For an example of Oracle database backup tuning for Ultrium technology under HP-UX, please consult the associated white paper "Performance tuning Oracle database backup on HP-UX using Ultrium tape technology".

Table 9 summarizes available ISV software features to enable improved performance in unix environments.

Table 9: Unix recommendations

Linux	SCO & UnixWare	HP-UX	Solaris
Has a 32K "soft" SCSI transfer limit. This can be increased by editing `st_options.h` We recommend this is edited to 64 K and the backup application blocksize is also set to 64 K for Ultrium. There are no ISV image options currently on Linux. Consider: - The Enhanced Software Technologies (EST) BRU product (Backup & Restore for Unix) - Yosemite Technologies "TapeWare", which supports limited multi-threading at a directory level - ArcServe Backup for Linux	Both these OS can be configured into fixed block size operation. Use the mtsetblk <> command to change block size. <> = 0 variable = X tape block size There are no ISV image options currently on SCO/UnixWare. Consider: - The Enhanced Software Technologies (EST) BRU product (Backup & Restore for Unix)	Has an additional Unix command called fbackup (fastbackup). Change the configuration file, as follows: Blocksperrecord = 128 Checkpointfrequency = 128 Readerprocess = 4 Filespersm upto 500 Use HP OmniBack II with concurrency support to get good performance OR Use Legato Networker with concurrency and optimized architecture (Power Edition only) OR Use Veritas NetBackup with concurrency. The Veritas NetBackup FlashBackup option is also available - for large database and file server image backup. This is guaranteed to give good performance and single file restore capability.	Can be changed to fixed mode by editing `st.conf` OmniBack II with Concurrency is the best solution. You can also use: - Veritas NetBackup with Concurrency - Veritas NetBackup FlashBackup Option for large database and file server image backup. This is-guaranteed to give good performance. - Legato Networker with concurrency and optimized architecture (Power Edition only)

Ways of testing RAID throughput to tape on Unix systems

To determine how fast the disk subsystem can supply data try:

- For "raw disk" partitions:
  ```
  time dd if=<rawdisk> of=/dev/null bs=64k count=<Nos blocks>
  ```
  where `Nos blocks` is the number of disk blocks to transfer (e.g. 65536).

- For file by file
  ```
  time tar cvf /dev/null <files>
  ```
  where `files` is the mount point for the RAIDed disk subsystem (e.g. `/mnt`)

  `tar` will create a default 10K tape block size

  OR
  ```
  time find .||cpio –ocB||dd ibs=5120 obs=64k of=<tape dev>
  ```
  will allow you to test throughput with larger block sizes.

  Alternatively:
  ```
  time tar cbf N /dev/null <files>
  ```
  will allow "N" to specify the tape block size and assess its impact on performance. "N" is the number of 512 byte blocks setting the tape block size.

Other Unix Performance Analysis Tools

- Glance - for HP-UX

- Top - for HP_UX

- Iostat - general input/output assessment tool (supplied with the operating system)

- A program called "ASIDD" is available from HP Data Protection Division Bristol supported on HP-UX, Sun Solaris, Linux and Free BSD to assess RAID performance. Available from http://www.hp.com/support/ultriumPAT

- LM Bench is a useful Unix tool for assessing I/O rates for disk transfers http://www.bitmover.com/lmbench/lmbench.html

Read this page for summary recommendations for scenario 4 - high-performance tape drives in a Unix Environment

Optimizing the performance of high performance tape drives in a Unix Environment

DO

- Ensure the source of the data to be backed up can supply the data at a minimum of 30 MB/sec, which probably means a 4 to 5 disk raid subsystem and 233 MHz or above RAID controller with a minimum 64K Cache. Use the Utilities mentioned above to determine this.

- Use a 3rd Party Backup Application - Tar & Cpio will give poor performance.

- Structure your filesystem so it can make use of the concurrency* feature offered in almost all Unix 3rd party backup applications.

- If you have lots of small files on a Solaris or HP-UX system consider using the FlashBackup option from Veritas.

- Some Applications (e.g. databases) need raw disk format - this offers the ability to do raw disk partition backup which will give higher performance but only full restores are possible in this mode.

- Use a tape drive that supports data rate matching to optimize performance.

DON'T

- Use the in-built backup applications.

Scenario 5: using high performance tape drives in databases and messaging systems

A growing number of mission critical applications, such as Email, databases, and ERP have a backup utility within the applications themselves. This section will investigate how the in-built applications and third party supplied "Agents" that communicate with the applications perform with today's high performance tape drives. For a quick summary of the information in this section, see page 39.

Backup and transaction processing

What is evident in this area is there are many ways to implement a backup strategy, and each has its own positive and negative attributes. We will use SQL server and Exchange server to illustrate the examples in this discussion. The user must decide on the compromise of speed over flexibility, and so on, depending on his unique circumstances.

What makes backup more challenging in these environments is that the application must be kept online to users (performing transactions) whilst the backup is taking place. Hence the application vendors have prepared special interfaces (APIs) to allow third party ISV software to put the databases into special modes whilst the data is backed up and the transactions continue. To optimize performance, the transactions are not committed immediately to the database but are held in "transaction" logs. The backup strategy must, therefore, include the transaction logs as well as the database. Should a total restore to a point in time be required, the database & transaction logs are recovered and then the transaction logs are executed to bring the application back to an integral condition at a given time.
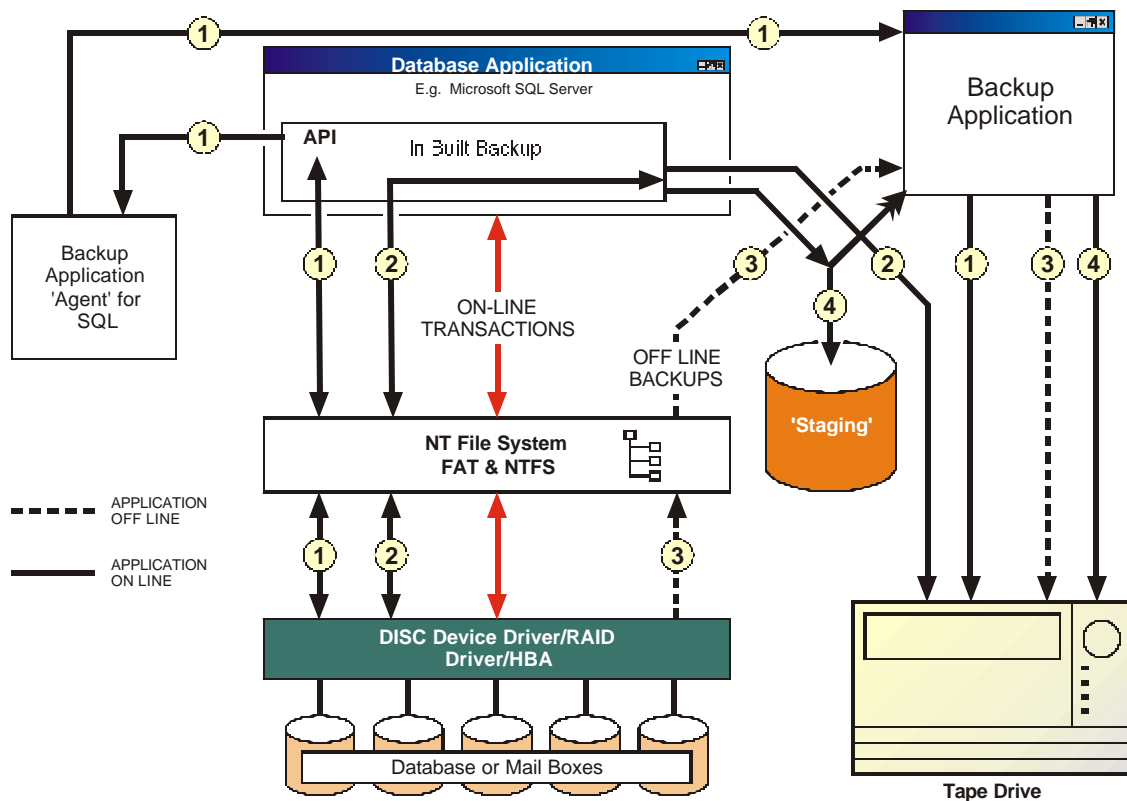

Figure 10: SQL backup options

Four ways to back up an SQL server!

The schematic (figure 10) shows possible paths for backing up an SQL server.

- Method 1 uses a 3rd party backup "agent" to talk to the SQL "API" within the inbuilt backup of SQL server. This allows access to the data within the "online" database for the 3rd party backup application, and allows more granular backup by being able to back up individual reports or mailboxes.

- Method 2 just uses the in-built backup application within SQL server itself. This has to back up the whole database.

- Method 3 performs the backup with SQL Server "off-line" by using a 3rd party backup application.

- Method 4 performs a "staged" backup of SQL server to a disk file (not to tape) using the in-built backup application. This is later backed up to tape as part of a network backup process.
  One benefit of using the in-built backup application is that it can make use of the multi-threading capabilities* within the application, because it is integrated with the application. We know this is greatly advantageous to backup performance. The use of 3rd party agents/backup software, however, allows increased flexibility (e.g. on Exchange the agent can back up individual mailboxes, on SQL it can back up specific records) and also allows the applications to be backed up over the network, managed from a remote terminal.

In terms of file sizes, the transaction logs in Exchange are fixed at 5 MB and the information store (database) is one large contiguous file (consistent with good backup performance) providing sufficient disk capacity is available. In SQL, the transaction logs grow in size with transactions and the committed log entries are cleared after the backup. The main database is a single large file - again conducive to a fast backup as long as disk space is not short, which is when fragmentation occurs. Microsoft strongly recommends that the transaction logs are on a different disk subsystem to the information store/database. Depending on the size of the Exchange or SQL server the backup strategy will include full online backups run as frequently as possible (no less than once per day) and incremental/differential backups (only available through online backup) used at regular periods throughout the day. After a full online backup is complete it triggers the deletion of the committed transaction logs and so the process sets the context for the next full backup.

The HP Tape Performance Lab found SQL 7.0 has significant improvements over 6.5 in terms of Backup/Restore performance

Use SQL 7.0 for improved backup and restore performance.

- SQL 7.0 has a second (improved) API called VDI (virtual device interface), which allows data from disk to go straight into backup application memory space without having to be held in temporary memory space.

- In the HP Tape Performance Lab testing we found that although 6.5 used multi-threading to improve backup performance, 7.0 uses more contiguous (sequential) data on disk - on our test system, both resulted in transfer rates of 30 MB/secs using Ultrium Technology. Because of this, HP now believes a 3 or 4 disk RAID 5 subsystem with high performance RAID controller would be capable of streaming today's high performance tape drives using SQL 7.0 with Agent/In-built backup

- SQL 7.0 supports differential backups of the transaction logs, allowing much more frequent backups and more flexibility in point of time recovery.

- SQL7.0 allows the database to be dynamically recreated on full restore, whereas with 6.5 you had to create "space" for the database to be restored into.

Table 10: advantages and disadvantages of different backup strategies

Method	+ attributes	- attributes	Performance implication with high performance tape drives
Use in-built utility to create Disk Dump of main database with database on-line. (SQL only). This process is called "staging".	Fast - minimum disruption Disk dump of database is backed up as part of normal "file server" backup routine. Only backs up used records. Incremental and differential backups possible. In our investigations this was the most popular option used by Database Administrators.	Uses additional Physical disk space, may be impractical on very large databases.	If tape attached to the application server itself, most backup applications will give *GOOD* performance, because the RAID performance is good and the file is large and contiguous. Backup over the Network of the "dump" files - performance will depend on network topology and dedicated backup server specification. *Follow the guidelines for network backup performance.*
Use In-Built Backup to backup direct to tape attached to the application server. (SQL & Exchange) Database on-line	Cost effective, no additional backup software required. Allows incremental and differential backup (SQL 7.0 only). Only backs up used records not empty records.(SQL only)	Limited manageability - cannot back up individual records or mailboxes - cannot monitor from a central console	Utilizes multi-threading capabilities within application which gives *GOOD* backup performance.(SQL) Exchange. When loaded, modifies NT Backup to provide an in-built backup capability.
Use Operating system and 3rd party Backup application with Database "offline" - full server backup Operating system + Applications + Database using Disaster Recovery options from ISVs.	Provides total peace of mind - bare metal server recovery. Should be used regularly as well as "on-line backups" to enable maximum uptime. Especially easy if the tape drive supports HP One Button Disaster Recovery* Available on HP Ultrium tape technology from March 2001.	Can only be done by taking the application offline. Backs up all database includung empty records.	With a reasonable performance RAID sub-system and relatively unfragmented disc, performance should be *GOOD.* *Look at tuning tips for direct attach Tape Drives.*

continued

Method	+ attributes	- attributes	Performance implication with high performance tape drives
Use 3rd party Agent to access the database. Database on-line	Efficient. Only backs up records which are used. Allows incremental and differential backup Allows individual mailbox backup (Exchange only) or records (SQL) In SQL 7.0 a new API interface called VDI (Virtual Direct Interface) allows data to be taken direct from disk and placed directly into the backup application buffer, reducing CPU load and vastly improving performance. Because Agent is part of a 3rd party application the backup can be Managed remotely and integrated into a Network Backup strategy if required. For Exchange - a 3rd party Agent is the only way to backup individual mailboxes.	Additional Software cost, but this may be justified by the increase in flexibility of backup and performance (if VDI interface on SQL 7 is used).	Direct Attach - full database/information store performance will be **VERY GOOD**, since SQL implements multi-threading on the transaction processing and the backup application makes use of this. SQL 7.0 is further improved by the more sequential nature of the data being backed up. Using VDI interface on SQL will be *VERY GOOD* HP noticed a marked difference in the processor usage by different Agents. *The mailbox level backup will significantly slow down the backup process* due to the requirement that Exchange server "walks-through" each mailbox to copy the data out of the information store. Similarly record by record backup of SQL. *POOR PERFORMANCE* *Drives with Data Rate Matching will perform best here.*

HP Data Points in SQL performance

Table 11 illustrates the results from the HP Tape Performance Lab, using an HP LT6000 (single 550MHz), 8 disk RAIID, HP NetRaid 4M 233MHz RAID controller 1GB SQL 7.0 database with data that was 3:1 compressible. The queue depth and max disk throughput were monitored using the Intel "ipeak" utility.

Table 11: SQL performance results

Method	Disk Access Size	Queue Depth	Max throughput from Disk	Max Tape Throughput @3:1	Actual achieved throughput
SQL Server 7.0 Native	8K-600K	1-5	60MBsec	45MB/sec	30MB/sec (stable)
OmniBack II SQL Server Agent	64K	1	45MB/sec	45MB/sec	25MB/sec (very peaky)
ArcServe IT SQL Server Agent	64K	1	45MB/sec	45MB/sec	25MB/sec (stable)
OmniBack II VDI with 260K block size	260K	1-5	54.5MB/sec	45MB/sec	36-40MB/sec

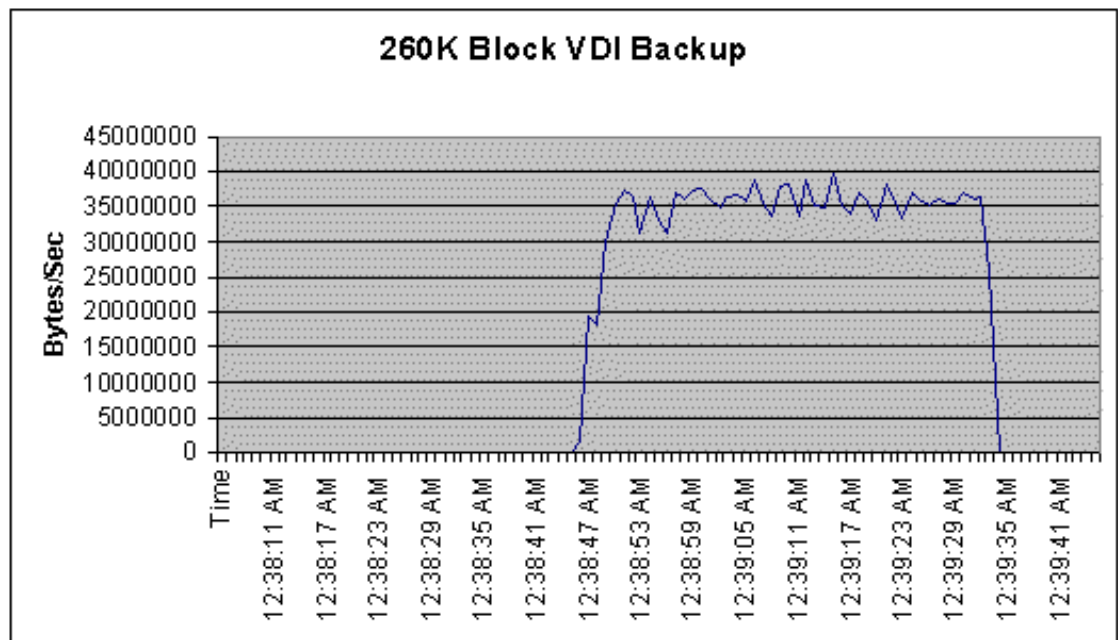Figure 11 illustrates the performance with OmniBack II VDI.

Figure 11: SQL Server 7.0 Backup with OmniBack II using VDI interface

Table 12: ISV features to enhance SQL and Exchange performance

ISV	Feature	Comment
Computer Associates	Exchange and SQL Agent options. DR Option with support for HP One- Button Disaster Recovery (OBDR).	Wide range of approaches. CA Exchange Agent allows individual mailbox backup.
HP OmniBack II	Exchange and SQL Agents. Uses VDI interface for SQL 7.0 Agent, which is very fast (version 3.5 onwards). DR option with support for HP One-Button Disaster Recovery (OBDR).	Setting a new level in SQL Agent Online backup performance. At the time of writing this product is attempting to prove a 1 terabyte per hour backup capability with SQL 2000 at the Microsoft Labs in Redmond US. Only supports full mailbox database backup on Exchange.
Veritas Backup Exec	Exchange & SQL Agents support individual mailbox and tablespace (6.5) or filegroup (7.0). Particularly good Exchange Agent co-developed with Microsoft, Much improved performance from version V7.3 onwards. IDR Option with support for HP One- Button Disaster Recovery (OBDR)	Excellent white paper on Exchange Backup Strategy (see "more information" section of this document). Exchange 2000 and SQL2000 support planed for Q4-00 IDR Option to support HP One Button Disaster Recovery (OBDR) on W2K availble in Q4-00
Veritas NetBackup	Veritas NetBackup module for SQL Server, Good graphical interface especially for restoring with simple "point and click maneuvers". Support for tablespaces, filegroups, transaction logs and differntial database backups. Veritas NetBackup for MS Exchange supports all the Exchange backup types - full backup, cumulative, copy, differential and incremental as well as brick level backup to allow recovery of individual mailboxes. Good graphical interface.	Uses the VDI API interface to SQL 7.0 for high performance. Supports concurrency on tape devices but with the new high performance tape drive, ability to stream multiple drives is unproven. Also allows performance tuning through transfer size and buffer usage. In Dec1999, this product was quoted in the SQL server magazine as offering "fastest combined backup and restore speed for SQL server".
Legato Networker	Supports "Staging" Sold as Legato Business Suite Modules in Network and Power editions. Available for both SQL & Exchange.	SQL & Exchange - supports up to 32 parallel data streams and 16 concurrent tape devices. Uses API "named pipes in SQL 6.5. In theory the ability to handle 32 parallel data streams should give this product a distinct performance advantage for Exchange. For SQL the product does not yet seem to support the VDI API interface for SQL 7.0.
Dantz	Currently developing a VDI API interface to SQL 7.0 The Dantz Microsoft Exchange Agent is free and allows the Exchange backup to be staged to a disk file, and backed up to tape at maximum speed (since no load transactions) at a later time. DR support for HP One-Button Disaster Recovery (OBDR).	Will give improved SQL 7.0 backup performance. Staging on Exchange is a good idea.

Differences between SQL Backup and Exchange Backup Performance

When Microsoft Exchange server is loaded it automatically creates an extension to "NT Backup" which allows NT Backup to act as the in-built backup for Exchange Server. Hence the NT Explorer tree is extended to show the Exchange database and NT Backup can be used to back up an open Exchange Mail server database.

Behind the scenes NT Backup for Exchange Server is actually talking to APIs within Exchange Server, and these APIs can be used by 3rd Party applications. Exchange Agents from 3rd parties talk to different APIs depending on whether the full database or individual mailboxes are to be backed up.

When Exchange mailboxes are created they are allocated a certain size. Some of the available space is for actual mail messages and other space is "empty". In terms of backup performance the actual data is already compressed so the native transfer rate of the tape drive is rarely exceeded. The "space area", however, is not compressed but contains no useful data, and is backed up at increased speeds.
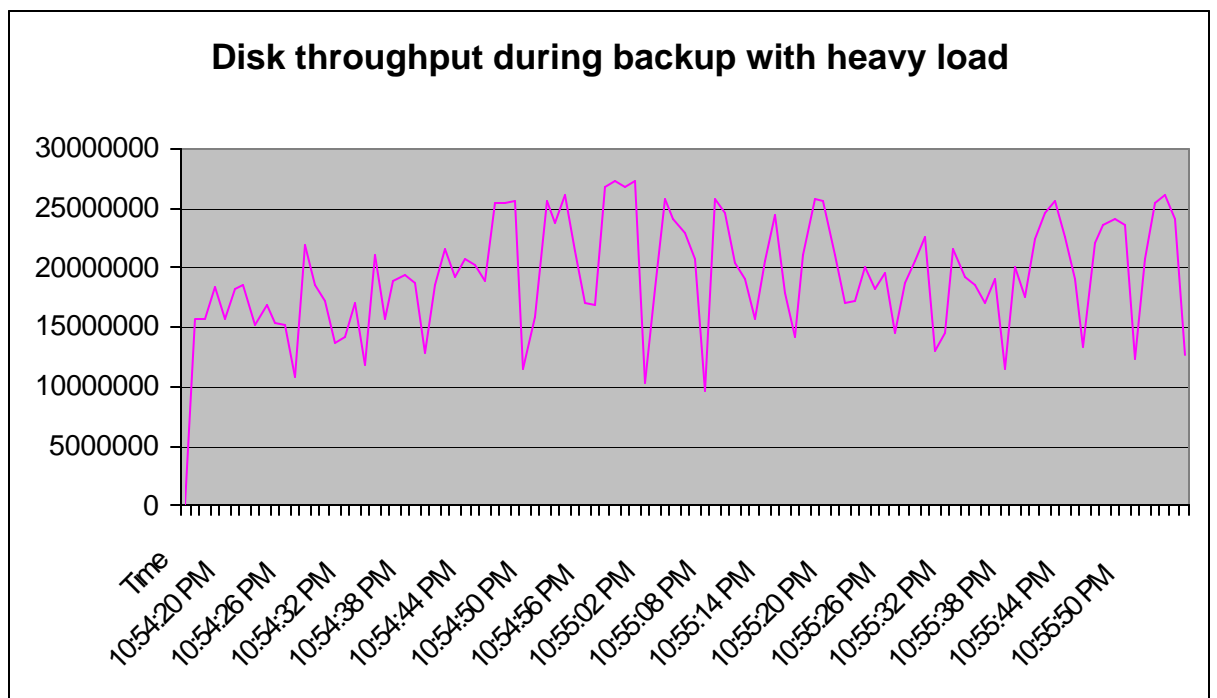

Figure 12: disk throughput during Exchange backup with heavy transaction load

Figure 12 shows HP OmniBack backing up a 300-user Exchange Server. The Exchange server is also running Microsoft LoadSIM to simulate real world usage. Each mailbox is 10 MB in size but has only 1 MB of mail in it. Note how the backup rate fluctuates between 15 MB/sec (compressed data) and 25 MB/sec "space in the mailbox". There are some occasional dips to 10 MB/sec caused by the loading, and it is here that drives that support data rate matching will have a performance advantage over drives that do not.

Some IT departments tend to back up the individual mailboxes of senior people within the company. Unfortunately, because of the way the API has to walk through the Exchange database, this leads to very poor backup performance as illustrated in Figure 13, using CA's ArcServer Agent for Exchange to back up individual mailboxes.
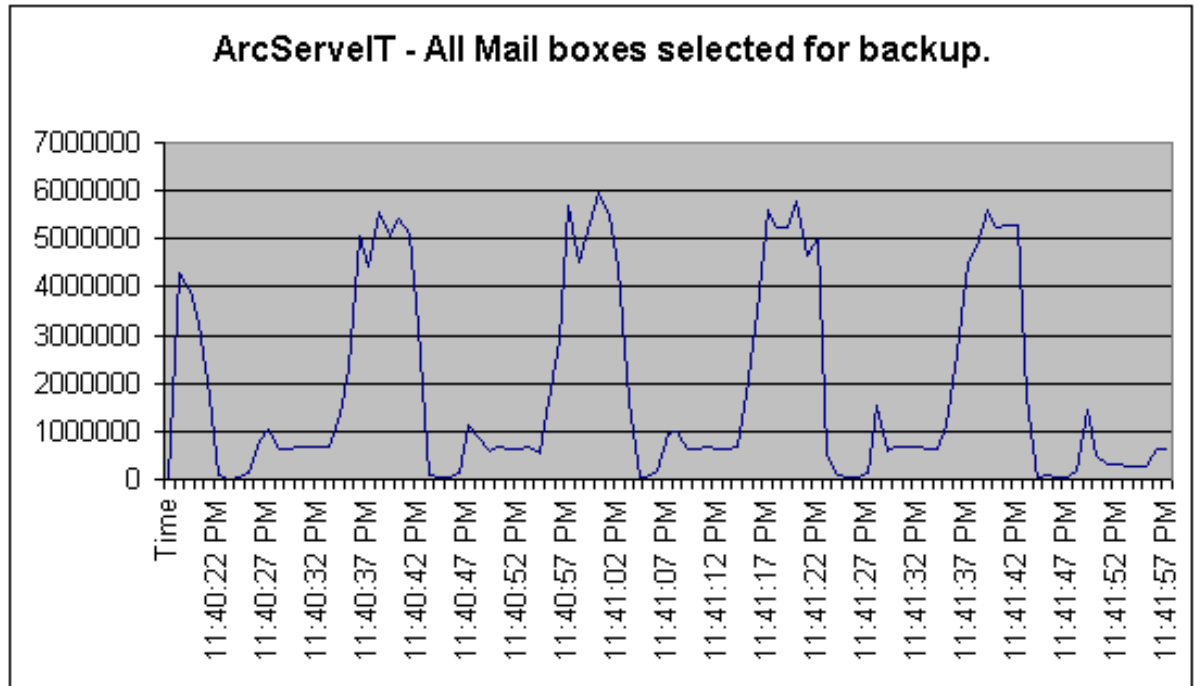

Figure 13: the effect on disk throughput of selecting all mailboxes

Performance averages at about only 2.5 MB/sec, partly because of the searching for the individual mailboxes, but also because in this method the disk reads are only 4k in size, even though the 8-disk RAID system is high performance. With such small read sizes the disk subsystem cannot deliver more that 4-5 MB/sec. The limited use of multi-threading in Exchange Server also restricts performance.

So for Exchange servers the capacity of the new generation of high performance tape drives is a very welcome development (with mail capacities rising exponentially) but because of the nature of the data within Exchange 15 MB/sec (54 GB/Hr) is probably the best that can be achieved. The use of staging to disk files and backup at a later date will prevent loading Exchange loading from further affecting backup performance.

Maximizing the performance of today's high performance tape drives in a database applications

DO

- Ensure the source of the data to be backed up can supply the data at a minimum of 30 MB/sec, which means a 4--5 DiskRAID subsystem and 233 MHz or above RAID controller with minimum 64K Cache.

- Consider "staging" the backup to a disk file and backing up as part of a more general network backup to alleviate the affect of system transactions whilst backing up.

- Use SQL 7.0 or Exchange 2000 as the backup/restore features in these applications are particularly well enhanced to provide better overall backup performance. With SQL Server 7.0 the data itself is stored more sequentially on disk improving the application and backup performance. Restore is also much improved on SQL 7.0.

- Because 3rd party ISV software generally gives more flexibility in backup and restore operations, choose an ISV who supports the faster API interface called VDI within SQL. The HP Tape Performance Test Lab found this could improve performance by as much as 40% and the CPU usage was as low as 20%.

- Deep databases (few columns, many rows) tend to give better backup performance than a wide database (many columns, fewer rows). Deep databases make better use of the paging features in SQL. Reading one page gives many rows and hence reduces overhead, which improves backup performance. The HP Tape Performance Lab saw a 50% improvement in backup performance when comparing deep databases with wide.

- Use a tape drive that supports data rate matching to allow maximum performance even when the system is under heavy load.

- Use a tape drive that supports HP One-Button Disaster recovery to enable fast recovery from a total system disaster.

DON'T

- Back up at peak transaction times.

- Expect record by record or mailbox by mailbox backup to give good performance. It won't.

Summary and Conclusion

New tape technologies, such as Ultrium, are setting a new paradigm in tape backup performance. Making optimum use of this performance is complex, because it involves many interactions of hardware, operating system, backup applications and network topologies. As systems are designed for specific applications, systems analysts must now think of backup performance at the system design stage.

This white paper has shown the various techniques and requirements to support these high performance tape drives in a number of different scenarios:

- drive attached directly to the server

- backup over the network

- drives connected to a Storage Area Network

- drives in Unix environments

- drives used for database backup

A summary of the conclusions and ISV performance features are shown in Appendix A and B.

In all these scenarios, HP is one of the few companies who can supply a total solution, be it in terms of:

- Servers - HP Netservers

- Storage - RAID subsystems

- Storage - Ultrium Tape Technology

- Networking Infrastructure (including Gigabit Ethernet)

- SAN components (HBAs, Fabric Switches etc)

- HP OmniBack NT & Unix Backup Application

- HP Consulting Services

- HP Service & Support

The latest information about HP tape products is available on the World Wide Web at http://www.hp.com/go/tape

For the latest hardware connectivity and software compatibility for HP Surestore products consult http://www.hp.com/go/connect

Linear Tape-Open, LTO, LTO Logo, Ultrium and Ultrium Logo are U.S. trademarks of HP, IBM and Seagate.

All brand and product names are trademarks or registered trademarks of their respective companies.

Information is subject to change without notice.

Copyright ©2000 Hewlett-Packard Company

Further Information

Table 13: useful reference information

URL	Contains:
http://www.hp.com/storage	The latest information about HP SureStore tape products
http://www.hp.com/go/connect	Hardware and Software Compatibility Guides - an up-to-date list of hardware platforms and Independent Software Vendors (ISV) who support HP High Performance Ultrium Tape Drives.
http://www.diskeeper.com	Disk De-Fragmentation Tools
http://www.gigabit.ethernet.org	Gigabit Ethernet
http://www.fibrechannel.co.uk	Fibre Channel
http://www.ultrium.com http://www.hp.com/tape/ultrium	Ultrium Tape Technology
http://www.cai.com/	Computer Associates
http://www.veritas.com/	Veritas Backup Exec and Veritas NetBackup
http://www.legato.com	Legato Networker & Celestra
http://www.openview.com/OVProduct/omniback.htm	HP OmniBack
http://www.dantz.com	Dantz
http://www.raid-advisory.com	RAID Books
http://www.mylex.com http://www.dpt.com	RAID Vendors All server vendors also support their own built in RAID controllers
http://netserver.hp.com/netserver/	HP Netservers NT/2000
http://www.hp.com/rnd/index.htm	HP Networking
http://www.enterprisestorage.hp.com/	HP SAN
http://developer.intel.com/design/servers/devtools/iometer/News.htm	Intel I/O meter performance assessment tool
http://www.hp.com/support/ultriumPAT	HP performance assessment tools
http://www.veritas.com/us/products/whitepapers.shtml	Planning a Microsoft Exchange server Data Protection strategy using Backup Exec 8.0 for NT/2000

Appendix A - server-based backup using high performance tape drives on an NT platform

Type of System	Typical HW Configuration	Hardware Considerations for best tape backup performance	Operating System & Application Parameters	Backup Application Parameters	What is Achievable today	Possible Future Enhancements	Recommendation
Entry Level File & Print Server NO RAID	Single Processor Non-RAID Disk subsystem	Place Tape Drive on separate HBA	Ensure Drive defragmented at regular intervals. File/Print servers can have a large number of small files.	Use recommended Tape Block Size Consider increasing SCSI transfer size above 64K	Dependent on File Sizes but around 5-10MB/sec with File By File backup of small files on Ultra 2 disks.	Multi-threading will help File by File Backup.	For Maximum performance consider using Image Backup Option. OR Consider using an Entry Level) tape drive from the same format family, such as the Ultrium Entry Level product with 7.5 MB/sec native transfer rate.
Messaging or Email Server BUDGET RAID (ROMB)	Dual Processor adequate. Minimum 512KB SDRAM) RAID5 for availability of message store. & Database 6 to 18 disks typically	RAID tuned to application. Experiment with RAID cache settings to maximize disk I/O Stripe Size typically 64K Tape drive on separate HBA For Network backup use Gigabit Ethernet backup LAN.	Ensure Drive defragmented at regular intervals. Transaction Logs held on separate disks to database/info store.	For SQL, NetBackup and OmniBack will give best SQL Backup because they support the VDI API interface. For remote backup consider the use of Agent accelerators and software that supports disk "staging"	Direct Attach tape has been shown to be able to deliver around 30MB/sec in SQL and upto 15 MB/sec on Exchange.	SQL server 2000 is rumored to have further improvements in Backup speed.	MS recommended backup strategy is to back up regularly on-line In-built backup (in SQL/Exchange) will give adequate performance. Agent backup gives more granularity and good backup performance EXCEPT for mailbox by mailbox or record by record backup, which will be slow.

| Data Warehouse or Large SQL Database.

E-Commerce

HIGH PERFORMANCE ADD-IN RAID | Compute Intensive 4 way is Medium.

Probable server to Disk Fibre channel connection.

24-72 drives RAID5 with mirroring.

Unlikely to back up large volumes over the Network | Use highest performance Disk RAID that is available.

High memory Requirements

Ideal candidate for SAN solution probably using switched Fabric Topology for maximum bandwidth.

Consultation with vendor recommended to ensure full plug & play of devices/switches etc. | Mixed NT/Unix environment

Unix more prevalent in this area. HP-UX, Solaris, AIX.

SAP front end.

Clustering also likely because of high-availability requirement. | HP OmniBack, NetBackup and Legato are the key players in this mixed environment space.

These applications allow concurrency and memory configuration to help improve throughput | Using an arbitrated loop SAN and full duplex HBAs will give the ability to support data rates of around 240 GB per Hour

On switched Fabric SAN then performance is limited only by the number of tape drives and Fabric Loops you can afford! | Serverless Backup using Extended copy command in Switches and drives (future) being pioneered in these environments because of high availability at maximum performance being a prime requisite of this type of system.

CA & Legato support at present. | If your data backup requirement requires rates of > 200GB per hour with low impact on transaction performance then SAN is the solution.

The backup performance is STILL highly dependent on the RAID performance but in these types of systems the online storage may already be transitioning onto the SAN so the tape migration is a natural progression.

Multiple tape drives in "shared" libraries is the most likely implementation here. |

Appendix B   ISV Performance Features for High Performance Tape Drives

	File by File Backup	Sequential Backup (Image)	Exchange Backup Agent	SQL Backup	Over the Network Backup	SAN implementations	Other Features
Computer Associates	✓memory tuning possible (see glossary)	✓ sold as separate option	✓ also has Agent for non-stop Exchange	✓ also has Agent for non-stop SQL	✓ Uses push agents	✓Announced May 2000. Limited Ext Copy switch support at present.	
Veritas Backup Exec	✓SCSI transfer size tunable (see glossary) Buffer Size and Block size tunable	☒ not available	✓ particularly well integrated and supports individual mailbox backup.	✓ supports SQL 6.5 and 7.0 Also can be used to backup SAP on SQL	✓ has optional Agent Accelerators to optimize network bandwidth for backup	✓ offer shared storage option (SSO) for Library sharing amongst NT/W2K and NetWare servers in the same SAN	
Veritas NetBackup	✓ Also supports checkpoint technology when used with Veritas VXFS filesystem	✓Use Flashbackup Option (Solaris/HP-UX only)	✓ supports individual mailbox backup.	✓Uses VDI (SQL7.0) interface for improved performance. Uses named pipes for SQL 6.5	✓ has optional Network Bandwidth Throttling by multiplexing remote clients' data onto tape.	✓ offer shared storage option (SSO) for Library sharing amongst servers on NT and Unix	Also supports checkpoint technology* when used with Veritas VXFS filesystem Supports HP's ZERO downtime backup solution
Legato	✓power edition has an optimized architecture for high speed tape devices and supports up to 64 parallel streams simultaneously	☒ not available	✓	✓	✓ "client-side parallelism " or Agent accelerator to optimize network bandwidth	✓ Celestra Power and Celestra Data Mover options allow Legato to perform server less backup with specific named switches .	
HP OmniBack	✓ supports concurrency (SCSI transfer size tunable (see glossary)	✓supports raw disk backup on unix filesystems	✓	✓Uses VDI (SQL7.0) interface for improved performance	✓ Uses push agents	✓SAN backup supported. No server less backup or Library sharing support at present.	Support s HP's ZERO downtime backup solution
Dantz	✓ Retrospect 5.2 for Windows has memory/threading and driver enhancements for high performance tape.	☒ not available	✓ Supplied free - supports staging.	✓ VDI version under development	✓ No specific enhancements.	☒ In development.	

Appendix C - HP Performance Assessment Tool

To help prospective customers to assess whether a high performance tape drive can be supported on their existing system, or whether some hardware modifications are necessary, HP has developed a performance assessment tool. The Tool currently only supports NT and W2K. It can be downloaded from our Support Web Site at:-

http://www.hp.com/support/ultriumPAT

What it does

The performance assessment tool analyzes what a customer's disk configuration is capable of delivering in both random disk access mode (file by file backup) and sequential disk access mode (image style backup). This will allow the customer to decide if the current hardware is capable of supporting transfer rates in excess of 30 MB/sec and what improvements could be made to enable the use of higher performance tape drives.

To relate accurately to the information provided by this tool the customer must have a good understanding of the application usage of the machine and the preferred backup strategy. For example, a file print server generally uses a file by file backup methodology to enable quicker file restore. An Exchange or SQL server with "on-line" backup tends to implement a more sequential style of backup.

How it does it

The tool has two modes of operation - a quick analysis mode and a user selectable data mode.

The quick analysis mode analyses file size distribution on the system and plots a profile of file size versus number of files of that file size. By performing transfers from disk with file sizes representative of those in the profile the tool can quickly assess what the likely backup speed of that file system is likely to be.

Alternatively the user can select which files to "read" through the filesystem and the assessment tool will actually read those specific files and produce an assessment of how quickly the files could be read.

Using these two approaches the tool will deliver a figure of what the disk subsystem is capable of providing in terms of throughput. To stream today's high performance tape drives, figures above 30 MB/sec are required.

GLOSSARY OF TERMS (in alphabetical order)

This section explains the terminology and concepts that are used throughout this document and have an impact on overall backup performance.

Agent Accelerator Technology

When doing remote backups over the network to a dedicated backup server, significant performance benefits can result from using "push" agent or "accelerator" agent technology. This concept is illustrated in the following two diagrams.
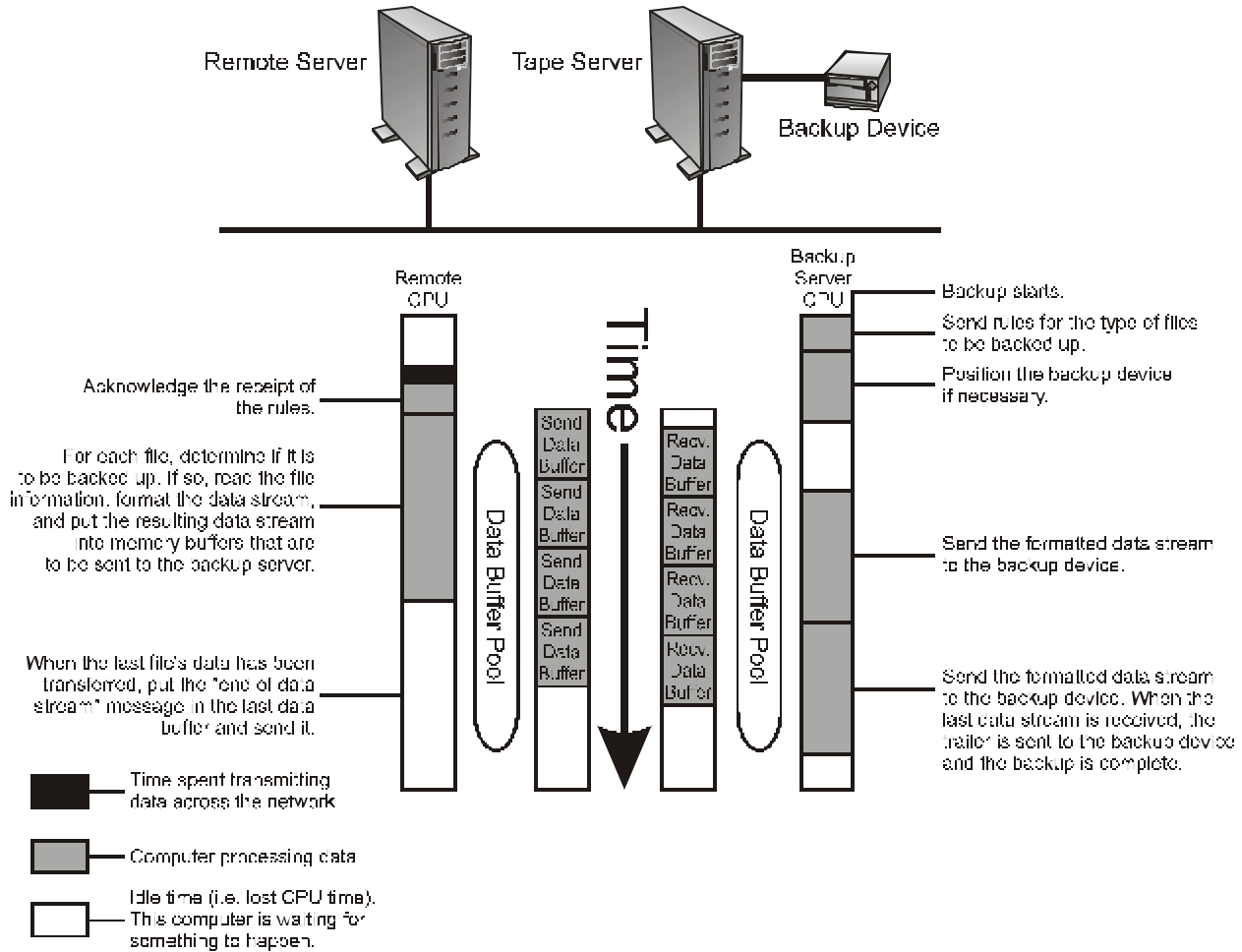
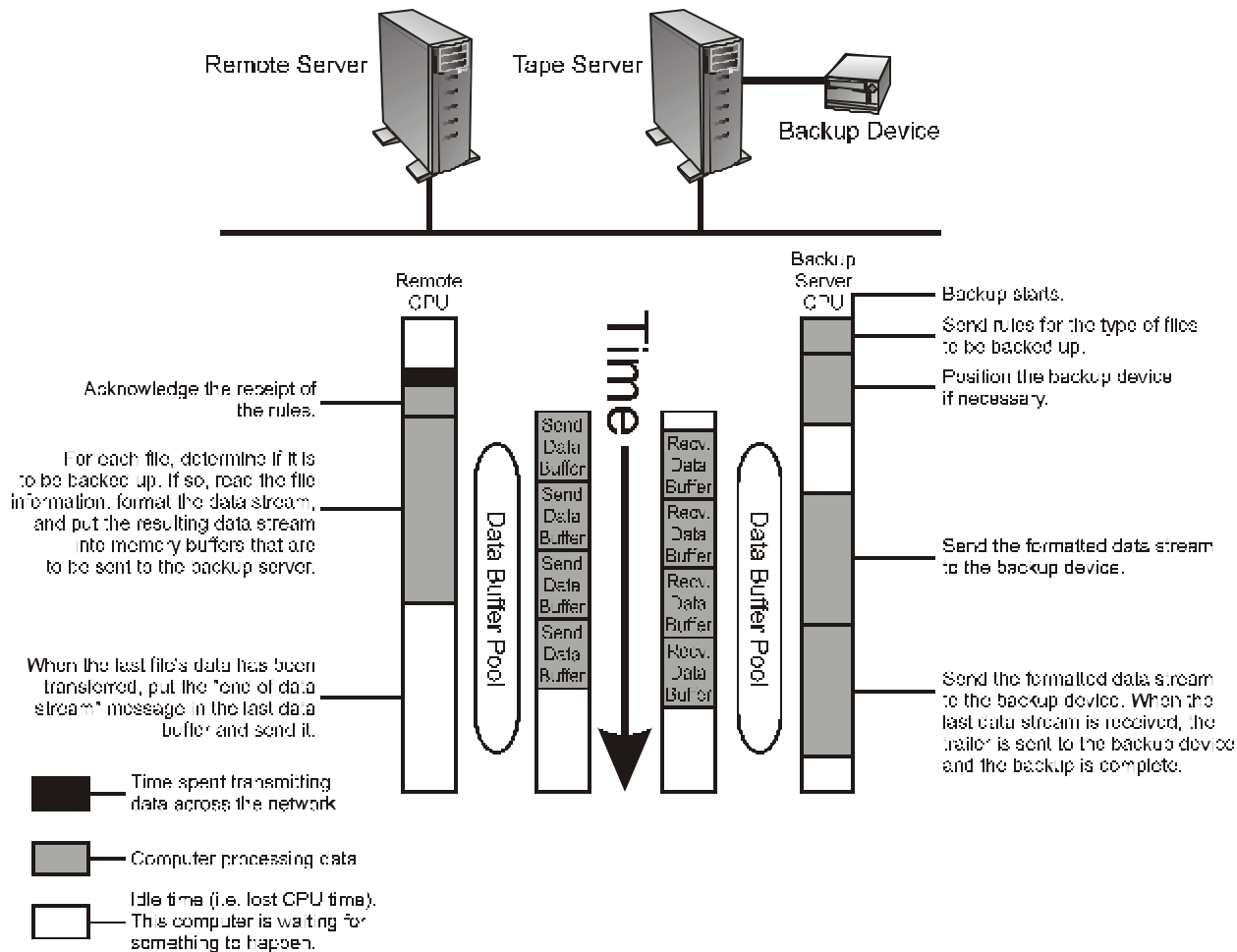

Figure 14: pull agent technology

Figure 15: Agent Accelerator Technology

| Agent Backup Technology | In some applications where 24X7 operation is essential, it is necessary to back up whilst transactions continue uninterrupted. Examples are MS Exchange Email server and SQL server. In these instances, backup agents running via an API in the application layer allow the database or mailboxes to be backed up "on-line". The major backup application ISVs, such as Computer Associates, Veritas and Legato, and HP OmniBack provide such agents at additional cost. |

| API | Application Programming Interface |
| | A standard interface supplied by the operating system vendor to allow application programs to control lower level protocols and system requests. |

CA ArcServeIT	Increasing SCSI Transfer size & Memory Buffer in NT
	In ArcServe 6.61 a default of 384K is allocated as a disk buffer for reading from disk, irrespective of the memory available on the server. If your system configuration has sufficient memory, you can increase the size of this buffer as follows:
	Prior to edit use task manager to check how much physical memory the `Tapeeng.exe` part of Arcserver IT consumes.

Use `regedit`

Select `HKEY_LOCAL_MACHINE\SOFTWARE\Computer Associates\ARCServeIT\Base\Tapeengine`

Right click on the mouse to enable you to enter a "NEW" parameter.

Enter parameter Key `SHARED MEMORY`

Right click again and assign a Value name : `Daemon_Buffer_Size`

Select `Daemon_Buffer_Size`

Right click again and assign a value to Daemon Buffer Size

The default is **24** x 16K = 384K.
So to assign 1 MB of RAM use 1024 x 1024/16 x 1024 = **64**

Make sure you select decimal in the radix window.

To check the edit has worked, stop and start the ArcServer IT job/tape and database engines

Using Task Manager again inspect the amount of memory the `tapeng.exe` is consuming, it should be increased from the previous value.

Use Task Manager also to check that when all applications are active and running there is still sufficient free memory space available.

This memory is per job, so remember if you plan to run concurrent jobs the amount of memory used increases accordingly.

This edit will have little effect if ArcServe is driving the tape as hard as possible already, but performance increases of 33% have been achieved with this technique

Arcserve 7 (July 2000) is expected to tune the memory automatically according to memory available and tape throughput.

Concurrency

One alternative way of streaming high-performance tape drives is to use Concurrency. This means backing up multiple data sources simultaneously to a single tape drive. The format on tape is then an interleaf of the data on the disks. You must check if your ISV backup software supports concurrency. HP OmniBack II, Legato Networker and Veritas NetBackup all support concurrency. This technique can also be applied to network backup where the file systems are interleaved as they are passed over the network and are written interleaved onto tape.
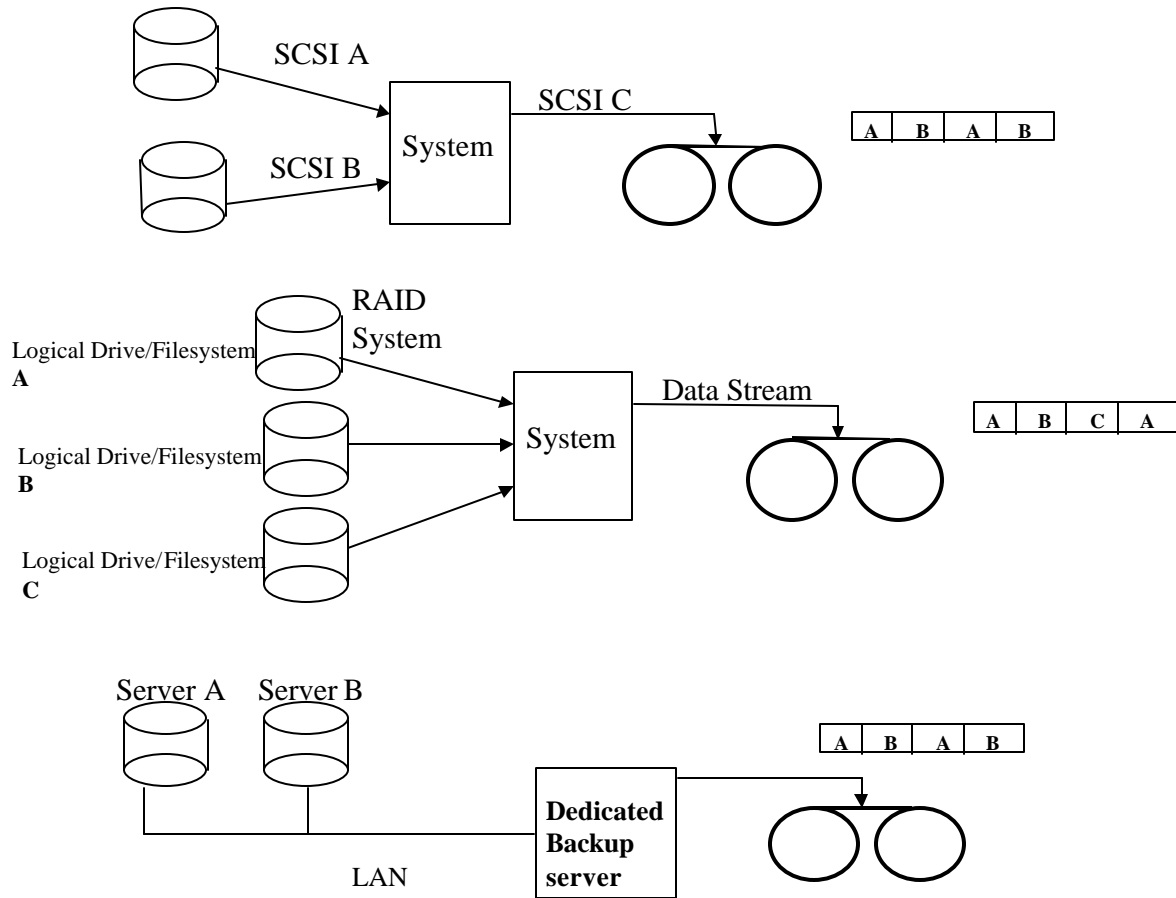
Figure 16: concurrency

Data Rate Matching

High performance tape drive manufacturers are implementing various techniques to match the incoming data rate (Data Rate matching). If the data input slows, the tape drive also slows to avoid re-positioning, which improves the overall throughput and increases reliability through less mechanical wear. HP Ultrium tape drives can data rate match incoming native transfer rates from 6 to 15 MB/sec. (12 MB/sec to 30 MB/sec with 2:1 compression).

Figure 17 illustrates this concept. As the incoming data rate changes, the intelligent firmware in the HP Ultrium tape drive dynamically adapts and changes:

1. the R/W channel to decode data at a faster or slower rate

2. the servo systems to slow down/speed up the tape motors.

The key point is that this is totally dynamic and changes on the fly for maximum performance effect. Other vendors simply use a large buffer or only change the speed in fixed increments at the end of a pass of tape, which is sub-optimal for performance.
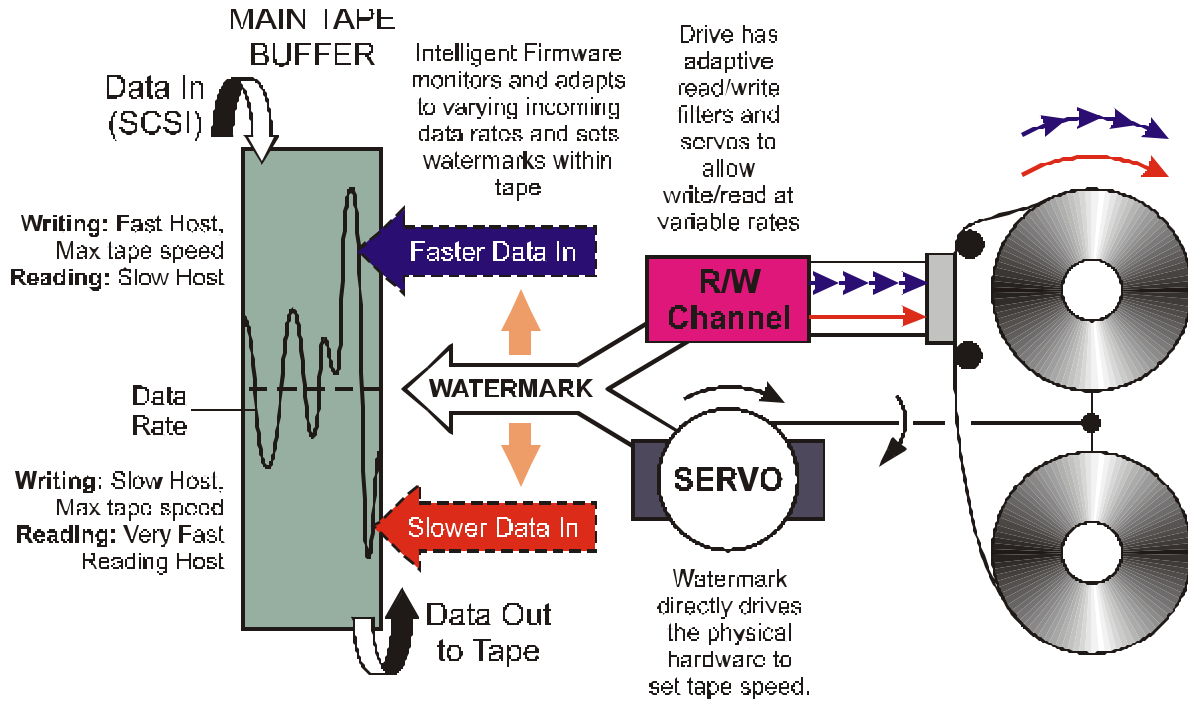
Figure 17: how data rate matching works

The graph in Figure 18 shows the tape response when 1:1 uncompressible data sets are constructed on disk to give varying input data rates to tape. With Data Rate Matching off, the tape response only achieves good performance as the data rate from disk approaches the native transfer rate of the drive (15MB/sec). With DRM on, the tape performance rises gradually because it is able to adapt the tape speed and write/read filter dynamically to align with the incoming rate off the disk.
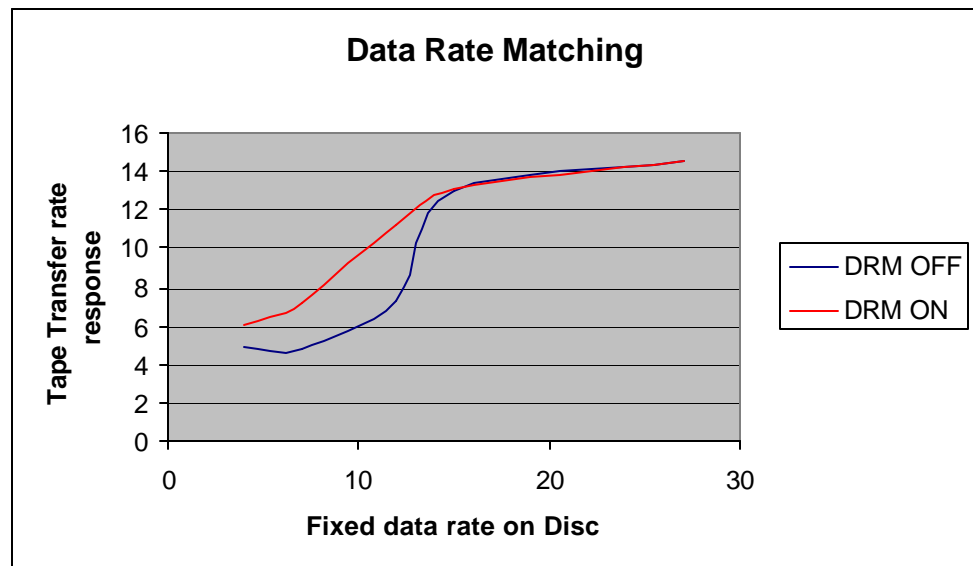

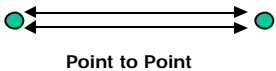*Figure 18: tape response using data rate matching*

Fibre Channel

Fibre Channel is a transport medium capable of multiplexing different protocols such as SCSI, TCP/IP, and so on, simultaneously. Fibre channel technology enables the ability to "share" tape backup devices amongst multiple servers over a high-speed link capable of around 100 MB/sec. The implementation of Fibre channel is creating the so called Storage Area Networks (SANs). New backup techniques such as "serverless backup" are now possible with this infrastructure and utilize the performance capability that Fibre Channel offers.

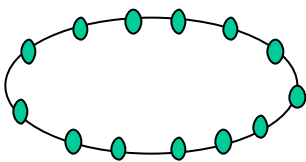*Table 14: features and specifications of fibre channel technology*

Feature	Specification
Bandwidth (Mbaud)	1062, 2124, 4248
Mode	Full Duplex
Average continuous data flow as % of full bandwidth	>80
Max No. of connections	126 Arb loop > 16 million switched
Typical No. of connections	10s to 100s
Link Distance	10m copper link 10km fiber link -1000s km total
Network Protocols	SCSI, IP, FICON, HIPPI etc

SAN Topology

Point to Point

Early implementations, generally for server attach storage or server to server. Maximum bandwidth 200 MB/sec full duplex. Tape FC generally will not support point to point.

FC - Arbitrated Loop

Arbitrated loop (FC-AL) up to 126 devices. Only two can communicate at any time through arbitration - bandwidth shared by devices. Loop resiliency circuit allows devices to be disconnected and still maintain the loop. Developed as a lower cost connect of disks and tape. With full duplex HBAs it is possible to access disks for application and still stream single tape at up to 30 MB/sec.

Switched Fabric

Switched fabric implementation. Costly at around $2.5K a port. Bandwidth is switched not shared- allows scaleable bandwidth with each connection capable of 100 MB/sec.

Fabric

Fabric Loop Attach

Fabric loop attach. Allows cost-effective connection of disk and tape to a main fabric without having to allocate a separate port per device. Up to 126 devices supported on the loop. Maximum loop bandwidth up to 100 MB/sec.

Arbitrated Loop Full Duplex Implementation

Server

HBAs support Full Duplex

30MB/sec In   **30MB/sec Out**

30MB/sec In

System can simultaneously support 60MB/sec on LOOP2 (Backup) and 30MB/sec on LOOP1 (System Applications)

LOOP 1

LOOP 2

Figure 19: arbitrated loop full duplex implementation

File by File Backup

Performing Backup through the file system using standard procedure calls and the API program layer

File Types and Size

File types and size affect backup performance in two ways:

- The compressibility of the data, higher compressible file data will yield higher performance backup, since all tape drives today use on-board hardware data compression. Some file types such as images, pictures and video are already compressed on disk and the tape drive can only operate in its native transfer rate for these types of files. Tape drives with better compression ratios such as Ultrium Technology (2.55:1) can offer typical transfer rates of up to 37.5MB/sec. The larger the files, the less overhead there is in assembling the data for transfer to tape.

- File/Print servers have a high distribution of small files and then a spread of other file sizes, as shown below. The small files represent a high workload for the backup application to access and concatenate all these small files into a data block for transfer to tape (see figure 20).

Figure 20: typical file size distribution on file/print server

- Servers such as Exchange Email servers. On these servers the transaction log file sizes are set at 5 MB, but the information store (Mailboxes) is one contiguous large file, typically 18 GB or more. When backed up (sequentially through an agent – see later), this will yield good throughput.

- Database servers, such as SQL. The record size of the SQL server is determined by design. Typical values range from 2-3 GB to Terabytes.

Fixed & Variable Block Sizes

The SCSI write command can handle 2 formats, it can transfer "n" number of blocks which are all the same size e.g. 20 X 64K blocks or it can transfer single blocks of variable sizes e.g. a single block of 1001 bytes.

Variable transfers allow for much more flexibility in tape format and efficiencies (less wasted data used in packing files to fit the fixed block size).

Flat vs Deep

If the directory structure has many sub directories (deep), then the file system overhead is less than many separate directories (flat). Flat directory structures tend to back up faster than deep directory structures. During the research of this white paper, customers with very deep and complex file system structures, particularly on NT, indicated this had been a major factor in poor backup performance. They were investigating the Veritas VxFS file system as a possible solution to their performance problem.

Fragmentation

On the NT operating system in particular, over a period of time as files are deleted or edited and new files created, files become non-contiguous, which means they are spread out all over the disk. For example a 30K file that was once intact, can over time be fragmented into 30 * 1K files dispersed all over the disk surface. This creates a negative impact on the tape backup performance since it takes much longer to assemble the file from disk.

To gain the best backup performance to high performance tape drives (and generally system performance) we strongly recommend running a defragmentation program about once a month to re-assemble files in a contiguous order. Windows 2000 comes with a built in de-fragmentation tool (called Diskeeper) supplied by Executive Software. Figure

20 below shows an analysis of the fragmentation of a system. with fragmented files in red. The fully functioned version of Diskeeper has a "set it and forget it" feature allowing regular automatic defragmentation to take place. Windows NT4 will require an additional 3rd party defragmentation tool (See further information section).

Caution: The Master File Table (MFT) performance is critical to the overall performance of an NTFS volume. Any defragmentation of the MFT reduces disk performance and defragmentation tools do not defragment the MFT. The MFT runs the risk of becoming fragmented as the disks near full capacity, or if the MFT grows larger than the 12.5% of disk space allocated to it. To prevent fragmentation of the MFT, a registry setting can be used to increase the MFT reservation zone on the disk before the NTFS volume is created.

Figure 20: sample of Diskeeper defragmentation analysis

HP One-Button Disaster Recovery

A feature which allows HP DAT and Ultrium (March 2001) tape drives to emulate a CD-ROM at boot time, thereby allowing a total system disaster recovery from tape. Prior to HP OBDR, system recovery consisted of the time consuming process of installing floppies, multiple CDs and previous backup tapes.

Gigabit EtherNet

With today's commonplace 100BaseT Network infrastructure, only around 7 MB/sec is possible even if the backup server uses a dedicated 100BaseT connection. Clearly this is not going to be good enough to stream today's high performance tape drives. Gigabit Ethernet (or 1000 Base FX) is now beginning to be introduced and is capable of an effective transfer rate of 80 MB/sec. Using Gigabit Ethernet to form a dedicated backup LAN can be more cost effective and lower risk than a full SAN implementation at the present time.

Media Type	Cable Type	Transceiver	Length-bandwidth product (MHz*Km)	Distance
1000Base-LX	8-10 Mm single-mode fiber	1300nm Laser	N/A	5000m
1000Base-LX	50 Mm multi-mode fiber	1300nm Laser	400	550m
1000Base-LX	50 Mm multi-mode fiber	1300nm Laser	500	550m
1000Base-LX	62.5 Mm multi-mode fiber	1300nm Laser	500	550m
1000Base-SX	50 Mm multi-mode fiber	850nm Laser	400	500m
1000Base-SX	50 Mm multi-mode fiber	850nm Laser	500	550m
1000Base-SX	62.5 Mm multi-mode fiber	850nm Laser	200	275m
1000Base-SX	62.5 Mm multi-mode fiber	850nm Laser	16	220m
1000Base-T	CAT 5U	Copper	N/A	100m
1000Base-CX	Shielded Twinax Cabling	Copper	N/A	25m

Image Backup   The sequential backup of data from a disk drive. Essentially accomplished by bypassing the file system, this is currently only available as an option in CA ArcServeIT or in the STAC Replica Backup products (for NT and NW). It is also available as the "Flashback" option for Veritas NETBackup for Solaris and HP-UX. Although giving faster backup, single file restore can take a long time because the file can be fragmented over a long area of tape. (See diag blow)

Multi-threading and Command Queuing   By far the biggest performance improvements on a fixed hardware configuration in the area of file by file backup could be achieved if backup applications implemented multi-threading. This is a technique where multiple data paths are established in parallel, and make use of the "command queuing" feature of SCSI. Command queuing especially in RAID controllers allows the controller to optimize the seek distances on the disk drives. Whilst it does not make any single transfer shorter it optimizes over a group of transfers resulting in significant overall improvement. Multi-threading will take some significant effort from the independent software vendors (ISVs) but the performance improvements can be significant. (See below.) A backup application using a queue depth of 1 on 64K reads with random seeks can only yield a 3 MB/sec transfer from the disk subsystem. If this application could implement a queue depth of 16 the transfer rate would increase to 15 MB/sec.
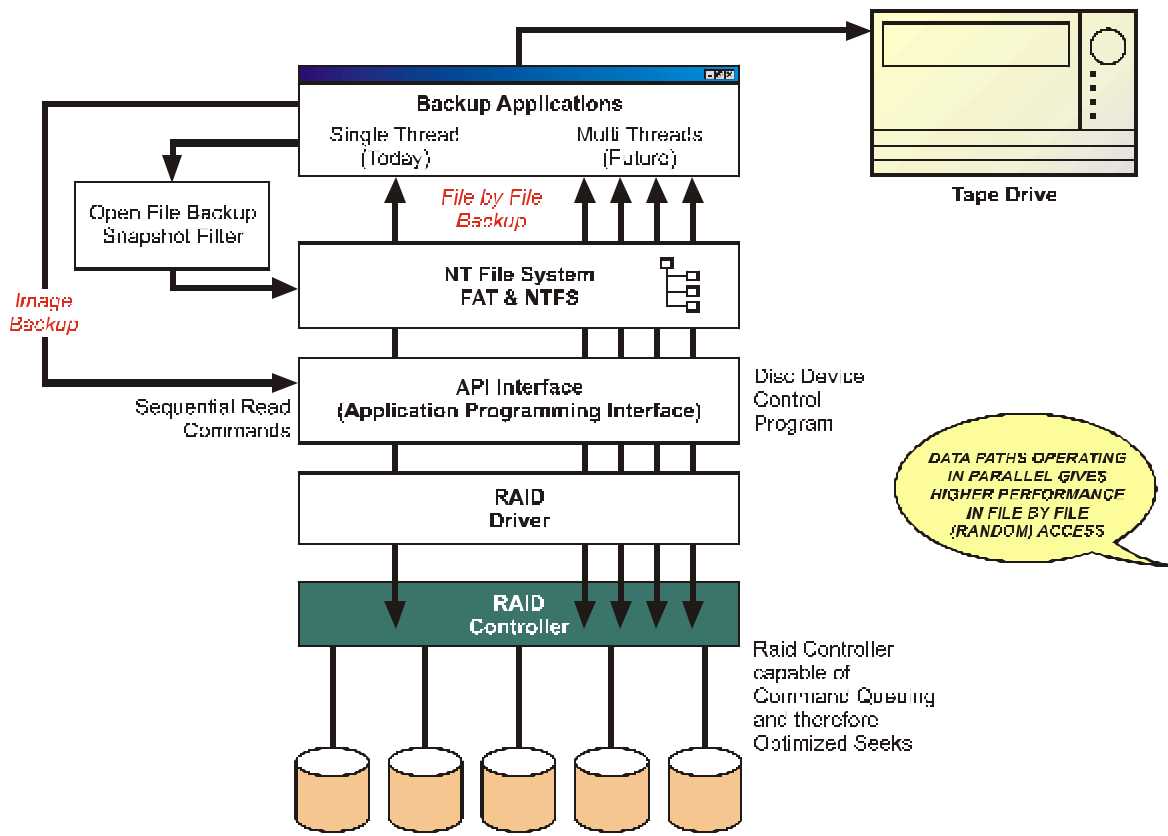
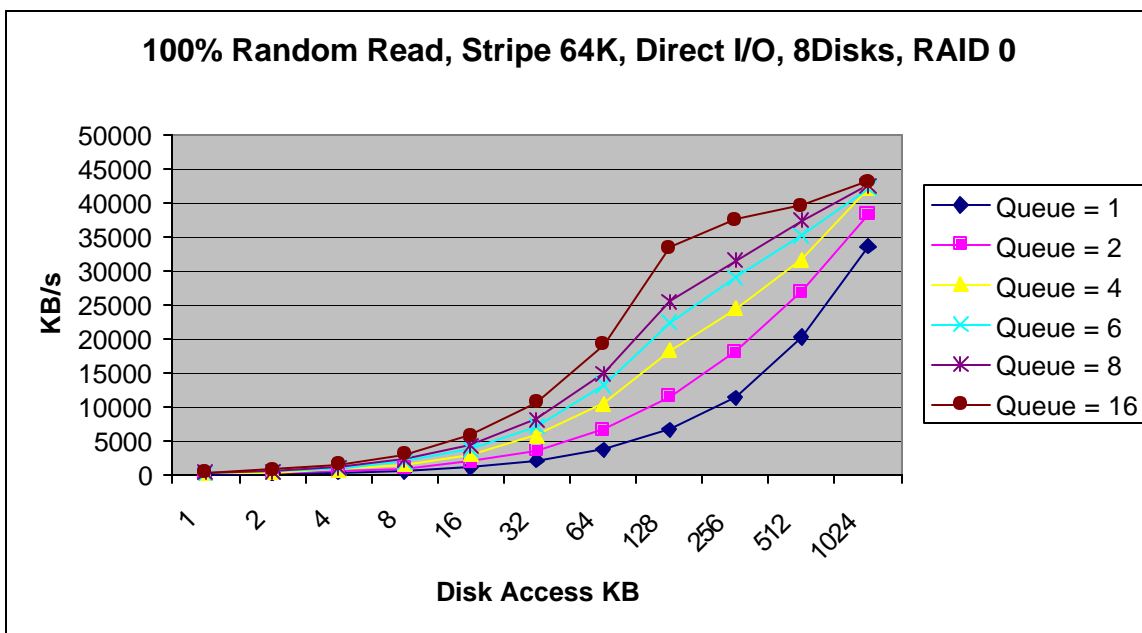*Figure 21: Image Backup, Snapshot and Multi-threading concepts*

Figure 22: Effect of Queue depth on disk transfer performance.

OmniBack Performance Tuning   As well as the raw disk backup and concurrency features in HP OmniBack, the user can also set the segment size using the logical device editor. The segment size increases the amount of data OmniBack II writes to tape in between writing file marks and file information to tape. Increasing this value will increase the importing speed of tapes. For Ultrium a value of 700 MB should be sufficient.

OmniBack II also allows the logical device editor to change the SCSI transfer size/tape block size as a variable. The default is 64K but can be changed to 128K or 256K, which will improve performance. When this value is changed tapes must be re-initialized or OmniBack II will issue a failure message.

PCI Bus Performance   This is shown below for information but PCI is not a limiting factor in backup performance. For the RAID controllers, however, we recommend that they are 32/66 or 64/66 so as to get maximum PCI throughput to/from memory.

Table 15: PCI bus performance

Bus	Width	Bus Freq MHz	Max Transfer Mbytes/sec
PCI-1	32	33	132
PCI-2	64	33	264
PCI-3	32	66	264
PCI-4	64	66	528

Network Transports   **Note:** Fibre Channel here refers to Fibre Channel in a SAN environment carrying only SCSI protocol.

Table 16: typical transfer rates for network topologies

Topology	Typical Transfer Rate Dedicated Backup LAN GB/Hr
10 BaseT	3.2
100 BaseT	32
FDDI	32
100VG Any LAN	32
GigaBit EtherNet (1000BaseT)	320
Fibre Channel (for comparison)	360

"snap-shot" Backup   This has the same effect as image backup but uses a filter inserted between the application and the file system to intercept application transfers to sectors on disk in case they are being updated whilst the snapshot image of the disk is being taken. The difference between image and snapshot is that image bypasses the file system whereas snapshot uses the file system. The key point is that snapshot access the disk sequentially. Both image and snapshot suffer from the restore time of a single file being long.

RAID   Redundant Array of Independent Devices. This is used to improve performance and fault tolerance by "ganging" together disk drives.

RAID 0

Data is striped across the available drives, using at least two drives. This method offers increased performance, but there is no protection against failure. It gives good read and write performance.

RAID 1 (Mirroring)

Data is written to two or more tapes simultaneously, providing a spare 'mirror' copy in case one disk fails. This implementation offers high assurance of the disk transfers completing safely, but adds no performance . It also adds cost.

RAID 5

In this mode, data is striped across all available drives, with parity information distributed across all of them. This method provides high performance, combined with failure protection, but requires at least three disk drives. Write performance on RAID 5 will be slower than RAID 0 because the parity has to be generated and written to disk.

RAID LEVEL 5 - DISTRIBUTED PARITY
(shown on 5 drives but can be more)

.

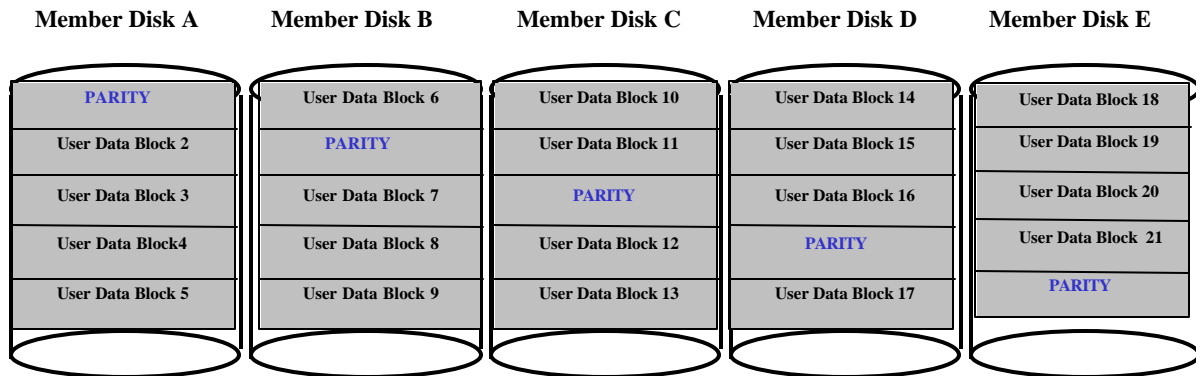| Member Disk A	Member Disk B	Member Disk C	Member Disk D	Member Disk E
PARITY	User Data Block 6	User Data Block 10	User Data Block 14	User Data Block 18
User Data Block 2	PARITY	User Data Block 11	User Data Block 15	User Data Block 19
User Data Block 3	User Data Block 7	PARITY	User Data Block 16	User Data Block 20
User Data Block4	User Data Block 8	User Data Block 12	PARITY	User Data Block 21
User Data Block 5	User Data Block 9	User Data Block 13	User Data Block 17	PARITY

Figure 22: RAID level 5

Stripe Size – the number of blocks of data in a stripe running across all member disks. A typical value is 64K, but can be as low as 8K to allow optimal sizing for record access in database applications.
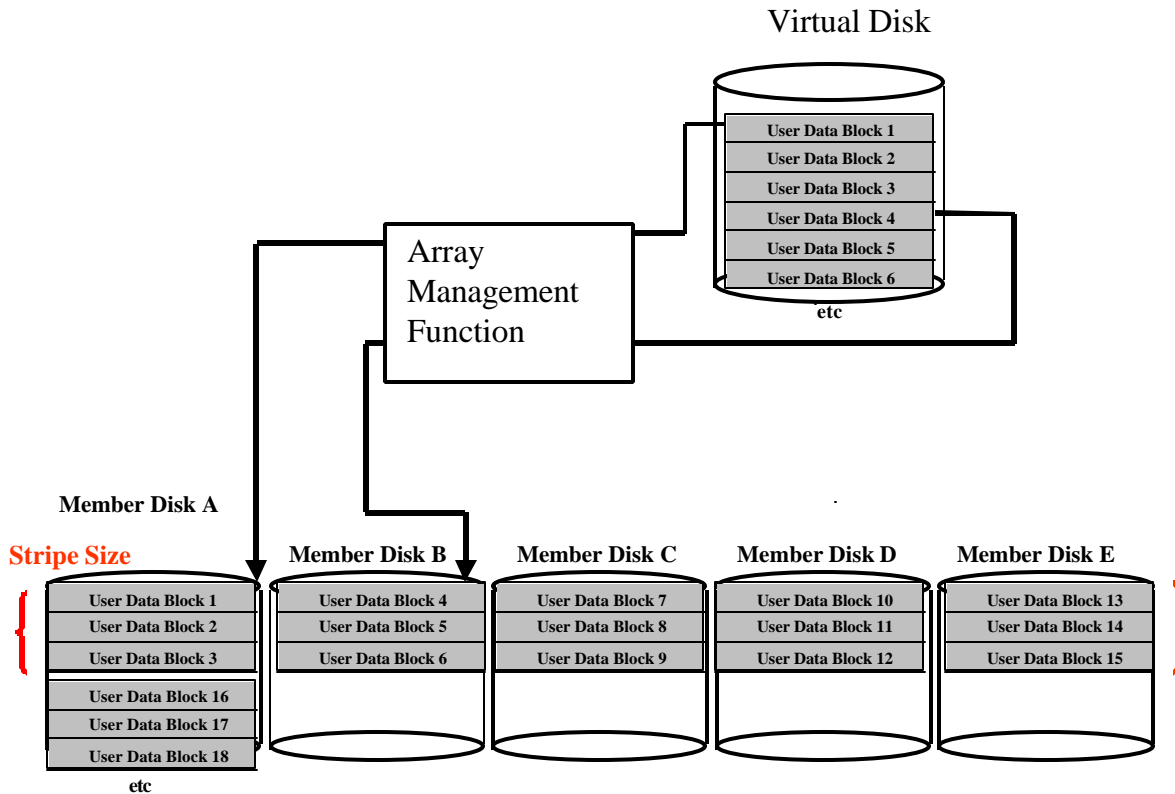
Virtual Disk

| User Data Block 1 |
| User Data Block 2 |
| User Data Block 3 |
| User Data Block 4 |
| User Data Block 5 |
| User Data Block 6 |

Array Management Function

Member Disk A

Stripe Size

Member Disk A	Member Disk B	Member Disk C	Member Disk D	Member Disk E
User Data Block 1	User Data Block 4	User Data Block 7	User Data Block 10	User Data Block 13
User Data Block 2	User Data Block 5	User Data Block 8	User Data Block 11	User Data Block 14
User Data Block 3	User Data Block 6	User Data Block 9	User Data Block 12	User Data Block 15
User Data Block 16				
User Data Block 17				
User Data Block 18				

etc

Figure 23: Disk Striping

RAID Caching Parameters

Most RAID cards today have on-board cache and the setting of the RAID cache parameters can have an important effect on overall RAID performance. Raid level, stripe size and physical RAID size are set at system install time (some later Raid controllers do allow them to be changed on-line) but caching parameters can be changed online by the utility programs supplied with the RAID controllers.

Terminology

Direct I/O – data sent to the RAID controller is first written onto disk(s) and then cached

Cached I/O – data sent to the RAID controller is first cached and then written to disk(s) when time allows.

Write Through Cache – a caching technique in which the completion of a write request is not signaled until data is safely stored on non-volatile media. Write through cache performance is roughly the same as non-cached system. But if direct I/O is enabled the data is subsequently cached and future accesses are much faster as the data comes from cache.

Write-Back Cache – Completion is signaled as soon as the data is in cache. Writing to non-volatile media happens at a later time. This carries an inherent risk that if following a write through cache operation an application or system error occurs then there will be an

inconsistency between actual media contents and what the application thinks has been written to disk. For this reason good write back cache implementations preserve cache contents (via battery backup) across system failures and flush cache contents at system restart time.

Read Options

normal : data is accessed as requested from disk.

read ahead: the controller reads more data than is requested on the expectation that the subsequent request will be a sequential request and the data will already be in cache.

adaptive read ahead: starts in normal mode and if several requests are sequential it switches into read ahead mode.

Note: All RAID card settings may use the same terminology but you have to physically experiment with the cache settings to evaluate their actual effect on performance, since implementations may vary widely across manufacturers.

SCSI Performance   The majority of high performance tape drives today use the Ultra2 SCSI LVD interface. Table 17 illustrates the various SCSI interfaces.

Table 17: SCSI interfaces

Name	Max Speed Mbytes/sec	Bus Width	Cable Length	LVD	HVD	Max Device Support
SCSI-1	5	8	6m		25m	8
Fast SCSI	10	8	3m		25m	8
Fast Wide SCSI	20	16	3m		25m	16
Ultra SCSI	20	16	1.5m		25m	8
Ultra SCSI	20	16	3m			4
Wide Ultra SCSI	40	16			25m	16
Wide Ultra SCSI	40	16	1.5m			8
Wide Ultra SCSI	40	16	3m			4
Ultra2 SCSI	40	8		12	25	8
Wide Ultra 2 SCSI	80	16		12	25	16
Ultra 3 SCSI(U160)	160	16		12		16
Ultra 320 (Future)	320	16		12		16

Storage Checkpoint Technology   Storage Checkpoint Technology available via the Veritas VXfs filesystem can instantly create an exact image of a file system or database and provides a consistent and persistent view of the database at the point in time the storage checkpoint was created. Certain backup products such as Veritas NetBackup and Backup Exec can take advantage of these storage checkpoints to perform high performance backup (in terms of small backup windows) by not having to back up all the actual physical data but only pointers and modified data blocks since the last checkpoint. The resulting checkpoints are hence files that can be backed up quickly by today's high performance tape drives. An explanation of how this works is shown below in Figure 24.
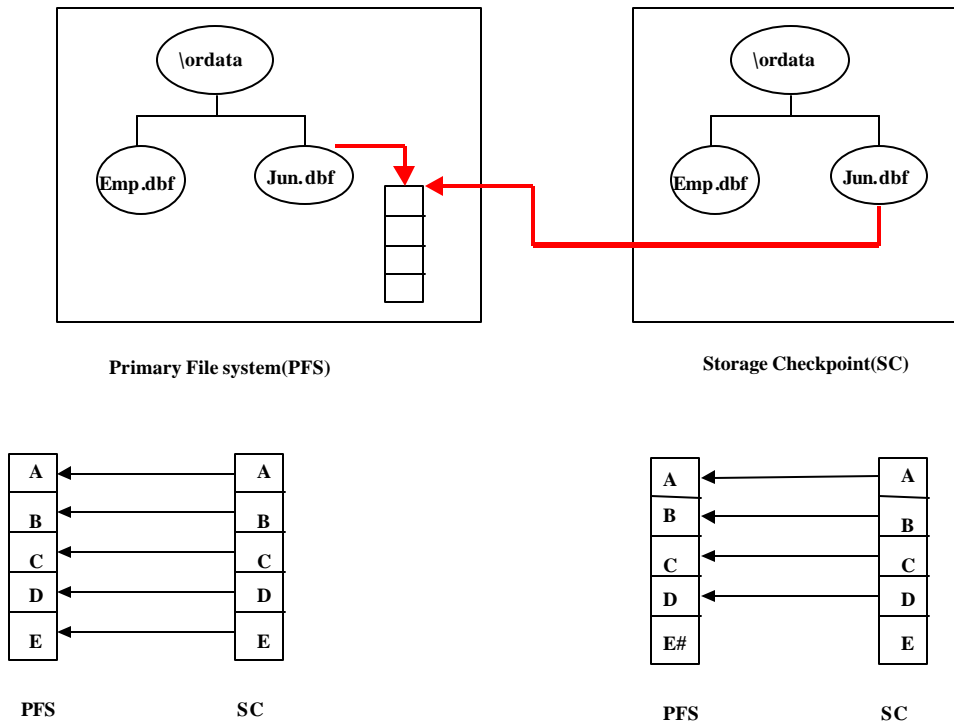
Storage Checkpoint Technology

Figure 24: storage checkpoint technology

The checkpoint only contains pointers to the primary file system at the moment the checkpoint is taken. After the checkpoint updates to the primary file system cause the "before image" data to be copied into the checkpoint in order to maintain its integrity as a snapshot taken at that moment in time. The old data is copied to the checkpoint before the new data is written to the primary file system. This allows for much smaller (and therefore faster) snapshot backups of the file system/data base to take place using today's high performance tape drives.

Tape Block Size

Tape backup applications can allow the size of the tape block to be varied. In general the larger tape block size the less overhead in the transfer to tape. If the block size were set to 512 bytes (disk sector size) then the overhead (time spent in non-data transfer) would be high compared to the amount of user data actually transferred. For this reason typical tape block sizes are 32K or 64K. However one important factor here is the tape drive architecture itself. Some drives (notably DLT) benefit from increased tape block size because their internal tape format is designed around a physical layout not a logical layout. For such devices small tape block transfers give very poor performance. For example, see the Veritas Backup Exec configuration example in the next entry in this glossary.

Veritas Backup Exec

Increasing SCSI Transfer size & Memory Buffer in NT

With Veritas Backup Exec it is possible to reconfigure

- Tape Block Size
- Buffer Size

- Buffer Count
- High Water count

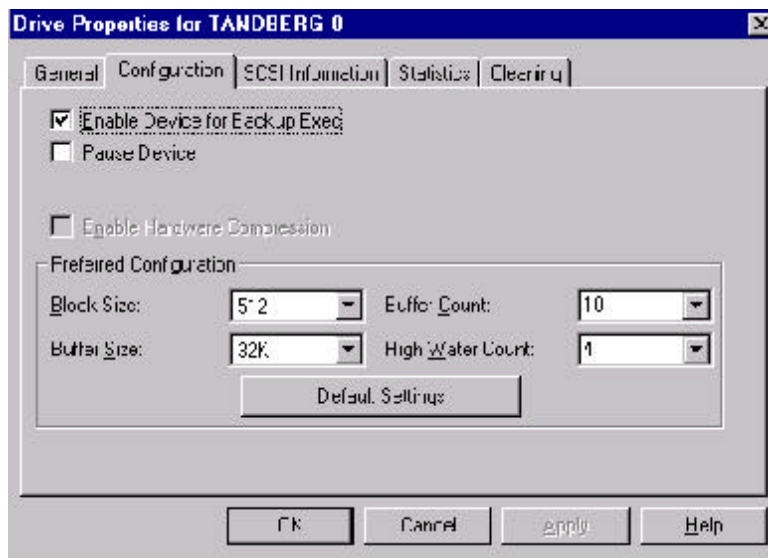By selecting the drive properties and configuration tab (see below)

Figure 25: configuring drive properties with Veritas

Caution: Preferred configuration settings are used to tune the performance of backup and restore operations. Changing preferred configuration settings is not generally recommended and may have a negative effect on your backup and system performance. Any changes should be thoroughly tested to make sure system performance does not deteriorate before being put into general use. All settings must be recorded for use in the disaster recovery process.

Block Size should be set to that recommended for optimal performance by the tape vendor. For example HP Ultrium recommends a tape block size of 32K or 64K.

Buffer Size per Device is the preferred amount sent to the tape on each read or write request and must be a multiple of the block size.

Buffer count is the number of buffers allocated to this device. The value that you can set this to depends on the amount of memory in your system and memory allocation to other tasks. Increasing this may increase performance but increasing it too much may cause the processor to start caching memory pages to disk which starts to decrease overall performance. Increasing this value means the High Water Count will also need increasing.

High Water Count is the preferred number of buffers to be filled before data is actually sent to tape or any time after if the drive underuns, it builds in data padding for "end of wrap" circumstances in linear tape drives.

Veritas Backup Exec also allows for the NT4 default of 64K MAX size transfers on the SCI bus to be increased up to 1020K. This is done by changing the following registry entry:

Use regedit

Select `HKEY_LOCAL_MACHINE`

Select `System\CurrentControSet\Services|DriverName\Parameters\Devicen\MaximumSGList`

The `DriverName` is the miniport driver, such as AIC78XX

`Devicen` is the bus number assigned at initialization

The `REG_DWORD` entry for `MaximumSGList` is the default maximum for the `scatter/gather` list which assembles the transfer. The maximum value is 255

VXFS File system for NT

Because the file system performance is the critical path in most file by file backup approaches, several developments are taking place in this area.

Veritas have developed a new file system for NT called VXFS that will

- Integrate into the NT/W2K NTFS file system

- Allow tuning of the file system for better overall performance

- Produce less defragmentation

- Support storage checkpoints for improved database backup windows (reduced backup times)