HP Data Protector Software

Performance White Paper

Executive summary

This white paper provides performance-related information for HP Data Protector software 6.0 together with some typical examples. The emphasis is on backup servers and two common backup and restore performance questions:

- Why are backups so slow?
- Why are restores so slow?

The first step towards estimating backup requirements and performance is a complete understanding of the environment. Many performance issues can be traced to hardware or environmental issues. A basic understanding of the complete backup data path is essential in determining the maximum performance that can be expected from the installation. Every backup environment has a bottleneck. It may be a fast bottleneck, but it will determine the maximum performance obtainable in the system.

There are many configuration options and procedures available that can help IT professionals to improve the performance of their backup environment. This white paper focuses on servers running HP-UX 11.11 (11iv1) and Windows Server 2003, which are backing up file servers and Microsoft Exchange Server 2003 in Local Area Network (LAN) and Storage Area Network (SAN) environments. All data was located on an HP StorageWorks 6000 Enterprise Virtual Array (EVA6000) and backed up or restored with an HP StorageWorks Ultrium 960 tape drive.

The following hardware and software was installed and configured:

- HP9000 rp3440 with HP-UX 11.11 (11iv1)
- HP ProLiant DL380 G5 with Microsoft Windows Server 2003 SP2 and Exchange Server 2003
- HP ProLiant DL380 G4 with Microsoft Windows Server 2003 SP2 and HP Data Protector software 6.0
- HP ProCurve LAN switch 20 10/100/1000-ports
- HP ProCurve SAN Switch 4/16
- HP StorageWorks 6000 Enterprise Virtual Array
- Two HP StorageWorks Ultrium 960 tape drives

The HP-UX test environment showed good performance results across all tests. The NULL device backup performance was 386 MB/s or 1,359 GB/h, and the HP StorageWorks Ultrium 960 tape backup performance 156 MB/s or 548 GB/h.

The Windows test environment also showed good performance results, except for the test with millions of small files. In such environments, the Windows NTFS file system responds very slowly if a file is restored and its attributes recovered. The restore performance from the HP StorageWorks Ultrium 960 tape drive was just 3.38 MB/s. See the section Local restore of small files on page 48.

As a result of these tests, several recommendations and rules of thumb have emerged:

1. HP Data Protector software tuning can help to improve the performance, for example, for file systems with millions of small files where the first tree walk is disabled and the backup concurrency increased. See the section Tuning Recommendations on page 77.
2. Data Protector’s default configuration parameters are well sized for most use cases.
3. Some changes of configuration parameters have almost no performance impact, for example, the Disk Agent buffers. See the section Disk Agent buffers on page 71.
4. Software compression causes high CPU loads and poorer backup performance than the Ultrium 960 built-in compression. See the section Software compression on page 50.

Overview

HP servers, storage, and software can help to provide a seamless enterprise backup and recovery solution. The solution starts with understanding what server and storage components work best for the required workload. Being able to determine performance baselines and uncovering potential bottlenecks in the solution can help to focus on areas that may need improvement, and can also provide information that helps with planning for data growth. Understanding the backup and recovery requirements for a data center can also help to maintain a consistent process and set expectations for data backup and recovery. Often these business expectations are documented in a Service Level Agreement (SLA).

HP has conducted testing in typical solution configurations for backup and recovery. While solutions can be designed by directly attaching storage devices, using SCSI or Fibre Channel, attaching to Backup Servers on the LAN, attaching to devices over a Fibre Channel SAN, or by offloading backup processes to a dedicated Backup Server for a non-disruptive backup solution, or Zero Downtime Backup, this paper focuses only on SCSI-attached tape devices in LAN and SAN configurations.

Objectives and target audience

The objective of this white paper is to educate and inform users of HP Data Protector software about what levels of performance are achievable in typical backup scenarios.

The emphasis is in showing mid-size environments and not very large installations with academic performance. This white paper highlights where the performance bottlenecks are and how these can be overcome. User loads are disregarded on the assumption that backups and restores are executed in idle environments, for example, executing backups at night when nobody is online.

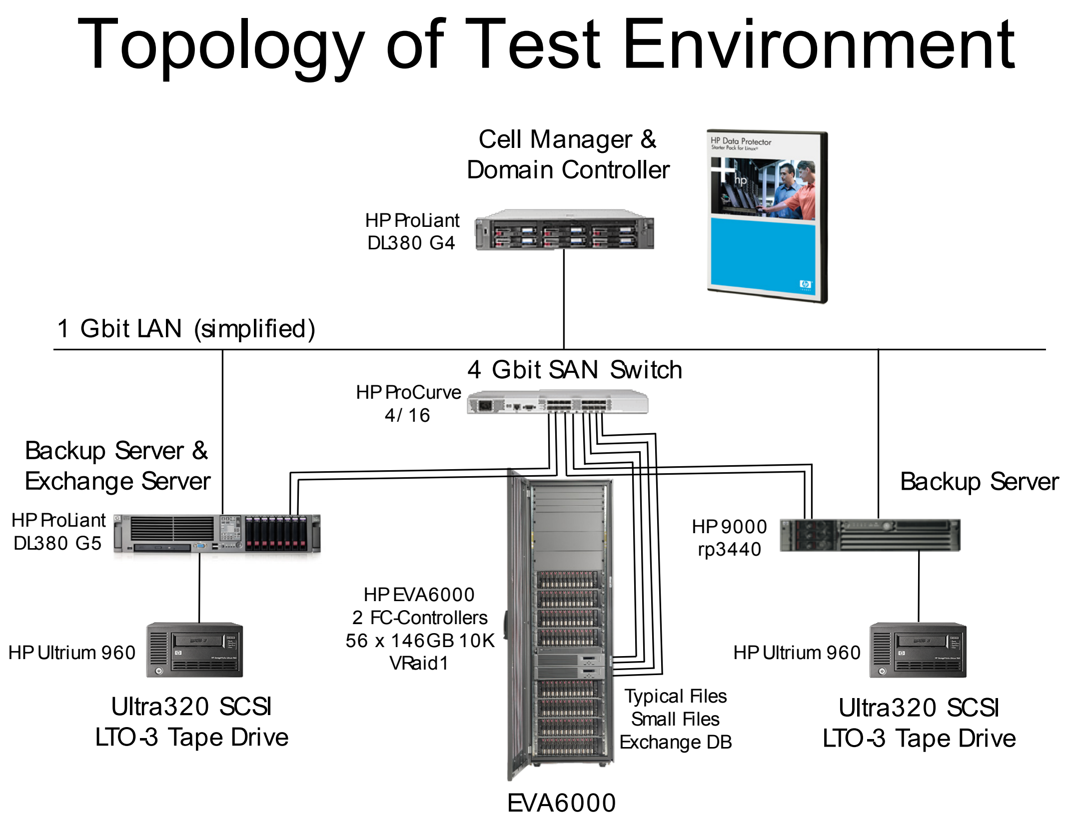
The target audience are system integrators and solution architects, and indeed anyone involved in getting the best backup performance out of their HP infrastructure investments.

Introduction and review of test configuration

Testing was conducted in a Microsoft Windows Server 2003 and HP-UX 11.11 environment using an HP StorageWorks 6000 Enterprise Virtual Array for the file server and Microsoft Exchange Server 2003 data. The HP StorageWorks Ultrium 960 tape drive was used for backups and restores via LAN and SAN. The configuration details are described in the following sections.

Figure 1 illustrates the topology layout of the test environment.

Figure 1. Topology layout of test environment



Storage array

The HP StorageWorks 6000 Enterprise Virtual Array is an enterprise class, high performance, high capacity, and high availability “virtual” array storage solution. The EVA6000 is designed to meet the needs of the data center to support critical application demands for consistently high transaction I/O.

This document assumes that you have previous knowledge of the EVA family of storage arrays. However, there are several terms that are exclusive to the EVA and referenced throughout the document. These terms are defined in the Glossary on page 80 to provide a knowledge baseline. For further information on the HP StorageWorks EVA6000 visit http://www.hp.com/go/eva6000.

The EVA6000 storage array was set up in a 2C4D configuration (2 controller and 4 disk shelves) with a total of 56 x 146 GB 10K Fibre Channel disk drives.

HP StorageWorks Enterprise Virtual Arrays perform excellently when all the disks are in a single disk group, so in this example the disk group consists of all 56 disks. 15 x 100 GB LUNs (Vdisks) were created on the array and presented to the two client servers. The RAID level applied to these LUNs was VRaid 1 to get the best read performance out of the available drives.

The firmware version was XCS 6.000.

Storage management

The HP StorageWorks Command View general purpose server running on an HP ProLiant DL580 G2 server was configured with Microsoft Windows Server 2003 Enterprise Edition. HP StorageWorks Command View EVA V5.0 was used to manage the EVA6000.

LAN infrastructure

The LAN consisted of one HP ProCurve switch with 20 x 10/100/1000-ports. All network links were configured with1 Gb.

SAN infrastructure

The fabric consisted of one HP ProCurve SAN switch 4/16. All links between servers and storage devices were 4 Gb.

The HP-UX and Windows server were equipped with one dual-channel Host Bus Adapter (HBA) each, and connected to the SAN switch for optimal performance.

The EVA6000 was equipped with two Fibre Channel controllers and 4 host ports (2 per controller).

HP StorageWorks Ultrium 960 tape drive

The backup devices consisted of two HP StorageWorks Ultrium 960 tape drives, each connected to one Backup Server via an Ultra320 SCSI interface.

Management server

The HP ProLiant DL380 G4 management server was configured as a Windows domain controller and installed with:

- Windows Server 2003 SP1, 32-bit Enterprise Edition
- Data Protector 6.0 Cell Manager including patches DPWIN_00243, DPWIN_00244, DPWIN_00245, DPWIN_00246, and DPWIN_00260

The HP ProLiant DL380 G4 management server was equipped with one Intel Xeon 3.60 GHz dual-core processor (hyper-threaded to two) and 1 GB RAM.

Both internal 146 GB disks were configured with RAID 1 which gave the Data Protector Internal Database (IDB) an excellent performance. For the LAN connection, only one adapter of the integrated dual-port NC7782 Gigabit Server Adapter was configured.

Windows backup and Exchange Server 2003

The HP ProLiant DL380 G5 Windows backup and Exchange server was installed with:

- Windows Server 2003 SP2, 32-bit Enterprise Edition
- Microsoft Exchange Server 2003 SP2

The HP ProLiant DL380 G5 Windows backup and Exchange Server 2003 was equipped with two Intel Xeon 3.00 GHz dual-core processors (hyper-threaded to four) and 16 GB RAM.

Both network adapters of the dual-embedded NC373i Multifunction Gigabit Server Adapter were configured for teaming with failover functionality. The total network performance was 1 Gb/s. Fibre connectivity was provided by way of one Emulex PLUS 4 Gb PCIe dual-channel Host Bus Adapter with the native Windows MPIO driver used for host connectivity for a total of 8 physical paths to the fabric (2 HBAs, 1 switch, 4 EVA controller ports). Ultra320 SCSI connectivity to the tape drive was provided by the HP SC11Xe PCIe HBA.

HP-UX Backup Server

The HP9000 rp3440 Backup Server was installed with HP-UX 11.11.

The HP9000 rp3440 Backup Server was equipped with two 1.0 GHz dual-core PA-8900 processors and 8 GB RAM. Fibre connectivity was provided by way of one 4 Gb dual-port host HBA.

Configuration guidelines

For general information please refer to the *HP StorageWorks Enterprise Backup Solutions (EBS) Design Guide* at http://www.hp.com/go/ebs.

Backup environment

The backup manager is the core of any backup environment that you should check for requirements and compatibility. Please check at http://www.hp.com/go/dataprotector the following documents:

- *HP Data Protector Concepts Guide*
- *HP Data Protector Installation and Licensing Guide*
- *HP Data Protector Product Announcements, Software Notes, and References*
- Data Protector support matrices

Note: Please check the Data Protector support matrices carefully and install the latest patches.

Backup server

You should also check the following components of the Backup Server environment for compatibility:

- Host bus adapter (HBA)
- OS/HBA drivers
- HBA BIOS
- Tape drive firmware

For detailed information please refer to the EBS compatibility matrix at http://www.hp.com/go/ebs.

For fast tape devices, check the PCI bus performance. Use separate PCI busses for disks and tape if possible. The total PCI bandwidth needs to be more than double the backup rate.

Windows 2000/2003 tape drive polling - Windows Removable Storage Manager service (RSM)

The Windows Removable Storage Manager service (RSM) polls tape drives frequently (every three seconds in Microsoft Windows 2000, and every second in Windows Server 2003). This polling, which involves sending a Test Unit Ready (TUR) command to each tape or library device with a loaded driver, is enabled via the AUTORUN feature of Windows Plug and Play. Windows built-in backup software (NTBACKUP) relies on the device polling to detect media changes in the tape drive.

Note: In SAN configurations, this polling can have a significant negative impact on tape drive performance. For SAN configurations, HP strongly recommends disabling RSM polling.

For additional information refer to the Microsoft Knowledge Base article number 842411.

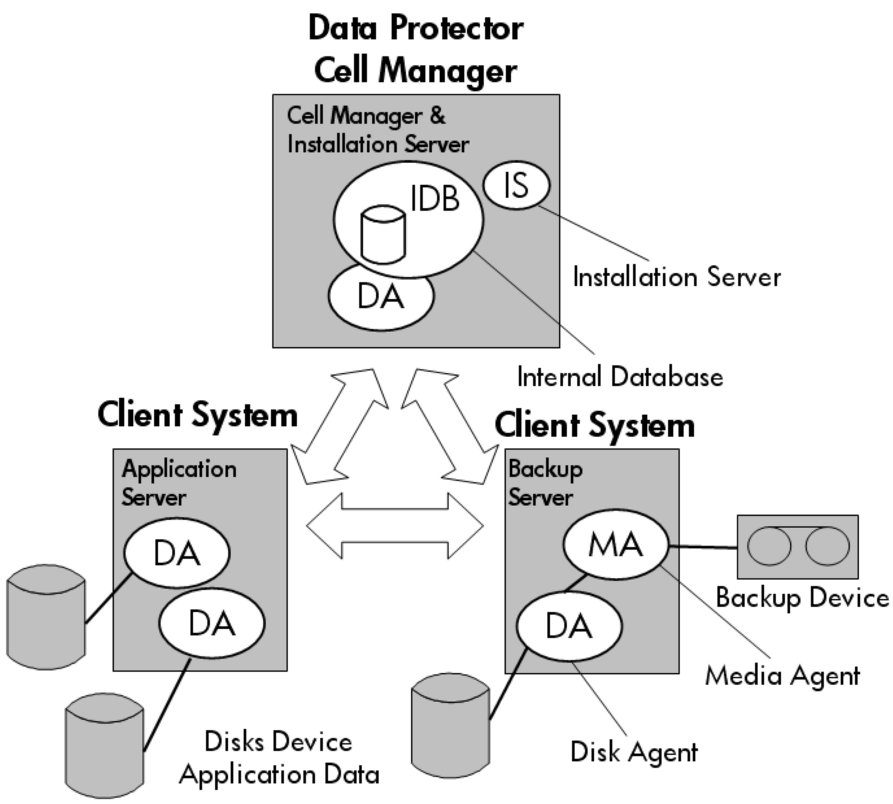
Data Protector overview and architecture

HP Data Protector automates high performance backup and recovery, from disk or tape, over unlimited distances, to ensure 24x7 business continuity and maximize IT resource utilization. Data Protector offers acquisition and deployment costs which are 30–70% less than competition, responding to the pressure for lower IT costs and greater operational efficiency. The license model is very simple and easy to understand, helping to reduce complexity. Data Protector lowers data protection costs while ensuring backup/recovery flexibility, scalability, and performance. Data Protector is a key member of the fast-growing HP storage software portfolio, which includes storage resource management, archiving, replication, and device management software. No software can integrate better with the HP market-leading line of StorageWorks disk and tape products, as well as with other heterogeneous storage infrastructures, than Data Protector.

Data Protector can be used in environments ranging from a single system to thousands of systems on several sites. The basic management unit is the Data Protector cell.

The Data Protector cell, as shown in Figure 2, is a network environment consisting of a Cell Manager, one or more Installation Servers, Clients Systems, Backup and Disk Devices. The Cell Manager and Installation Server can be on the same system, which is the default option, or on separate systems.

Figure 2. Data Protector architecture



Cell Manager

The Cell Manager is the main system that controls the Data Protector cell from a central point, where the Data Protector core software with the Internal Database (IDB) is installed. The Cell Manager runs Session Managers that control backup and restore sessions and write session information to the IDB. The IDB keeps track of the backed-up files as well as of the configuration of the Data Protector cell.

Installation Server

The Installation Server (IS) is the computer where the software repository of Data Protector is stored. At least one Installation Server for UNIX and one for the Windows environment are required to enable remote installations through the network and to distribute software components to client systems in the cell.

Client Systems

After installing Data Protector's Cell Manager and Installation Server, components can be installed on every system in the cell. These systems become Data Protector clients. The role of a client depends on the Data Protector software installed on the system.

Systems to be backed up

Client systems (application servers) to be backed up must have the Data Protector Disk Agent (DA) installed. The Disk Agent reads data from and writes data to disks on the system, and exchanges data with a Media Agent (MA).

The Disk Agent is installed on the Cell Manager as well, allowing the administrator to back up data on the Cell Manager, the Data Protector configuration, and the IDB.

Systems with backup devices

Client systems with connected backup devices must have a Data Protector Media Agent installed. A Media Agent reads data from and writes data to media in the device, and exchanges data with the Disk Agent. A backup device can be connected to any system and not only to the Cell Manager. Client systems with backup devices are also called Backup Servers.

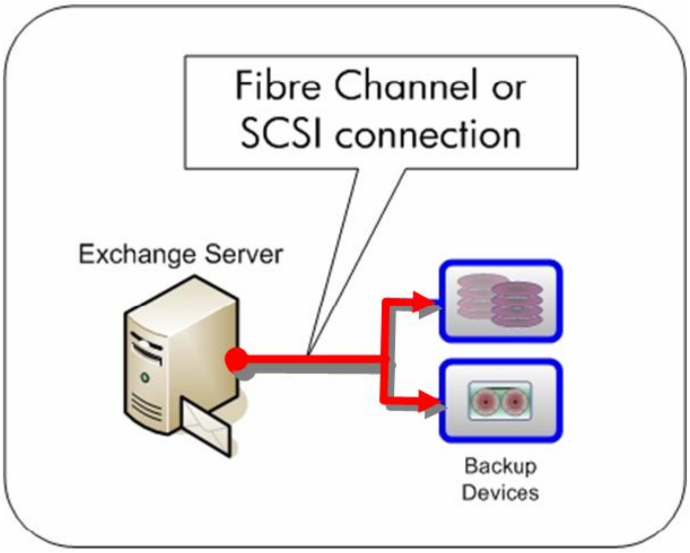
Backup and restore designs

Backup environments can be designed using a variety of configurations. The following sections describe three designs typically used in backup and restore solutions.

Direct attached storage (DAS)

Direct attached storage (see Figure 3) implies that the storage you are backing up to and restoring from is directly attached to the server through a SCSI or Fibre Channel bus. This can be a good solution for smaller data centers with only a few servers that do not require the sharing of a large tape or disk resource. While this solution can provide good performance, it cannot be leveraged across multiple servers.

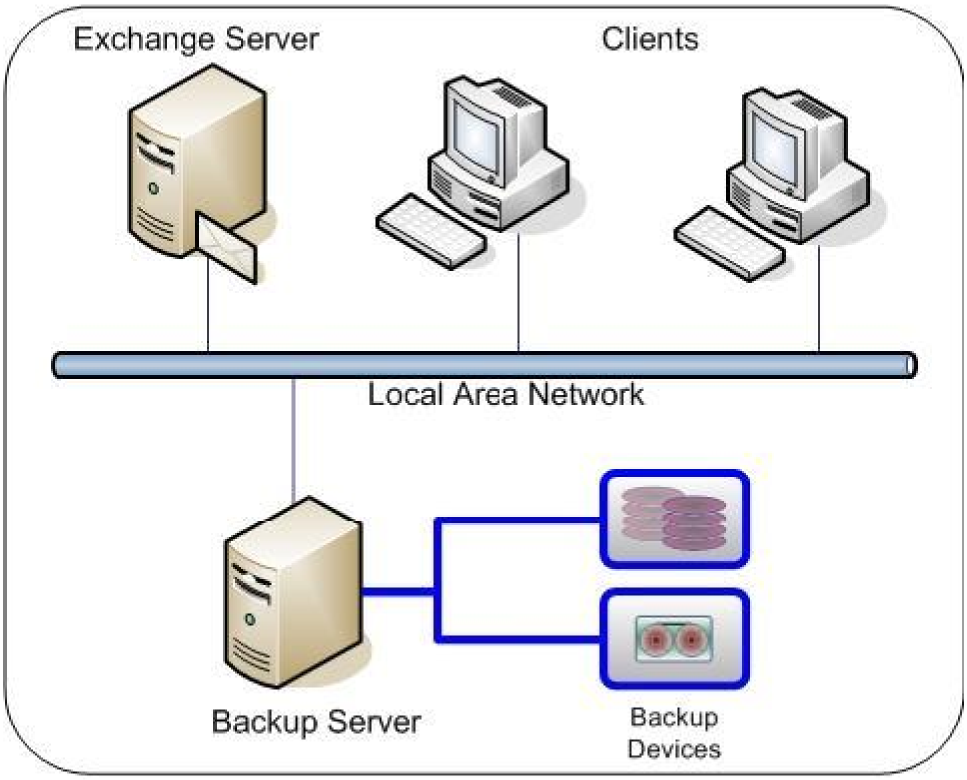
Figure 3. DAS



Local area network (LAN)

When combined with a server that has a DAS design, this solution allows servers and clients to back up and restore over the LAN to the machine with the DAS tape library or disk array (see Figure 4). Nowadays, this is the standard method that is employed throughout most data centers. It allows backup devices to be shared by other servers on the network. However all backup data must be passed over the public LAN, or a dedicated backup LAN, which can reduce the overall performance.

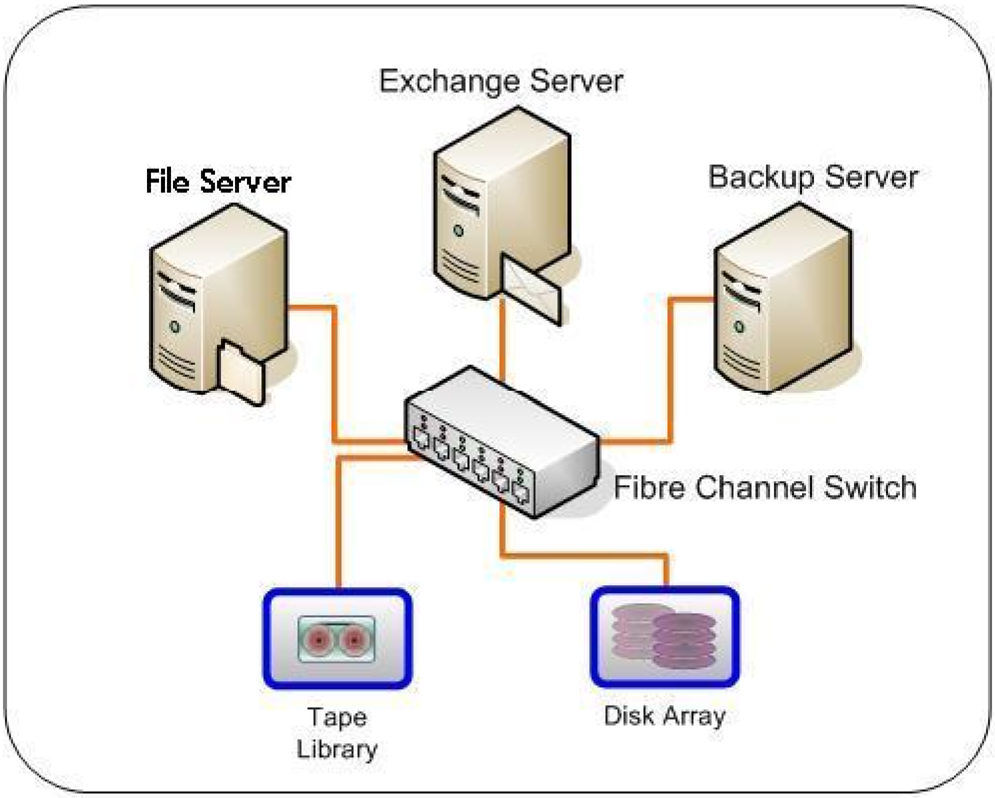
Figure 4. LAN



Storage Area Network (SAN)

SAN solutions (see Figure 5) are very similar to DAS solutions. However, the storage on the SAN may be shared between multiple servers. This shared storage can be both disk arrays and tape libraries. To the server it will appear as if it owns the storage device, but the shared Fibre Channel bus and the controllers for the disk and tape devices allow the devices to be shared. This configuration is very advantageous for both resource sharing and performance.

Figure 5. SAN



Other backup and restore designs

The following technologies go beyond the scope of this white paper:

- Direct backup (a severless backup solution for SAN environments)
- Disk backup
- Split mirror backup
- Snapshot backup
- Microsoft Volume Shadow Copy Service (VSS)

For further information, see the *HP Data Protector Zero Downtime Backup Concept Guide* at http://www.hp.com/go/dataprotector.

Performance bottlenecks

The goal is that a tape drive becomes a performance bottleneck. In that case, backup and restore times are predictable.

Backup performance

Backup performance will always be limited by one or more bottlenecks in the system, of which the tape drive is just one part.

The flow of data throughout the system must be fast enough to provide the tape drive with data at its desired rates. High-speed tape drives, such as the HP StorageWorks Ultrium 960 tape drive, are so fast that making them the bottleneck can be very challenging.

All components must be considered and analyzed for getting the best backup performance. Standard configurations with default values are mostly not sufficient. In enterprise environments, performance estimations are very difficult to provide without benchmarks.

Restore performance

Many backup environments are tuned for backup performance because frequent backups require most resources. Large restores are less common and so runtimes are less predictable. Backup runtimes cannot be used for calculating exact restore runtimes. Therefore, runtimes can only be proved by restore tests.

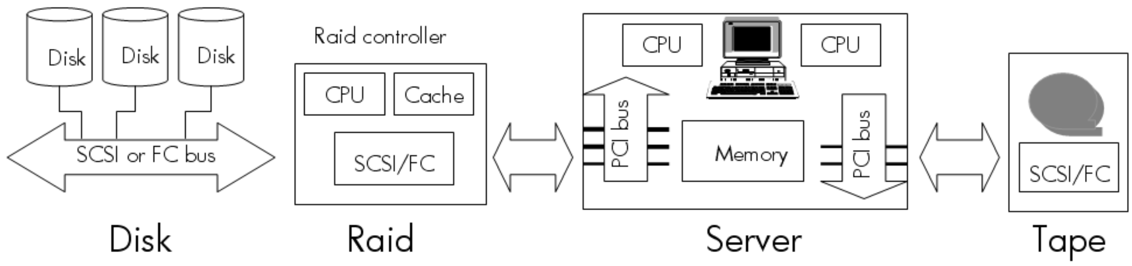
Backup and restore data path

The limiting component of the overall backup and restore performance is the slowest component in the backup environment. For example, a fast tape drive combined with an overloaded server results in poor performance. Similarly, a fast tape drive on a slow network also results in poor performance.

The backup and restore performance depends on the complete data transfer path.

Figure 6 shows that the backup path usually starts at the data on the disk and ends with a copy on tape. The restore path is the other way round. Ideally, the tape drive should be the limiting factor in the data path. All other components should process data faster than the tape drive so the tape drive has never to wait for data.

Figure 6. Backup and restore data path



Manager server (Data Protector Cell Manager)

The Cell Manager should be sized according to its installation requirements.

Note: Additional applications or backups on the Cell Manager will increase the installation requirements. Please check carefully any application installation requirements. In larger backup environments, it is not recommended to share this server with other applications or to use it as a backup server.

The server should be clustered for having no Single Point of Failure (SPOF) in the backup environment. The network connections to each client are essential and should be protected against single link failures.

Backup and application server (Data Protector Client System)

Backup and application servers should be sized according to the installation requirements.

During backup, these servers should not be under heavy load from other applications that run I/O and CPU intensive operations, such as virus scans or a large number of database transactions.

Backup servers demand special attention for proper sizing because they are central to the backup process as they run the required agents. The data is passed into and out of the server's main memory as it is read from the disk subsystem or network and written to tape. The server memory should be sized accordingly, for example, in the case of an online database backup, where the database uses a large amount of memory. Backup servers that receive data from networks rely also on fast connections. If the connection is too slow, a dedicated backup LAN or moving to a SAN architecture could improve the performance.

Application servers without any backup devices depend basically on a good performance of connected networks and disks. In some cases, file systems with millions of small files (such as Windows NTFS) could be an issue.

Backup application

For database applications (such as Oracle, Microsoft SQL and Exchange), use the backup integration provided by those applications, as they are tuned to make best use of their data structures.

Use concurrency (multi-threading) if possible; this allows multiple backups to be interleaved to the tape, thus reducing the effect of slow APIs and disk seeks for each one. Note that this can have an impact on restore times as a particular file set is interleaved among other data.

File system

There is a significant difference between the raw read data rate of a disk and the file system read rate. This is because traversing a file system requires multiple, random disk accesses whereas a continuous read is limited only by the optimized data rates of the disk.

The difference between these two modes becomes more significant as file size reduces. For file systems where the files are typically smaller than 64 KB, sequential backups (such as from raw disk) could be considered to achieve the required data rates for high-speed tape drives.

File system fragmentation could be also an issue. It causes additional seeks and slower throughput.

Disk

If the system performing the backup has a single hard disk drive, the major factor restricting backup performance is most likely to be the maximum transfer rate of the single disk. In a typical environment, maximum disk throughputs for a single spindle can be as low as 8 MB/s.

High capacity disk

A high capacity disk is still one spindle with its own physical limitations. Vendors tend to advertize them as "best price per MB," but a single spindle can cause serious problems in high performance environments.

Two smaller spindles provide twice the performance of one large spindle. The backup performance of large disks may be acceptable without any application load. But if an application writes in parallel to that disk, the total disk performance can go below 5 MB/s and the hit ratio of a disk array read cache below 60%.

Disk Array

Benchmarks have shown that theoretical disk array performance cannot be achieved with standard backup tools. The problem lies in the concurrency of read processes, which cannot be distributed equally among all I/O channels and disk drives. The disk array can be seen as a bunch of disks, where the internal organization and configuration is hidden for the backup software. High capacity disks can cause additional problems that intelligent disk array caches cannot overcome. They are not able to provide reasonable throughput for backup and restore tasks; the number of sequential reads and writes is too high.

The 50% backup performance rule became a standard for disk array sizing.

RAID

The use of RAID for a disk backup should be carefully considered. There are five main levels of RAID each with their own strengths. The raw I/O speed of the disk backup device significantly affects the overall backup performance.

There are five main levels of RAID that are commonly referred to:

- RAID 0 (Striping)
  Data is striped across the available disks, using at least two disk drives. This method offers increased performance, but there is no protection against disk failure. It will allow you to use the maximum performance the disks have to offer, but without redundancy.
- RAID 1+0 (Disk Drive Mirroring and Striping)
  The disks are mirrored in pairs and data blocks are striped across the mirrored pairs, using at least four disk drives. In each mirrored pair, the physical disk that is not busy answering other requests answers any read request sent to the array; this behavior is called load balancing. If a physical disk fails, the remaining disk in the mirrored pair can still provide all the necessary data. Several disks in the array can fail without incurring data loss, as long as no two failed disks belong to the same mirrored pair. This fault-tolerance method is useful when high performance and data protection are more important than the cost of physical disks. The major advantage of RAID 1+0 is the highest read and write performance of any fault-tolerant configuration.
- RAID 1 (Mirroring)
  Data is written to two or more disks simultaneously, providing a spare ”mirror copy“ in case one disk fails. This implementation offers high assurance of disk transfers completing successfully, but adds no performance benefit. It also adds cost, as you do not benefit from the additional disk capacity.
- RAID 3 (Striping and Dedicated Parity)
  This type of array uses a separate data protection disk to store encoded data. RAID 3 is designed to provide a high transfer rate. RAID 3 organizes data by segmenting a user data record into either bit- or byte-sized chunks and evenly spreading the data across several drives in parallel. One of the drives acts as a parity drive. In this manner, every record that is accessed is delivered at the full media rate of the drives that comprise the stripe group. The drawback is that every record I/O stripe accesses every drive in the group. RAID 3 architecture should only be chosen in a case where it is virtually guaranteed that there will be only a single, long process accessing sequential data. Video servers and graphics servers are good examples of appropriate RAID 3 applications. RAID 3 architecture is also beneficial for backups but becomes a poor choice in most other cases due to its limitations.
- RAID 5 (Striping and Distributed Parity)
  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. If one disk fails, all data will be recoverable from the remaining disks due to the parity bit, which allows data recovery. The disk write performance on RAID 5 will be slower than RAID 0 because the parity has to be generated and written to disk.

Tape drive

The mechanical impact on tape drives of providing data at too slow a rate is generally underestimated and can result in slow performance and broken-down tape drives.

The tape drive should be operating at or above its lowest streaming data rate to achieve the best life for the head, mechanism, and tape. If the data is not sent fast enough, the internal buffer will empty and the drive will not write a continuous stream of data. At that point, the drive has to perform head repositioning. This is also known as “shoe shining” and causes excessive wear to the tape, the tape drive heads, and the mechanical tape drive components. Tape drives have buffers in the data path that are large enough to prevent head repositioning from explicitly slowing the backup further, however, the increased error rate from worn heads and media causes more tape to be used and additional retries to be performed. This will slow the backup down and it will get worse over time.

Storage Area Network (SAN)

The standard topology for mid-size and large environments is SAN-based. The SAN has its own components and data paths. Each of them could become a bottleneck:

- Any host bus adapter (of server, disk, tape and tape library bridge)

- SAN switch (total and single port bandwidth)
- Cabling (each cable type has unique performance and distance specifications)

Finding bottlenecks

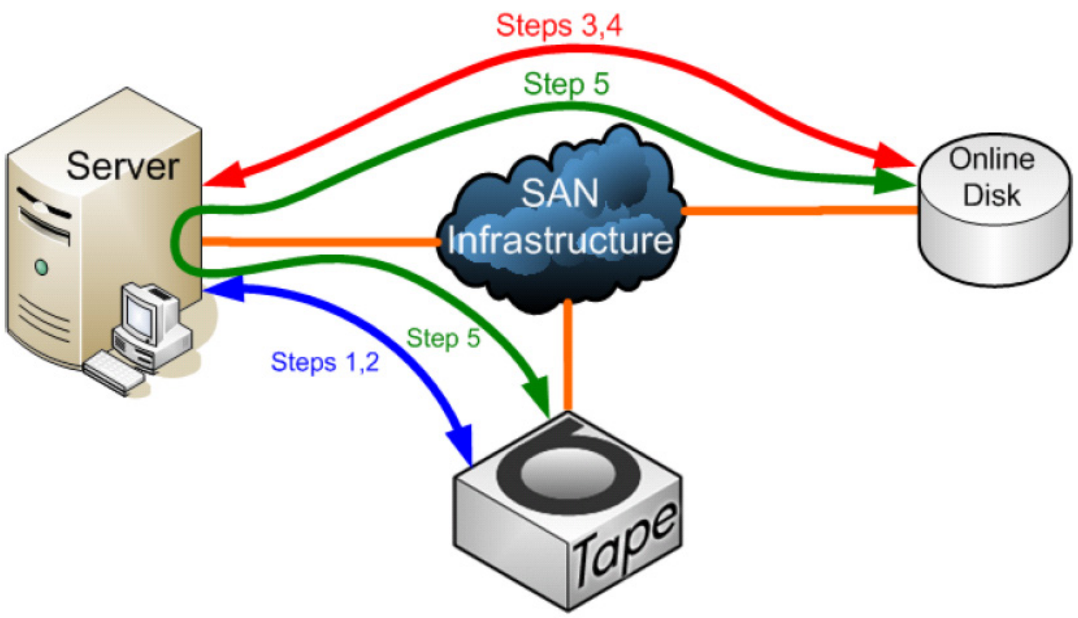
Process for identifying bottlenecks

In order to get the maximum performance of a backup and restore environment, it is important to understand that many elements influence backup and restore throughput.

A process is needed that breaks down a complex SAN infrastructure into simple parts, which can be analyzed, measured, and compared. The results can then be used to plan a backup strategy that maximizes performance.

Figure 7 illustrates an example for identifying bottlenecks in SAN environments.

Figure 7. Example for identifying bottlenecks



The following steps are used to evaluate the throughput of a complex SAN infrastructure:

1. Measure the tape subsystem's WRITE performance.
2. Measure the tape subsystem's READ performance.
3. Measure the disk subsystem's WRITE performance.
4. Measure the disk subsystem's READ performance.
5. Evaluate the backup and restore application's effect on disk and tape performance.

Details of each of the steps are demonstrated in the test environment. See Evaluating tape and disk drive performance on page 31.

With the tests and subsequent analysis on a component level, it is possible to identify bottlenecks in the SAN.

Performance tools

Understanding the performance of a server, its HBAs, CPU, memory, and storage, can be vital in determining the expected performance of a backup solution. Much of the theoretical performance information may be known when sizing the server and its components, but the true performance may not be known until baseline testing is conducted on the system. HP offers tools that help in determining the raw performance of each component involved in the backup and restore solution.

The following performance tools were used:

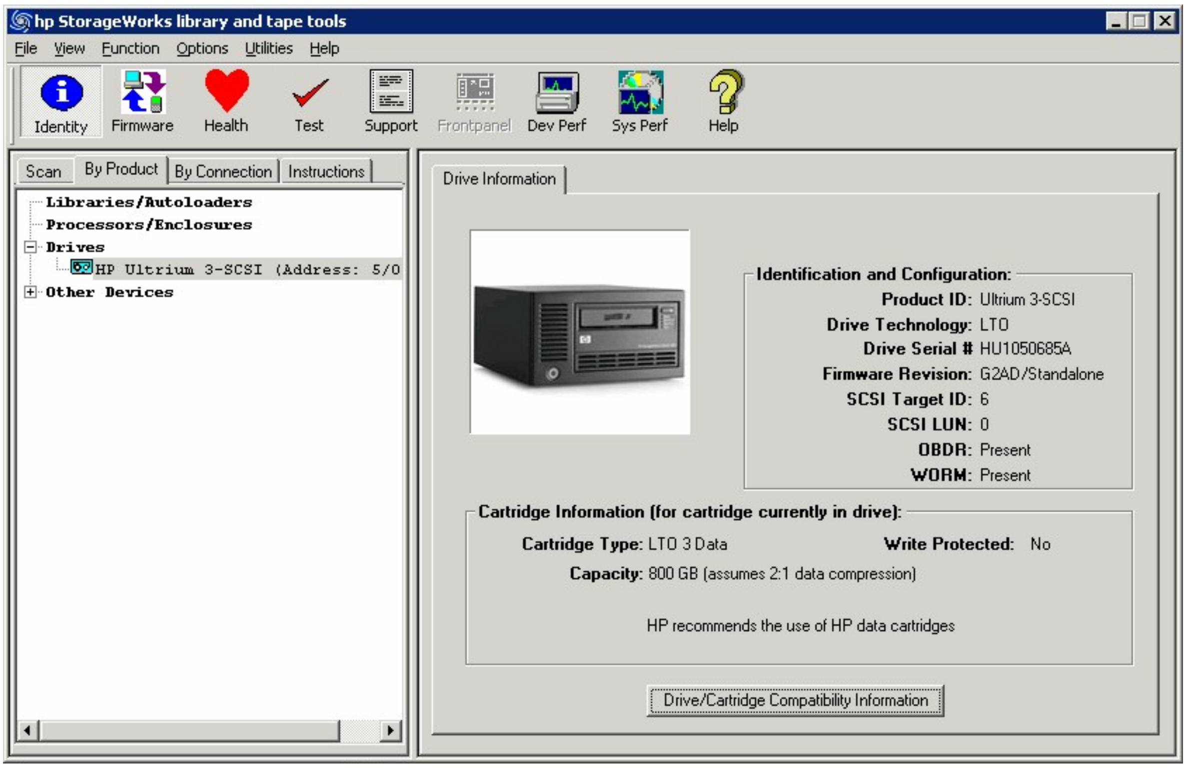
- HP StorageWorks Library and Tape Tools diagnostics (L&TT)

- HPCreateData
- HPReadData
- LoadSim 2003

HP StorageWorks Library and Tape Tools diagnostics (L&TT)

HP's industry-leading Library and Tape Tools diagnostics assist in troubleshooting backup and restore issues in the overall system. They include tools to help identify where bottlenecks exist in a system, and valuable tools for HP tape drive performance and diagnostics needs. The Windows version of L&TT uses a graphical user interface (GUI) as shown in Figure 8. All other versions of the program (such as HP-UX) use a command screen interface (CSI).

Figure 8. Library and Tape Tools Diagnostics GUI – Drive Information

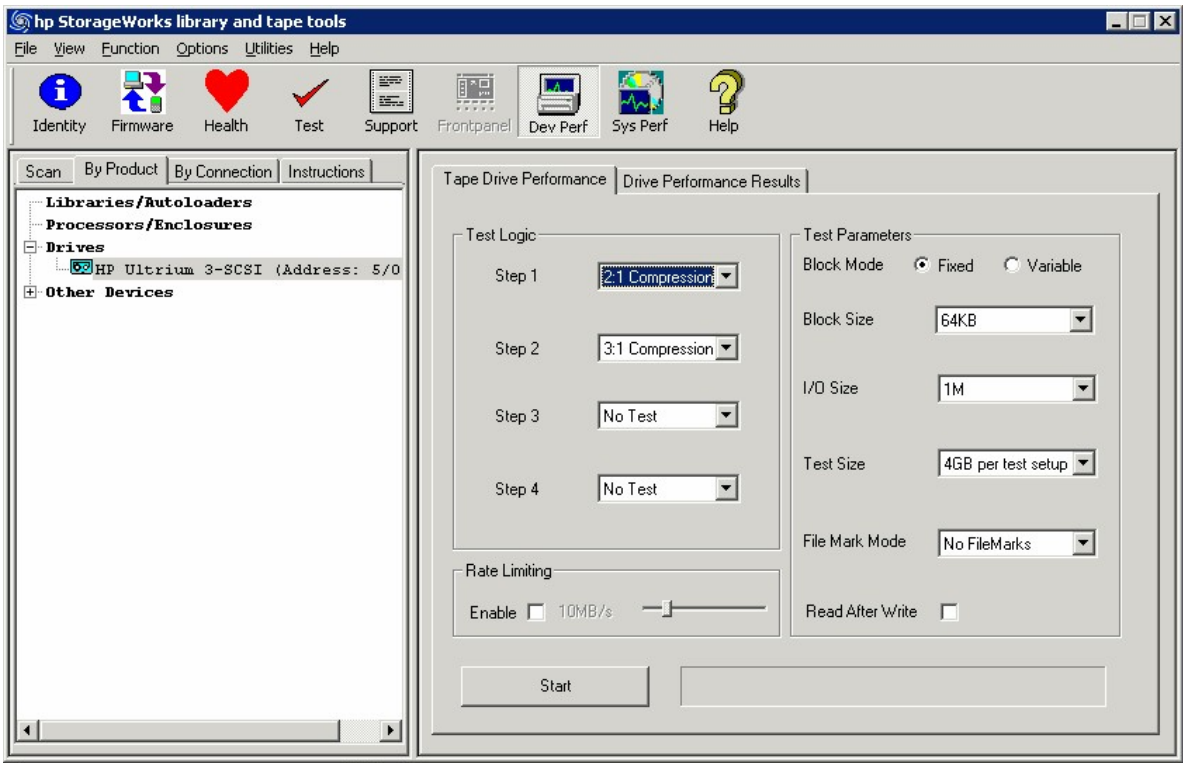


L&TT version 4.2 SR1 was used in this white paper for:

- Determining the rate at which the tape drive reads and writes data
- Determining the rate at which the disk subsystem can write and supply data

The features of L&TT can be seen in the Tape Drive Performance GUI as shown in Figure 9. The embedded tools behind it are HPCreateData and HPReadData.

Figure 9. Library and Tape Tools Diagnostics GUI – Tape Drive Performance



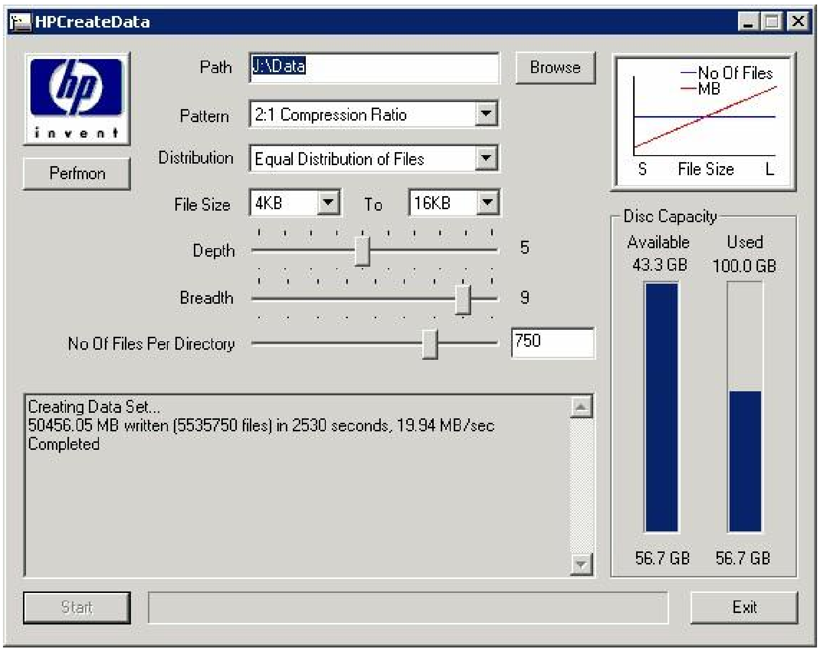
L&TT can be downloaded free from http://www.hp.com/support/tapetools.

HPCreateData

The HPCreateData PAT utility is a file system generator for measuring restore performance. It is useful in assessing the rate at which your disk subsystem can write data, which is ultimately what will limit the restore performance. To write more than one stream, initiate multiple instances of HPCreateData.

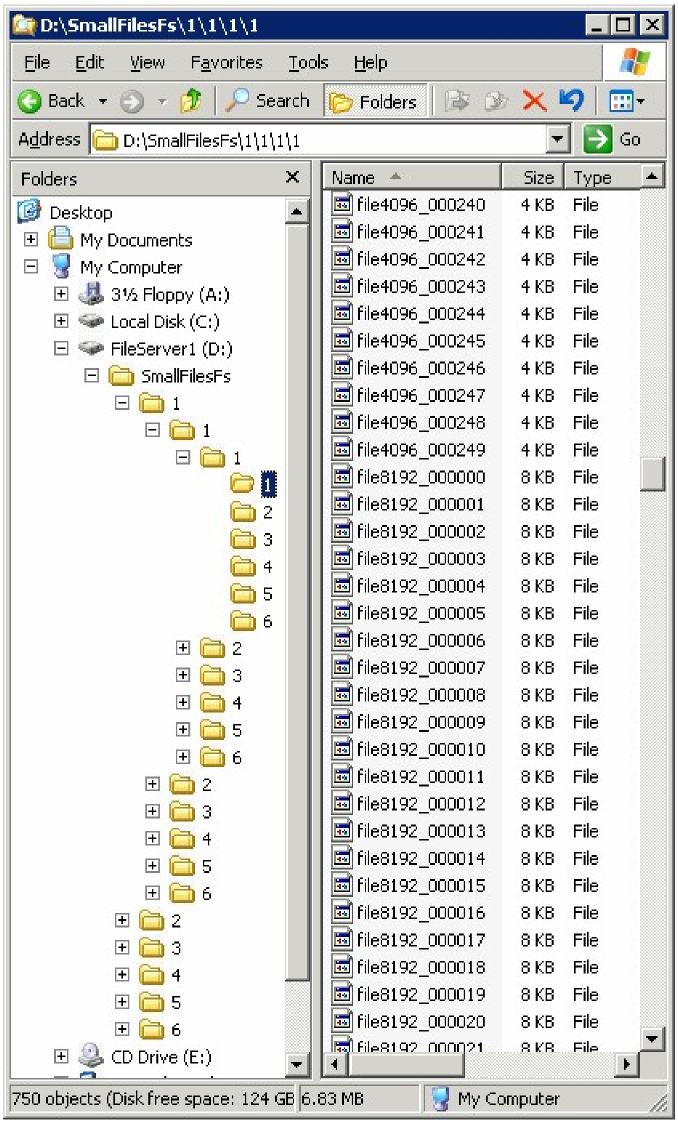
HPCreateData version 1.2.3 was used for testing. Figure 10 shows an example for creating Windows test data.

Figure 10. HPCreateData for Windows example



Executing HPCreateData results in a directory and file structure similar to that shown in Figure 11. In each directory, 750 files were created with a depth of 5 and breadth of 6.

Figure 11. Executing HPCreateData for Windows – directory and file structure



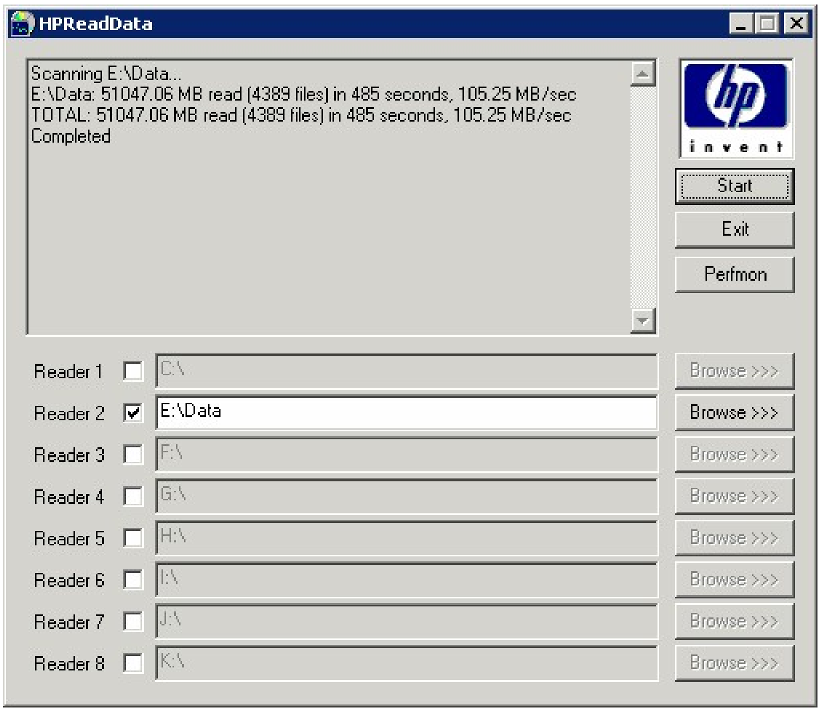
HPCreateData can be downloaded free from http://www.hp.com/support/pat.

HPReadData

The HPReadData PAT utility is useful in assessing the rate at which your disk subsystem can supply data, which is ultimately what will limit backup performance. It simulates the way Data Protector reads files. A single instance of HPReadData can read eight streams simultaneously from your array. To read more than eight streams, initiate multiple instances of HPReadData. HPReadData is available for Windows, HP-UX, Solaris, and Linux.

HPReadData version 1.2.4 was used for testing. Figure 12 shows an example for reading Windows test data.

Figure 12. HPReadData for Windows example



The screenshot shows HPReadData reading 1 LUN from a file system in a manner similar to the way a backup application will read files. You can see that the maximum read rate from this configuration is 105.25 MB/s, so we cannot expect any higher back-up transfer performance to tape than this figure.

Note: For simulating Data Protector disk read agents, start the equivalent number of HPReadData instances, in which each instance should only read from one file system.

HPReadData can be downloaded free from http://www.hp.com/support/pat.

LoadSim 2003

Microsoft Exchange Server 2003 Load Simulator (LoadSim) is a benchmarking tool to simulate the performance load of MAPI clients.

For this white paper it was very useful to use its additional features for data creation:

- Create Topology
- Initialize Test

Figure 13 shows the initial screen of LoadSim 2003.

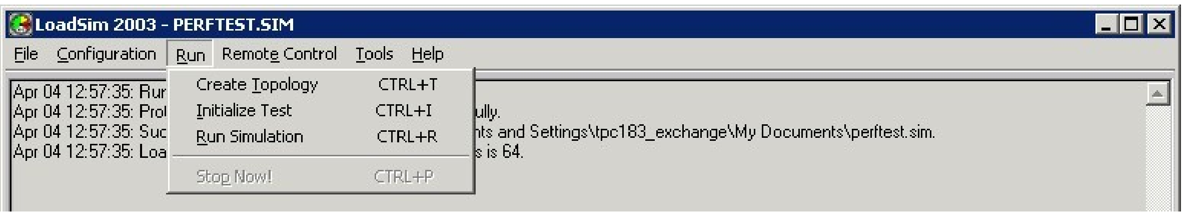


Figure 13. LoadSim 2003 – initial screen

LoadSim provides many configuration parameters for simulating Microsoft Exchange Server environments. If 500 users are initialized with the default initialization values for either the MMB3, Heavy, or Cached Mode profile, the Exchange database will grow to approximately 54 GB.

Figure 14 illustrates the default LoadSim configuration parameters for distribution lists.

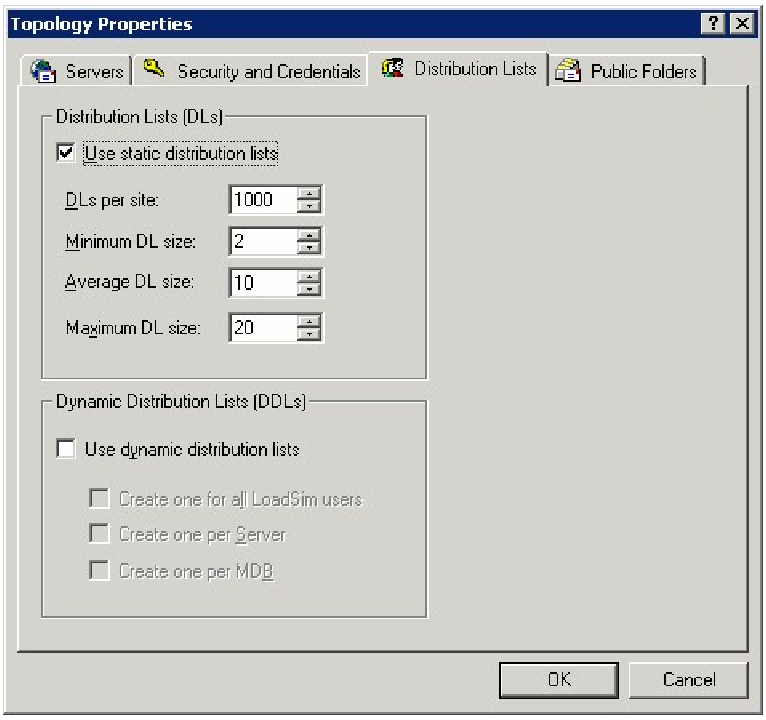


Figure 14. LoadSim – Distribution Lists screen of Topology Properties

You can download LoadSim 2003 from http://go.microsoft.com/fwlink/?LinkId=27882.

Test data

In the following proof points, two different types of file systems and one Microsoft Exchange Server 2003 were created, so that the results shown are realistically achievable in similar situations:

1. Typical file server data with fewer files and a broad range of size (KB/MB)
2. Problematic file server data with millions of small files
3. Typical Microsoft Exchange Server 2003 data

Data creation tools

All test data was created using the following public tools:

- HP L&TT version 4.2 SR1a
- HP HPCreateData version 1.2.3
- Microsoft Exchange Server 2003 Load Simulator (LoadSim 2003)

The datasets were developed using the following utilities:

- HPCreateData for Windows
- L&TT for HP-UX

HPCreateData for HP-UX was not used because it is only available as a CLI and cannot create more than one directory. All files would be created in a single directory, which does not correspond to what occurs in real file server environments.

HPCreateData and L&TT generated different file sizes with different data contents (fixed, random, up to 4:1 compression ratio) and different distribution (file-based, MB-based). All created files had names with a maximum of 16 characters to avoid corner cases.

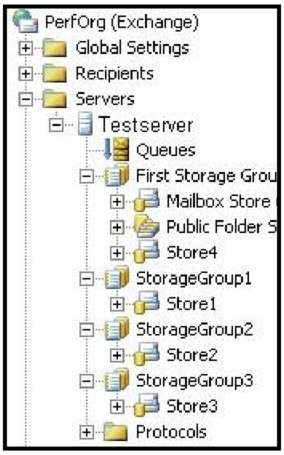
HPCreateData and L&TT are downloadable from http://www.hp.com/support/pat.

LoadSim 2003 was used for creating the Exchange test data. Before it was used, the Exchange organization had been prepared with the Exchange System Manager. For optimal performance, all databases files were moved to separate disk volumes. When data is read or written, the storage system should not become a bottleneck for the test.

LoadSim 2003 is downloadable from http://go.microsoft.com/fwlink/?LinkId=27882.

Figure 15 illustrates an example of an Exchange Server that is configured for LoadSim testing. In this example, the Exchange System Manager was used to add three additional storage groups named StorageGroup1 – StorageGroup3. Also, one mailbox store was added to each storage group. These mailbox stores were named Store1 – Store4.

Figure 15. Exchange System Manager with additional storage groups and additional mailbox stores



Two LoadSim features were used with their default initialization parameters:

1. Create Topology
   After topology parameters were specified, the topology was created. This step creates the LoadSim users and distribution groups (also referred to as DLs or distribution lists) in the active directory on the Exchange Server.
2. Initialize Test
   When the test was initialized, LoadSim added or deleted messages in the Inbox and folders of each user in the test so that each mailbox has the number of messages specified in the Initialization tab of the Customize Tasks window. LoadSim also created new folders, smart folders, rules in Inbox, appointments and contacts.

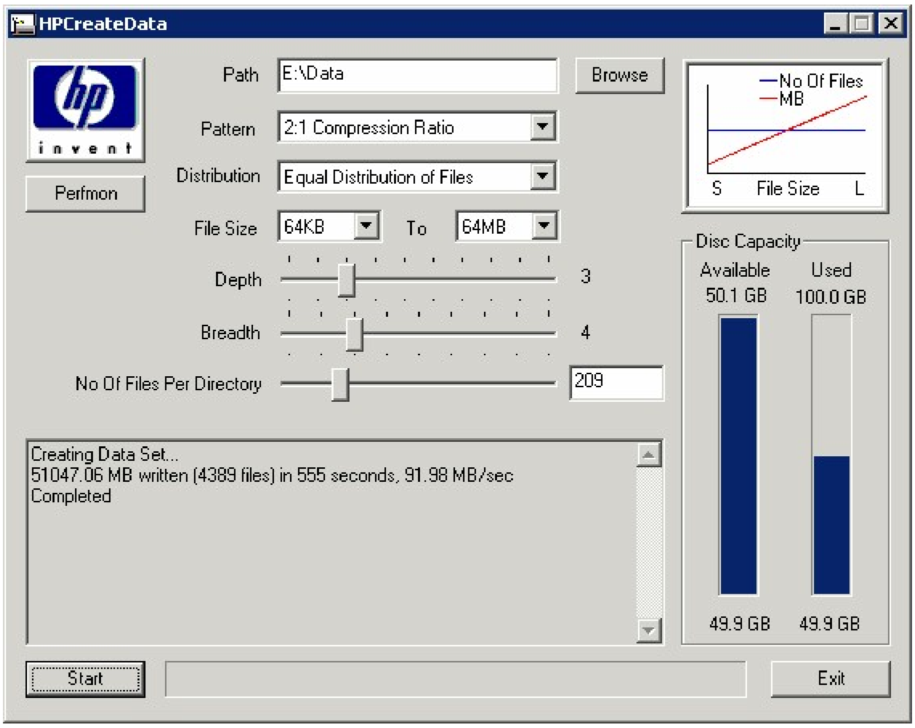
You can download LoadSim 2003 from http://go.microsoft.com/fwlink/?LinkId=27882.

Creating typical files for Windows NTFS

The typical file system was created with file sizes between 64 KB and 64 MB and the compressibility of the data 2:1. The utility created an equal distribution of files in each directory.

Figure 16 shows the HPCreateData input parameters and results.

Figure 16. HPCreateData for Windows – creating typical files



Finally, the file system contained 49.85 GB with 4,389 files in 20 folders.

Creating typical files for HP-UX VxFS

The typical file system for HP-UX VxFS was created with file sizes between 64 KB and 64 MB and the compressibility of the data 2:1. The utility created an equal distribution of files in each directory.

Figure 17 illustrates the L&TT input parameters and Figure 18 the results.

Figure 17. Library and Tape Tools Diagnostics for HP-UX – creating typical files – input parameters

```
Main > System Performance Tools > Restore Pre-Test >
------------------------------------------------------------------------------
The Performance of the disk subsystem can be measured using this tool. This
creates a directory structure based on file size, directory depth, blocksize
and compression ratio and measures the data rates it achieves. Type the
corresponding number to select/set each option.

  1.     Set minimum file size in bytes (min-filesize) - 65536
  2.     Set maximum file size in bytes (max-filesize) - 67108864
  3.     Set no. of files per directory (nofile) - 209
  4.     Set dir where to create dataset (dirpath) - /typical1/data
  5.     Set compressibility pattern to create files.1) Fixed Byte. 2) Incr...
  6.     Set file tree depth (file tree depth) - 3
  7.     Set file tree breadth (file tree breadth) - 4
```

Figure 18. HP StorageWorks Library and Tape Tools Diagnostics (L&TT) for HP-UX – creating typical files – results

```
Main > System Performance Tools > Restore Pre-Test > Viewer >
-------------------------------------------------------------
51047 MB written (4389 files) in 307 seconds at 166 MB/sec
Restore Pre-Test successfully created files.
166.28 MB/sec
```

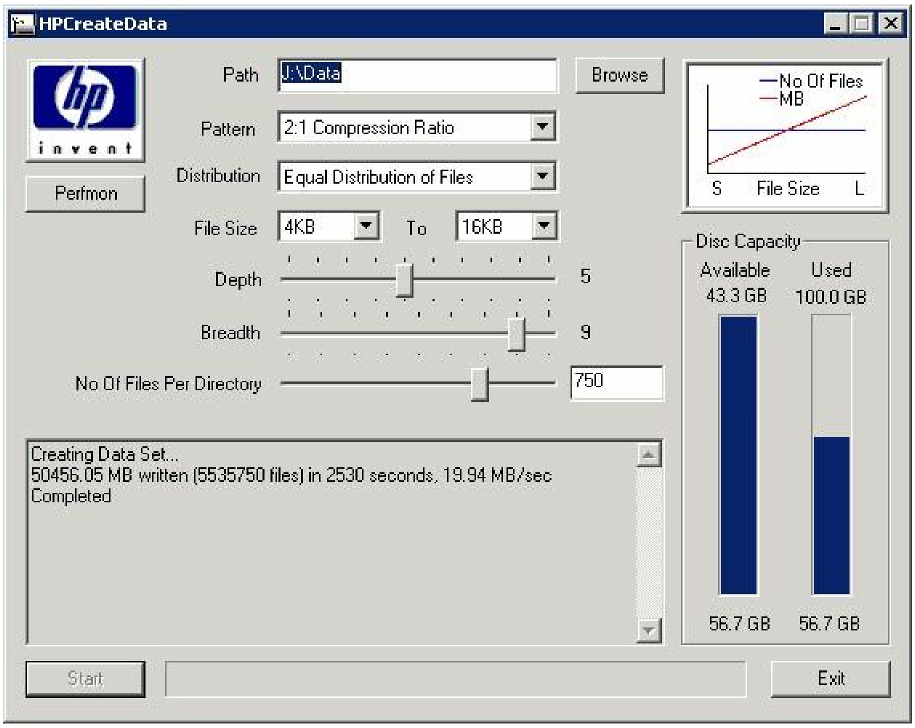
Finally, the file system contained 49.85 GB with 4,389 files in 20 folders.

Creating millions of small files for Windows NTFS

The file system with millions of small files for Windows NTFS was created with file sizes between 4 KB and 16 KB and the compressibility of the data 2:1. The utility created an equal distribution of files in each directory.

Figure 19 illustrates the HPCreateData input parameters and results.

Figure 19. HPCreateData for Windows – creating small files



Finally, the file system contained 49.27 GB with 5,535,750 files in 7,380 folders.

Creating millions of small files for HP-UX VxFS

The file system with millions of small files for HP-UX VxFS was created in a similar manner to Windows. The file sizes were between 4 KB and 16 KB and the compressibility of the data 2:1. The utility created an equal distribution of files in each directory.

Note: The number of files per directory was set to 120 due to limitations of L&TT version 4.2 SR1a. The restore pre-test was executed six times in order to achieve a HP-UX VxFS layout which was similar to the Windows NTFS layout as described in the previous section Creating millions of small files for Windows NTFS. In later L&TT releases, the number of files should be increased to 750.

Figure 20 illustrates the L&TT input parameters and Figure 21 the results.

Figure 20. Library and Tape Tools Diagnostics for HP-UX – creating small files – input parameters

```
Main > System Performance Tools > Restore Pre-Test >
--------------------------------------------------------------------------------
The Performance of the disk subsystem can be measured using this tool. This
creates a directory structure based on file size, directory depth, blocksize
and compression ratio and measures the data rates it achieves. Type the
corresponding number to select/set each option.

  1.     Set minimum file size in bytes (min-filesize) - 4096
  2.     Set maximum file size in bytes (max-filesize) - 16384
  3.     Set no. of files per directory (nofile) - 120
  4.     Set dir where to create dataset (dirpath) - /small1/data1
  5.     Set compressibility pattern to create files.1) Fixed Byte. 2) Incr...
  6.     Set file tree depth (file tree depth) - 5
  7.     Set file tree breadth (file tree breadth) - 9
```

Figure 21. Library and Tape Tools Diagnostics for HP-UX – creating small files – results

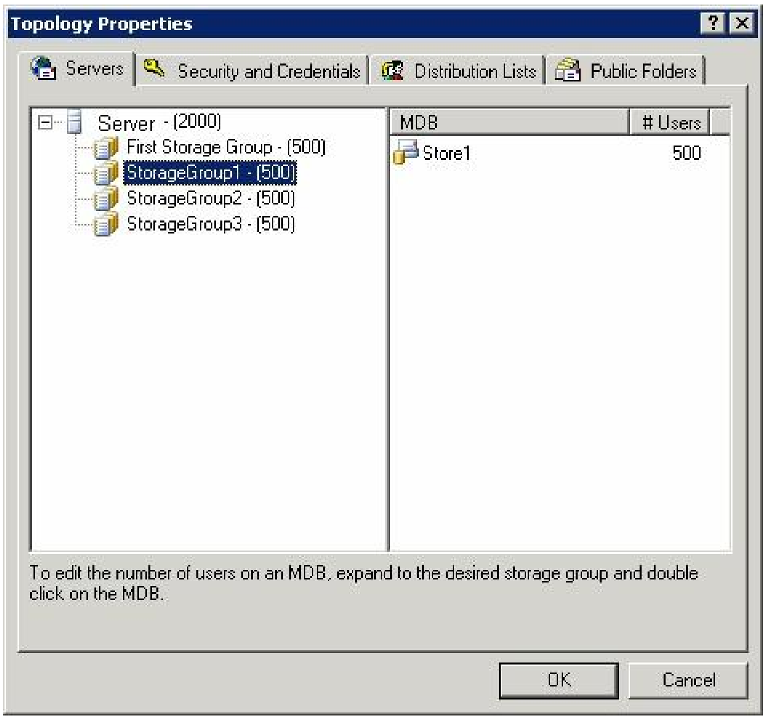
```
Main > System Performance Tools > Restore Pre-Test > Viewer >
--------------------------------------------------------------
8072 MB written (885720 files) in 161 seconds at 50 MB/sec
Restore Pre-Test successfully created files.
8072 MB written (885720 files) in 156 seconds at 51 MB/sec
Restore Pre-Test successfully created files.
8072 MB written (885720 files) in 157 seconds at 51 MB/sec
Restore Pre-Test successfully created files.
8072 MB written (885720 files) in 158 seconds at 51 MB/sec
Restore Pre-Test successfully created files.
8072 MB written (885720 files) in 156 seconds at 51 MB/sec
Restore Pre-Test successfully created files.
8072 MB written (885720 files) in 154 seconds at 52 MB/sec
Restore Pre-Test successfully created files.
```

Finally, the file system contained 47.29 GB with 5,314,320 files in 44,286 folders.

Creating data for Microsoft Exchange Server 2003

LoadSim was configured for simulating 2,000 users. Four equal storage groups were created with one store each as illustrated in Figure 22.

Figure 22. LoadSim Topology for simulating 2,000 users



The LoadSim initialization resulted in four storage groups with one 50 GB database each. The total size, which is relevant for backups and restores, was 200 GB.

Evaluating tape and disk drive performance

The performance test of tape and disk drives gives a good overview of what the source and target devices are able to provide. Backup applications cannot perform better than these basic tools.

Tape write and read performance

The tape drive performance was determined with the HP StorageWorks Library and Tape Tools Diagnostics (L&TT).

L&TT for Windows was configured for writing and reading:

- Zeros with fixed block mode, 256 KB block size, 1M I/O size, 32 GB test size and no file marks
- 2:1 compressible data with fixed block mode, 256 KB block size, 1M I/O size, 32 GB test size and no file marks

Figure 23 illustrates the L&TT for Windows input parameters and Figure 24 some test results.

Figure 23. Library and Tape Tools Diagnostics for Windows – input parameters

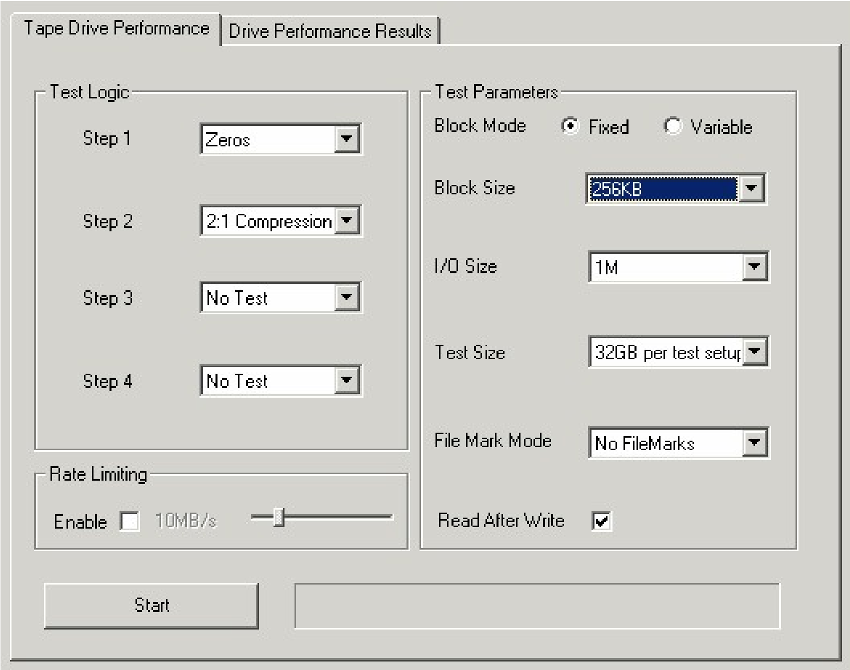
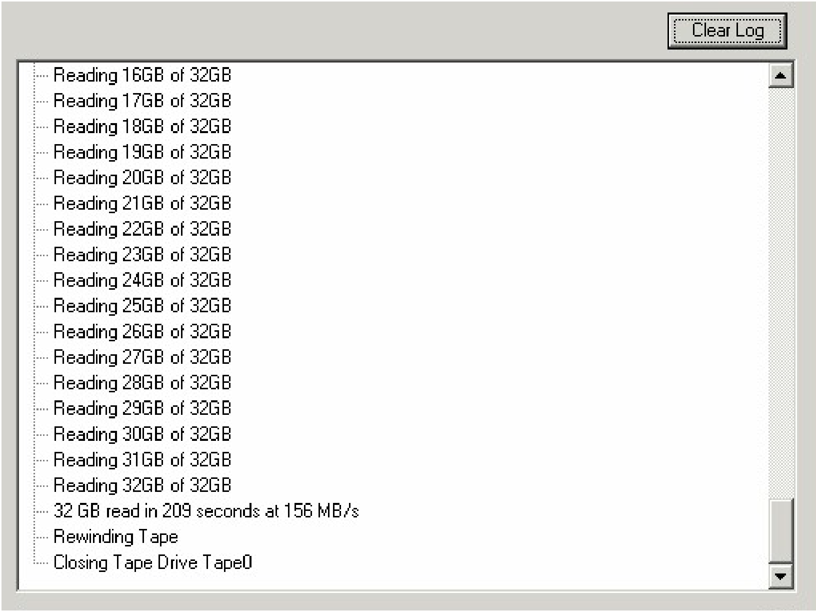


Figure 24. Library and Tape Tools Diagnostics for Windows – test results



L&TT for HP-UX was configured for writing and reading with the following parameters:

- 4.3:1 compressible data with 128 KB block size, 32768 repeats and 8 blocks
- 2:1 compressible data with 128 KB block size, 32768 repeats and 8 blocks

Note: L&TT for HP-UX does not offer zero patterns. Therefore, the highest available compression pattern of 4.3:1 was specified.

Figure 25 illustrates the L&TT for HP-UX input parameters and Figure 26 some test results.

Figure 25. Library and Tape Tools Diagnostics for HP-UX – input parameters

```
Main > Drive Information > Device Performance Tool >
--------------------------------------------------------------------------------
The purpose of this tool is to measure the tape drive performance by checking
how fast the server can access the tape drive independent of the disk
subsystem. It works by transferring pre-loaded data with different compression
ratios directly from the memory of the server to the tape drive.

  1.     Set the blocksize to write to tape (blocksize) - 131072
  2.     Set the no. of writes/reads to perform (reps) - 32768
  3. [X] Enables the readback test after writing data to the device (readback)
  4.     Set the no. of blocks per read/write to tape (count) - 8
  5.     Pattern (1)-uncompressible (random), (2)-2/1 compressible, (3)-3/1...
```

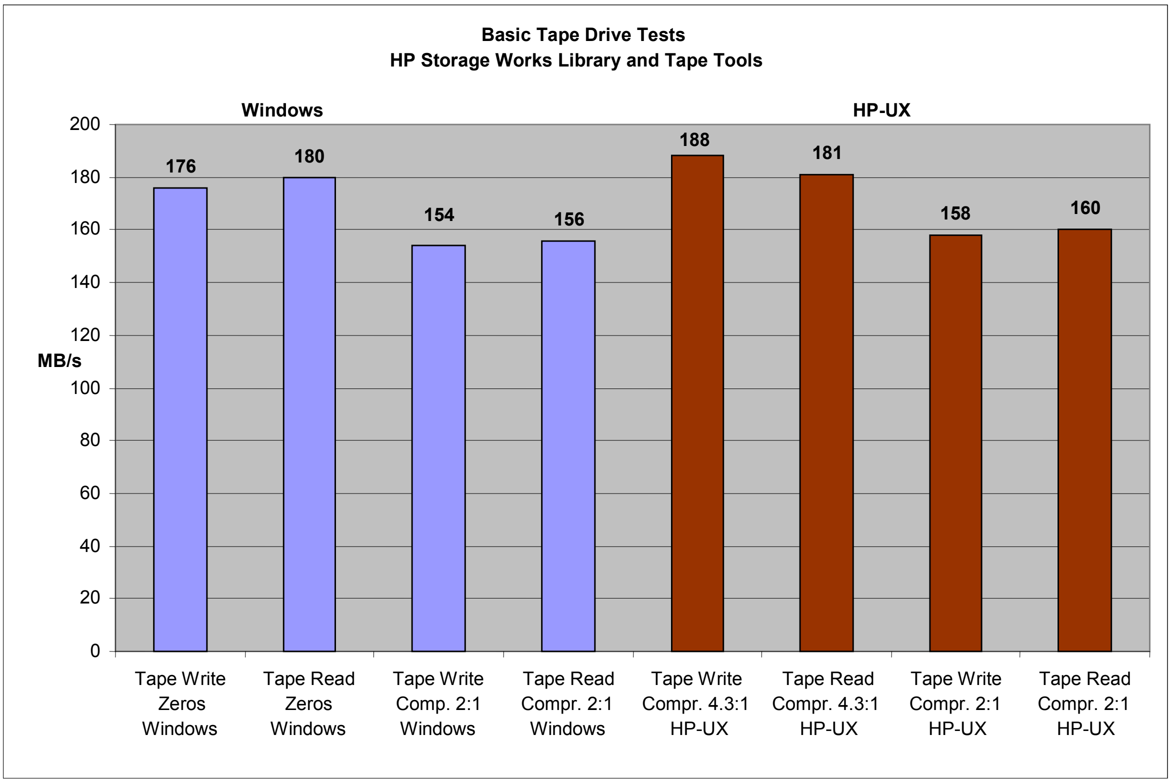
Figure 26. Library and Tape Tools Diagnostics for HP-UX – test results

```
Compression 4.3/1
Opening Tape Drive 2.5.0[0-20/0/2/1/0.5.0]
Successfully opened the Tape Drive /dev/rmt/0mnb
34359 MB written in 182.7644 seconds at 188.0002 MB/s
34359 MB read in 189.7036 seconds at 181.1233 MB/s
Device performance tool has successfully completed execution.

Compression 2/1
Opening Tape Drive 2.5.0[0-20/0/2/1/0.5.0]
Successfully opened the Tape Drive /dev/rmt/0mnb
34359 MB written in 216.7234 seconds at 158.5419 MB/s
34359 MB read in 213.7984 seconds at 160.7110 MB/s
Device performance tool has successfully completed execution.
```

The results of all L&TT tape drive tests are illustrated in Figure 27, which shows the performance limits of a direct attached Ultrium 960 tape drive. This figure is a reference for any later Data Protector test. The tape drive cannot be faster as with L&TT. This applies to any backup application.

Figure 27. Library and Tape Tools – results of tape drive tests



The tests revealed that the Ultrium 960 tape drive performs slightly better in the HP-UX than in the Windows environment, but the difference is only marginal. Both environments demonstrated that they are fully able to use the Ultrium 960 tape drive.

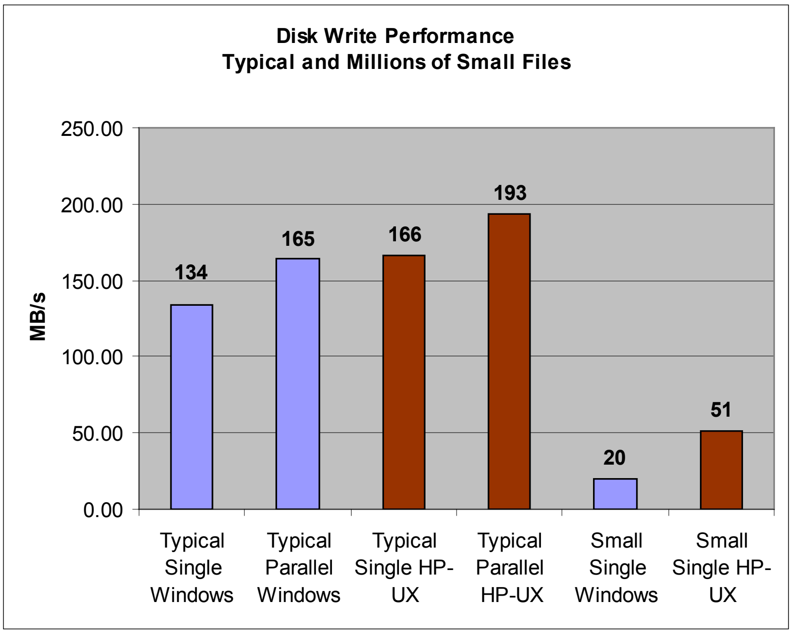
Disk write performance

The disk write performance was determined with L&TT for HP-UX and with the HPCreateData utility for Windows.

Note: The performance data of parallel writes is determined by the data volume divided by the time of the slowest write process. This approach simulates a backup application that does not finish until the last byte is written.

Figure 28 demonstrates that single writes of typical files (one stream to one disk volume) already perform well compared to multiple writes (five streams to five disk volumes). With Windows, single writes of small files (4 – 16 KB) are very slow because the NTFS file system overhead is much higher than with HP-UX VxFS.

Figure 28. Results of disk write tests



For Windows, parallel writes (five streams to five disk volumes) of small files failed due to problems with an overflow of the Windows system paged pool. This cannot be solved without tuning the Windows kernel, which is beyond the scope of this white paper.

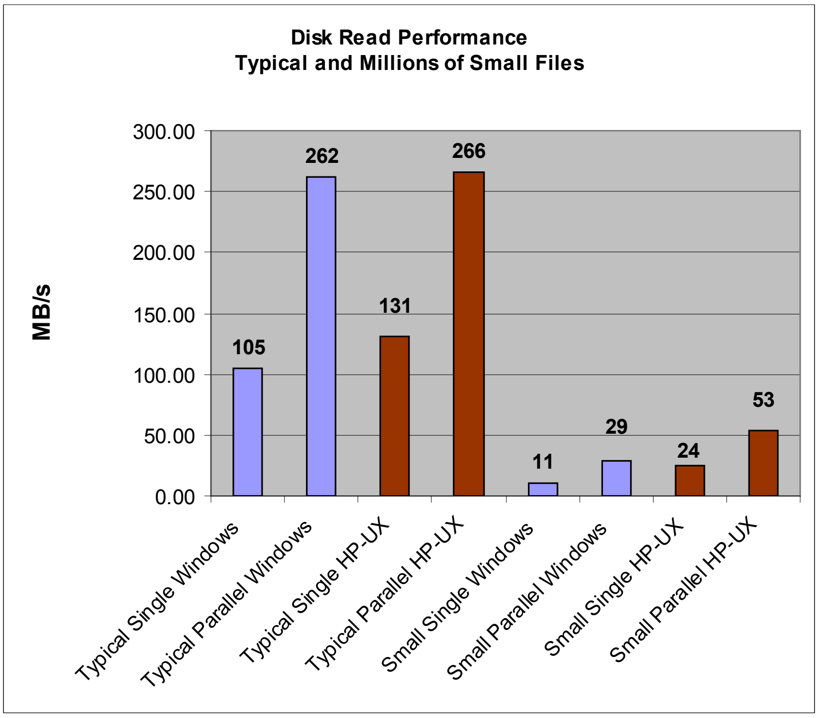
Note: For Windows, it is recommended that file systems with millions of small files are only restored in a single stream.

Disk read performance

The disk read performance was determined for HP-UX with L&TT, and for Windows with the HPReadData utility.

Figure 29 demonstrates that multiple reads (five streams from five disk volumes) perform much better than single reads (one stream from one disk volume). With Windows, single reads of small files (4 – 16 KB) are very slow because the NTFS file system overhead is much higher than with HP-UX VxFS.

Figure 29. Results of disk read tests



Note: The performance data of parallel reads is determined by the data volume divided by the time of the slowest read process. This approach simulates a backup application which does not finish until the last byte is read.

Evaluating Data Protector performance

This section highlights Data Protector performance, and uses the results of the previous sections for finding the bottleneck and giving recommendations. The procedures of this section should be transferable to many customer environments.

The Data Protector performance was tested and analyzed for two different types of file server and a classic Microsoft Exchange Server 2003:

- Typical file server data with a common amount of files and a broad range of size (KB/MB)
- Problematic file server data with millions of small files (KB)
- Typical Microsoft Exchange Server 2003 data

The performance tests were executed by configuring single and multiple backup/restore streams – also called *concurrency* or *multiplexing*. The results of single and multiple stream tests were compared and analyzed. In some cases, multiple streams could be slower than single streams.

The tests included the measurement of performance (MB/s), CPU load (%) and memory usage (MB) of the backup server and, if applicable, the client. Windows performance data was measured with the built-in tool Perfmon and HP-UX performance data with the built-in tool vmstat.

Note: Memory usage is *not* documented in the following sections because all tests showed that the memory usage by Data Protector itself was very little and never exceeded 46 MB. For example, for the HP-UX local backup of typical files, Data Protector did not use more than 19 MB memory, and for the worst case with the Windows local backup of typical files, not more than 46 MB memory. These days, 46 MB usage is not relevant for servers that have gigabytes of memory.

Note: The Cell Manager performance was *not* logged because it is beyond the scope of this white paper. During testing, some quick Cell Manager performance checks confirmed that the Cell Manager was far away from becoming a bottleneck. The CPU and IDB load was very little, even while saving and restoring the file server with millions of small files. The test scenarios of this paper were not sufficient to put the Cell Manager under pressure.

Data Protector configuration

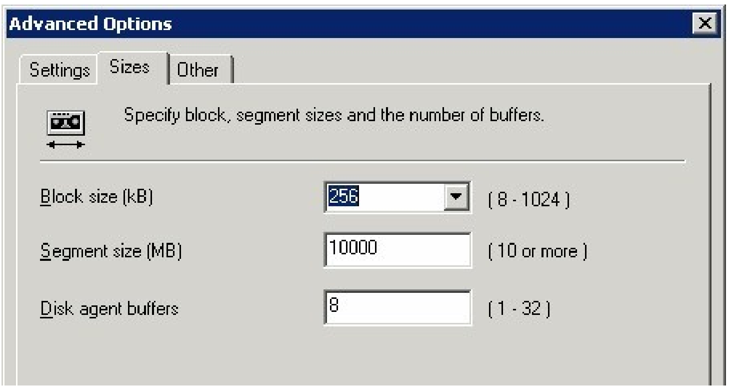
The backup and restore tests were executed with the following changes of Data Protector's default configuration:

- 256 KB tape drive block size as described in the section HP StorageWorks Ultrium 960 tape drive configuration on page 38
- Only one file system tree walk for Windows NTFS file systems with millions of small files as described in the section File system tree walk on page 38

HP StorageWorks Ultrium 960 tape drive configuration

For all tests with the Ultrium 960 tape drive, the block size was configured in Data Protector with 256 KB (64 KB default) as shown in Figure 30. See *Getting the most performance from your HP StorageWorks Ultrium 960 tape drive white paper* (downloadable from http://h18006.www1.hp.com/storage/tapewhitepapers.html) for further details.

Figure 30. HP StorageWorks Ultrium 960 configuration – drive block size 256 KB (64 KB default)



File system tree walk

File system tree walks are typically problematic in Windows NTFS file systems with millions of small files. UNIX file systems are less sensitive and respond much faster.

For each file system backup, Data Protector executes two tree walks by default:

1. The first file system tree walk is required for:
   a. Running backup statistics
      By default Data Protector creates backup statistics during runtime. The tree walk scans the files selected for the backup and calculates its size, so that the progress (percentage done) can be calculated and displayed in the Data Protector Monitor GUI.
   b. Detecting Windows NTFS hard links
      NTFS hard links are detected if the advanced WINFS file system option Detect NTFS Hardlinks is selected (*default:* OFF).
   c. Detecting UNIX POSIX hard links
      POSIX hard links are detected and backed up as links if the advanced UNIX file system option Backup POSIX Hard Links as Files is not selected (*default:* OFF).
2. The second file system tree walk is required for cataloging, indexing and saving the selected files.

Note: Disabling the first file system tree walk is only recommended in the following scenarios:

- Running unattended backup sessions
- Backing up millions of small files from Windows NTFS or UNIX file systems with the requirement of an overall runtime decrease
- If the Windows file system does not keep any NTFS or POSIX hard links
  Otherwise Data Protector would back up the entire file contents for each hard link, which would occupy more space on the backup media.

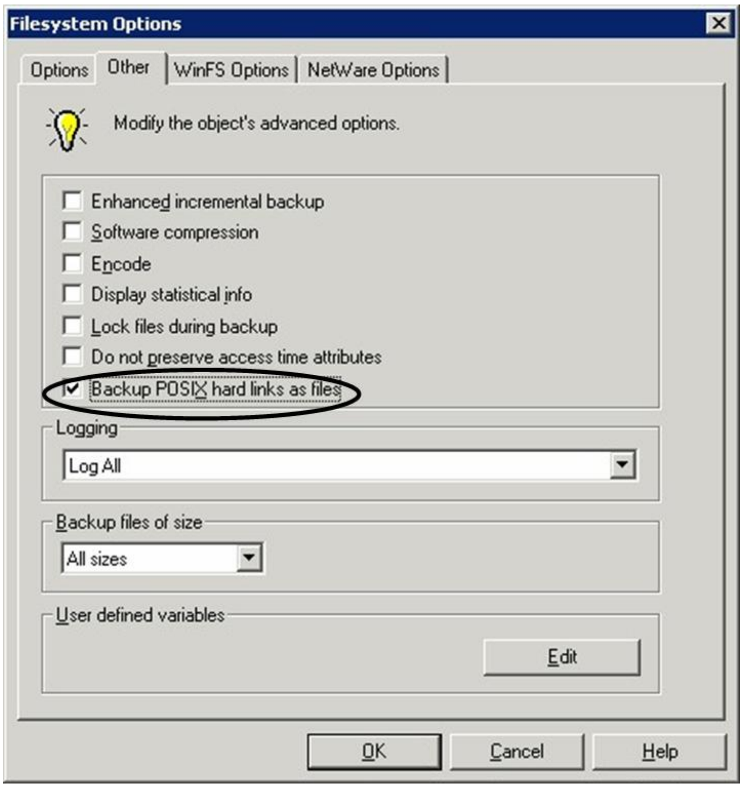
How to disable the first tree walk for Windows NTFS:

If the advanced WINFS file system option Detect NTFS Hardlinks is not set (OFF), the first tree walk can be disabled in the client's OMNIRC file (`Data_Protector_home\omnirc`):
DP 6.0: `NoTreeWalk=1`
DP 6.1: `OB2NOTREEWALK=1`

How to disable the first tree walk for UNIX file systems:

If the advanced UNIX file system option Backup POSIX Hard Links as Files is selected (ON), the first tree walk of UNIX file systems is disabled and POSIX hard links are backed up as files. Figure 31 illustrates the correct configuration for disabling the tree walk.

Figure 31. Disabling the first tree walk for UNIX file systems – Backup POSIX hard links as files



Data Protector IDB considerations

Note: The IDB logging level was not changed from its default value of "Log All". This logging level is the worst case for the IDB while each file is tracked. See the section IDB logging level on page 54 for further details.

The tests in the following section showed that more tape space was consumed for small files than for typical files. The backup of 5,535,750 small files and 249.36 GB total volume resulted in 267.32 GB tape space usage. The overhead was caused by the high number of small files and the required catalog information for each single file written to tape (its file name and attributes).

Backup of typical and small files

This section covers the local and network backup of typical and small files. For Windows NTFS file systems with millions of small files, the first tree walk was disabled as described in the section File System Tree Walk on page 38. For UNIX file systems, the first tree walk was always enabled (default) because it is less critical.

Figure 32 illustrates the effect of the first file system tree walk during the HP-UX network backup of small files. At the beginning, the Windows server loaded the tape and waited for data from the remote HP-UX server, which was busy with the execution of the HP-UX file system tree walk.

Figure 32. CPU Load of the Windows backup server during the HP-UX network backup of millions of small files

Local backup of typical files

The typical files were saved to the SCSI-attached Ultrium 960 tape drive.

For Windows, test 1 in Table 1 shows that that the tape device wrote at 154.00 MB/s. This was sufficient for backing up a single file system as shown in test 2 with 105.25 MB/s. But when backing up multiple file systems, as shown in test 3, the file systems were able to provide 262.24 MB/s, which was faster than the Ultrium 960 tape drive. Test 5 showed that the Data Protector performance of 151.69 MB/s was very close the to the tape drive performance of test 1 with 154.00 MB/s. In test 5, the tape drive was able to stream with its highest performance and therefore it became the bottleneck. This is a very good example of a well-balanced environment for an Ultrium 960 tape drive.

For HP-UX, test 1 in Table 1 shows that the results were much better than for Windows but that the same rules applied. The Ultrium 960 tape drive was the bottleneck.

The CPU load was low except during the HP-UX NULL device backup of test 4, which resulted in 43% and the excellent backup performance of 368.76 MB/s. Test 5, with the backup to the Ultrium 960 tape drive, did not stretch the CPU resources. The CPU load was only 13% for Windows and 16% for HP-UX.

Table 1. Local backup of typical files – bottleneck determination

Test	*Performance (MB/s)*	*CPU Load*	*Bottleneck?*
1. Windows L&TT Tape Write[1]	154	-	Yes (Tape)
2. Windows HPReadData Single[1]	105	-	No
3. Windows HPReadData Parallel[1]	262	-	No
4. Windows DP NULL Parallel	197	18%	No
5. Windows DP Ultrium 960 Parallel	152	13%	No
1. HP-UX L&TT Tape Write[1]	159	-	Yes (Tape)
2. HP-UX HPReadData Single[1]	131	-	No
3. HP-UX HPReadData Parallel[1]	266	-	No
4. HP-UX DP NULL Parallel	3696	43%	No
5. HP-UX DP Ultrium 960 Parallel	156	16%	No

Recommendation: When backing up typical files locally to the SCSI-attached Ultrium 960 tape drive, parallel backups (multiplexing/concurrency) are recommended because one single stream cannot fully use the tape drive.

[1] Tested in the section Evaluating tape and disk drive performance on page 32

Local backup of small files

The small files were saved to the SCSI-attached Ultrium 960 tape drive.

Table 2 shows that file systems with millions of small files were not able to fully use the Ultrium 960 tape drive. The file systems were the bottleneck for both operating systems. This also had some impact on Data Protector because the large amount of file information that had to be written to tape and into Data Protector's IDB resulted in an additional performance loss of 6.04 MB/s for Windows (test 3 with 29.02 MB/s, test 5 with 22.98 MB/s = 6.04 MB/s). See the section Data Protector IDB Considerations on page 39 for further details.

The CPU load of tests 4 and 5 was higher for Windows (61%) than for HP-UX (25% and 16%). This is the result of the inefficient Windows NTFS file system when dealing with millions of small files.

Note: For Windows tests 4 and 5, the first file system tree walk was disabled which improved the overall performance. Without this, the performance would have been even worse.

Table 2. Local backup of small files – bottleneck determination

Test	*Performance (MB/s)*	*CPU Load*	*Bottleneck?*
1. Windows L&TT Tape Write[2]	154.00	-	No
2. Windows HPReadData Single[2]	10.59	-	Yes (File System)
3. Windows HPReadData Parallel[2]	29.02	-	Yes (File System)
4. Windows DP NULL Parallel	22.54	61%	No
5. Windows DP Ultrium 960 Parallel	22.98	61%	No
1. HP-UX L&TT Tape Write[2]	158.54	-	No
2. HP-UX HPReadData Single[2]	24.41	-	Yes (File System)
3. HP-UX HPReadData Parallel[2]	53.08	-	Yes (File System)
4. HP-UX DP NULL Parallel	50.50	25%	No
5. HP-UX DP Ultrium 960 Parallel	48.76	16%	No

Recommendation: If backing up file systems with millions of small files, parallel backups with a higher concurrency are recommended. These kinds of file systems are very slow and should be multiplexed. Instead of the fast Ultrium 960 tape drive, consider using a slower Ultrium tape drive or backup-to-disk technology.

[2] Tested in the section Evaluating tape and disk drive performance on page 32

Network backup of typical files

The typical files were saved from the client server via the network (Gigabit Ethernet) to the remote backup server and its SCSI-attached Ultrium 960 tape drive.

Table 3 shows that the tape device in test 1 and the disk device in test 3 were faster than Data Protector in Tests 4 and 5. In this scenario, both operating systems showed the same backup performance via network, which was the bottleneck. The Gigabit Ethernet itself has a 1,000 Mb/s or 120 MB/s limitation which is very close to the results of test 5 (Windows 108.29 MB/s and HP-UX 111.84 MB/s).

Table 3. Network backup of typical files – bottleneck determination

Test	*Performance (MB/s)*	*CPU Load Client*	*CPU Load Backup Server*	*Bottleneck?*
1. Windows L&TT Tape Write[3]	154.00	-	-	No
2. Windows HPReadData Single[3]	105.25	-	-	No
3. Windows HPReadData Parallel[3]	262.24	-	-	No
4. Windows DP NULL Parallel	108.16	12%	12%	Yes (Network)
5. Windows DP Ultrium 960 Parallel	108.29	16%	15%	Yes (Network)
1. HP-UX L&TT Tape Write[3]	158.54	-	-	No
2. HP-UX HPReadData Single[3]	130.58	-	-	No
3. HP-UX HPReadData Parallel[3]	265.75	-	-	No
4. HP-UX DP NULL Parallel	112.83	26%	26%	Yes (Network)
5. HP-UX DP Ultrium 960 Parallel	111.84	24%	26%	Yes (Network)

Recommendation: If you are backing up typical files from a fast disk via Gigabit Ethernet to a remote Ultrium 960 tape drive, you could consider backups without multiplexing and concurrency. In this case, the network is the bottleneck.

[3] Tested in the section Evaluating tape and disk drive performance on page 32

Network backup of small files

The small files were saved via the network (Gigabit Ethernet) to the remote backup server and its SCSI-attached Ultrium 960 tape drive.

Table 4 shows that file systems with millions of small files were not able to fully use the remote Ultrium 960 tape drive with the network (Gigabit Ethernet) limit of 120 MB/s. The file systems were the bottleneck for both operating systems.

Note: For Windows tests 4 and 5, the first file system tree walk was disabled which improved the overall performance. Without this, the performance would have been even worse.

Table 4. Network backup of small files – bottleneck determination

Test	*Performance (MB/s)*	*CPU Load Client*	*CPU Load Backup Server*	*Bottleneck?*
1. Windows L&TT Tape Write[4]	154.00	-	-	No
2. Windows HPReadData Single[4]	10.59	-	-	No
3. Windows HPReadData Parallel[4]	29.02	-	-	No
4. Windows DP NULL Parallel	22.02	3%	62%	Yes (Network)
5. Windows DP Ultrium 960 Parallel	22.25	3%	62%	Yes (Network)
1. HP-UX L&TT Tape Write[4]	158.54	-	-	No
2. HP-UX HPReadData Single[4]	24.41	-	-	No
3. HP-UX HPReadData Parallel[4]	53.08	-	-	No
4. HP-UX DP NULL Parallel	49.22	32%	21%	Yes (Network)
5. HP-UX DP Ultrium 960 Parallel	48.89	31%	21%	Yes (Network)

Recommendation: Parallel backups with a higher concurrency are recommended for file systems with millions of small files – also if you are backing up via the network (Gigabit Ethernet). These kinds of file systems are very slow and should be multiplexed. In the case of the fast Ultrium 960 tape drive and Gigabit Ethernet, you could consider a slower Ultrium tape drive or backup-to-disk technology.

[4] Tested in the section Evaluating tape and disk drive performance on page 32

Restore of typical and small files

This section covers local and network-based restore of typical and small files.

Local restore of typical files

The files were restored directly from the SCSI-attached Ultrium 960 tape drive.

For Windows, test 1 in Table 5 shows that the tape device did not read faster than 156.00 MB/s. Test 3 shows that the disk device resulted in164.67 MB/s write performance and test 4 in 145.07 MB/s restore performance for Data Protector. If the performance of two devices is so close, the tape has sometimes to wait for I/O which explains the slightly slower Data Protector performance. Note that a disk device does not show a completely consistent performance from the operating system's point of view. It will always fluctuate, and monitoring tools just display the average value in a configured timeframe. This can be verified, for example with the built-in performance tool Perfmon for Windows.

For HP-UX, test 1 in Table 5 shows that the tape device result of 160.71 MB/s was less than the disk device result in test 3 (193.07 MB/s). Here, the tape device was the bottleneck.

Table 5. Local restore of typical files – bottleneck determination

Test	*Performance (MB/s)*	*Bottleneck?*
1. Windows L&TT Tape Read[5]	156.00	Yes (Tape)
2. Windows HPCreateData Single[5]	133.98	No
3. Windows HPCreateData Parallel[5]	164.67	No
4. Windows DP Ultrium 960 Parallel	145.07	No
1. HP-UX L&TT Tape Read[5]	160.71	Yes (Tape)
2. HP-UX HPCreateData Single[5]	166.28	No
3. HP-UX HPCreateData Parallel[5]	193.07	No
4. HP-UX DP Ultrium 960 Parallel	153.30	No

[5] Tested in the section Evaluating tape and disk drive performance on page 32

Local restore of small files

The files were restored directly from the SCSI-attached Ultrium 960 tape drive.

For Windows, test 1 in Table 6 shows that the tape device did not read faster than 156.00 MB/s. Test 2 shows that the disk device resulted in 19.94 MB/s single write performance. Test 3, with Data Protector, resulted in just 3.38 MB/s single restore performance. Here, the bottleneck was the file system which was very busy during the recovery. Millions of small files were written back with their original names and file attributes. The tape device was always in start/stop mode.

Note: With Windows, test 4 with parallel writes to file systems was not possible due to problems with an overflow of the Windows system paged pool. See the section Disk write performance on page 35. For that reason, HP-UX test 4 was skipped.

For HP-UX, test 2 in Table 6 shows a much better disk write performance (51.41 MB/s) than for Windows (19.94 MB/s). If you compare the results of test 3 between Windows (3.38 MB/s) and HP-UX (20.76 MB/s), the difference is even bigger. This shows how efficient the HP-UX file system is.

Table 6. Local restore of small files – bottleneck determination

Test	*Performance (MB/s)*	*Bottleneck?*
1. Windows L&TT Tape Read[6]	156.00	No
2. Windows HPCreateData Single[6]	19.94	Yes (File System)
3. Windows DP Ultrium 960 Single	3.38	Yes (File System)
4. Windows DP Ultrium 960 Parallel	-	-
1. HP-UX L&TT Tape Read[6]	160.71	No
2. HP-UX HPCreateData Single[6]	51.41	Yes (File System)
3. HP-UX DP Ultrium 960 Single	20.76	Yes (File System)
4. HP-UX DP Ultrium 960 Parallel	-	-

[6] Tested in the section Evaluating tape and disk drive performance on page 32

Network restore of typical files

The typical files were restored via the network (Gigabit Ethernet) from the remote backup server and its SCSI-attached Ultrium 960 tape drive.

Table 7 shows that the tape device of test 1 and the disk device of test 3 were faster than with Data Protector in test 4. In this scenario, both operating systems showed the same performance for the tests via the network, which was the bottleneck. The Gigabit Ethernet itself has a 1,000 Mb/s or 120 MB/s limitation which is very close to the results of test 4 (Windows with 104.96 MB/s and HP-UX with 104.22 MB/s).

Table 7. Network restore of typical files – bottleneck determination

Test	*Performance (MB/s)*	*Bottleneck?*
1. Windows L&TT Tape Read[7]	156.00	No
2. Windows HPCreateData Single[7]	133.98	No
3. Windows HPCreateData Parallel[7]	164.67	No
4. Windows DP Ultrium 960 Parallel	104.96	Yes (Network)
1. HP-UX L&TT Tape Read[7]	160.71	No
2. HP-UX HPCreateData Single[7]	130.58	No
3. HP-UX HPCreateData Parallel[7]	265.75	No
4. HP-UX DP Ultrium 960 Parallel	104.22	Yes (Network)

[7] Tested in the section Evaluating tape and disk drive performance on page 32

Network restore of small files

The small files were restored via the network (Gigabit Ethernet) from the remote backup server and its SCSI-attached Ultrium 960 tape drive.

For Windows, Table 8 shows that the tape device did not read faster than 156.00 MB/s, and test 2 shows that the disk device can write with 19.94 MB/s. But test 3 resulted in just 3.66 MB/s restore performance for Data Protector. The bottleneck was the file system, which was very busy during the recovery. Millions of small files were written back with their original names and file attributes. The tape device was always in start/stop mode.

Note: For Windows, test 4 with parallel writes to file systems was not possible due to problems with an overflow of the Windows system paged pool. See the section Disk write performance on page 35 for further details. For that reason, HP-UX test 4 was skipped.

Table 8. Network restore of small files – bottleneck determination

Test	*Performance (MB/s)*	*Bottleneck?*
1. Windows L&TT Tape Read[8]	156.00	No
2. Windows HPCreateData Single[8]	19.94	Yes (File System)
3. Windows DP Ultrium 960 Single	3.66	Yes (File System)
4. Windows DP Ultrium 960 Parallel	-	-
1. HP-UX L&TT Tape Read[8]	160.71	No
2. HP-UX HPCreateData Single[8]	51.41	Yes (File System)
3. HP-UX DP Ultrium 960 Single	20.18	Yes (File System)
4. HP-UX DP Ultrium 960 Parallel	-	-

Backup and restore of Microsoft Exchange Server 2003

The Microsoft Exchange Server was principally tested for backup performance. Four storage groups with one 50 GB database each were saved to the SCSI-attached Ultrium 960 tape drive. Other scenarios were not tested, for example with multiple databases per storage group, because they would not significantly change the performance results (MB/s). Multiple databases within one storage group are backed up one after the other.

Note: Multiple storage groups are backed up in parallel, but multiple databases (stores) within a storage group only sequentially.

Basic Microsoft Exchange Server 2003 test tools for backup and restore were not available during the creation of this white paper, so only L&TT and Data Protector were used for testing.

Data Protector configuration parameters

The maximum device concurrency for backing up the Exchange Server data is two. A higher concurrency would use too much of the Exchange Server's resources and would not improve the backup performance.

Recommendation: The recommended device concurrency for backing up the Exchange Server data is two for devices connected directly to the server, and one for devices connected remotely.

The buffer size is based on the formula *buffer size = concurrency* x 16 KB. The minimum buffer size is 32 KB, which is the default buffer size as well.

Local backup of Microsoft Exchange Server 2003

All storage groups were saved to the SCSI-attached Ultrium 960 tape drive with the default configuration values for concurrency (2) and backup buffer size (32 KB).

The performance was only calculated based on the backup time of storage groups and their databases. The backup time for the subsequent transaction log backup was excluded.

In Table 9, test 2 (226.99 MB/s) and test 3 (289.55 MB/s) demonstrate that the Exchange Server integration of Data Protector was able to provide more performance than the Ultrium 960

[8] Tested in the section Evaluating tape and disk drive performance on page 32

tape drive can handle (154 MB/s). The Exchange Server backup of test 4 (138.70 MB/s) did not completely reach the Ultrium tape drive performance of test 1 (154 MB/s). This is because the data was not always streamed with the same performance. The performance ups were higher, and if the database was switched (one database backup had finished and the next one started), the downs much lower.

Table 9. Local backup of Exchange Server 2003 – bottleneck determination

Test	*Performance (MB/s)*	*Bottleneck?*
1. Windows L&TT Tape Write[9]	154.00	No
2. DP Null Single (conc.=1)	226.99	No
3. DP Null Parallel (conc.=2)	289.55	No
4. DP Ultrium 960 Parallel	138.70	Yes (Tape)

Local restore of Microsoft Exchange Server 2003

All storage groups were restored from the SCSI-attached Ultrium 960 tape drive.

The performance was only calculated based on the restore time of the storage groups and its databases. The time for the subsequent transaction log restore and recovery was excluded.

As shown in Table 10, the Exchange Server restore of test 2 (109.61 MB/s) was slower than the Ultrium tape drive performance of test 1 (156 MB/s). This is because the data was not always streamed with the same performance. The performance ups were higher and, for example if a storage group was initialized, the downs much lower.

Table 10. Local restore of Exchange Server 2003 – bottleneck determination

Test	*Performance (MB/s)*	*Bottleneck?*
1. Windows L&TT Tape Read[9]	156.00	No
2. DP Ultrium 960 Parallel	109.61	Yes (Tape)

Tuning Data Protector performance for typical files

Data Protector's backup and restore performance can be improved by modifying its configuration for backups, devices and media.

Because most users tune their environment for backup performance, only restores are discussed here. The focus of this white paper is on the most important backup parameters that are configurable in the GUI. The client-based parameter for the File system tree walk (see page 38) is not considered.

For simplicity, all tests in this section were executed on HP-UX with the typical files dataset as created in the section Creating typical files for HP-UX VxFS on page 27.

Note: In this section, all tests were based on HP-UX typical files with compressibility of data 4:1.

Backup options

Data Protector offers a comprehensive set of backup options for fine tuning. The most relevant options for performance are:

- Load balancing
- Software compression
- IDB logging level
- Detect Windows NTFS hardlinks

Load balancing

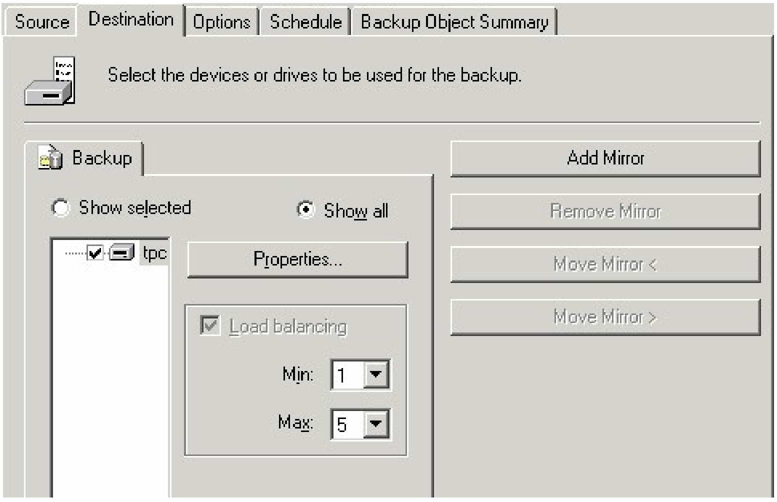
By default, Data Protector automatically balances the load (usage) of devices so that they are used evenly. Since load balancing is done automatically during backup time, it is not necessary to manage the assignment of objects to devices, so that all assigned devices stay busy during

[9] Tested in the section Evaluating tape and disk drive performance on page 32

the backup session. If you selected the load balancing option in the Create New Backup dialog (the default), it cannot be deselected.

Figure 33 illustrates a backup configuration with the default load balancing switched ON.

Figure 33. Backup configuration – load balancing



If load balancing was not selected (OFF) in the Create New Backup dialog, you can choose single devices to be used for each object in a backup specification. This could be very beneficial for bundling objects (such as file systems) based on their speed rather than their size. Data Protector does not track the backup speed of each object and so cannot automatically balance based on that. But the manual load balancing option provides an alternative solution for fine tuning the distribution of backup objects to devices.

Software compression

This option enables compression of the data that is read by the Disk Agent. It is based on the Lempel-Ziv compression algorithm, which is compatible with the standard UNIX compress utility.

Software compression could be an advantage in network environments with small bandwidths when the network is the bottleneck. The Disk Agent compresses the data and subsequently sends it across the network to the remote Media Agent. Sometimes this procedure improves the backup performance. The default compression value is OFF.

Note: Most modern backup devices provide built-in hardware compression that you can configure when adding the device to a client. If this is the case, do not use the software compression option, since double compression only decreases performance without giving better compression results.

Figure 34 shows the default configuration parameters for file systems. The software compression is set to OFF by default.

Figure 34. Backup Configuration – default file system options – software compression OFF

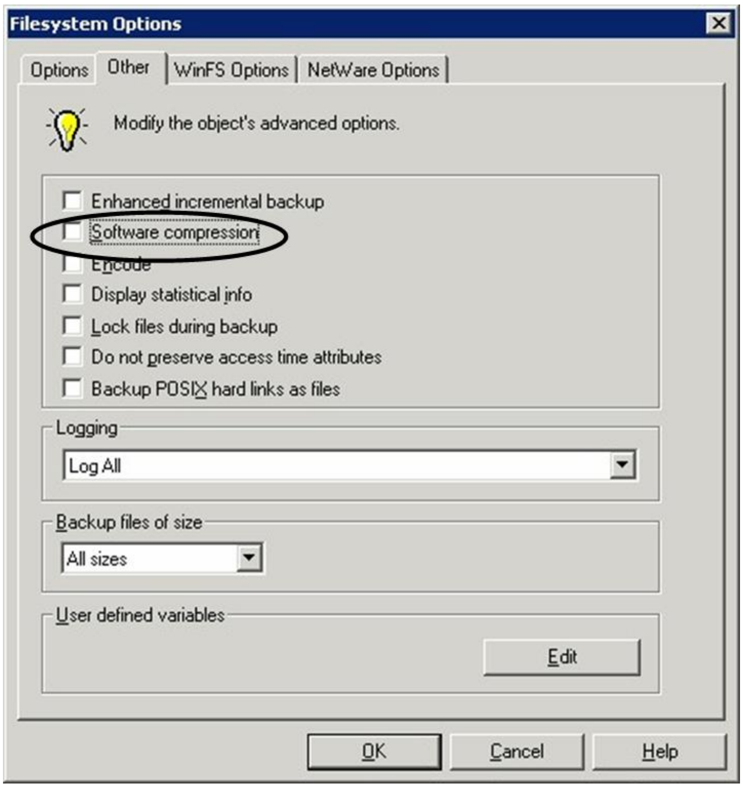


Figure 35 demonstrates how performance efficient the Ultrium 960 tape drive compression is. With the software compression disabled (default), the backup performance was 156 MB/s. With the software compression enabled, the performance was just 46 MB/s.

Figure 35. Performance of HP-UX local backup to Ultrium 960 with software compression OFF/ON

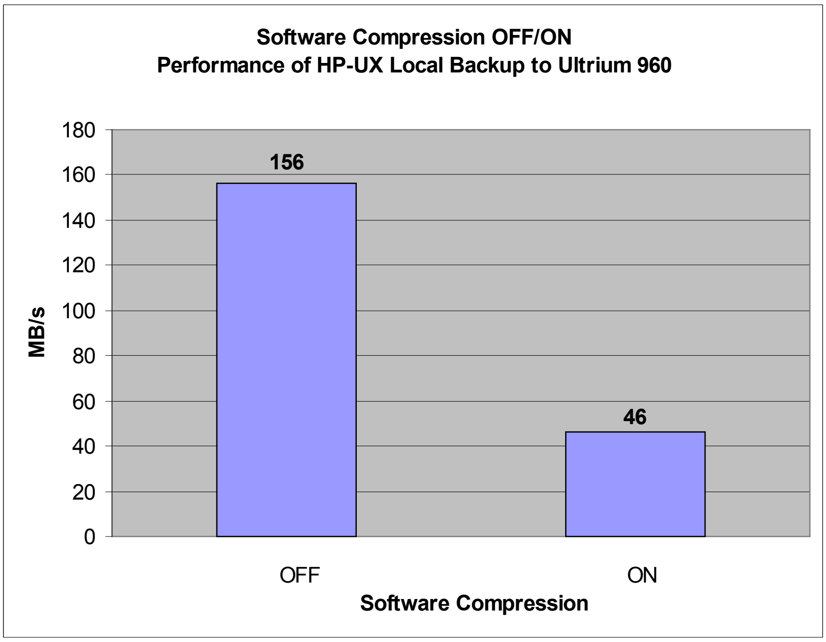
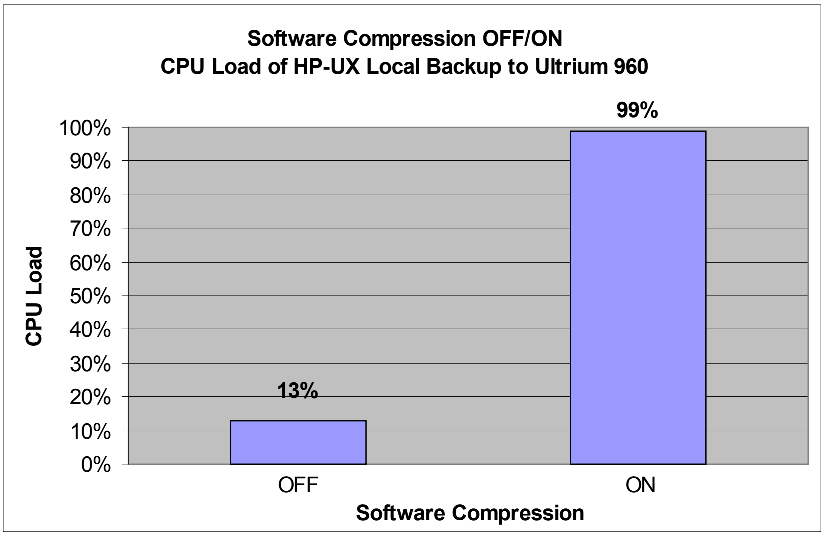


Figure 36 shows that enabling software compression increased the CPU load from 13% to 99%. The CPU load was very high because Data Protector compressed five file systems in parallel.

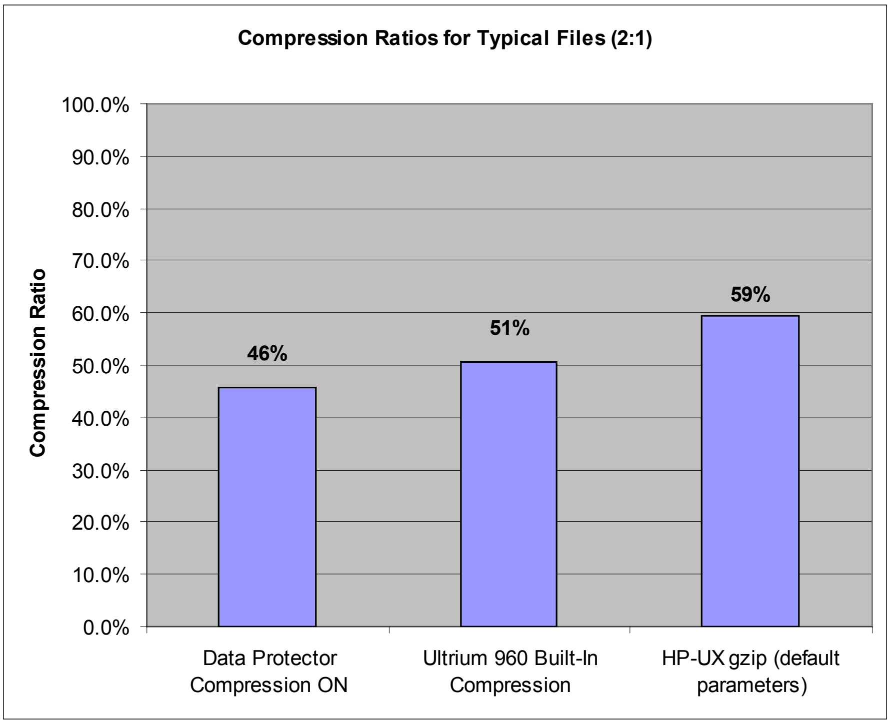
Figure 36. CPU load of HP-UX local backup to Ultrium 960 with software compression OFF/ON



Note: As a rule of thumb, each software compression process takes all the available resources of a single CPU. For instance, one Disk Agent would use only one out of four CPUs, which would result in an overall 25% CPU load.

Figure 37 shows the different ratios for software and hardware compression. HP-UX gzip (GNU zip) was additionally tested because it is the current standard compression utility for UNIX and a good indicator for possible ratios. The compression ratio of HP-UX gzip resulted in 59%, the Ultrium 960 built-in compression in 50% and Data Protector software compression in 45%.

Figure 37. Compression ratios for HP-UX



There are some reasons why Data Protector's software compression ratio was lower:

1. The Data Protector compression is based on the Lempel-Ziv compression algorithm which is less space-efficient than for instance the newer gzip compression utility. On the other hand, better compression has its price in terms of speed.
2. gzip offers different compression levels – between less and best compression. This test was executed with the default level, which is biased towards high compression at expense of speed.
3. Ultrium tape drives use the Advanced Lossless Data Compression (ALDC) algorithm for data compression. ALDC is an implementation of the Lempel-Ziv method of compressing data and can be implemented in hardware very efficiently. ALDC has the ability to switch into a non-compressed mode according to the structure of the data pattern. This means that highly random data does not actually expand when compressed. This method of compression is very fast and space-efficient at the same time.

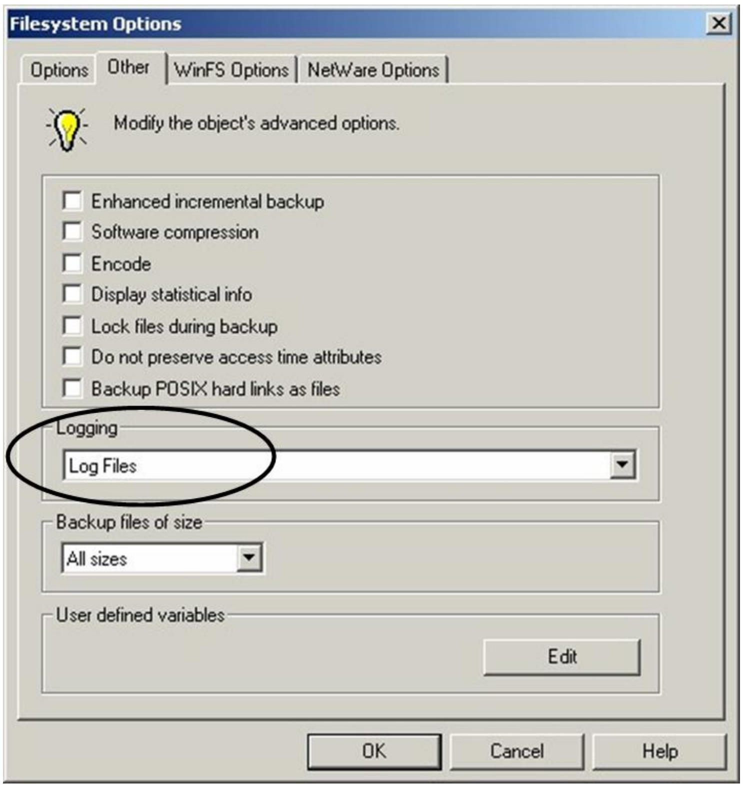
IDB logging level

The logging level determines the amount of detail on files and directories that is written to the IDB during backup, object copying, or object consolidation. Regardless of the logging level, the backed up data can always be restored. Data Protector provides four logging levels: Log All, Log Directories, Log Files, and No Log. The different logging level settings influence the IDB growth, backup speed, and the convenience of browsing data for restore. Since the impact is mostly relevant for the Cell Manager which is not focus of this white paper, it was not tested.

Note: Any Data Protector test was executed with the default configuration parameter Log All which represents the worst case in terms of performance.

Figure 38 shows the default configuration parameters for file systems. The logging level is set to Log All.

Figure 38. Backup Configuration – default file system options – logging level Log Files



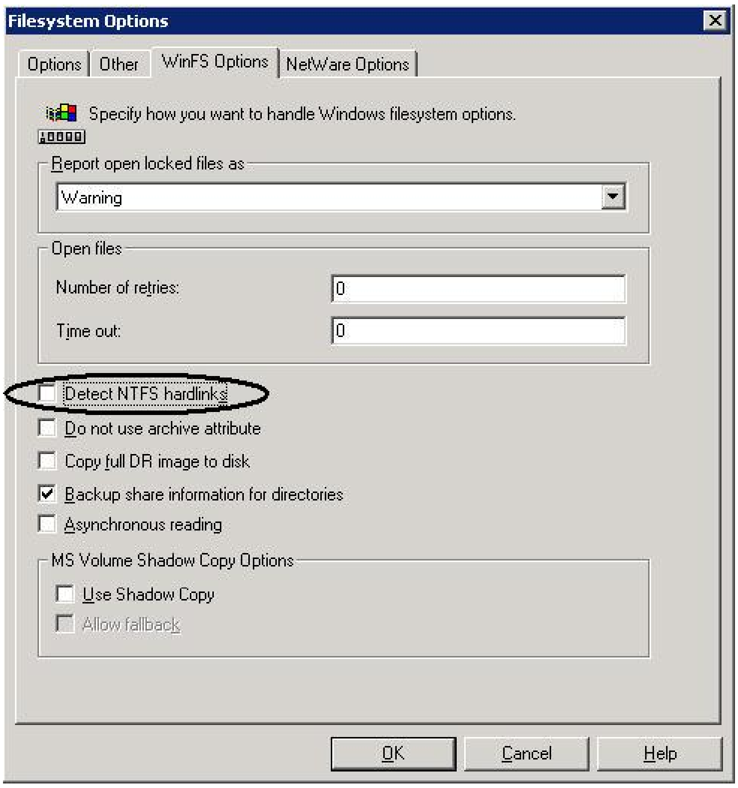
Detect Windows NTFS hardlinks

This option enables detection of NTFS hard links. By default, Data Protector does not detect NTFS hard links, and backs them up as files. This significantly improves backup performance, but the files occupy more space on the media. The original structure is not preserved and, when restored, hard links are restored as files.

Note: NTFS hard links are not commonly used in Windows environments so they are outside the scope of this white paper.

Figure 39 shows the default configuration parameters for Windows file systems. The NTFS hardlinks detection is OFF.

Figure 39. Backup configuration – advanced WinFS filesystem options – detect NTFS hardlinks



Use native Change Log Provider

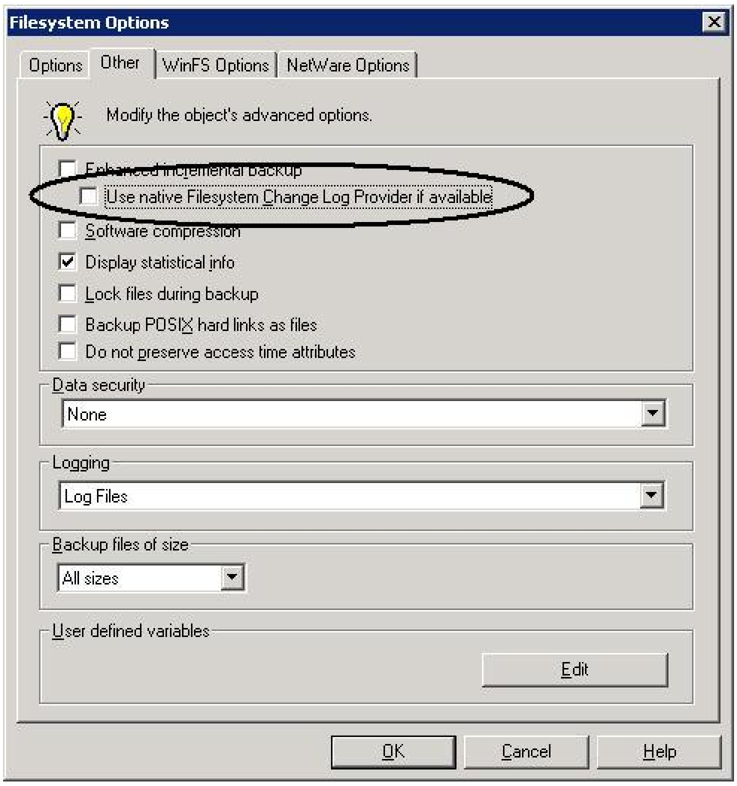
This option was introduced with Data Protector 6.1. If selected, enhanced incremental backup will be performed using the Change Log Provider (if it is available). Currently the Change Log Provider is only available for the Windows NTFS File System and is based on the Windows Change Journal Provider.

The Change Log Provider generates a list of files that have been modified since the last backup and does not perform a file tree walk.

Note: On Windows Vista and Windows Server 2008 the Change Journal is active by default.

Figure 40 shows the default configuration parameters for filesystems. Use of the native Filesystem Change Log Provider is set OFF.

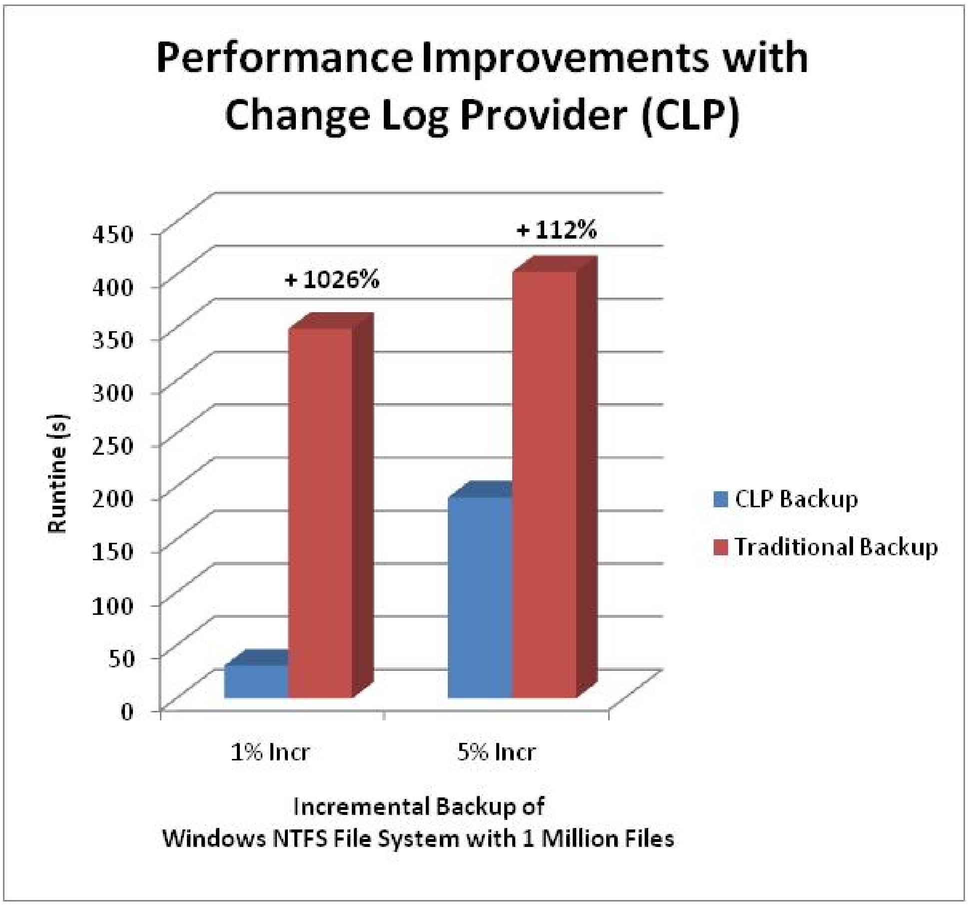
Figure 40. Backup configuration – default filesystem options – use Change Log Provider



A file tree walk can take a considerable amount of time, especially when the directory structure is large and contains millions of files. The Windows NTFS Change Log Provider, based on the Windows Change Journal, addresses this issue by querying the Change Journal for a list of changed files rather than performing a file tree walk. The Change Journal reliably detects and records all changes made to the files and directories on an NTFS volume, which enables Data Protector to use the Change Journal as a tracking mechanism to generate a list of files that have been modified since the last backup. This is very beneficial for the environments with large file systems, where only a small percentage of files change between backups. In this case, the process of determining changed files completes in a much shorter period of time.

Figure 41 demonstrates how efficiently the File System Change Log Provider option performs. Best results can be achieved if file systems have few changes, like in this chart shown with 1% incremental backup.

Figure 41. Dependencies of Change Log Provider and Enhanced Incremental Database (EIDB)



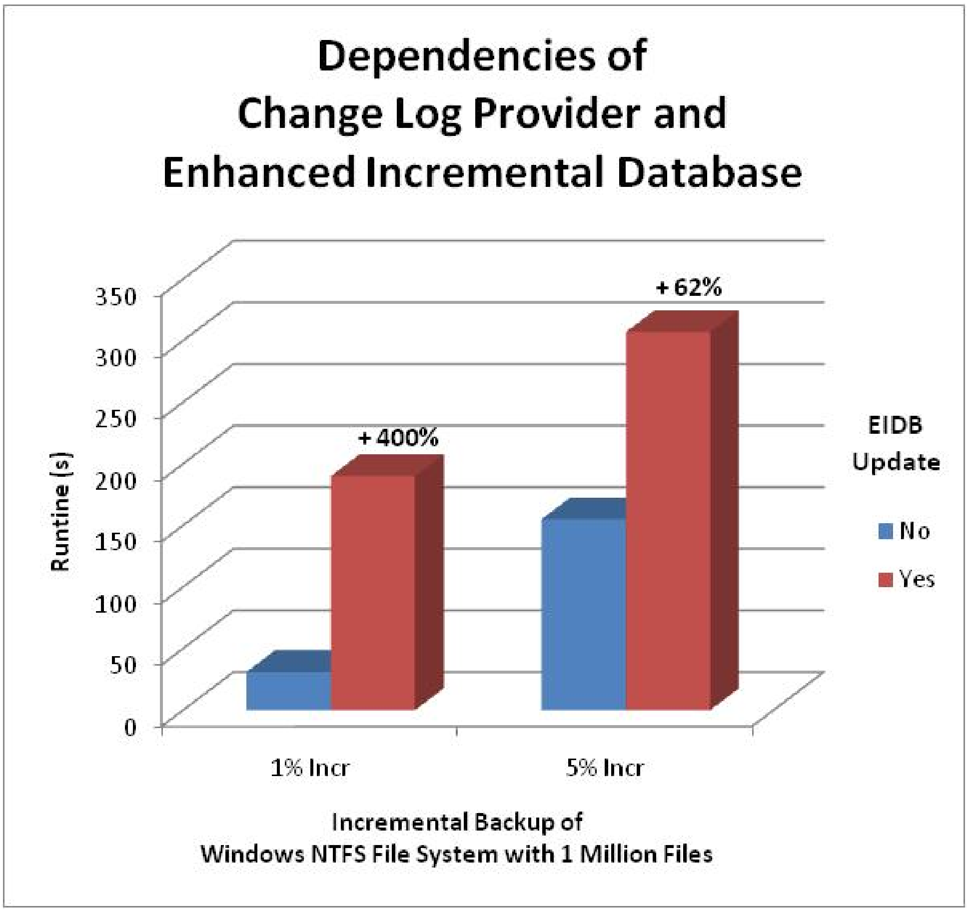
Updating the Enhanced Incremental Database (EIDB) has a large performance impact. The current default configuration is to update the EIDB if it exists for the volume in question (this will change in the near future). The Change Log Provider does not depend on the EIDB, but updates it so that it can be used in case of a Change Log Provider error. In other words, if the EIDB is kept up-to-date, it can be used in order to provide the enhanced incremental functionality in the case of a Change Log Provider error.

The Change Log Provider performance can be improved by:

1. Deleting the EIDB for the volume in question. If it is not there, it will not be updated.
2. Setting a parameter in the Data Protector template `Data_Protector_home\omnirc`:
   DP 6.1: `OB2_CLP_UPDATE_EI_REPOSITORY =0`

Figure 42 demonstrates the different runtimes if the EIDB is updated or not. The Windows NTFS file system was initially loaded with 1 million small files.

Figure 42. Dependencies of Change Log Provider and Enhanced Incremental Database (EIDB)



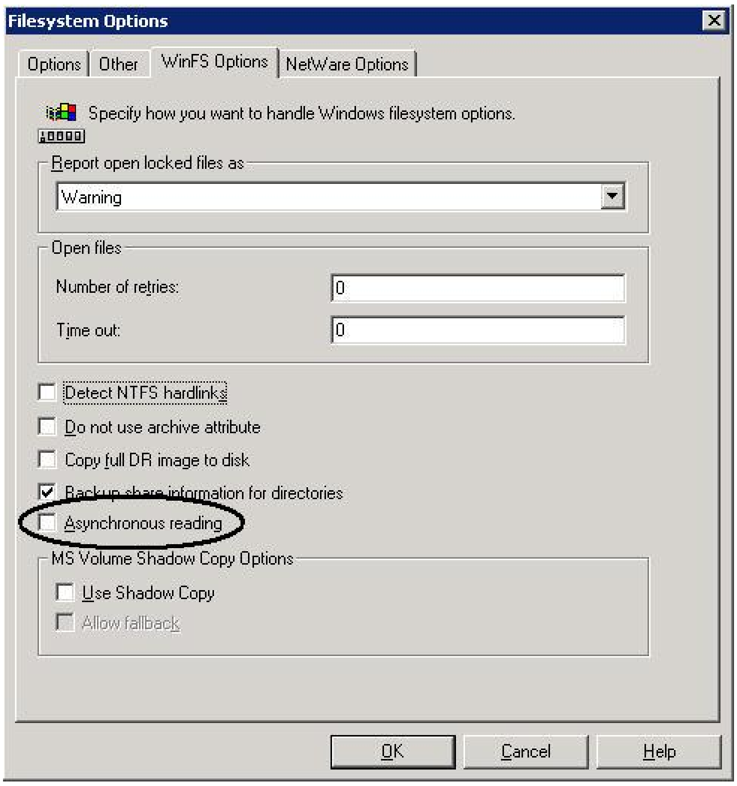
Asynchronous reading

This option was introduced with Data Protector 6.1 and enables faster reading for data that is striped across multiple physical disks. It applies to disk arrays if their logical drives are configured in striped mode.

Note: Asynchronous reading was not designed for reading data from the same physical disk.

Figure 43 shows the default configuration parameters for Windows file systems. The asynchronous reading is set OFF.

Figure 43. Backup configuration – advanced WinFS filesystem options – asynchronous reading



On the client, the Data Protector template `Data_Protector_home\omnirc` keeps parameters for asynchronous reading:

`OB2DAASYNC=not set|0|1` (*default:* not set)

If set to one, the Disk Agent will always use asynchronous reading. If set to zero, the Disk Agent will never use asynchronous reading. By default, asynchronous mode is not forced and is governed by the option in the backup specifications. If this variable is set, the setting takes precedence over the option in the backup specification or the command line.

`OB2DAASYNC_READERS=2-256` (*default:* 64)

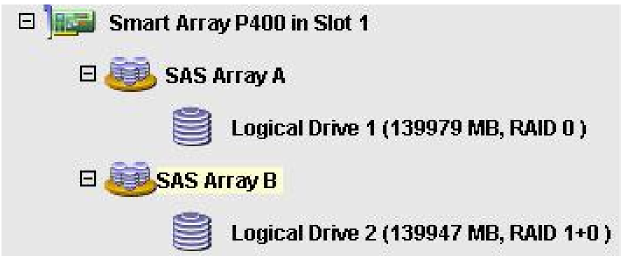
This variable controls the number of concurrent reads from the disk (APCs) when asynchronous reading is used. Generally, performance improves with higher values, but a plateau is typically quickly reached. Note that good performance will be achieved more easily with contiguous files, while higher values may be beneficial with fragmented files. However, setting the value too high may cause some performance degradation with certain types of disks or RAID arrays due to an excessively high number of outstanding I/O requests.

`OB2DAASYNC_SECTORS=1-1024` (*default:* 128)

This variable tells how many sectors are read in one read request when asynchronous reading is used. The sector size of the disk being backed up is determined at runtime (typically 512 bytes). Generally, performance improves with higher values. Some types of disks are less sensitive to this parameter and reach the maximum throughput with lower values than others. Some types of disks do not handle reads well with sizes greater than a certain value (for example, 256), which degrades performance. Setting the value slightly lower (such as 248) may yield the best performance for such a disk.

Figure 44 shows how backup server disks were reconfigured for testing. One SAS array was created based on four internal 72 GB 15K SAS disks and all space was allocated for a single logical drive in RAID 1+0 mode.

Figure 44. Smart array configuration of backup server – one logical drive in RAID 1+0 mode and 4 SAS disks



Array B	
Controller	Smart Array P400
Bus Interface	PCI Express
Controller Location	Slot 1
Array	B
Array Type	SAS
Number of Logical Drives	1
Physical Drives Attached to Port 1I	Box 1 : Bay 5, 72 GB (SAS)
Physical Drives Attached to Port 2I	Box 1 : Bay 2, 72 GB (SAS) Box 1 : Bay 3, 72 GB (SAS) Box 1 : Bay 4, 72 GB (SAS)
Online Spare Drive(s)	None
Unused Space on this Array	6 MB

As shown in Table 11, the backup performance of typical files in test 1 (125.42 MB/s) was considerably improved with asynchronous reading enabled (163.09 MB/s). The backup performance of small files in test 2 (9.23 MB/s) was slightly degraded with asynchronous reading enabled (8.42 MB/s).

Table 11. Synchronous and asynchronous read test

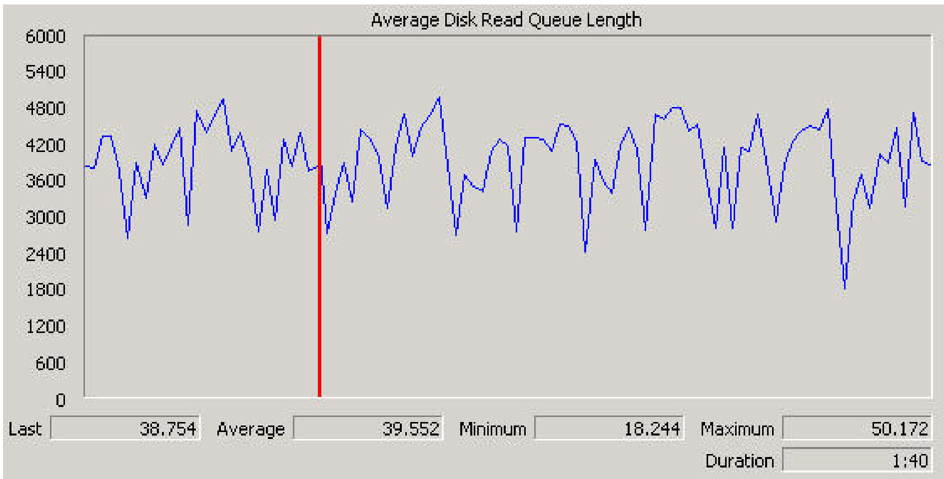
Test	*Performance (MB/s) Async Read OFF*	*Performance (MB/s) Async Read ON*	*Async Read ON Effect*
1. Typical files	125.42	163.09	+ 30.04%
2. Small files	9.23	8.42	− 8.78%

Note: The backup performance could decrease if asynchronously reading a large number of small files (file size less than 100 KB).

Note: Non-asynchronous reading usually results in a disk queue length less than 1. On the other hand, asynchronous reading increases the disk queue length and could impact the performance of an application if the same physical disks are accessed.

Figure 45 shows that asynchronous reading resulted in a disk queue length between 18 and 50.

Figure 45. Smart array configuration of backup server – one RAID 1+0 logical drive with 4 SAS disks



Device and media options

You can set advanced options for devices and media when configuring a new device, or when changing device properties. The availability of these options depends on the device type.

You can also set some of these options when configuring a backup, for example, the concurrency. Device options set in a backup specification override options set for the device in general.

The most relevant options for performance are:

- Concurrency
- CRC check
- Block size
- Segment size
- Disk Agent buffers
- Hardware compression

Concurrency

The number of Disk Agents started for each Media Agent is called Disk Agent (backup) concurrency and can be modified using the advanced options for the device or when configuring a backup. The concurrency set in the backup specification takes precedence over the concurrency set in the device definition.

Data Protector provides a default number of Disk Agents that are sufficient for most cases. The fast HP StorageWorks Ultrium 960 drive is configured with a default of four.

For example, if you have a library with two Ultrium 960 drives, each controlled by a Media Agent, and each Media Agent receives data from four Disk Agents concurrently, data from eight disks is backed up simultaneously.

Figure 46 shows the default configuration parameters for Ultrium tape drives. The concurrency is set to four. Note that other tape drives or backup devices could have different values.

Figure 46. Ultrium 960 advanced options – concurrency

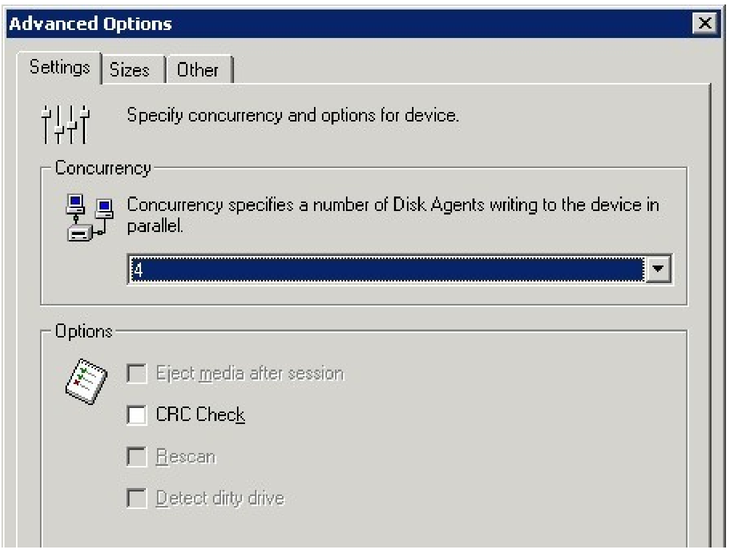


Figure 47 illustrates the performance of a local backup to the Ultrium 960 tape drive with different concurrencies. It shows that a Ultrium 960 concurrency value of three was sufficient for this test environment with the SAN-connected disk volumes. The default value of four or the higher value of five did not improve performance but did also not degrade it. Higher values for concurrency are useful for environments with slower disks and file systems.

Figure 47. Performance of HP-UX local backup to Ultrium 960 with concurrency 1 - 5

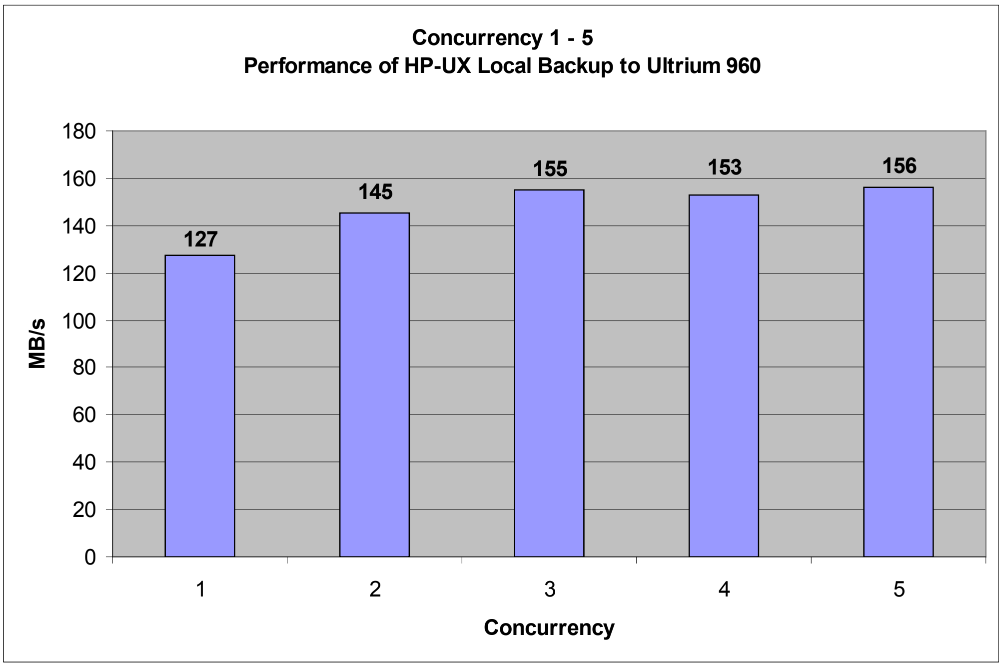
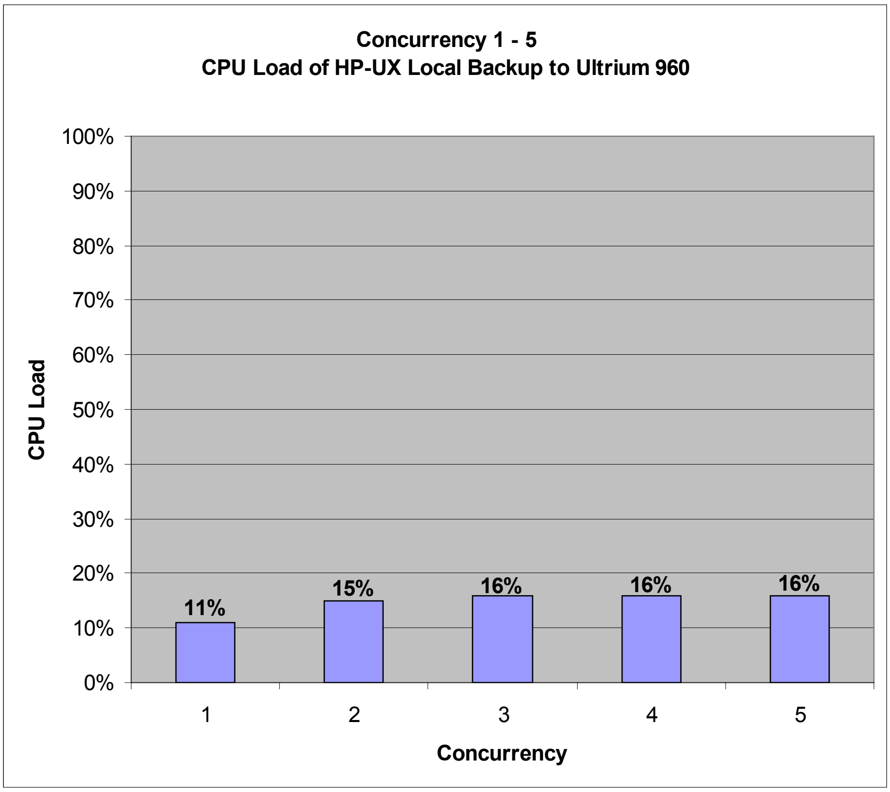


Figure 48 demonstrates that the CPU load was low during all backup tests. Concurrency values of three to five resulted in the same CPU load. This is similar to the backup performance as shown in Figure 45.

Figure 48. CPU Load of HP-UX local backup to Ultrium 960 with concurrency 1 – 5



Cyclic redundancy check (CRC)

Cyclic redundancy checking (CRC) is an enhanced checksum function. When this option is selected, CRC checksums are written to the media during backup. CRC checks allow you to verify the media after the backup. Data Protector re-calculates the CRC during a restore and compares it to the CRC on the medium. It is also used while verifying and copying the media. This option can be specified for backup, object copy, and object consolidation operations. The default value is OFF.

Figure 49 shows the default configuration parameters for Ultrium tape drives. CRC checking was OFF.

Figure 49. HP StorageWorks Ultrium 960 advanced options – CRC checking OFF

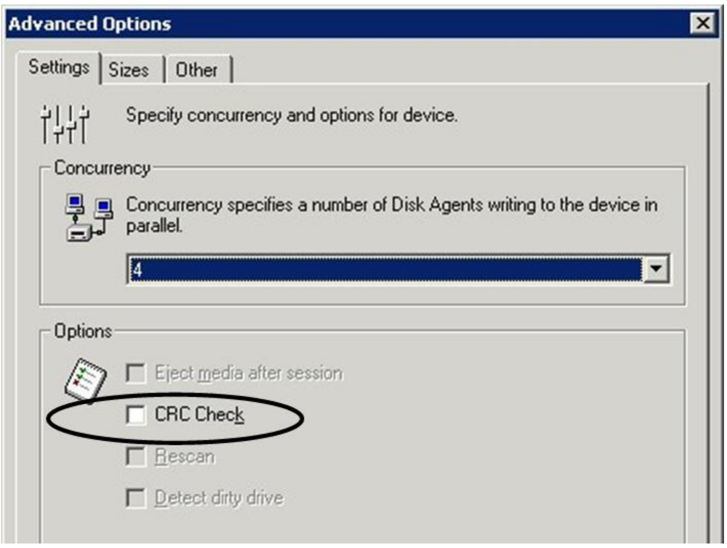


Figure 50 demonstrates that enabling CRC checking required additional system resources. The CRC ON test resulted in approximately 20% performance decrease.

Figure 50. Performance of HP-UX local backup to Ultrium 960 with CRC checking OFF/ON

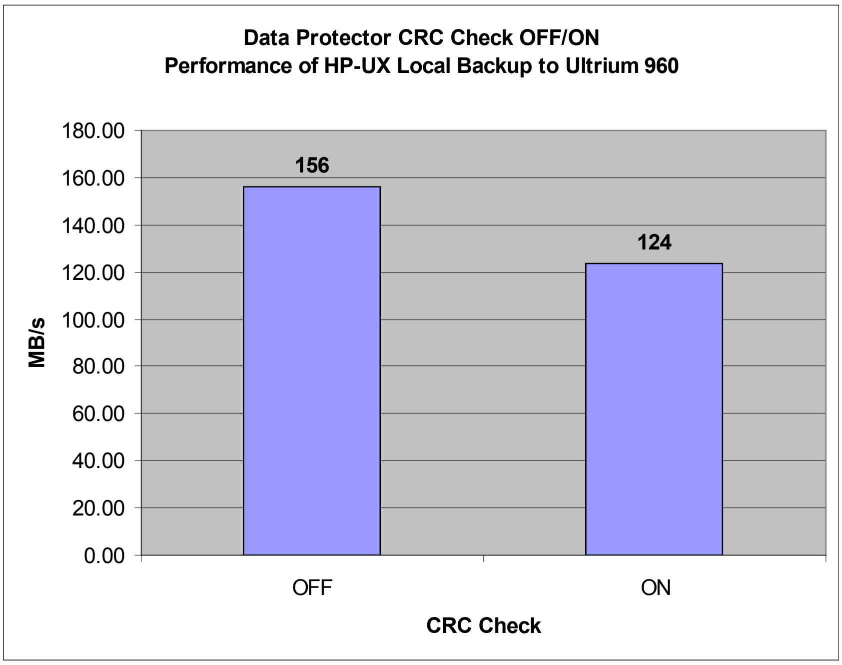
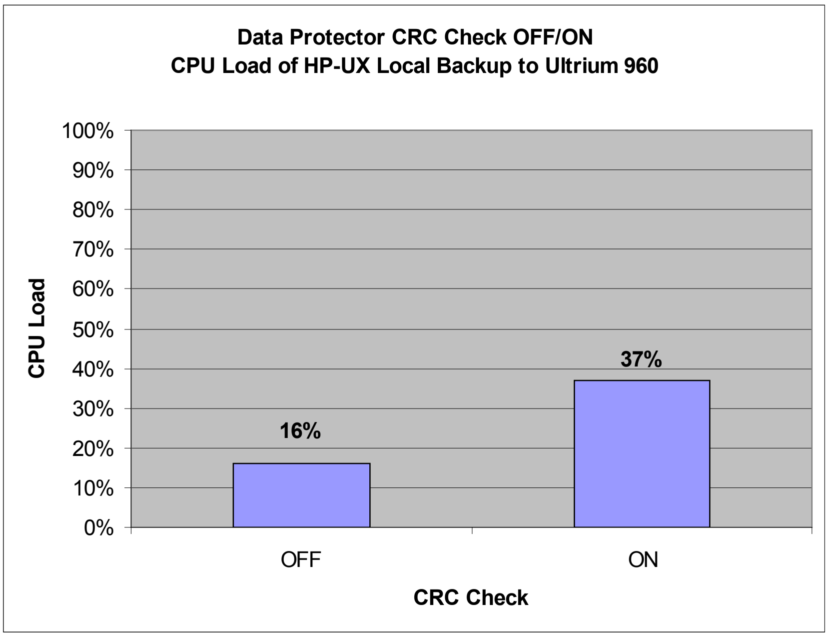


Figure 51 shows that enabling CRC checking required additional CPU resources. The CRC ON test resulted in more than twice the CPU load.

Figure 51. CPU load of HP-UX local backup to Ultrium 960 with CRC checking OFF/ON



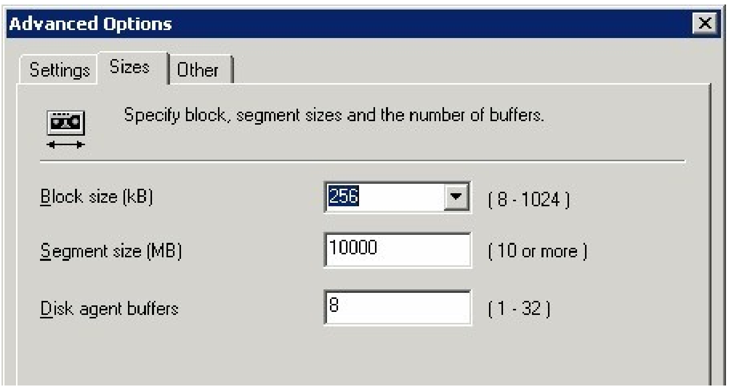
Block size

Segments are not written as whole units, but in smaller subunits called *blocks*. Data Protector uses a default device block size regarding different device types. The block size applies to all devices created by Data Protector and to Media Agents running on the different platforms.

Increasing the block size can improve performance. You can adjust the size when configuring a new device or when changing the device properties using the Advanced options for the device. A restore automatically adjusts the block size.

For the Ultrium 960, the block size was configured with a fixed value of 256 KB as shown in Figure 52 and as described in the section Data Protector configuration on page 37.

Figure 52. Ultrium 960 Configuration – block size, segment size and disk agent buffers

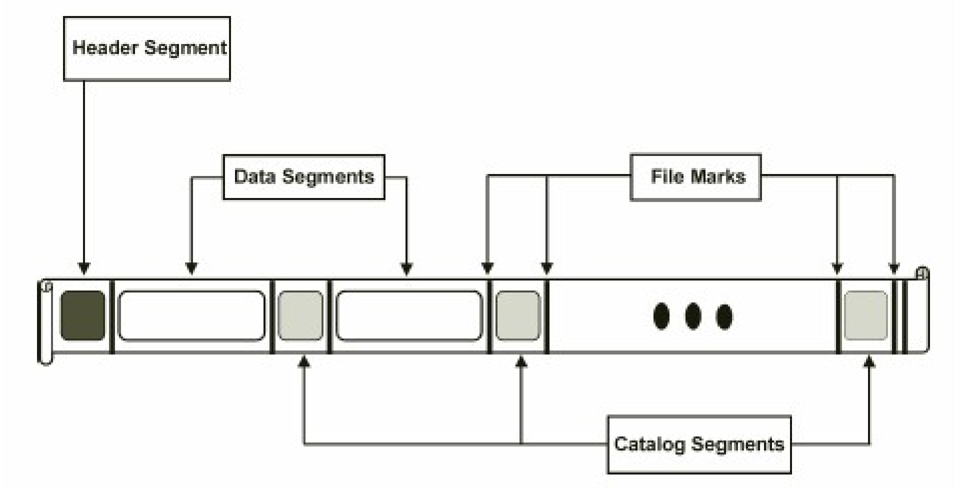


Segment size

A medium is divided into data segments, catalog segments, and a header segment. Header information is stored in the header segment, which is the same size as the block size. Data is stored in data blocks of data segments. Information about each data segment is stored in the corresponding catalog segment. This information is first stored in the Media Agent memory and then written to a catalog segment on the medium as well as to the IDB.

All segments are divided by file marks as shown in Figure 53.

Figure 53. Data Protector Medium – data segments, catalog segment, header segment, and file marks



Segment size, measured in megabytes, is the maximum size of data segments. If you back up a large number of small files, the actual segment size can be limited by the maximum size of catalog segments. Segment size is user configurable for each device and influences the restore performance. You can adjust the segment size when configuring a new device or when changing the device properties using the Advanced options for the device.

Optimal segment size depends on the media type used in the device and the kind of data to be backed up. The average number of segments per tape is 50. The default segment size can be calculated by dividing the native capacity of a tape by 50. The maximum catalog size is limited to a fixed number (12 MB) for all media types.

Data Protector finishes a segment when the first limit is reached. When backing up a large number of small files, the media catalog limit could be reached faster, which could result in smaller segment sizes.

Figures 54 and 55 illustrate the results of backups with different segment sizes. Larger segment sizes improved the backup performance, but resulted in some additional CPU load.

Figure 54 Performance of HP-UX local backup to Ultrium 960 with segment sizes 10 – 10,000 MB

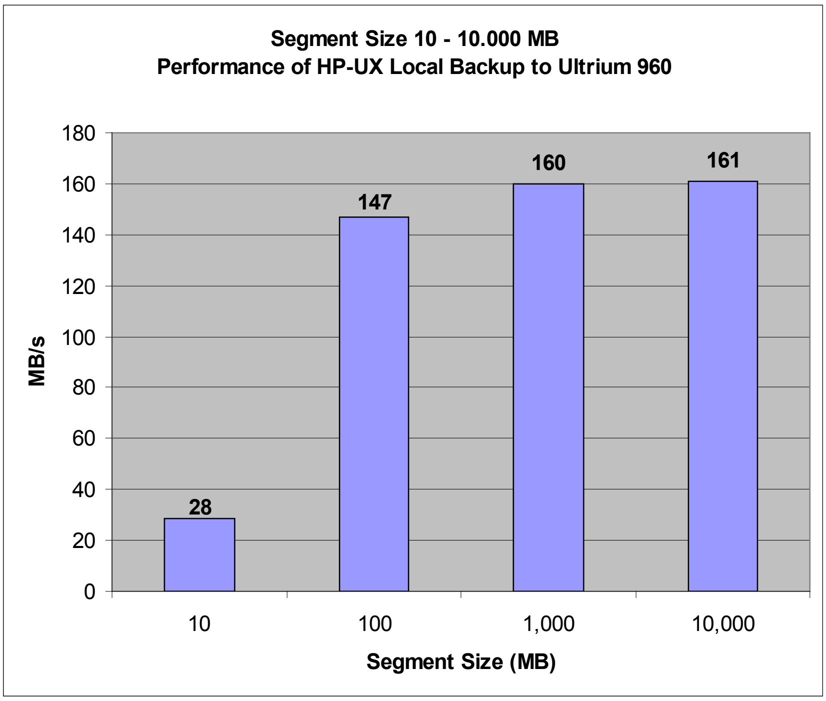
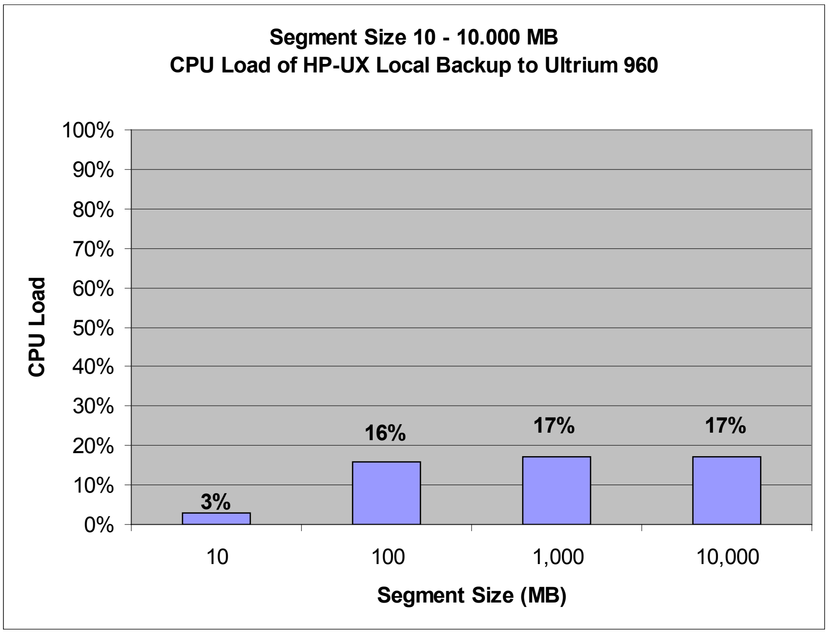


Figure 55. CPU load of HP-UX local backup to Ultrium 960 with segment sizes 10 – 10,000 MB



Note: Larger segments sizes improve the backup performance but could have negative impact on restore performance. Data blocks are found faster if restored from a backup with smaller segment sizes.

Disk Agent buffers

Data Protector Media Agents and Disk Agents use memory buffers to hold data waiting to be transferred. This memory is divided into a number of buffer areas (one for each Disk Agent and depending on device concurrency). Each buffer area consists of eight Disk Agent buffers (of the same size as the block size configured for the device).

This value can be changed while configuring a new device or when changing the device properties using the advanced options for the device, although this is rarely necessary. There are two basic reasons for changing this setting:

1. Shortage of memory
   If there is a shortage of memory, the shared memory required for a Media Agent can be calculated as follows:
   Media Agent Shared Memory = Disk Agent Concurrency x *Number of Buffers* x *Block Size*
2. Streaming
   If the available network bandwidth varies significantly during backup, it is important that a Media Agent has enough data ready for writing to keep the device in the streaming mode. In this case, increasing the number of buffers could help.

Figures 56 and 57 demonstrate that different numbers of Disk Agent buffers had no significant impact on performance in this local backup scenario. The Media Agent always had enough data ready for writing.

Figure 56. Performance of HP-UX local backup to Ultrium 960 with 1 – 32 Disk Agent buffers

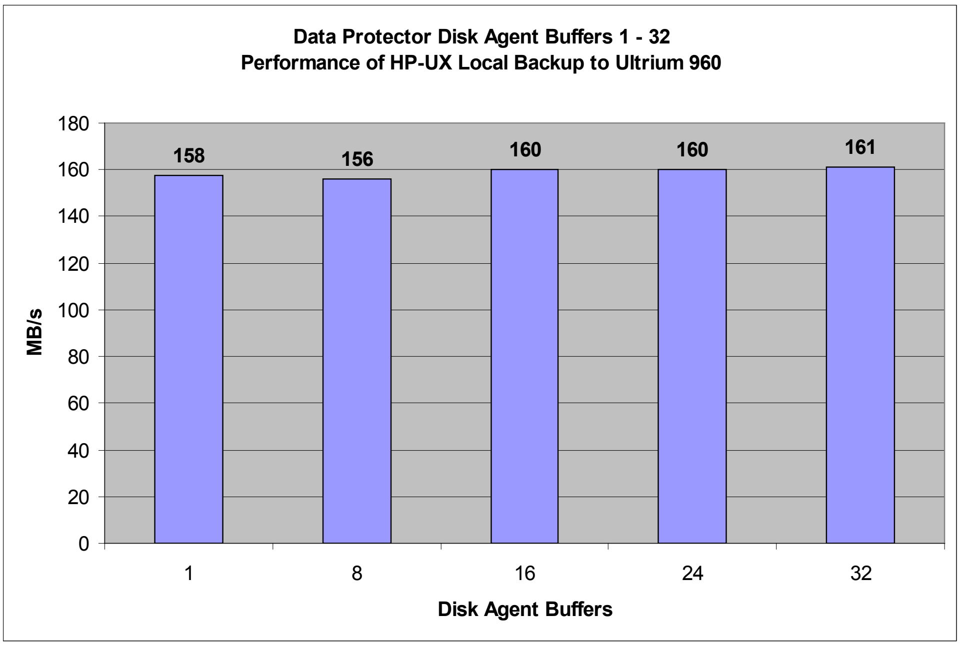
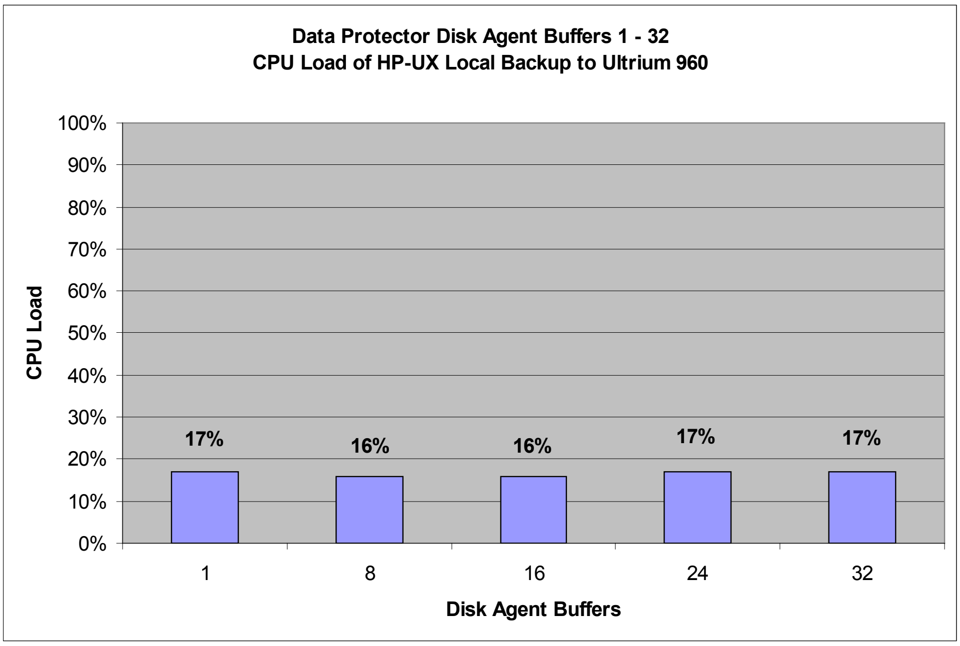


Figure 57. CPU load of HP-UX local backup to Ultrium 960 with 1 – 32 Disk Agent buffers



Hardware compression

Most modern backup devices provide built-in hardware compression that can be enabled by selecting the appropriate device file or SCSI address in the device configuration procedure. Hardware compression increases the speed at which a tape drive can receive data, because less data is written to the tape.

Recommendations:

1. Do not use software and hardware compression at the same time because double compression decreases performance without giving better compression results.
2. Keep the software compression option disabled when an Ultrium drive is configured with Data Protector.

When configuring a device, you can select the SCSI address from the drop-down list. Data Protector automatically determines whether the device can use hardware compression or not.

Figure 58 illustrates which devices are detected during the automatic configuration process for an Ultrium 960 drive. In this example, the hardware compression is selected. The end of the device's SCSI address is extended with the "C" option: `Tape0:0:6:0C`.

Figure 58. HP StorageWorks Ultrium 960 automatic configuration – SCSI address for hardware compression

Figure 59 shows that with the correct SCSI address selection the Hardware Compression option (bottom left) is grayed out.

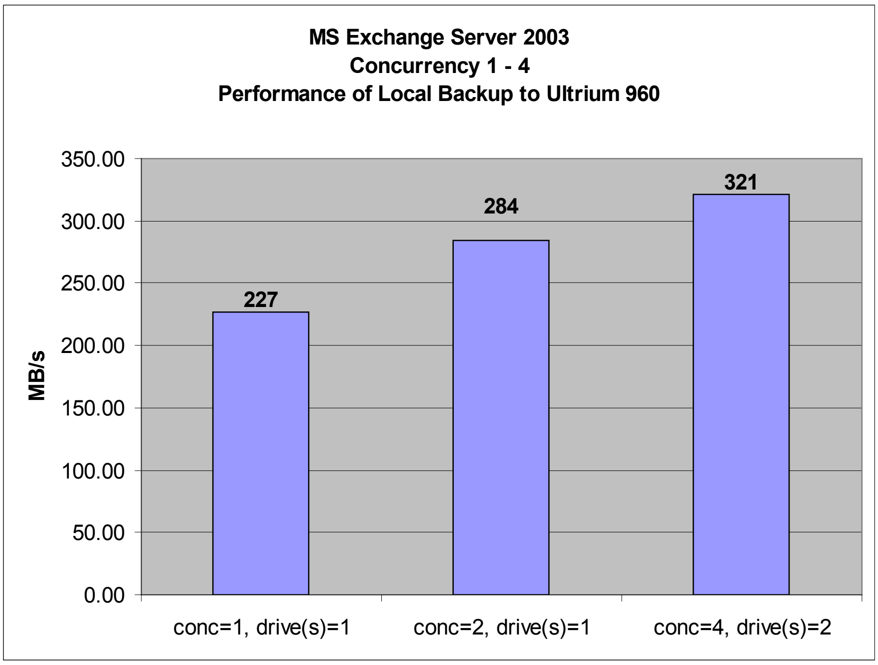
Figure 59. HP StorageWorks Ultrium 960 configuration – hardware compression enabled

Tuning Data Protector performance for Microsoft Exchange Server 2003

Data Protector's backup and restore performance for Microsoft Exchange Server 2003 can be improved by modifying its configuration parameters for backups. All tests were executed with the NULL device to remove the tape drive as a bottleneck.

Figure 60 demonstrates how efficient a higher concurrency value could be. All tests were executed with the same buffer size of 64 KB because this is the minimum value for a concurrency of 4 (4 x 16 KB). The minimum buffer for each stream (concurrency = 1) is 16 KB.

Figure 60. Performance of MS Exchange Server 2003 local backup to Ultrium 960 with concurrency 1 - 4



Note: For the test with the concurrency of four, an additional NULL device had to be configured. Note that for the Exchange Server integration the maximum concurrency of one device is two.

Figure 61 shows that a higher concurrency caused a higher CPU load. However, this slight increase can be disregarded when compared with the remarkable increase of the performance as shown in the previous Figure 60.

Figure 61. CPU load of MS Exchange Server 2003 local backup to Ultrium 960 with concurrency 1 - 4

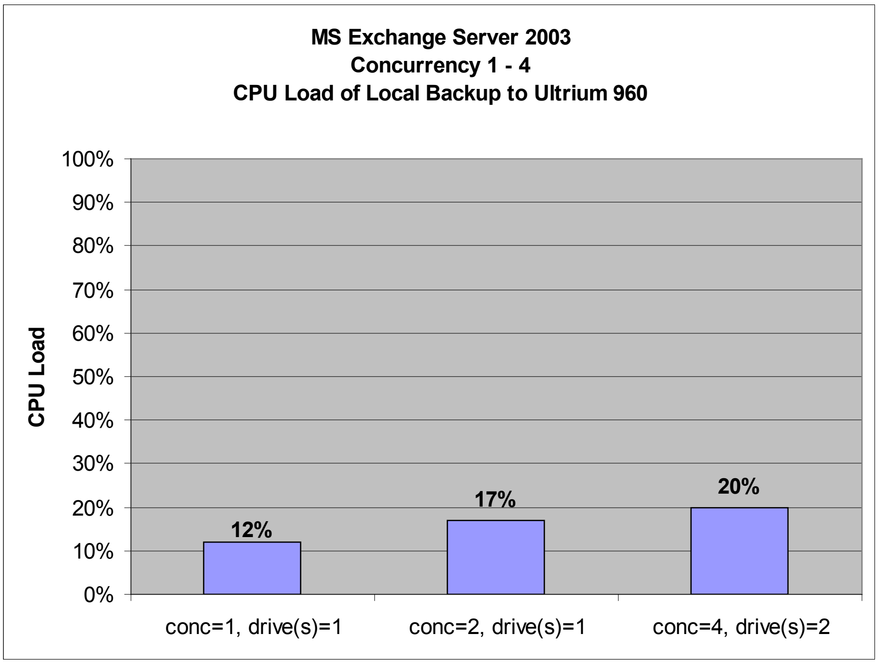


Figure 62 demonstrates that different buffer sizes resulted in similar backup performance. The default buffer size of 32 KB (per backup device) was already a good choice.

Figure 62. Performance of MS Exchange Server 2003 local backup to Ultrium 960 with buffer size 32 – 1,024 KB

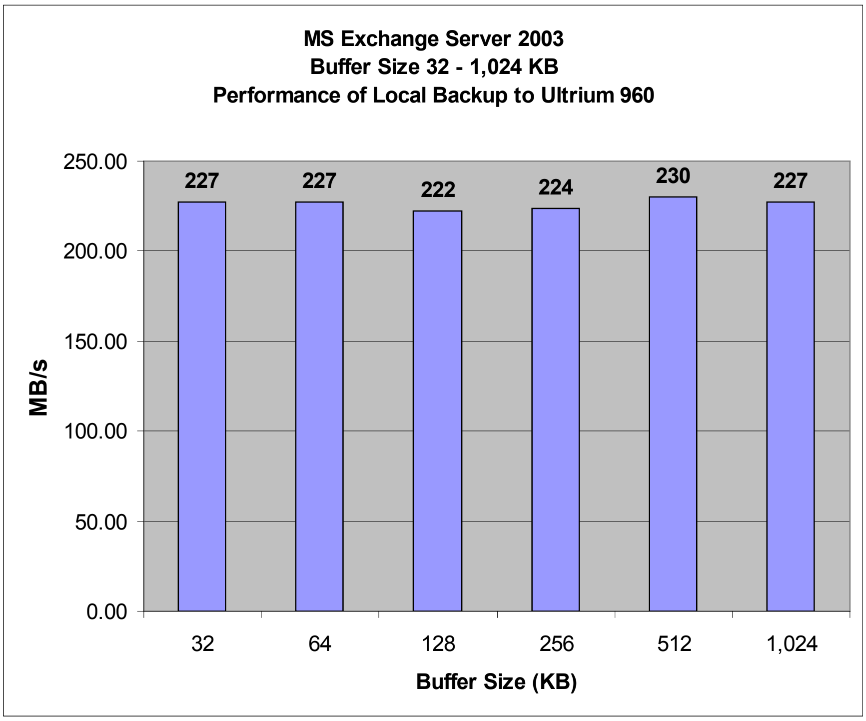
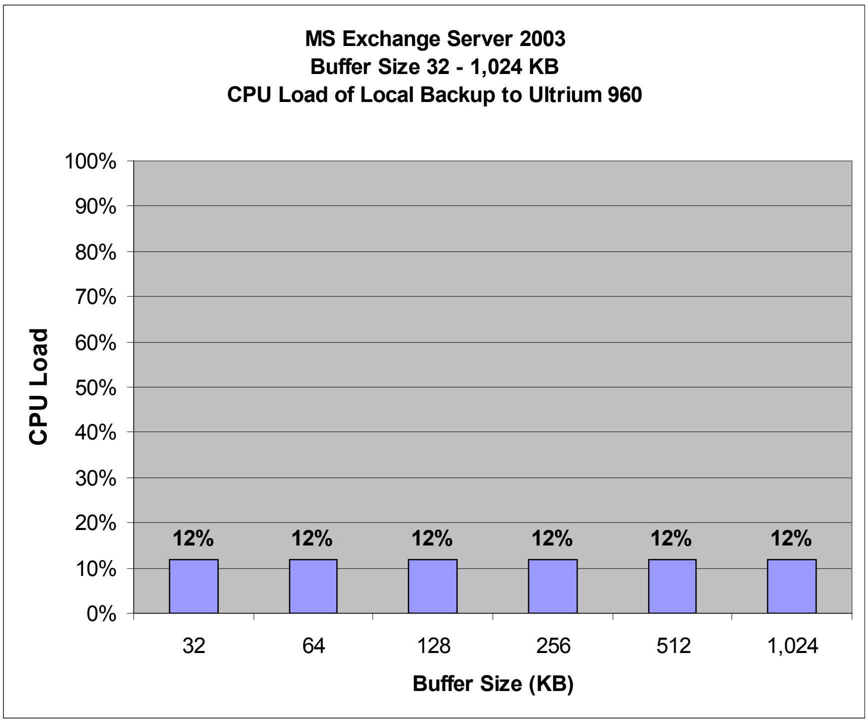


Figure 63 shows that different buffer sizes resulted in the same CPU load (12%). This was expected because there was only a small performance difference, as shown in Figure 62.

Figure 63. CPU load of Exchange Server 2003 local backup to Ultrium 960 with buffer size 32 – 1,024 KB



Tuning recommendations

Due to the high number of variables and permutations, it is not possible to give specific recommendations that will fit all user requirements and affordable investment levels. However, consider the following when trying to improve backup or restore performance:

- Ensure that the server is sized for the backup requirements. For example, fast tape devices like an HP StorageWorks Ultrium 960 tape drive should be placed on a dedicated 133-MHz PCI-X bus and not share the bus with other HBAs, such as a Gigabit network adapter.
- During high performance backups, Windows memory problems could occur. This is a general Windows kernel problem and applies to almost all backup applications — including NTBACKUP.

  The Data Protector error message looks like:
  ```
  [Major] From: VBDA@tpc131.bbn.hp.com "G:"  Time: 29.03.2005 19:08:57
  [81:78]   G:\1\file67108864_000003
  Cannot read 57256 bytes at offset 0(:1): ([1450] Insufficient system
  resources exist to complete the requested service. ).
  ```
  Note: It is advisable to follow the instructions in Microsoft article Q304101, which explains the circumstances that can lead to this error and settings for avoiding memory problems. The article is here:
  http://support.microsoft.com/default.aspx?scid=kb;en-us;Q304101

- Ultrium 960 tape drives are best used with a block size of 256 KB as described in *Getting the most performance from your HP StorageWorks Ultrium 960 tape drive white paper* (downloadable from http://h18006.www1.hp.com/storage/tapewhitepapers.html).
- For all backups using Ultrium drives, it is recommended to keep the software compression option disabled when configured with Data Protector.
- If backup devices are SCSI- or SAN-attached, software compression should be disabled. On the other hand, software compression could make sense when executing network backups

across slow LAN (100 Mb/s) environments. This could provide a better backup performance, but it will also cause high CPU loads on the client server.

- When backing up typical files directly to a SCSI-attached Ultrium 960 tape drive, parallel backups (multiplexing/concurrency) are recommended because one single stream cannot fully use the tape drive. Data Protector's default concurrency is four.
- When backing up typical files via Gigabit Ethernet HBAs to a remote Ultrium 960 tape drive, consider performing backups without multiplexing/concurrency. The network is often the bottleneck and enabling multiplexing/concurrency does not significantly improve the performance.
- Backup performance can be generally improved if setting the following options to ON:
  - Do not preserve access time attributes
  - Do not use archive attribute (Windows specific option)
- When backing up Windows NTFS file systems with typical files striped across multiple physical disks, enable the asynchronous reading option (introduced with Data Protector 6.1).
- For Windows NTFS file systems with millions of small files, parallel backups with a high concurrency are recommended. These kinds of file systems are very slow and should be multiplexed. Instead of the fast Ultrium 960 tape drive, consider using a slower Ultrium tape drive, or backup-to-disk technology.
- For Windows NTFS file systems with millions of small files, disable double tree walks. The first tree walk briefly scans the files selected for the backup and calculates its size, so that the percentage done can be calculated during the backup. The second tree walk is executed during the actual file backup. On these particular systems, it is advisable to set the following parameter in the Data Protector template `*Data_Protector_home*\omnirc`:
  - *DP 6.0:* `NoTreeWalk=1`
  - *DP 6.1:* `OB2NOTREEWALK=1`
- For Windows NTFS file systems with millions of small files, consider using the Windows Change Log Provider for incremental backups. It will also address the tree walk issue, but only for incremental backups.
- For restores of Windows NTFS file systems with millions of small files, it is advisable to start only one restore session. Parallel restores will cause an overflow of the Windows system paged pool. You could also tune the Windows kernel, but this is not focus of this white paper.
- For Exchange Servers, the recommended device concurrency is two for devices connected directly and one for devices connected remotely (backup server).

Appendix A. Reference documentation

HP documents and links

Storage

- *HP StorageWorks Enterprise Virtual Array configuration best practices*
  ftp://ftp.compaq.com/pub/products/storageworks/whitepapers/5982-9140EN.pdf
- *HP StorageWorks SAN Design Reference Guide*
  http://h18000.www1.hp.com/products/storageworks/san/documentation.html

Backup

- *Enterprise Backup Solution (EBS) Design Guide* and compatibility matrix
  http://www.hp.com/go/ebs

Tools

- Performance troubleshooting and using performance assessment tools
  http://www.hp.com/support/pat
- Library and tape tools
  http://www.hp.com/support/tapetools

Microsoft documents and links

- Microsoft support and knowledge base
  http://support.microsoft.com

Glossary

cell

A set of systems that are under the control of a Cell Manager. The cell typically represents the systems on a site or an organizational entity, which are connected to the same LAN.

Cell Manager

The main system in the cell where the essential software of Data Protector is installed and from which all backup and restore activities are managed.

CSI (Command Screen Interface)

A non-graphical user interface which is based on a menu structure.

DAS (Direct Attached Storage)

DAS is a digital storage system directly attached to a server or workstation, without a storage network in between.

Disk Agent (DA)

A Data Protector component needed on a client to back it up and restore it. The Disk Agent controls reading from and writing to a disk. During a backup session, the Disk Agent reads data from a disk and sends it to the Media Agent, which then moves it to the device. During a restore session the Disk Agent receives data from the Media Agent and writes it to the disk.

EIDB (Enhanced Incremental Database)

Conventional incremental backup backs up files that have changed since a previous backup, but has certain limitations in detection of changes. Unlike conventional incremental backup, enhanced incremental backup reliably detects and backs up also renamed and moved files, as well as files with changes in attributes.

GUI (Graphical User Interface)

The Data Protector GUI is a cross-platform (HP-UX, Solaris, Windows) graphical user interface, for easy access to all configuration, administration, and operation tasks.

HBA (Host Bus Adapter)

A host bus adapter connects a host system (the computer) to other network and storage devices.

IDB (Internal Database)

The Data Protector IDB is an internal database, located on the Cell Manager, that keeps information regarding what data is backed up, on which media it resides, the result of backup, restore, copy, object consolidation, and media management sessions, and which devices and libraries are configured.

LAN (Local Area Network)

A computer network covering a small geographic area, like a home, office, or group of buildings.

MA (Media Agent)

A Data Protector process that controls reading from and writing to a device, which reads from or writes to a medium (typically a tape). During a backup session, a Media Agent receives data from the Disk Agent and sends it to the device for writing it to the medium. During a restore session, a Media Agent locates data on the backup medium and sends it to the Disk Agent. The Disk Agent then writes the data to the disk. A Media Agent also manages the robotics control of a library.

PCIe (PCI Express)

PCI Express, officially abbreviated as PCIe (and sometimes confused with PCI Extended, which is officially abbreviated as PCI-X), is a computer system bus/expansion card interface format.

SAN (Storage Area Network)

An architecture to attach remote computer storage devices such as disk array controllers and tape libraries to servers in such a way that to the operating system the devices appear as locally attached devices.

SPOF (Single Point of Failure)

Any part of the system that can, if it fails, cause an interruption of required service. This can be as simple as a process failure or as catastrophic as a computer system crash.

For more information

- HP Data Protector software
  http://www.hp.com/go/dataprotector
- HP HP9000 and ProLiant servers
  http://h18000.www1.hp.com/products/servers
- HP StorageWorks Ultrium tape drives
  http://h18006.www1.hp.com/storage/tapestorage/ultriumdrives.html
- HP StorageWorks Enterprise Virtual Arrays (EVA)
  http://h18006.www1.hp.com/products/storageworks/eva
- Enterprise Backup Solution (EBS)
  http://www.hp.com/go/ebs
- Microsoft Exchange Server
  http://www.microsoft.com/exchange

4AA1-3836ENW, July 2009