HP Data Protector 6.0 software Advanced Backup to Disk performance white paper

Executive summary

This white paper provides performance-related information for HP Data Protector 6.0 software and the Advanced Backup to Disk feature.

This white paper covers HP ProLiant Windows® 2003 servers connected to HP StorageWorks Modular Smart Array (MSA) disk arrays and HP StorageWorks Ultrium 960 (LTO3) tape drives.

The proof points are all Windows-based for simplification of equipment needs, but the lessons still hold good for heterogeneous environments.

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

- HP Data Protector software tuning may help to improve the performance, for example, by modifying tape block size and file depot size. See the Tuning recommendations section.
- HP StorageWorks Ultrium 960 tape drives are best utilized with a block size of 256 KB as described in the "Getting the most performance from your HP StorageWorks Ultrium 960 tape drive white paper," available at http://h18006.www1.hp.com/storage/tapewhitepapers.html.
  - If you are running the Media Agent on Windows 2003 SP1, there exists a limitation in the involved tape driver (tape.sys) on the data transfers to a block size of no greater than 64 KB. To solve this issue, see the hotfix from Microsoft® that removes the 64-KB limitation on block sizes (see the Tape drive section).
- For entry-level JBOD and low-end disk array usage, it is important to understand what your disk subsystem is capable of delivering. This can be done using HP performance assessment tools (downloadable from http://www.hp.com/support/pat). The performance tools are also embedded within the HP industry-leading Library and Tape Tools diagnostics (downloadable from http://www.hp.com/support/tapetools).
- The configuration of disk arrays can have a remarkable impact on the backup and restore performance. Important parameters are the configured number of logical arrays, logical volumes, and type of RAID levels. The configuration of logical arrays and volumes should reflect the internal disk array layout. For example, disk arrays with two SCSI buses could perform best by creating two logical arrays with one logical volume each. RAID levels must be chosen carefully. RAID 0 provides the best performance but should not be considered due to missing fault tolerance. Therefore, it is recommended to configure RAID 5, which fits best for staging areas. It is space and cost efficient and provides a good read performance.
- Serial ATA (SATA) disks used as secondary disk storage arrays can have slower rotational speeds than their SCSI counterparts (for example, the HP StorageWorks MSA SATA drives are only 7.2K rpm) and although they are high capacity (250 GB), their performance reflects their pricing. Therefore, backing up to tape from a staged (secondary) disk array can be slower than backing up directly from the primary storage to tape in some circumstances. Ironically, tape is now faster than disk. Disk staging is useful, however, for gathering several small files into a single object, or backing up slow networked hosts before the data is sent to tape. Both small files and slow hosts can cause very slow backups.
- The restore of many small files (20 million in this setup) could cause serious file system bottlenecks. Data Protector must wait until Windows and the belonging NTFS responds.
- Finally, the test environment with the HP ProLiant DL380 G3 server and HP StorageWorks 1000/1500 Modular Smart Arrays (MSA1000/1500) is capable to manage data with low CPU utilization. The typical file backup (from MSA1000) directly to the Ultrium 960 tape drive showed a good transfer rate (74 MB/s) together with a low average CPU load (10%). The comparable disk backup (from MSA1500) to tape still showed an acceptable transfer rate (33 MB/s) together with a little higher impact on average CPU load (22%).

Objective and target audience

The main objective of this white paper is to educate and inform users of the HP Data Protector 6.0 software Advanced Backup to Disk feature about what levels of performance are achievable in different backup scenarios.

The emphasis is in showing what is typical and not what best-case scenarios are. This white paper highlights where the current performance bottlenecks are and how these might be overcome.

The target audience for this white paper are system integrators and solution architects and anyone involved in getting the best backup performance out of their infrastructure investments.

Introducing HP Data Protector 6.0 software

HP Data Protector software is designed for the most demanding 24x7 environments and offers an automated high-performance backup and recovery from disk and tape.

HP Data Protector software simplifies the use of complex backup and recovery procedures with the fastest installation, automated routine tasks, and easy-to-use features. It is the ideal solution to reduce IT costs and complexity while remaining reliable and scalable to grow from single server environments to the largest distributed enterprise infrastructures, providing broad compatibility of operating systems, applications, drives, libraries, and disk arrays.

The Advanced Backup to Disk feature improves the backup process with continuous backup of transaction log files, backup of slow clients without multiplexing, easy resource access and sharing, plus backup in tape-less branch offices, while offering fast and easy configuration and licensing. Furthermore, it allows single file restore directly from disk or tape.

The following sections, File library and Object copy, give a short overview about the features required for Advanced Backup to Disk.

For further information on Advanced Backup to Disk, refer to the Disk-Assisted Backup white paper at: http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00668397/c00668397.pdf

File library

Data Protector provides a disk-based device type called "file library." Disk-based devices are designed to do backup and restore to and from disk. The file library device is the most sophisticated disk-based device. It focuses on low-cost disk arrays, such as the MSA1000, which is positioned mainly as a backup device. In case the file library is running out of space, new backup capacity can be assigned automatically.

Configuration of a file library is very easy. You must define mount points where Data Protector will create its "media" and optionally, the number of simultaneous "writes" that will be used. Data Protector will utilize this like any other physical backup device and will auto-create the media files on the fly as required. This capability is called "Advanced Backup to Disk."

Object copy

Object copy provides the capability to copy a backup object from the source destination to a target destination, where the target could be another media type. With object copy you have the possibility to create copies of backed-up objects and keep them in several locations. You are able to migrate the backed-up data to another media type, so you may use high-speed disk backup for initial backup and then replicate (migrate) the data to tape for offsite storage.

Basics of 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 goal is to make the tape drive the bottleneck. That way the system will achieve the performance figures as advertised on the drive's specification sheet.

Note that backup jobs can stress hardware resources up to their highest limit, which would never happen during normal application load. This puts the emphasis on the rest of the system and causes failures, which are based in the involvement of many components and their time-critical handshake of data.

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 Ultrium 960, are so fast that this can be a difficult goal to achieve. If the backup performance is not matching the data sheet of the tape drive, then there is a bottleneck somewhere else in the system.

One single component, like a 100-Mb/s network, can decrease an SDLT or LTO tape drive performance to a very low transfer rate (this would be a very good use case for first staging the data on disk and then backing it up to tape).

All components must be considered for getting the theoretical backup performance. Practical performance data can only be obtained from benchmarks.

Factors that critically affect the backup speed include:

- Multiplexing

  This allows better bandwidth utilization of the tape drive during backup but can slow down restore performance because all the data is interleaved all the way down the tape. Therefore, the time spent to perform a single stream restore is higher due to other streams having to be read (and potentially ignored).

- Disk and tape buffers

  Data Protector offers a set of advanced options for backup devices and disk agents. The default settings are device-based and match most of all backup environments. Ultrium 960 is an exception and requires a modification as described in the Tuning recommendations section.

- Data file size

  The larger the number of smaller files, the larger the overhead associated with backing them up. The worst-case scenario for backup is large numbers of small files due to system overhead of file access.

- Data compressibility

  Incompressible data will back up slower than compressible data. JPEG files, for example, are not compressible, whereas database files can be highly compressible. The accepted standard for quoting tape backup specifications revolves around an arbitrary figure of 2:1 compressible data.

- Disk array performance

  It is often overlooked that you cannot put data on tape any faster than you can read it from disk. Backup is more sequential in nature than random (from a disk array access perspective). Disk array performance depends on the number of disks, RAID configuration, the number of Fibre Channel ports to access the array, queue depth available, and so on. HP has written several utilities to "read data" from disk arrays and deliver a performance value. This enables users to determine the throughput of their disk arrays when operating in backup mode and performing file system traversals typical of this activity.

These Performance Assessment Tools (PAT utilities) can be downloaded from http://www.hp.com/support/pat. The performance tools are also embedded within the HP industry-leading Library and Tape Tools diagnostics, which can be downloaded from http://www.hp.com/support/tapetools.

- Transfer size

  Transfer size is the overall total size of the SCSI transfer within a single SCSI command. In some operating systems, there is a limit set on this. For example, in Microsoft Windows the default transfer size is 64 KB and to increase the overall transfer size above 64 KB in Windows, a registry entry called "MaximumSGList" (associated with the HBA) must be changed. Many modern HBAs already install their drivers with this registry value set appropriately. Check the registry entries, and see the Tuning recommendations section.

- Fragmentation

  The more fragmented the files are on disk, the more random will be the disk access method, hence the backup will take longer. If your system does have a de-fragmentation utility, it is advisable to run it before full backups or on a regularly scheduled basis to ensure files are largely contiguously arranged on disk.

- SAN inter-switch links

  In the switched fabric, SAN inter-switch links (ISLs) ensure that SAN connections have sufficient bandwidth to support the backup traffic going through them. Trunk multiple ISLs together where possible.

- 2-Gb Fibre Channel

  With a 2-Gb/s Fibre Channel connection the theoretic maximum transfer rate is 200 MB/s, so this is the maximum transfer rate you can expect.

- SCSI burst rate

  Beware of disk drives quoted as Ultra 320. Ultra 320 refers to the burst rate not the sustained rate. The typical sustained rate from a 15K rpm SCSI Ultra 320 disk drive is approximately 80 MB/s for raw sequential I/Os (that is, without file system read overhead).

- SATA disks

  These types of disks are lower cost, lower performance, and lower reliability than the SCSI disk previously listed but are useful for staging backups because they offer high capacity. A typical 7200K rpm SATA disk drive has lower seek times than an equivalent SCSI disk drive, a burst rate of approximately 150 MB/s, and a sustained transfer rate of approximately 50 MB/s.

Disk–to-Disk to-Tape (D2D2T) data protection architecture

With the high capacity and lower cost offered by SATA and Fibre Channel ATA (FATA) disk technologies, many customers are now considering implementing backup to low-cost disk arrays before backup to tape.

The use of secondary disk arrays for backup is best suited to environments where:

- The business dictates rapid single file restore capabilities (seconds to minutes). It is generally quicker to restore from a secondary disk subsystem than it is from tape (minutes to hours). A single file on a disk system can be accessed easier because the disk has not to be positioned and searched sequential like the tape. However, the secondary storage array is not infinite in capacity and only the most recent backups may still reside on the disk array. In addition, disk-to-disk backup is no substitute for offsite media. You need the tape in such a disk-to-disk-to-tape (D2D2T) setup to be able to do offsite vaulting and to put the data in a safe, offsite place.
- The hosts can only supply data at a relatively modest rate (10–20 MB/s). The backup image can be gradually produced on the secondary array without having to interleave multiple streams to tape and without the need for a large powerful dedicated backup server.
- Small file backup to tape has always been a performance limiter. With backup to disk, a complete backup image of small files can be constructed and then passed to tape at much higher transfer rates than if the small files were transferred directly to tape.

Table 1. Comparing SATA and SCSI disks

Feature	SATA disks	SCSI disks
Mean time to failure (MTTF)	500,000 hrs @ 20% duty cycle	1,200,000 hrs @ 100% duty cycle
Burst transfer rate	150 MB/s	320 MB/s
RPM	7.2K	10K or 15K
Queuing	Non-Tagged Serial execution	Tagged Optimized seeks Better performance
Capacity	250 GB	146 GB
Warranty	1 year	3 years

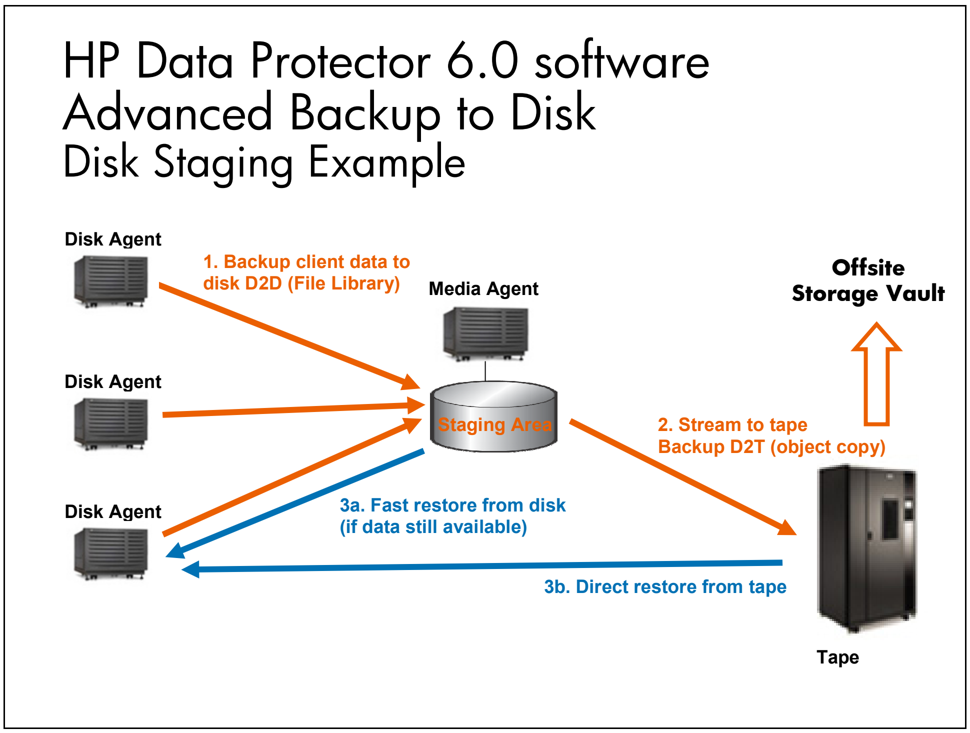
While the SCSI disk MTTF of 1,200,000 looks high, this figure decreases when there are many disk drives bound together in an array. By comparison, HP StorageWorks Ultrium 460 tape drives have a MTTF of 250,000 at 100% duty cycle but are used in smaller volumes within an automated tape library. Therefore, do not assume that disk is automatically more reliable than tape.

The key point is D2D2T has its place and with proper administration can improve the data protection process **but it is not a replacement for tape**. Tape is still the foundation of a robust data protection strategy.

HP Data Protector 6.0 software provides a comprehensive implementation of D2D2T called "Advanced Backup to Disk." The following example shows a scenario with a single central backup server and multiple network connected clients:

- The first step is to back up clients over the slow network to a central backup server on to its staging area.
- The second step is to copy the consolidated backups from the staging area to tape. Typically the central backup and restore environment is designed to provide high-speed backup from disk to tape.
- If a restore is required, data can be accessed either from disk (if still available) or from tape over the network.

Figure 1. Data Protector 6.0 Advanced Backup to Disk example



Use cases

- Continuous backup of transaction log files (avoids tape drives being in start/stop mode)
- Backup of slow clients without multiplexing
- Tape-less branch office backup
- Faster small file backup
- Resource sharing

On the other hand, database backups are typically not first candidates for a D2D2T backup solution. The direct backup of large database files to tape can usually be done very efficiently, utilizing the full performance of the tape drive.

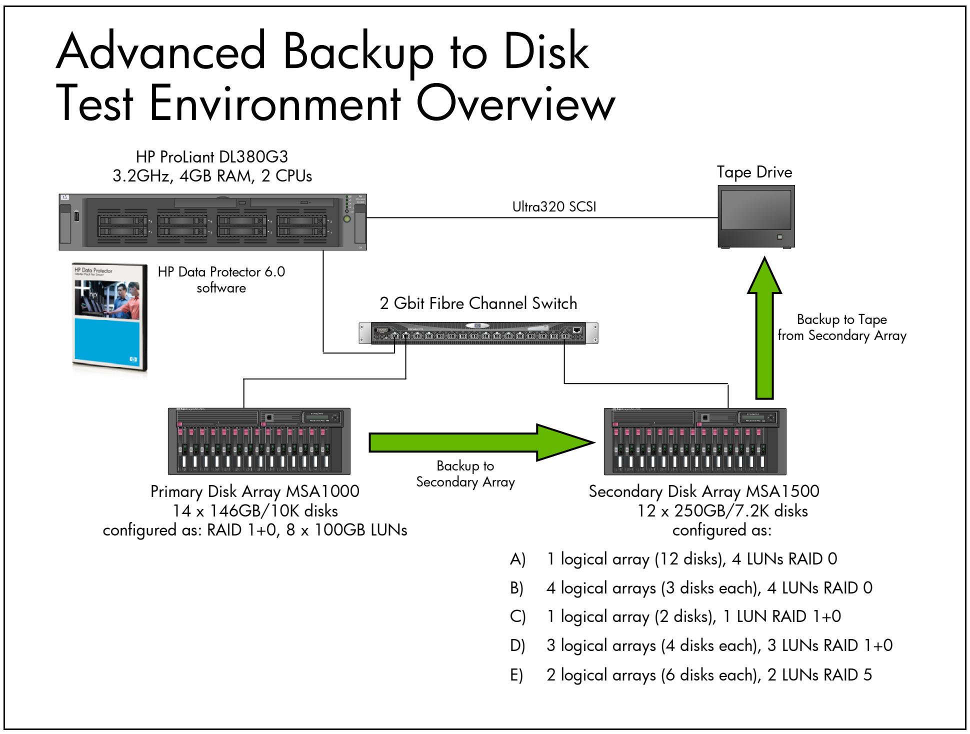
Configuration

Backup performance will **always be limited** by one or more bottlenecks in the system, of which the tape drive is just one part. The goal is to make the tape drive the bottleneck. That way the system will achieve the performance figures as advertised on the drive's specification sheet.

Test environment overview

Figure 2 shows an overview of the test environment, which will be described in detail in the next sections.

Figure 2. Test environment overview



HP ProLiant DL380 G3 server

The HP ProLiant DL380 G3 server features 2 x 3.2-GHz processors, 4 GB of RAM, 1 x 36-GB local disk, and Windows Server 2003 SP1. All file systems are configured as NTFS. The latest HP ProLiant Support Pack (PSP) 7.60 has been installed on the system.

PSP is available at: http://h18013.www1.hp.com/products/servers/management/psp/index.html

Figure 3. System information

System:
Microsoft Windows Server 2003
Enterprise Edition
Service Pack 1

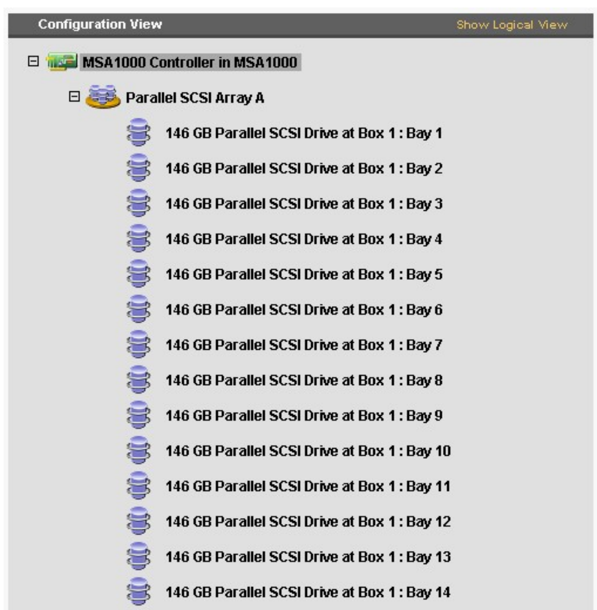
Primary Disk Array—HP StorageWorks MSA1000

The Primary Disk Array that keeps the test data (application data) is an HP StorageWorks MSA1000.

It is composed of 14 x 146-GB 10K disks in an enclosure and driven by an MSA1000 controller with four Ultra3 SCSI (160 MB/s per channel) interfaces. The MSA1000 controller is running on Firmware Version 4.48.

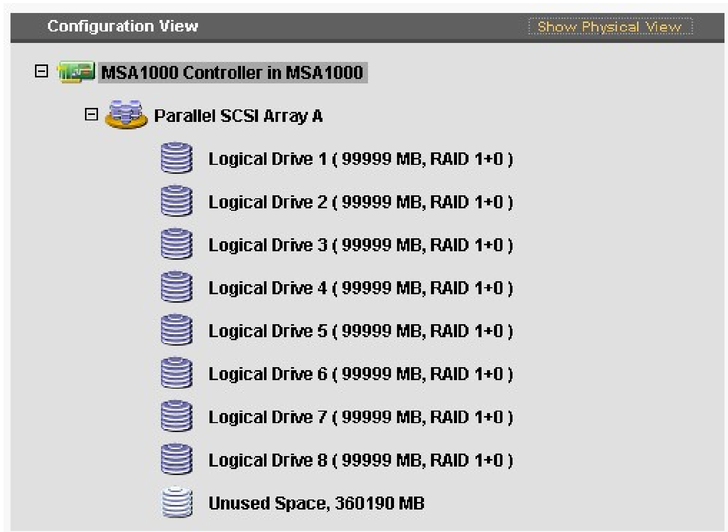
The disk enclosure is configured as one logical disk array with 14 disks.

Figure 4. HP StorageWorks MSA1000 physical array configuration



Eight LUNs protected by RAID 1+0 are distributed across all 14 disks (only LUNs 1–4 will be utilized for tests) as shown in Figure 5.

Figure 5. HP StorageWorks MSA1000 logical drive configuration



Secondary Disk Array—HP StorageWorks MSA1500

The Secondary Disk Array where the file library (backup-to-disk data) for the backup is located is an HP StorageWorks MSA1500.

It is composed of 12 x 250 GB 7.2K disks in an enclosure and driven by a MSA1500 controller with 4 Ultra3 SCSI (160 MB/s per channel) interfaces.

The MSA1500 controller is running on active/passive Firmware Version 5.20 and the connected MSA20 is running on Firmware Version 2.02.

MSA1500 cs	P7428X8X3PR02A	MSA1500cs	5.20	2.02	2.02

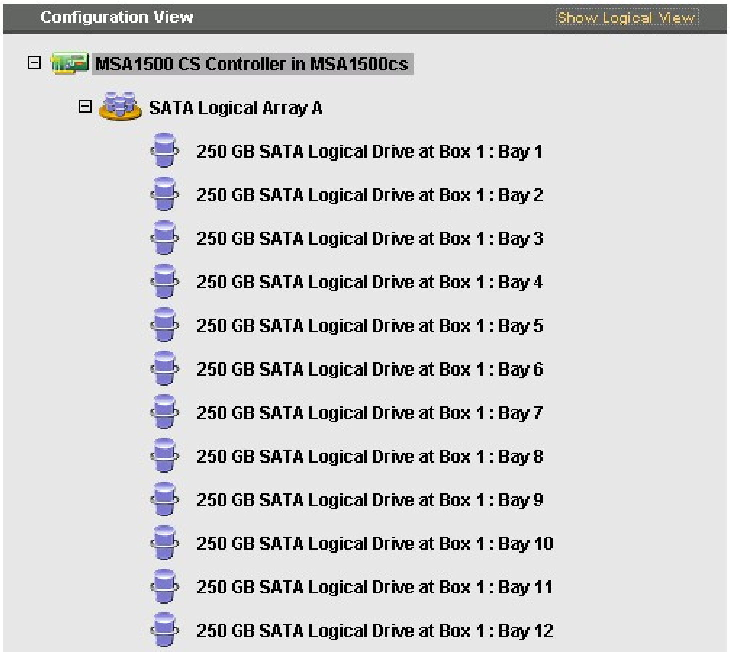
For the tests different configurations are created:

- 1–4 logical disk arrays
- 1–8 LUNs
- RAID levels 0, 1+0, 1, 5

RAID 0 and one logical array

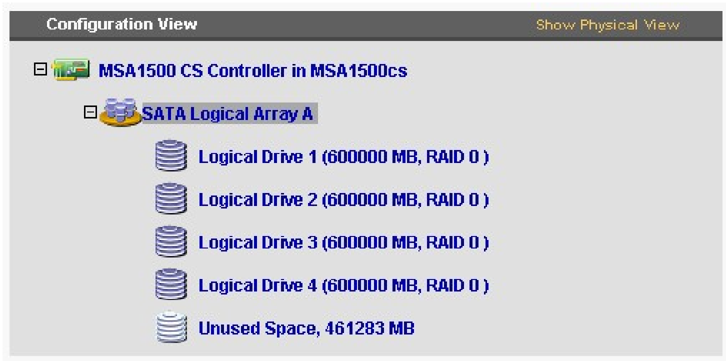
The disk enclosure is configured as one logical disk array with 12 disks.

Figure 6. MSA1500 physical array configuration—RAID 0 and one logical array



Four LUNs protected by RAID 0 are created as shown in Figure 7.

Figure 7. MSA1500 logical drive configuration—RAID 0 and one logical array



RAID 0 and four logical arrays

The disk enclosure is configured as four logical disk arrays with three disks each.

Figure 8. MSA1500 physical array configuration—RAID 0 and four logical arrays

Four LUNs protected by RAID 0 are created as shown in Figure 9.

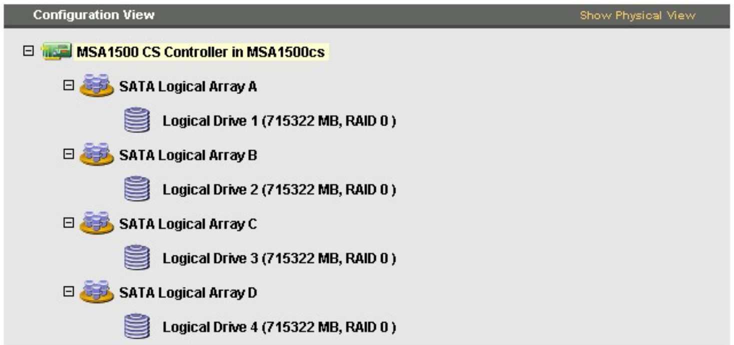


Figure 9. MSA1500 logical drive configuration—RAID 0 and four logical arrays

RAID 1 and one logical array

The disk enclosure is configured as one logical disk array with two disks.

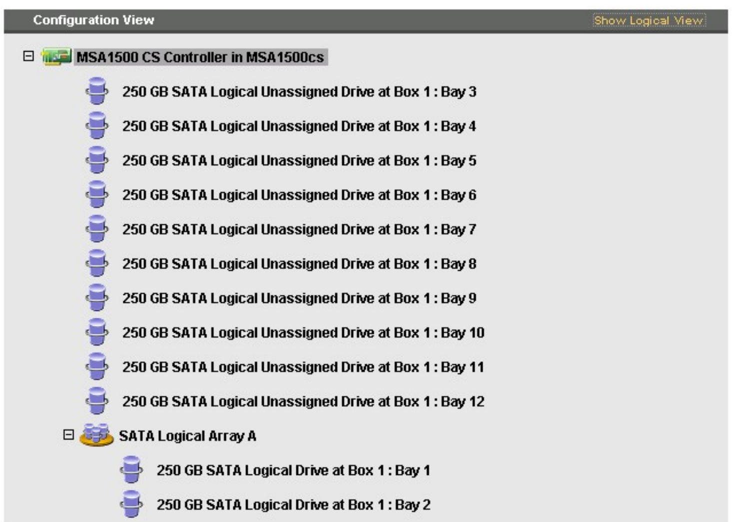
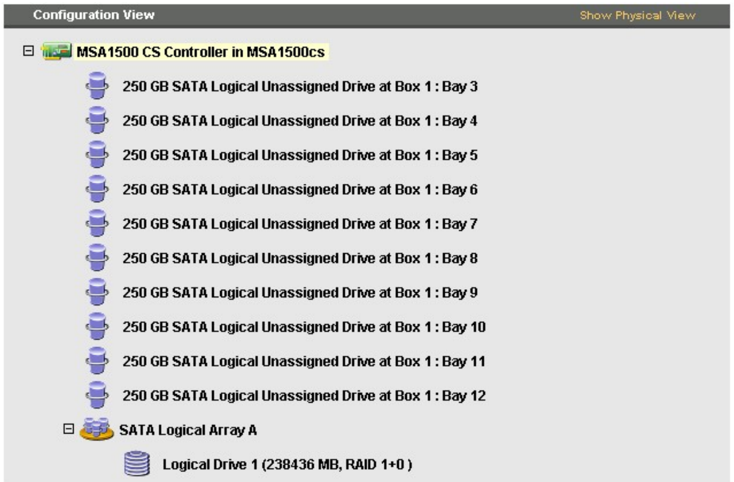


Figure 10. MSA1500 physical array configuration—RAID 1 and one logical array

One LUN protected by RAID 1+0 is created. Due to having only two disks assigned, this is equivalent to RAID 1 (no striping).

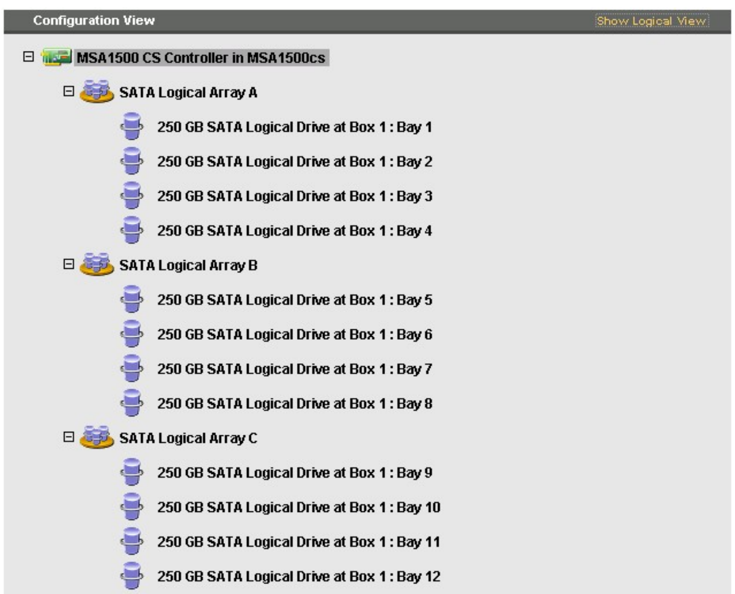
Figure 11. MSA1500 Logical Drive Configuration—RAID 1 and one logical array



RAID 1+0 and three logical arrays

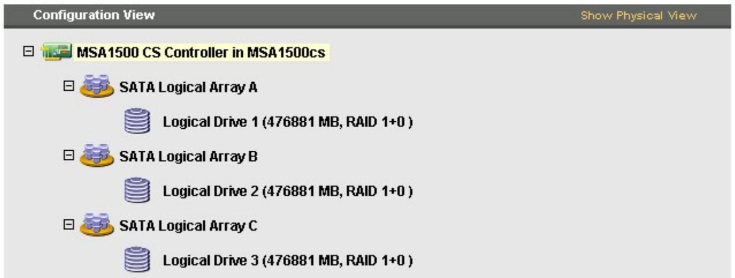
The disk enclosure is configured as three logical disk arrays with four disks each.

Figure 12. MSA1500 physical array configuration—RAID 1+0 and three logical arrays



Three LUNs protected by RAID 1+0 are created (only LUNs 1–2 will be utilized for tests) as shown in Figure 13.

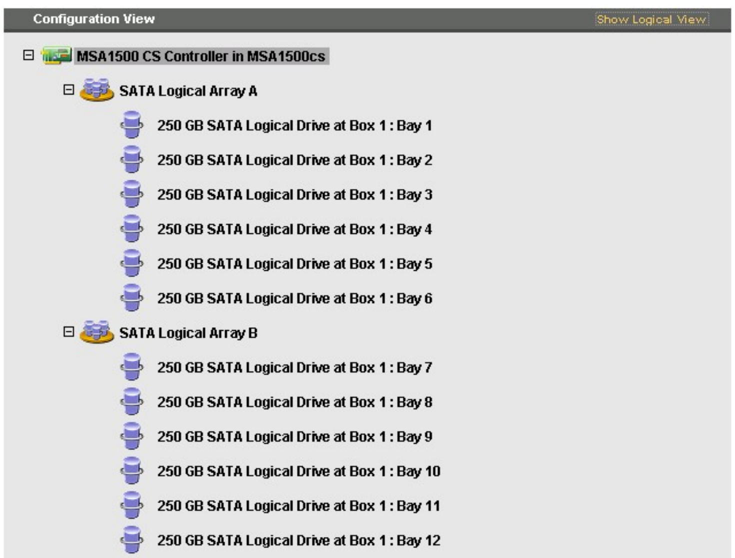
Figure 13. MSA1500 logical drive configuration—RAID 1+0 and three logical arrays



RAID 5 and two logical arrays

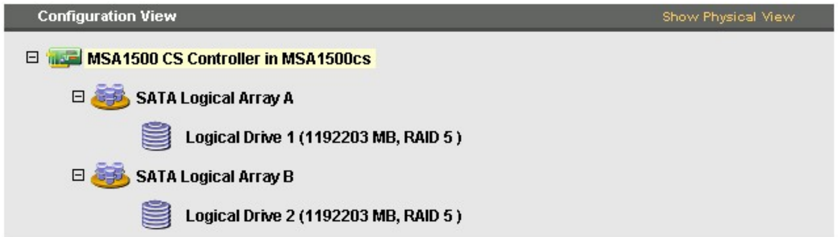
The disk enclosure is configured as two logical disk arrays with six disks each.

Figure 14. MSA1500 physical array configuration—RAID 5 and two logical arrays



Two LUNs protected by RAID 5 are created as shown in Figure 15.

Figure 15. MSA1500 logical drive configuration—RAID 5 and two logical arrays



SAN infrastructure

The Primary and Secondary Disk Arrays are connected by way of an FC switch to the server. The FC switch is an HP FC 1-Gb/2-Gb switch 16B (P/N: A7340A) with firmware revision w3.0.2f. The system is connected to the SAN by way of an FC2408 HBA with Emulex driver version 5.5.20.10. With the newer driver version the performance decreases rapidly (see Tuning recommendations).

Figure 16. FC2408—Driver version

Tape drive

The HP StorageWorks Ultrium 960 tape drive performance is best utilized for high-performance backup in a disk and tape SAN environment, backing up directly from a high-performance disk subsystem (from primary disk storage, split mirrors, snapshots, or clones). The Ultrium 960 drive is not ideal (performance wise) for network backups (for example, 100 Mb/s) or backup from lower performance disk arrays. HP recommends matching appropriate LTO technology generations to your server capabilities. However, if you have capacity needs met by a higher generation product, you can utilize it in the knowledge that it will give you performance headroom in the future.

Table 2. Tape drive characteristics

Specification	HP StorageWorks Ultrium 960
Capacity (native)	400 GB
Transfer rate (native)	Up to 80 MB/s
Data rate matching	27–80 MB/s
Host interfaces	Ultra 320, 2/4 Gigabit FC
Head channels	16
Number of tracks	704
Media length	680 m
WORM media support	Yes
Media compatibility	
Read	Ultrium Gen. 1, 2, 3
Write	Ultrium Gen. 2, 3

Figure 17. HP StorageWorks Ultrium 960 tape drive (Standalone)

Figure 18. HP StorageWorks Ultrium 960 tape drive configuration

For all Ultrium 960 drive backups, HP recommends the tape block size be configured to 256 KB. If you are running the Media Agent on Windows 2003 SP1, there exists a limitation in the involved tape driver (tape.sys) on the data transfers to a block size of no greater than 64 KB. To solve this, see the hotfix from Microsoft that removes the 64-KB limitation on block sizes.

With the hotfix (http://support.microsoft.com/kb/907418/en-us) the tape.sys driver is updated.

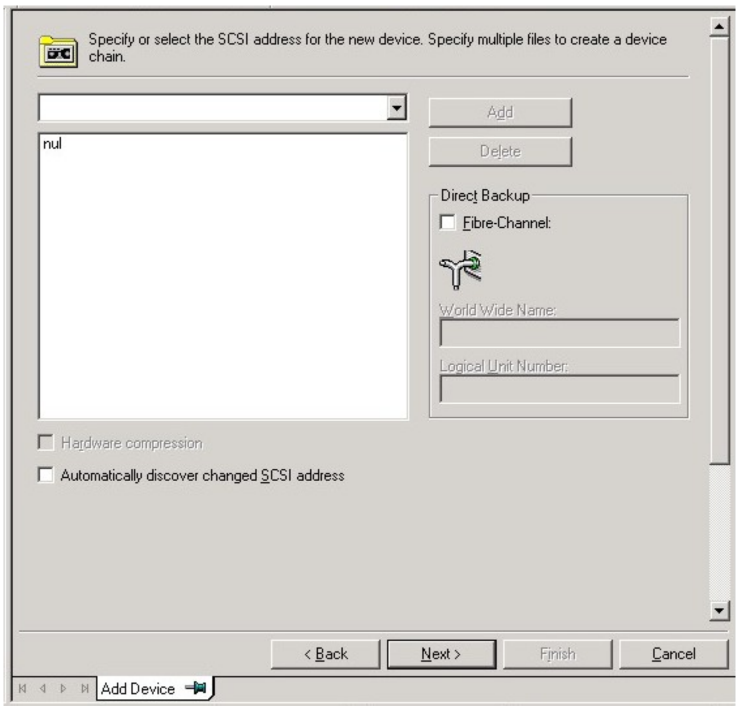
Figure 19. tape.sys

HP Data Protector 6.0 software

HP Data Protector 6.0 software is configured with default values if not specified in the following sections. When more than one LUN is backed up, the data is multiplexed with the described concurrency parameter (see the result tables).

A "NULL" device is created for determining the maximum read rate. This can be achieved by specifying a "standalone" device and defining "nul" in the SCSI address field of the device drive window.

Figure 20. Data Protector NULL device



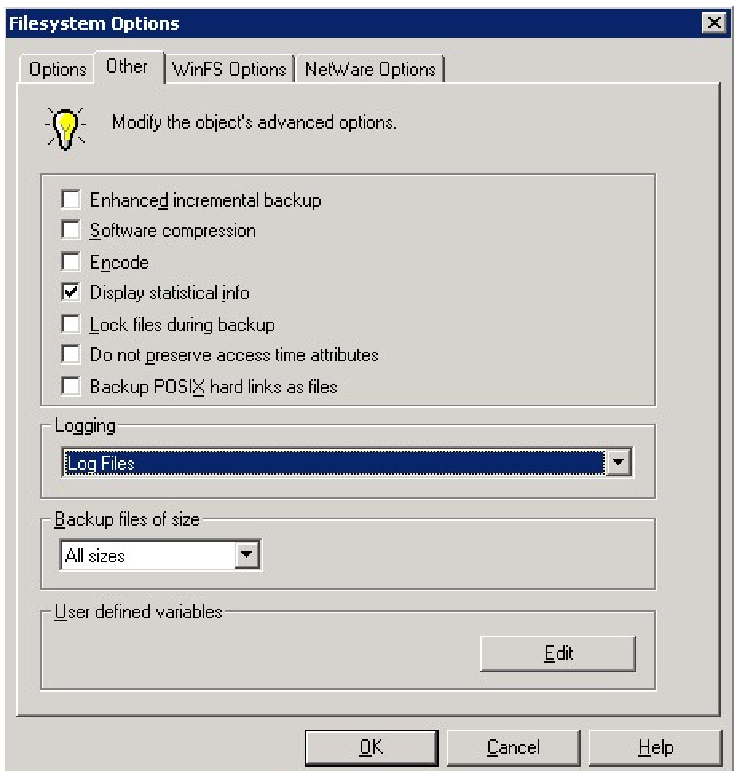
Internal Database (IDB) logging level

The IDB logging level determines how much detailed information about backed-up files and directories (name and versions) is written to the IDB.

The tests are executed with option value “Log Files,” which is appropriate for many customer environments. With this level, directories and files can be browsed before restoring. Data Protector can also fast position on the tape when restoring a specific file or directory. The information written to the IDB does not occupy much space (as with “Log All”), since not all the file details (file attributes) are logged to the database.

Note that with “Log Files,” a filename is only logged once and basically during the initial backup, which requires more CPU and disk resources than the following backups, when the filename is already known.

Figure 21. IDB logging level



File library

The HP Data Protector software Advanced Backup to Disk license (terabyte-based) enables the MSA1500 to be configured as a "file library," so the initial backup from MSA1000 to MSA1500 takes place as a normal backup job with the file library being specified as the destination. The file library can be configured in much the same way as a tape device, with block size, segment count, disk buffers, and so on. It is important to realize that the backup file created on the MSA1500 is for all intent and purpose a tape image file. The benefit of this will be seen later when examining small file backup performance.

Using HP Data Protector software Advanced Backup to Disk, the migration of the backup on the MSA1500 to tape is executed by means of a "copy" style function. It can be interactive (manual) or scheduled to occur at a certain time or even directly after the initial backup to the MSA1500 has completed.

Two file libraries ("FileLibrary," "FileLibrary10GB") are created with different maximum file depot sizes (10, 50 GB). Each file library is created with two writers.

Figure 22. File library configuration—Directories and drives

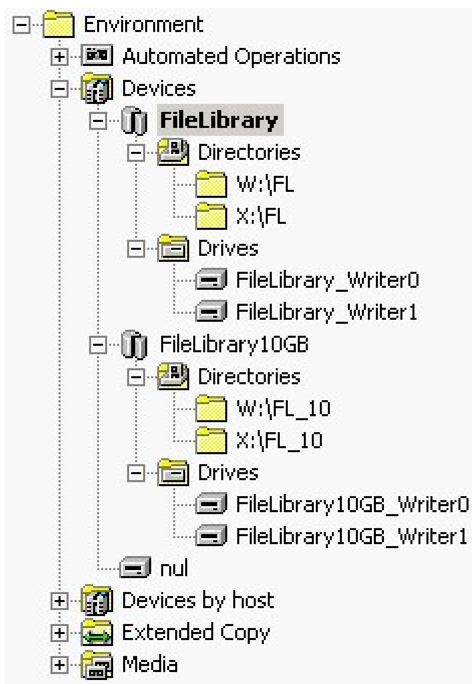
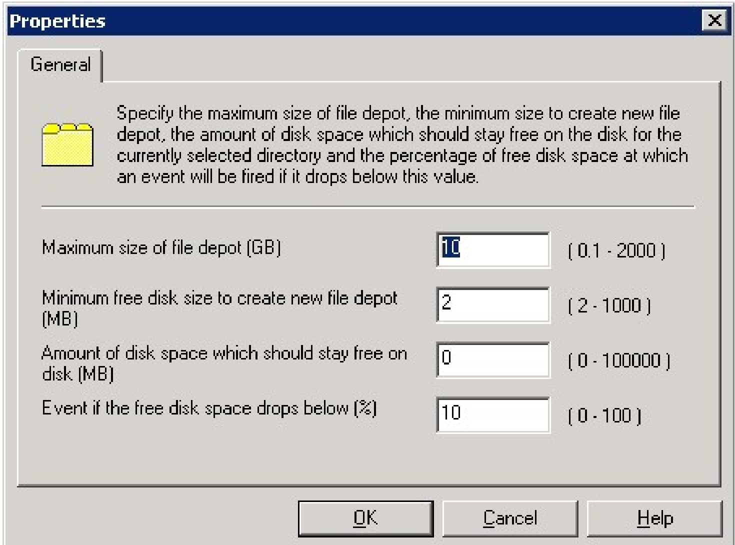
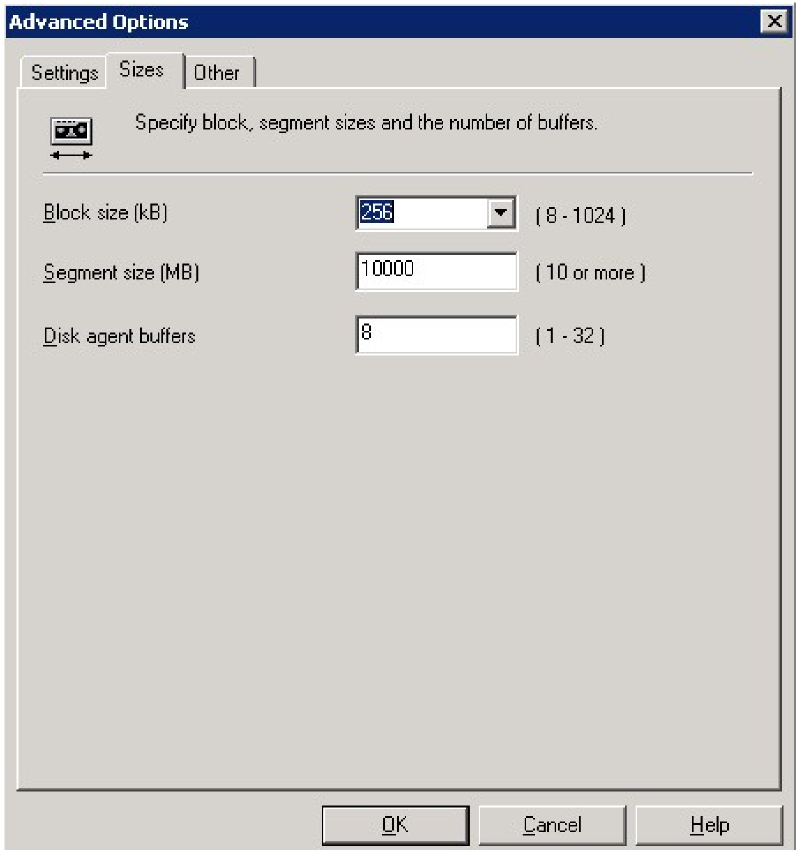


Figure 23. File library configuration—Maximum file depot size 10 GB (50 GB default)



Because we are performing the tape tests on an Ultrium 960 tape drive, the block sizes of file library drives are configured with 256 KB (64 KB default), as described in the "Getting the most performance from your HP StorageWorks Ultrium 960 tape drive white paper" (http://h18006.www1.hp.com/storage/tapewhitepapers.html).

Figure 24. File library configuration—Drive block size 256 KB for Ultrium 960 drive (64 KB default)



Note
For avoiding resource-intensive block repacking during backup from file library to tape drives, it is recommended to configure same block sizes for file library drives and the tape drives. For example, Ultrium 960 tape drives perform best with the block size of 256 KB, which should be also configured for the belonging file library drives.

File system tree walk

During runtime, Data Protector creates backup statistics by a first tree walk, which briefly scans the files selected for the backup and calculates its size, so that the progress (percentage done) can be calculated. In a second tree walk, the data is written to the backup device.

Note
If millions of small files are backed up, the first tree walk could take considerable time and increase the overall runtime. The tree walk can be disabled by setting “NoTreeWalk=1” in the configuration file “<Data_Protector_home>\omnirc.”

Test data

In the following proof points, two different file systems were created to cover large file servers and typical clients, so that the results shown are realistically achievable in similar situations:

- File server data with millions of small files
- Typical client data with fewer files and a broad range of size (KB/MB)

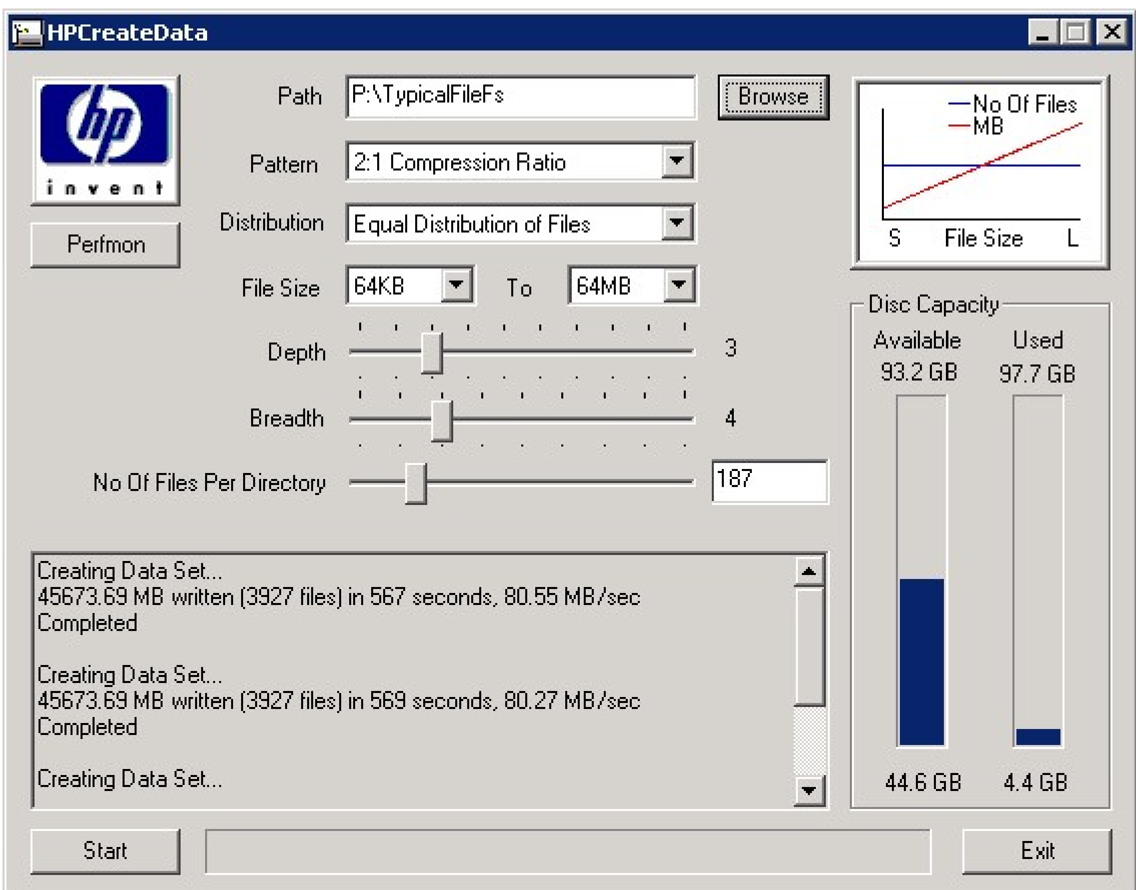
HPCreateData

The datasets were developed using the HPCreateData PAT utility. It generates different file sizes with different data contents (fixed, random, up to 4:1 compression ratio) and different distribution (file-based, MB-based).

Typical files

The “typical” file system is created with file sizes between 64 KB and 64 MB and the compressibility of the data 2:1. The utility creates an equal distribution of files in each directory. Finally, the file system contains 44.6 GB with 3,927 files in 20 folders.

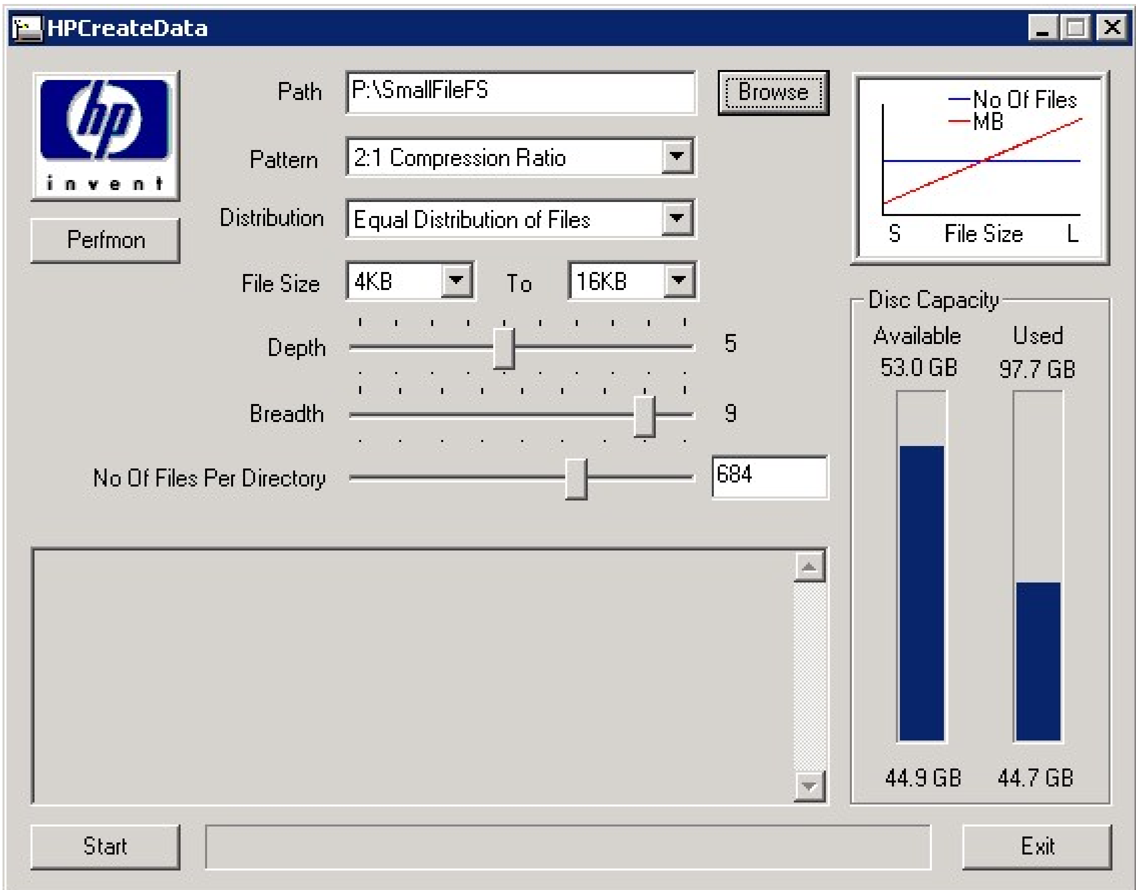
Figure 25. HPCreateData for typical files



Small files

The "small" file system is created with file sizes between 4 KB and16 KB and the compressibility of the data 2:1. The utility creates an equal distribution of files in each directory. Finally, the file system contains 44.9 GB with 5,084,604 files in 7,380 folders. The created files have a name with a maximum of 16 characters for avoiding corner cases.

Figure 26. HPCreateData for small files



HPReadData

The HPReadData PAT utility is useful in assessing the rate at which your disk subsystem can supply data, and this is ultimately what will limit the 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. It can be downloaded free from http://www.hp.com/support/pat. The performance tools are also embedded within the HP industry-leading Library and Tape Tools diagnostics which are downloadable from http://www.hp.com/support/tapetools.

Figure 27. HPReadData for determination of read performance

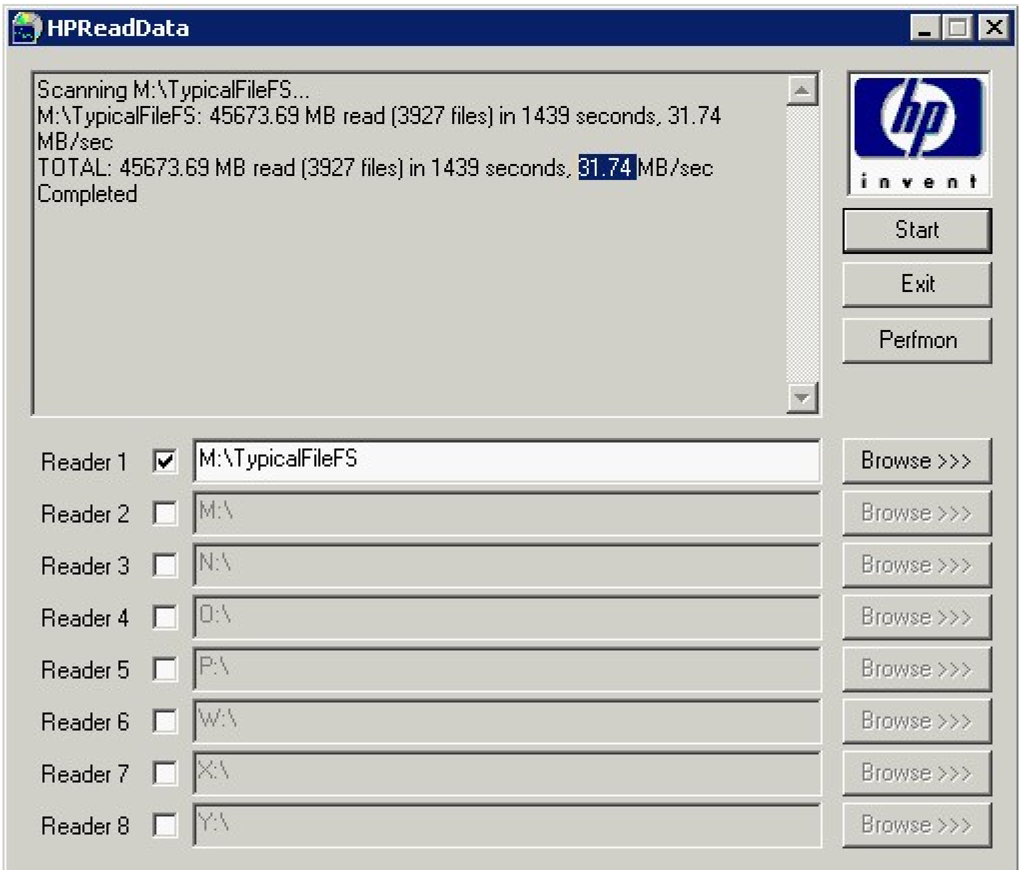


Figure 27 shows HPReadData reading one single LUN in a manner similar to the way a backup application will read files. We can see that the maximum read rate from this configuration is 32 MB/s, so we cannot expect any higher backup transfer performance to tape than this figure.

Library and Tape Tools

The tape drive write tests are executed with the HP industry-leading Library and Tape Tools diagnostics.

The tool is configured to create:

- Zeros with 64-KB block size
- 2:1 compressible data with 64-KB block size
- 2:1 compressible data with 256-KB block size

Figure 28. Library and Tape Tools—Zeros with 64-KB block size

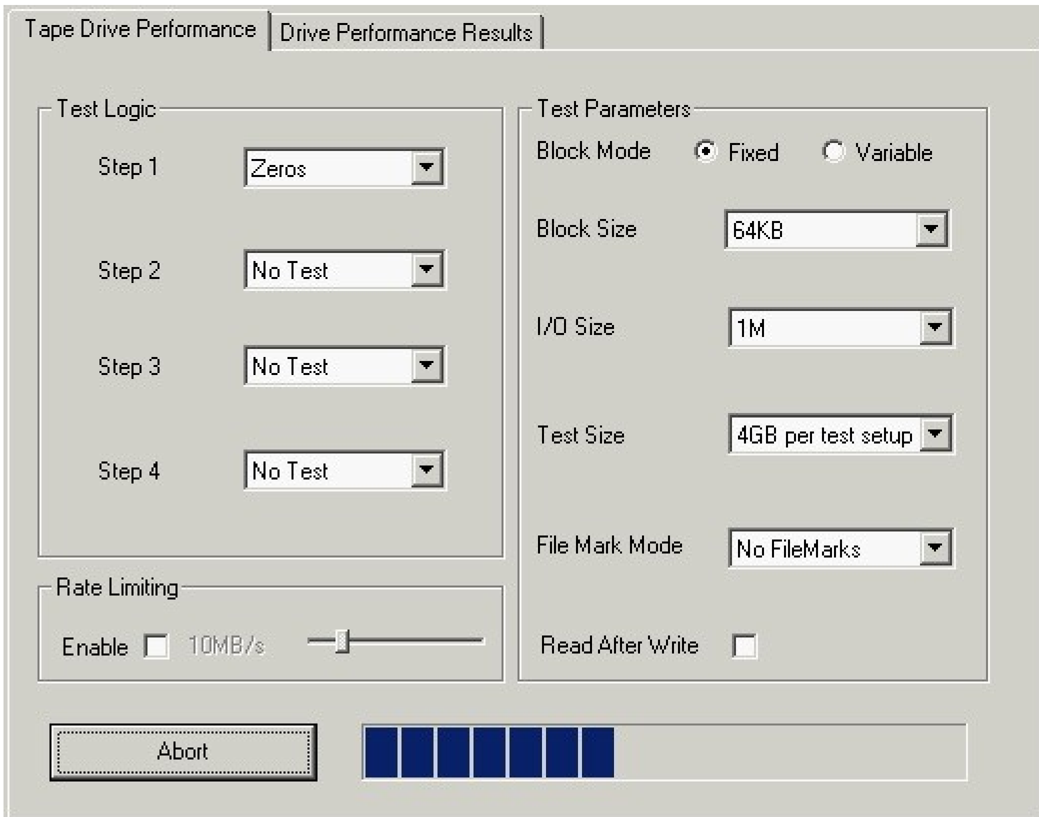


Figure 29. Library and Tape Tools—2:1 compressible data with 64-KB block size

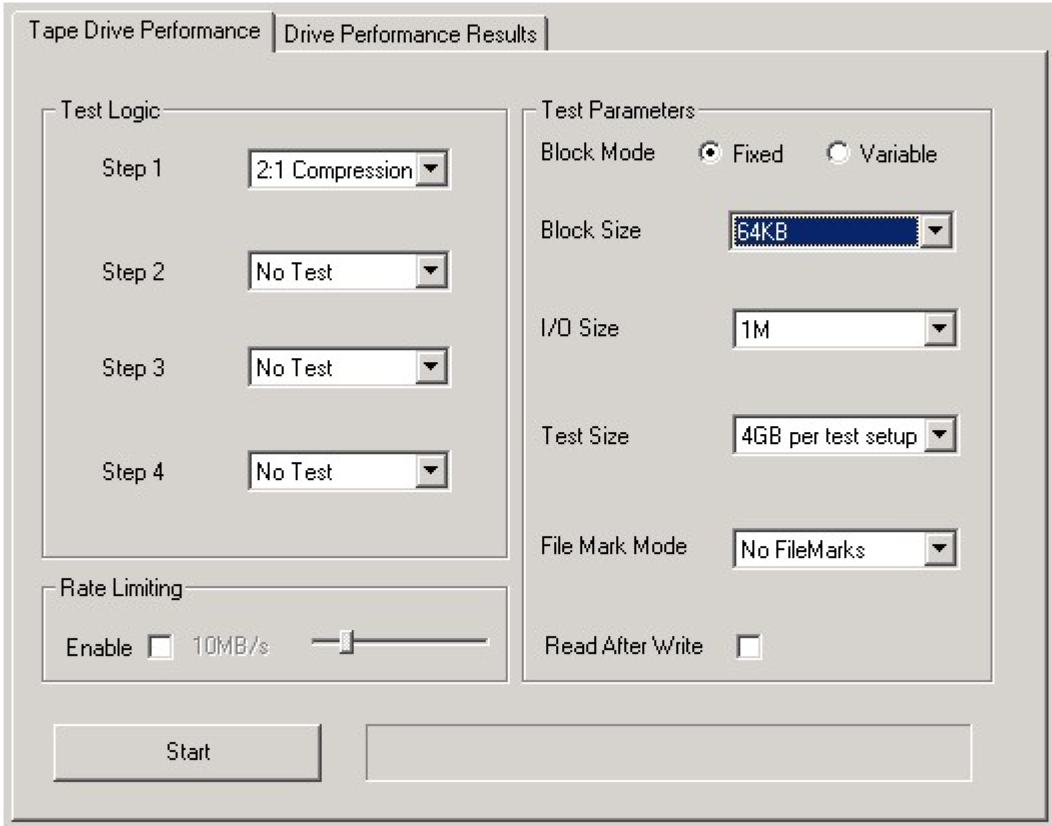
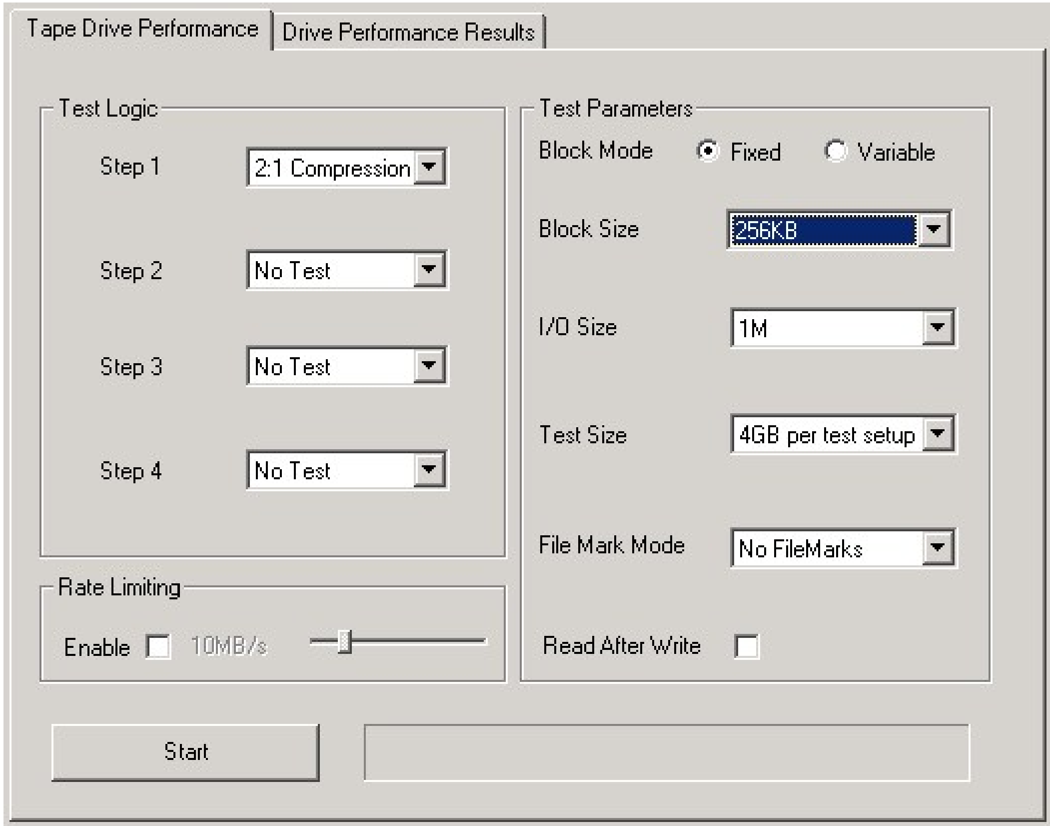


Figure 30. Library and Tape Tools—2:1 compressible data with 256-KB block size



Results

Read and write tests

Table 3. MSA1000 write test—HPCreateData

Specification	MSA1000 transfer rate (MB/s)
Typical Files LUN #1	81
Small Files LUN #1	11

Table 4. MSA1000 read test—HPReadData

Specification	MSA1000 transfer rate (MB/s)
Typical Files LUN #1	33
Typical Files LUN # 1–4	80
Small Files LUN # 1	5
Small Files LUN # 1–2	6
Small Files LUN # 1–4	6

Table 5. MSA1500 write test—HPCreateData (all tests with typical files)

Specification	MSA1500 transfer rate (MB/s)
RAID 0 1 Logical Array LUN #1 (12 Phys. Disks)	79
RAID 0 4 Logical Arrays LUN #1 (3 Phys. Disks)	22
RAID 1 1 Logical Array LUN #1 (2 Phys. Disks)	9
RAID 1+0 3 Logical Arrays LUN #1 (4 Phys. Disks)	14
RAID 5 2 Logical Arrays LUN #1 (6 Phys. Disks)	32
RAID 5 2 Logical Arrays LUN #1/2 (2x6 Phys. Disks)	44

Table 6. MSA1500 Read Test—HPReadData (all tests with typical files)

Specification	MSA1500 transfer rate (MB/s)
RAID 0 1 Logical Array LUN #1 (12 Phys. Disks)	28
RAID 0 4 Logical Arrays LUN #1 (3 Phys. Disks)	28
RAID 0 4 Logical Arrays LUN #1-4 (3 Phys. Disks)	70
RAID 1 1 Logical Array LUN #1 (2 Phys. Disks)	28
RAID 1+0 3 Logical Arrays LUN #1 (4 Phys. Disks)	24
RAID 5 2 Logical Arrays LUN #1 (6 Phys. Disks)	32
RAID 5 2 Logical Arrays LUN #1/2 (2x6 Phys. Disks)	49

Table 7. Tape drive write test—Library and Tape Tools

Specification	Ultrium 960 transfer rate (MB/s)
Zeros, 64 KB	175
Compr. 2:1, 64 KB	155
Compr. 2:1, 256 KB	156

Table 8. Backup to NULL device—MSA1000 to NULL

Specification	NULL device transfer rate (MB/s)	Conc.	Average CPU load %
Typical Files LUN #1	31	1	14
Typical Files LUN #1–4	52	2	23
Typical Files LUN #1–4	77	4	25
Small Files LUN #1	4	1	27
Small Files LUN #1–4 (Initial Run)	5	4	77
Small Files LUN #1–4	6	4	67

Table 9. Backup to tape (Ultrium 960)—MSA1000 to tape

Specification	Ultrium 960 transfer rate (MB/s)—64 KB	Conc.	Average CPU load %
Typical Files LUN #1–4	73	4	47

Specification	Ultrium 960 transfer rate (MB/s)—256 KB	Conc.	Average CPU load %
Typical Files LUN #1	30	1	11
Typical Files LUN #1–4	50	2	19
Typical Files LUN #1–4	74	4	24
Small Files LUN #1	4	1	27
Small Files LUN #1–4	6	4	28

Advanced Backup to Disk

Table 10. Backup to Disk (D2D2T)—MSA1000 to MSA1500 (File Library)

Specification	File Library two drives transfer rate (MB/s)	Conc.	Average CPU load %
Typical Files LUN #1–4 Block Size 64 KB File Depot Size 50 GB	29	4	11
Typical Files LUN #1–4 Block Size 64 KB File Depot Size 10 GB	30	4	14
Typical Files LUN #1–4 Block Size 256 KB File Depot Size 50 GB	31	4	22
Typical Files LUN #1–4 Block Size 256 KB File Depot Size 10 GB	26	4	10
Small Files LUN #1–4 Block Size 256 KB File Depot Size 50 GB	5	3	N/A
Small Files LUN #1–4 Block Size 256 KB File Depot Size 10 GB	6	1	77

Table 11. Disk to Tape (Object copy)—D2D2T—MSA1500 (File library) to Tape (Object copy)

Specification	Ultrium 960 transfer rate (MB/s)—64 KB	Conc.	Average CPU load %
Typical Files LUN #1–4 Block Size 64 KB File Depot Size 50 GB	27	4	22

Specification	Ultrium 960 transfer rate (MB/s)—256 KB	Conc.	Average CPU load %
Typical Files LUN #1–4 Block Size 256 KB File Depot Size 10 GB	33	4	22
Small Files LUN #1–4 Block Size 256 KB File Depot Size 50 GB MSA RAID0 1 Logical Array	25	4	22

Restore

Table 12. Restore from disk—MSA1500 (File library) to MSA1000

Specification	File library transfer rate (MB/s)	Conc.	Average CPU load %
Typical Files LUN #1–4 Block Size 64 KB File Depot Size 10 GB	29	4	12
Typical Files LUN #1–4 Block Size 256 KB File Depot Size 10 GB	34	4	11
Small Files Single 8-KB file	<30s		

Table 13. Restore from tape—Tape to MSA1000

Specification	Ultrium 960 transfer rate (MB/s)	Conc.	Average CPU load %
Typical Files LUN #1	84	1	18
Typical Files LUN #1–4	78	2	20
Typical Files LUN #1	37	4	15
Typical Files LUN #1–4	75	4	21
Small Files Single 8-KB file	110 sec		
Small Files LUN #1–4	Aborted after 30 hours	4	

Observations

HP ProLiant DL380 G4 server

The highest CPU load (80%) occurred during the small file backup.

Tape drives and LTT

The tape drive performance was successfully verified with the HP industry-leading Library and Tape Tools diagnostics.

Compressible (2:1) test data was saved with expected transfer rates as described in the drive specifications: 175 MB/s with Ultrium 960

MSA read and write tests

The MSA read and write tests were executed using the HPReadData/HPCreateData PAT utilities. The results provide information about the disk array performance and finally good estimates for backups and restores.

MSA1000

The MSA1000 RAID 1+0 configuration meets the requirements of production environments (fast and secure). Therefore, only this setup was created and tested. See the Primary Disk Array—HP StorageWorks MSA1000 section.

For typical files, one single LUN was written with 81 MB/s and read with 33 MB/s. For four LUNs simultaneously, the read performance was 80 MB/s. This results show that this disk array configuration cannot completely utilize the high-performance Ultrium 960 drive (175 MB/s, 2:1 comp.).

For small files, one single LUN was written with 11 MB/s and read with 5 MB/s. For four LUNs simultaneously, the read performance was 6 MB/s. These results show that the direct backup to tape would suffer from the poor read performance. On the other hand, it is a good reason to use the Advanced Backup to Disk feature.

MSA1500

The MSA1500 results show that not only different RAID levels configurations influence the performance, but the number of configured logical arrays and LUNs played a major role to achieve a better performance. Some combinations were created and tested. See the Secondary Disk Array—HP StorageWorks MSA1500 section.

The maximum performance for writing was 79 MB/s and for reading 70 MB/s.

RAID 0 was tested only for determining maximum performance values and showing the impact of configuring multiple logical arrays. The RAID 0 configuration showed a single LUN write transfer rate of 79 MB/s with one logical array and 22 MB/s with four logical arrays. The big write difference is based on the horizontal LUN organization. The configuration of one logical array distributes the LUN across all physical 12 disks but the configuration of four logical arrays across only three physical disks. The read transfer rate was almost the same (28 MB/s with one array and 28 MB/s with four arrays). If simultaneously reading from four LUNs, the read transfer rate doubled (70 MB/s with four arrays). This demonstrates that configuring single logical arrays with many physical disks in RAID 0 could improve the write performance. It is important to understand how the data is organized and when physical disks are accessed at the same time. Note that RAID 0 is not recommended due to missing fault tolerance.

RAID 1 was only tested in the one logical array configuration, which resulted in poor performance values (write 9 MB/s and read 28 MB/s). Furthermore, it must be considered that RAID 1 reduces the usable disk space by 50%.

RAID 1+0 achieved better results than RAID 1 (write 14 MB/s and read 24 MB/s) due to striped disk data but the problem of inefficient disk usage remains.

RAID 5 was tested in the configuration with two logical arrays to follow the disk array technology with two SCSI/SATA ASICs. Six physical disks were grouped into one logical array. If writing to one LUN, the transfer rate showed 32 MB/s. If writing simultaneously to two LUNs, the transfer rate increased to 44 MB/s. If reading from a single LUN, the read transfer rate was 32 MB/s. If simultaneously reading from two LUNs, the read transfer rate was 49 MB/s.

The RAID 5 level is chosen for further testing. The results give a good estimate as to what the MSA1500 staging area (file library) is able to deliver. The maximum read performance of 49 MB/s shows similar to the MSA1000: that this best-case scenario cannot completely utilize the high-performance Ultrium 960 drive (157 MB/s, 2:1 comp.).

Backup to NULL device

The backup to the NULL device is a good proof point for the maximum read performance and which tape drive technology would fit to it.

The best typical file backup was achieved with the concurrency of four. The transfer rate of 77 MB/s matches almost the MSA1000 read test result of 80 MB/s. The results for small files are similar.

Backup to tape

The Ultrium 960 drive was tested with different block sizes (64 and 256 KB). The data transfer rate was comparable (73 MB/s and 74 MB/s) but the average CPU load was much different (47% and 24%). Finally, the increased block size of 256 KB reduced the CPU utilization by 50%.

These results confirm the block size recommendation of the Ultrium 960 drive white paper "Getting the most performance from your HP StorageWorks Ultrium 960 tape drive white paper" (http://h18006.www1.hp.com/storage/tapewhitepapers.html).

Advanced Backup to Disk

The Advanced Backup to Disk performance relies basically on the disk arrays and tape drives—source and target. It must be taken into consideration that the MSA1500 provides less performance than the MSA1000 and both less than one Ultrium 960 tape drive. Note that the MSA1500 and its SATA disks have also advantages. It enables you to improve small file backups and fast single file restores.

If typical files were saved from MSA1000 to MSA1500, the maximum transfer rate was 31 MB/s. If saved from MSA1500 to the Ultrium 960 drive, the transfer rate was 33 MB/s. The transfer rate was higher, if the same files were directly saved from the MSA1000 to the Ultrium 960 drive (74 MB/s). In this case, tape was faster than disk.

If small files were saved from MSA1000 to MSA1500, the maximum transfer rate was (6 MB/s). The transfer rate was the same, if the same files were directly saved from the MSA1000 to the Ultrium 960 drive (6 MB/s). But if saved from MSA1500 to the Ultrium 960 drive, the transfer rate increased to 25 MB/s. These results confirm that **Advanced Backup to Disk improves small file backups and tape drive utilization** by first staging to disk and then copying to tape.

The file depot size plays an important role for the CPU utilization and the resulting performance. For the typical file backup (to file library), the default maximum file depot size of 50 GB resulted in a transfer rate of 31 MB/s with an average CPU utilization of 22%. The reduced maximum file depot size of 10 GB resulted in a transfer rate of 30 MB/s but with an average CPU utilization of only 14%. The transfer rate with 50 GB and 10-GB depot size is more or less the same, but the CPU utilization doubles when using a depot size of 50 GB. The creation of a 50-GB file depot took more CPU then creating a smaller 10-GB file depot.

Restore

The data was restored from the file library (MSA1500) or tape (Ultrium) to the original location (MSA1000).

The typical file restore of four LUNs (concurrency = 4) showed a transfer rate of 34 MB/s for the file library, which was beaten by one Ultrium 960 drive (75 MB/s).

The typical file restore of one single LUN (concurrency = 1) resulted in 84 MB/s for the Ultrium 960 drive.

The typical file restore of one single LUN (concurrency = 4) was tested for the Ultrium 960 drive to demonstrate its fast speed. The Ultrium 960 drive showed 37 MB/s, which is half of the full restore performance of four LUNs (84 MB/s). This can be explained by the maximum Ultrium 960 drive performance of 160 MB/s (drive specification, 2:1 comp.), which theoretically provides 40 MB/s for 1:4 multiplexing.

Single File Restore

The restore of single 8 KB files from the file library was tested multiple times. This was executed due to the fact that different files have different positions in file depot segments which affect the runtime. All restores lasted less than 110 seconds each and were even possible within 30 seconds. Multiple single file restores (one after the other) strongly benefit from these little runtimes. **Single restores of small files are much faster from disk than from tape.** In this scenario, tape drives cannot compete due to the time-consuming tape load/unload and positioning procedure.

Tuning recommendations

General

1. Ensure the Ultra 320 SCSI HBA for the tape drive is placed on a dedicated 133-MHz PCI-X bus and is not sharing the bus with other HBAs. For this test environment, the Emulex driver version 5.5.20.20 (10/26/2005) gave the best and expected performance. Tests with driver version 5.1.11.101 (11/7/2005) do not result in the expected performance.
2. For all HP StorageWorks Ultrium 960 drive backups, HP recommends the Tape Block Size be configured to 256 KB. If you are running Windows Server 2003 SP1, install the Microsoft hotfix 907418, which solves the limitation of the tape.sys to 64 KB (http://support.microsoft.com/kb/907418/en-us).
3. Match file library block size to tape drive block size. When using a secondary subsystem as a file library for backups, it is important to give it the same block size as the tape device that will be used in the final part of the data protection process. So in the case of HP StorageWorks Ultrium 960 tape drive, this would be 256 KB. Different block size between source and target medium will cause block repackaging, which has a performance impact in regards to additional CPU load, CRC processing (if enabled), and shared memory utilization.
4. Disable double tree walks for file systems with millions of small files. 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 recommended to set the option "NoTreeWalk=1" in the Data Protector configuration file "<Data_Protector_home>\omnirc."
5. In Windows environments, configure smaller file depot sizes for better performance. Creating larger files requires more CPU and memory resources. For similar environments, the recommended maximum file depot size is 10–50 GB.

Fragmentation of file library file systems

Fragmentation of Windows NTFS file systems can decrease the file library performance. For best performance, each file system should belong to one file library writer only. No other writer, process, or application should write to it.

The following examples show one optimal and one problematic configuration.

The optimal configuration with one writer results in very little fragmentation and good read performance as shown in the next figures.

Figure 31. Optimal file library configuration—One writer per file system

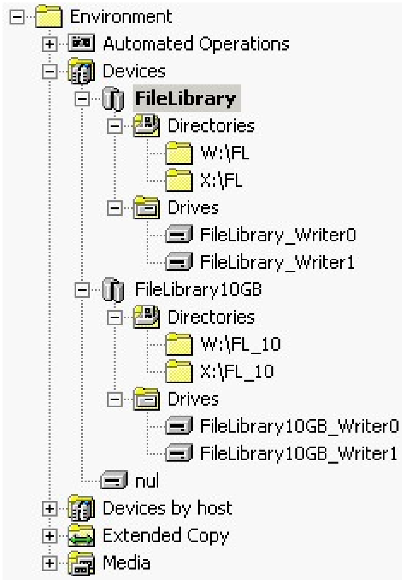
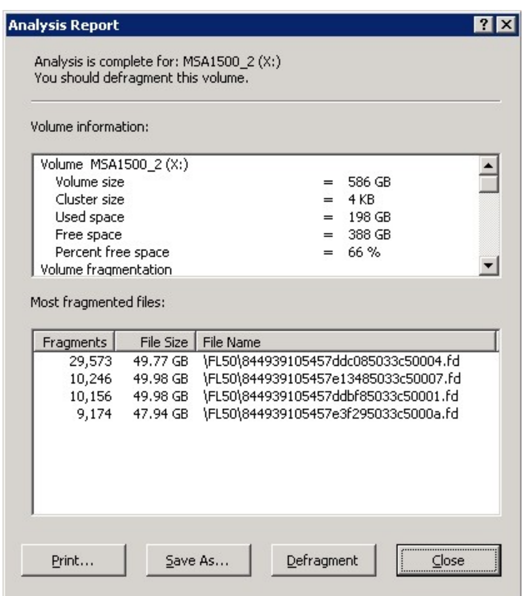


Figure 32. Problematic file library configuration—Large fragmentation



Summary and conclusions

This white paper describes performance-related information for HP Data Protector 6.0 software and the Advanced Backup to Disk feature:

- The described test environment is able to provide a good performance with a low CPU usage and low usage of memory resources. If you reach a high number of small files, I/O performance problems will occur.
- Disk staging acts as a buffer allowing media drives to operate at maximum speeds and provide the option to do automatic data replication during off-peak hours. This technique is highly recommended when backing up numerous small files to prevent poor transfer rates to tape drive.
- Disk technologies (SATA, SCSI) and RAID levels could have major performance differences.
- RAID 0 results in the best performance in writing but should not be considered due to missing fault tolerance features. The backup device—in this case disk—should be very reliable. In general, backups should be available for disasters and selective restores. If a disk fails (disks always spin, tapes not), backup data would be lost.
- Backup and restore of typical files (64 KB–64 MB) could be faster with tape technology than with disk technology (file library).
- Single file restores are executed with an excellent performance by disk technologies. This is very helpful for selective file restores (particularly multiple times) where time is an important issue. No tape must be loaded and positioned, which is a major advantage against tape technologies.
- The performance can be determined by the performance tools described in the HPReadData and Library and Tape Tools sections. It is recommended to check the disk performance before Data Protector is configured. In most cases, an optimized disk array configuration has more impact than default parameter changes of Data Protector.
- For performance data regarding different tape drives (SDLT 320, Ultrium 460), see the HP OpenView Storage Data Protector 5.5 Advanced Backup to Disk Performance Whitepaper.

For more information

- HP Data Protector software
  www.hp.com/go/dataprotector
- HP OpenView Storage Data Protector 5.5 Advanced Backup to Disk Performance white paper
  http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00669480/c00669480.pdf
- Getting the most performance from your HP StorageWorks Ultrium 960 tape drive white paper
  http://h71028.www7.hp.com/ERC/downloads/5982-9971EN.pdf
- HP Performance Assessment Tools
  www.hp.com/support/pat
- Library and Tape Tools
  www.hp.com/support/tapetools
- HP Storage Works Enterprise Backup Solutions Design Guide
  www.hp.com/go/ebs
- EBS Design Guide
  http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00775232/c00775232.pdf

4AA1-2978ENW, June 2007