Coolspirit®
the data storage company

Data Sheets : **White Papers** : Case Studies

> "We never hesitate at introducing Coolspirit to any customer looking for a Quantum product or solution regardless of whether it is a few dozen media or a complete backup system. Their focus on customer needs is second to none."

Ewan Johnson
Quantum Storage UK

Our address
24 The Bridge Business Centre
Beresford Way
Chesterfield
S41 9FG

Get in touch
Call us on: 01246 454222
Email us: web@coolspirit.co.uk
Find us: View location map
Web: www.coolspirit.co.uk

Office hours

mon - thurs	8:30am - 5:30pm
fri	8:30am - 5pm
sat - sun	Closed

Buy with confidence from Coolspirit your authorised Quantum Partner

Quantum. ALLIANCE
Platinum Partner

4th Generation DLT® Technology

openBench Labs

A Division of Strategic Communications

Analysis: DLT-S4 Performance Scaling and Value Proposition for Automation

Quantum

Analysis: DLT-S4 Performance Scaling and Value Proposition for Automation

Author: Jack Fegreus, Ph.D.
Chief Technology Officer
Strategic Communications
http://www.StratCommInc.com

April 12, 2006

Jack Fegreus is Chief Technology Officer at Strategic Communications, which consults with a number of independent publications. He currently serves as Editorial Director of Open magazine, Labs Director of HP World and contributes to InfoStor. He has served as Editor in Chief of Data Storage, BackOffice CTO, Client/Server Today, and Digital Review. Previously Jack served as a consultant to Demax Software and was IT Director at Riley Stoker Corp. Jack holds a Ph.D. in Mathematics and worked on the application of computers to symbolic logic.

Table of Contents

Executive Summary

"Sporting a unique set of manageability features and twice the native data capacity of an LTO-3 tape cartridge, the DLT-S4 is ideally suited for the role of archival device of choice at the widest possible range of IT operations."

As CEOs rely on e-business for revenue growth and regulatory bodies expand the compliance burdens being put on records retention, the explosive growth in corporate storage is unlikely to abate. From 1999 to 2003, IT at global companies increased online storage requirements at an astonishing 79.6% compound annual rate, according to the Storage Networking Industry.

openBench Labs Test Briefing:

Quantum DLT-S4 versus IBM LTO-3 Super Tape Drives

1) **Quantum DLT-S4 completed a 10GB backup 30% more quickly than the IBM LTO-3 and a restore 33% faster.**
2) **Measured Native Throughput:**
   61MB/s for the DLT-S4 versus 76MB/s for the IBM LTO-3.
3) **Data Throughput, 64KB Data Blocks, 3X Compressibility:**
   163MB/s on the DLT-S4 versus 123MB/s on the IBM LTO-3.
4) **For 2X and 3X compressible data, the DLT-S4 approached its maximum throughput speed more rapidly and at smaller tape block sizes than the IBM LTO-3.**
5) **In backup tests using real-world data files, DLTtape® S4 Media was projected to provide a 323% advantage in capacity over an LTO-3 cartridge using an IBM drive.**
6) **Throughput-based Drive Cost:**
   Quantum DLT-S4 $35/MB/s versus IBM LTO-3 $41/MB/s
7) **Capacity-based Media Cost:**
   DLTtape S4 $0.065/GB versus Ultrium LTO-3 $0.16/GB

With no respite in sight for storage growth, IT has embarked on a number of resource management strategies, including SAN/NAS adoption, storage consolidation, and Information Lifecycle Management, with profound implications for tape storage. Robert Abraham, author of Freeman Report's *2005 Tape Library Outlook,* states, "The continuing adoption of SAN and NAS network storage solutions and the insatiable demand for storage will be the dominant driving forces in the future growth of automated tape libraries." Nonetheless, the nature of the data being collected is even more important than the volume of data.

Dealing with issues from shrinking product life cycles to expanding global markets has led businesses to pursue real-time sense-and-respond operational strategies. For IT, implementing such a strategy fosters a growing dependence on database-resident transactional data. which makes production systems particularly susceptible to performance degradation as the volume of data expands. As a result, the primary role of tape media is beginning to shift from off-line backup to near-line archival operations for read-only archived transactions. This is makpng cartridge capacity equal in importance-if not more important-to tape storage throughput.

Looking to capitalize on the growth of online data, Quantum unveiled the 4th generation of SDLT tape drives and media, the DLT-S4 and DLTtape® S4. Sporting a unique set of manageability features and twice the native data capacity of an LTO-3 system, the DLT-S4 is ideally suited for the role of archival device of choice at the widest possible range of IT sites. To assess the DLT-S4, Quantum Corporation commissioned Strategic Communications to examine the DLT-S4 in terms of its performance and functionality. Of particular interest was the new drive's potential to lower the Total Cost of Ownership (TCO)in robotics-assisted environments.

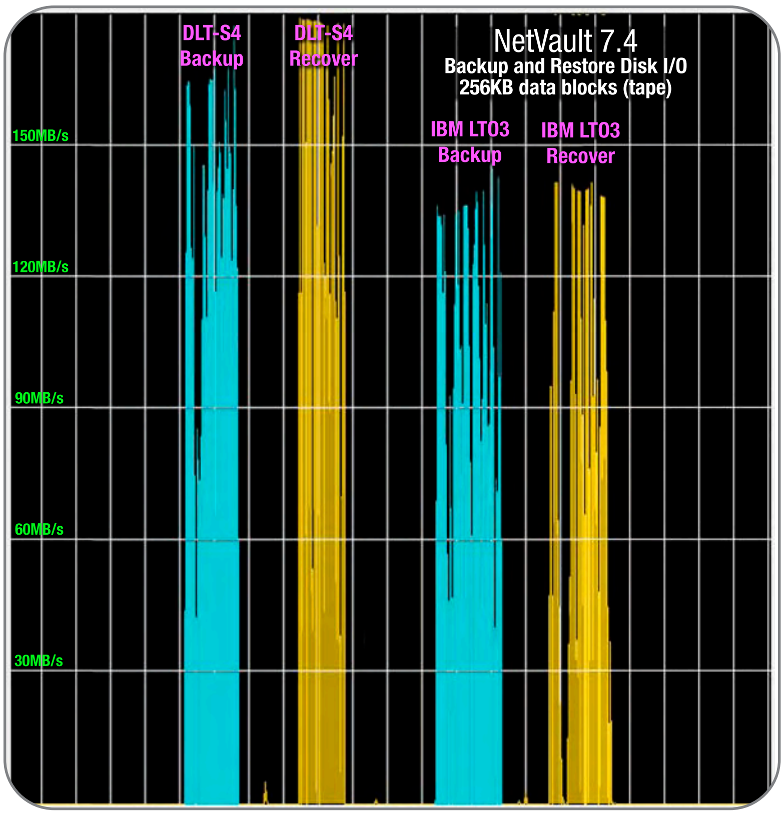
Performance Spectrum

"The DLT-S4 sports a slower native throughput rate than the IBM LTO-3, but completes a 10GB backup 30% more quickly and a restore 33% faster."

To examine potential performance of Quantum's DLT-S4 in a datacenter environment, openBench Labs installed the tape drive via an Ultra320 SCSI interconnect on a high-end server featuring a PCI Express (PCIe) I/O architecture. By replacing the old shared-bus I/O architecture of a PC with a silicon-based switched I/O architecture modeled on super computers and mainframes, PCIe represents a significant advance in I/O throughput.

For comparison, we also installed an IBM LT03 drive, which inexplicably sports an Ultra160 SCSI interface. We chose an XtremeServer from Appro to run all of our tests. This 3U server features four AMD dual-core Opteron CPUs and nVidia's nForce Professional 2200 and 2050 ASICs for PCIe. The nVidia nForce Professional 2200 also acts as a bridge for internal I/O connections including SATA, for which this server provides six hot-swap bays. We populated these bays with Western Digital 15K Raptor drives—one was reserved for SUSE Linux 10 Professional and the remaining five were combined as a high-throughput, software-RAID device.

Running our benchmark for disk throughput, oblDisk, we pegged the SATA array as capable of sustaining extraordinary levels of continuous throughput, which averaged 263MB per second on reads and 187MB per second on writes. With the Appro XtremeServer's internal SATA RAID array able to sustain four times the native throughput rates of the tape drives we were testing, we would not have to stream data from multiple disks simultaneously to satisfy the input needs of the Quantum DLT-S4. As a result, we would be able to log backup throughput—and hence data compression—on a file-by-file basis.



We monitored I/O throughput for backup and restore operations run with NetVault v7.4. Compare to our benchmark data, the results were strikingly counterintuitive.

Using 256KB data blocks. we measured native (uncompressed) throughput for the IBM LTO3 drive, at 76MB per second and 61MB per second for the Quantum DLT-S4. Nonetheless, on a backup, throughput peaked at 137 MB per second using the IBM drive, but reached 175MB per second using the DLT-S4. The DLT-S4 also considerably outpaced the IBM LTO-3 when running restore operations. That leaves quite a paradox: The DLT-S4 sports a native throughput rate that's 20% slower than that of the IBM LTO-3, but completes a 10GB backup 30% more quickly and a restore 33% faster.

We ran backup and restore operations using NetVault v7.4 on our standard 10GB backup test set. This set includes a mix of Microsoft® Outlook™, Access™, HTML, JPEG, and GIF data files. The compressibility of these files falls into a range that typically runs from incompressible to highly (3x) compressible. Average data compression on a backup for the entire test set is typically 1.8-to-1. **Using the IBM LTO-3, we estimated average data compression to be 1.3-to-1; however, using the DLT-S4, we calculated average data compression to be 2.1-to-1 .**

06

Since the introduction of the DLT 7000, all of our real-world file sets for backup testing have consistently demonstrated a narrow range for average compression: from 1.7-to-1 up to 1.9-to-1. We leveraged that fact in our oblTape benchmark, with which we have been able to consistently and accurately determine the probable upper and lower bounds for backup and restore throughput. To make that determination, we streamed a mix of incompressible random data with data that was calibrated to produce a 2-to-1 compression ratio using a limited set of characters in a frequency determined using a normal distribution pattern. Invariably, backup and restore tests would match oblTape results within a narrow range around a mix of 80% 2X data and 20% incompressible data.

In all of those tests, the determining factor was the attempt to compress incompressible data. Compression schemes need to add metadata about the compressibility of the original data in order to reconstitute that data in a restore operation. Ideally, the total number of bits of metadata and compressed data are less than the number of bits in the original data. In that case, fewer bits will be written to tape than are read from disk. Throughput is then calculated using the number of raw bits read and the time spent writing the compressed data. When the original data cannot be compressed, that metadata becomes pure overhead. Both the metadata and original data are written to tape and the perceived throughput rate drops to less than the drive's native throughput specification.

Now the Quantum DLT-S4 joins LTO drives with electronics that are able to compare buffered data before and after compression without any perceptible impact on throughput performance. In effect, this scheme, which LTO drives dub Adaptive Loss-less Data Compression (ALDC), allows DLT-S4 and LTO-3 drives to turn data compression on and off inline. With both drives able to pick the optimal data set and eliminate the possibility of the drive slowing to less than its native throughput rate, the key area for differentiating these drives shifts to high-end throughput.

A number of strategies are available to accomplish that task. One is to increase linear tape speed. That strategy raises the native throughput rate of the drive, but it also increases the overhead of repositioning tape. Alternatively, today's faster ASICs provide the opportunity to compress data more quickly, which is precisely what Quantum did with the DLT-S4.

To test this new aspect of tape drive performance with our oblTape benchmark, we needed to make the compressibility of the patterned data generated by oblTape a variable with discrete values: 1X (incompressible), 2X, 3X. Once we associated a discrete input set with the values used to generate specific distributions of characters for the data pattern streamed to the tape drive, we were able to examine tape throughput as a function of both formatted block size and data compressibility.

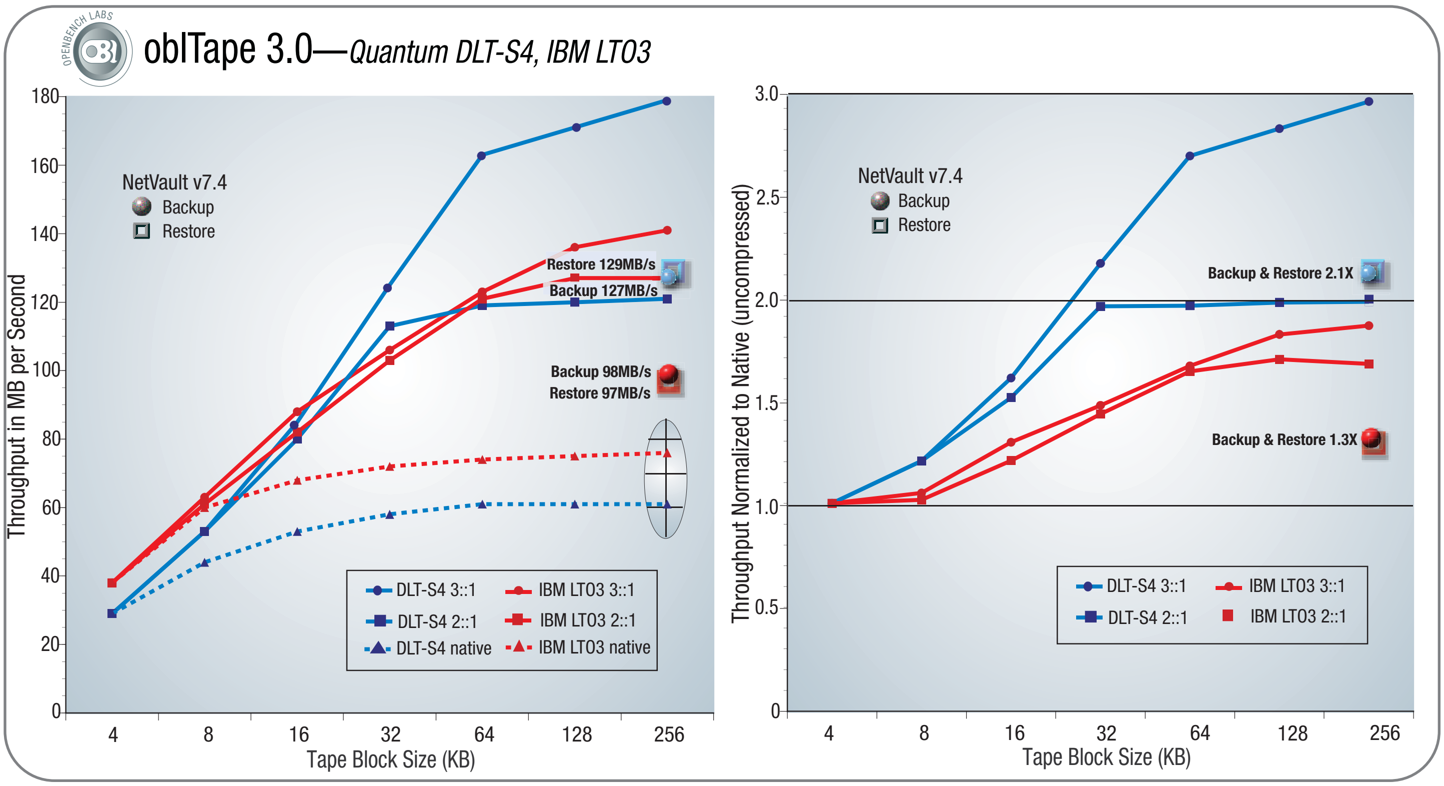
With the advent of firmware that changes a drive's tape speed based on data flow, the formatted size of data blocks directly affects uncompressed data throughput, which represents the effective native throughput rate. Using small data blocks increases the number of interrupts in the flow of data and triggers a slowing of the drive's speed. A number of backup

oblTape v3.0

To test the throughput potential of a tape drive, we used our openBench Labs tape benchmark, oblTape. The benchmark generates two types of data: purely random and patterned. The patterned data is generated from a fixed set of characters in a normal distribution. The standard deviation of the distribution can be set to provide either a 2-to-1 or a 3-to-1 compression ratio on a tape drive using the Digital Liv Zempel (DLZ) algorithm. All data is streamed directly to the device from memory to avoid any issues of bandwidth.

The data can be streamed in block sizes of 2n KB, where n ranges from 0 to 8. This simulates the differences in the way backup applications write data to a tape drive. In particular, many mid-range packages that run on multiple operating systems only allow 64KB tape data blocks—the maximum default size for Windows—for compatibility. On the other hand, enterprise-class backup applications often provide for writing data in 128KB and 256KB blocks.

packages with versions for both Windows and Linux have a fixed block size of 64KB, the default maximum I/O size for Windows. As a result, the rate at which the drive's electronics can ramp up throughput to its maximum potential vis à vis block size is very significant for performance on Microsoft® Windows Server™ 2003.



Plotting the actual benchmark results provides a good macro view of relative drive throughput. Both the Quantum DLT-S4 and the IBM LTO-3 essentially reached their maximum native throughput rates with a block size of 64KB. With 2X compressible data, raw throughput writing data to tape was quite similar for both drives with the IBM LTO-3 typically holding a 5% edge. With 3X compressible data, however, raw throughput diverged dramatically. Interestingly, the Ultra160 SCSI connection never impinged on throughput for the IBM LTO-3, which never reached its 2X-compression limit even with 3X-compressible data.

That throughput data was very consistent with the results of our backup and restore tests using NetVault v7.4 and our standard 10GB backup test set. This set includes a mix of Microsoft® Outlook™, Access™, HTML, JPEG, and GIF data files, which fall into a typical range for compression—from

08

Normalizing raw throughput relative to native reveals a clear edge for the DLT-S4. **Throughput for the DLT-S4 rapidly converged on theoretical limits for 2X- and 3X-compressible data.** In contrast, throughput for the IBM LTO-3 never converged on its 2X limit. **Based on the backup tests, throughput translated into an average compression rate of just 1.3-to-1 for the IBM LTO-3 and 2.1-to-1 for the DLT-S4.**

incompressible to highly (3x) compressible. In those tests, we monitored throughput on a file-by-file basis. Peak file throughput using the DLT-S4 reached 175MB per second and 137MB per second using the IBM LTO-3 drive, which approximated the oblTape results for both drives using 3X compressible data. Average backup throughput using the DLT-S4 was pegged at 127MB per second and 98MB per second using the IBM LTO-3.

Real insight into the differences between the drives really becomes evident, however, when the raw benchmark data is normalized to native throughput. When the data is put in that form, a detailed micro view of how each drive handles data compression can be garnered.

With 2X- or 3X-compressible data, throughput for the DLT-S4 rapidly converges on its theoretical limits: 120MB or 180MB per second. Notably, 32KB transfers on the DLT-S4 with 2X-compressible data appear to be anomalous on the raw throughput graph as the DLT-S4 shows distinctly better throughput than the IBM LTO-3. When viewed on the normalized throughput graph, that data point is predictable rather than anomalous. In a normalized context, that data point is completely in line with the DLT-S4's characteristic rapid convergence on its theoretical limit. In contrast, the IBM LTO-3 drive increases throughput at a far slower rate and never approaches either it's 2X- or 3X-throughput limit.

In addition to data throughput, the efficiency of a tape drive's data compression scheme also affects another important aspect of the drive's value proposition: cartridge capacity. Since the IBM drive converges on a compression rate of only 1.75-to-1 with data that is 2X-compressible using the DLT-S4, an IBM LTO-3 cartridge will at best hold 700GB rather than 800GB of data as compared to 1,600GB for a DLTtape S4 cartridge.

Value Proposition

"At a cost of only $0.06 per GB, the extended capacity of a DLTtape S4 cartridge makes it possible to configure very cost-effective Petabyte-scale storage subsystems."

The explosion in storage growth has been no small factor in the rising use of tape automation. The Freeman Reports' *2005 Tape Library Outlook* projects that the tape library market, as measured in units shipped, will maintain a compound annual growth rate of 8.75% through 2010. Robert

C. Abraham, the report's author, sees standardized media, large capacities, high transfer rates, and well articulated migration paths as the factors that account for the popularity of technologies in the highly competitive tape automation market.

Those characteristics articulated by Abraham are a perfect match for the DLT-S4 and highlight just why tape automation devices will be a key market for the DLT-S4. Whether it's the 2U autoloaders that increasingly replace stand-alone tape drives in the SMB arena or the silo-class libraries that attract the attention of large companies, the key factors used to characterize these automated storage subsystems are the characteristics of the tape drives that they contain.

Super Tape Drive Characteristics

Tape Drive	Drive Street Price	oblTape Uncompressed Throughput	Native Throughput Specification	Benchmark Backup/Restore Throughput	Throughput-based Cost ($/MB/Second)
Quantum DLT-S4	$4,495	61 MB/sec	60 MB/sec	127/129 MB/sec	**$35**
IBM LTO-3	$4,050	76 MB/sec	80 MB/sec	98/97 MB/sec	**$41**

This reliance on tape drive characteristics, such as transfer rate and cartridge capacity, arises out of the many library and autoloader designs that now promote the interchangeable use of either DLT or LTO drives. For these automated tape subsystems, it is the choice of the drive type that determines the distinguishing characteristics.

Super Tape Cartridge Characteristics

Tape Drive Cartridge	Cartridge Street Price	Native Cartridge Capacity	oblTape 2X-Compressibility Test Compression	Benchmark Backup Test Compression	Backup-based Capacity Projection	Capacity-based Cost ($/GB)
DLTtape S4	$100	800GB	2.0-to-1	2.1-to-1	**1,680GB**	**$0.06/GB**
Ultrium LTO-3	$85	400GB	1.7 -to-1	1.3-to-1	**520GB**	**$0.16/GB**

The first generation of LTO Ultrium drives, introduced a hybrid scheme of digital and analog electronics to give LTO drives an edge in throughput performance over early SDLT drives. As a result, the market share of LTO drives grew rapidly in the tape automation arena. Nonetheless, advances in both the electronics and mechanics of high-end tape drives now shifts that performance edge to the pure digital circuitry of the DLT-S4 drive. What's more, the edge in compression performance translates directly into an all-important edge in cartridge capacity.

Starting with the SDLT 320, Quantum has been at the forefront of extending the capacity of tape cartridges. For capacity, the key metric for any magnetic medium is areal density. For tape, this means writing more bits per inch (bpi) on data tracks and packing more tracks per inch (tpi) on the tape.

To pack and accurately access more data tracks on a tape cartridge, Quantum uniquely married optical laser servo technology with magnetic read/write technology to create Laser Guided Magnetic Recording (LGMR). Traditional magnetic tape designs reserve 10-20% of the recording surface for embedded servo track information. With LGMR, Quantum laser-etches optically readable servo tracks on a specially formulated back coating of the media. The DLT-S4 utilizes a three-beam hologram to precisely follow those tracks. As a result, excluding a small guard band area, 100% of the magnetic tape surface is available for reading and recording data tracks.

LGMR provides the means to format DLTtape S4 cartridges at 2,988 tpi. In comparison, the IBM LTO3 drive sports 1,408 tpi. That gives DLTtape S4 cartridges a 2-to-1 advantage in native (uncompressed) capacity. When the added advantage of the DLT-S4 drive's compression efficiency is factored into the capacity equation, the DLT-4 advantage grows beyond 3-to-1. At a cost of only $0.06 per GB, the extended capacity of a DLTtape S4 cartridge makes it possible to configure very cost-effective Petabyte-scale storage subsystems.

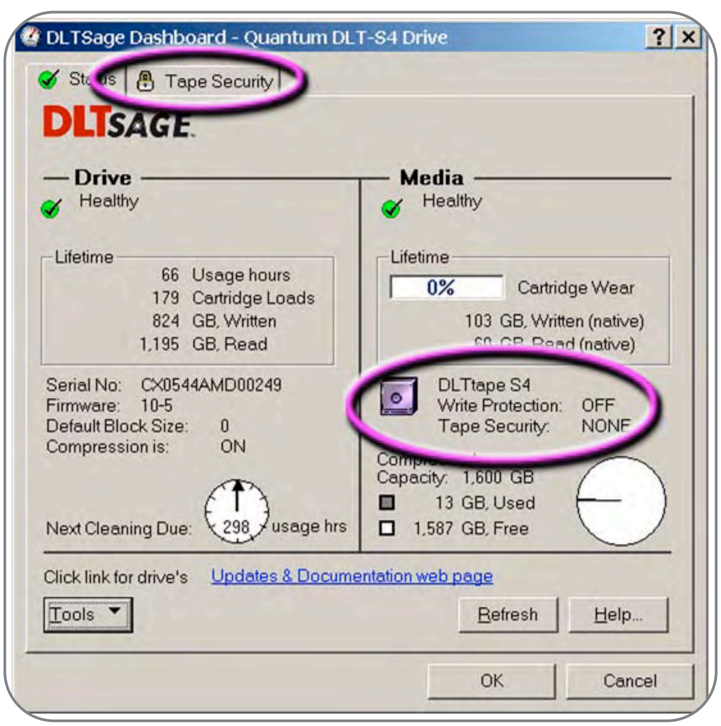
What's more, the laser servo tracks on the DLTtape S4 media cannot be magnetically erased. This indelible servo information eliminates the need to magnetically pre-format tapes and makes the bulk erasure of DLTtape S4 cartridges possible. Traditional media must have the servo tracks re-recorded after a bulk erasure. That process is highly susceptible to environmental variables and highly discouraged.

A further complicating library comparison is the tight coupling of features with software. Library functionality depends on software just as much as it does on the robotics hardware. For most sites, this software will come as embedded Web/Java utilities resident on the device or as modules that are integrated into the backup and archival software packages utilized on site. For DLT-S4 drives, this opens the door to further distinguish the reliability of any library using the drive via Quantum's DLTSage™ achitecture.

IT's notions about reliability, however, are rapidly growing beyond simple media readability and data reproducibility. For larger and more sophisticated IT sites, reliability is beginning to encompass the wider notion of media integrity. In a growing number of instances, restoring backup data is often no longer sufficient: New legislative mandates on data integrity are making it necessary to demonstrate that the restored data is accurate. In this area, requirements call for a compliant electronic storage medium to support integrity protection, accessibility, duplication, and auditing.

Naturally, IT wants any solution to integrate into existing infrastructure. Using DLT drives, any site can easily extend traditional backup operations to secure archival operations. The DLTSage tool suite includes DLTSage WORM, which exploits the firmware in DLT drives to convert standard DLTtape cartridges into WORM compliant media for archival operations.

Creating a WORM cartridge involves writing a unique electronic key that cannot be altered on a standard DLTtape cartridge. This unique identifier creates a tamper-proof archive cartridge that meets stringent requirements for integrity protection while providing full accessibility for reliable duplication. Now the The DLT-S4 builds on that electronic key mechanism to create a tape security system that will protect the data in the event that the tape cartridge is lost or stolen.



The new DLTSage™ Tape Security system uses the notion of electronic keys to prevent unauthorized access to data on tape cartridges. Using a new simplified interface called the DLTSage Dashboard, a secure user can add, change, and remove keys that lock and unlock cartridges with respect to a standalone drive or a library with multiple drives. In this way, the DLTSage Tape Security system provides a means to secure a tape from unauthorized access without the added overhead of encrypting and decrypting the data on a backup or restore

The DLTSage Dashboard presents a user-friendly, high-level summary of the state of the tape media in a DLTtape cartridge, which includes the amount of tape that remains free. The Dashboard also provides the means to apply electronic security keys to devices and cartridges. Once a security key is assigned to a DLTtape S4 cartridge, its data will only be accessible on a device that has a matching key. More complex tasks, such as firmware updates and WORM conversion require DLTSage xTalk, which is a sophisticated diagnostic utility that provides the means to test all of the performance aspects of a DLT drive.

operation.

As businesses become more global and more 24X7, they become more dependent on digital data. That has triggered an evolution in IT practices, which is changing the role of tape from an off-site backup medium to a near-line archive. This emphasizes the need for capacity over that of performance, shifts the focus of attention to secure self-managing devices, and leaves the DLT-S4 perfectly positioned to leverage all of these changes.