LAB VALIDATION REPORT

NetApp VTL

Enterprise Class, Storage Efficient, Disk-Based Backup

By Tony Palmer
With Brian Garrett

February, 2009

Table of Contents

ESG Lab Reports

The goal of ESG Lab reports is to educate IT professionals about emerging technologies and products in the storage, data management and information security industries. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by NetApp.

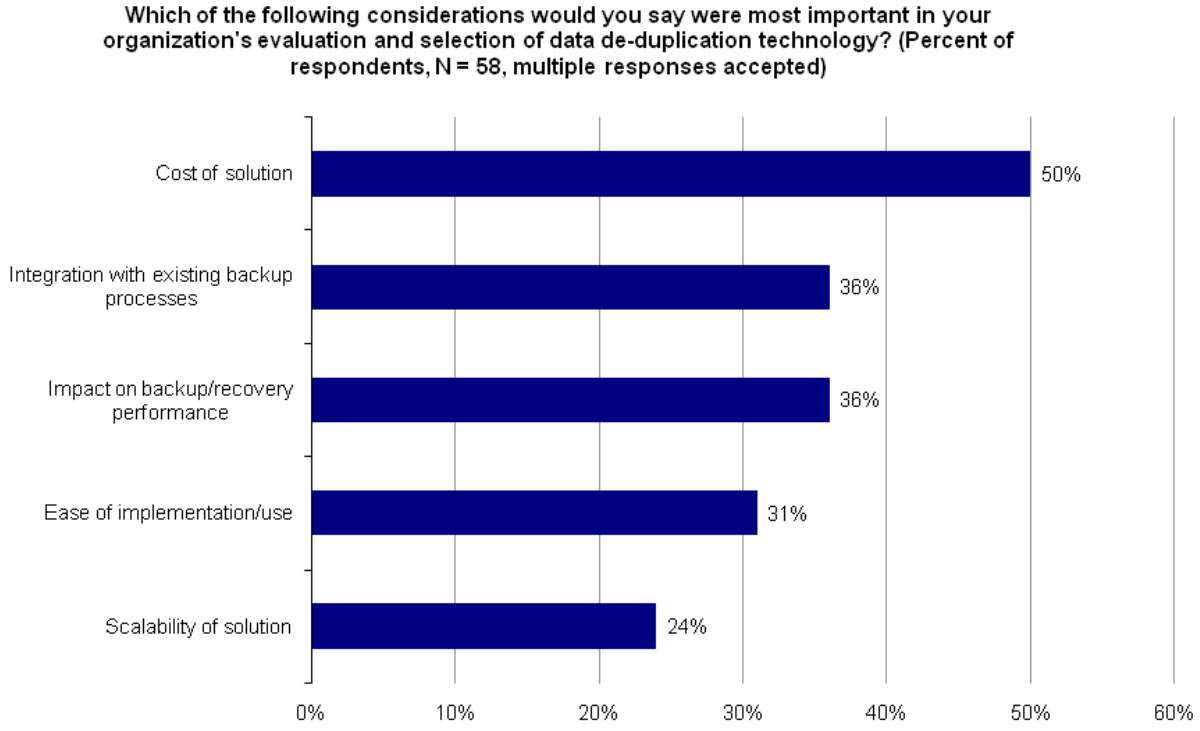
Introduction

In 2007, ESG Lab tested NetApp VTL appliances and found enterprise-class performance and scalability, combined with easy implementation and integration. This ESG Lab Validation report presents the results of a new round of hands-on testing of the NetApp VTL appliances with a focus on recently added data deduplication functionality. Enhanced enterprise-class performance and architectural scalability as well as the ease of deployment, painless integration with physical tape and simplified management are also explored.

Background

ESG has observed that data deduplication technology for disk-based backup systems and virtual tape libraries has been gaining in popularity and is being integrated into increasing numbers of backup solutions. ESG research indicates that no single concern dominates when users are considering data deduplication technology.[1] As shown in Figure 1, cost, compatibility, ease of integration and management, performance and scalability all have significant mindshare with enterprise users selecting deduplication technology. In order to succeed at the enterprise level, data deduplication appliances must provide cost-effective scalability, performance and flexibility.

FIGURE 1. DATA DEDUPLICATION SELECTION CRITERIA

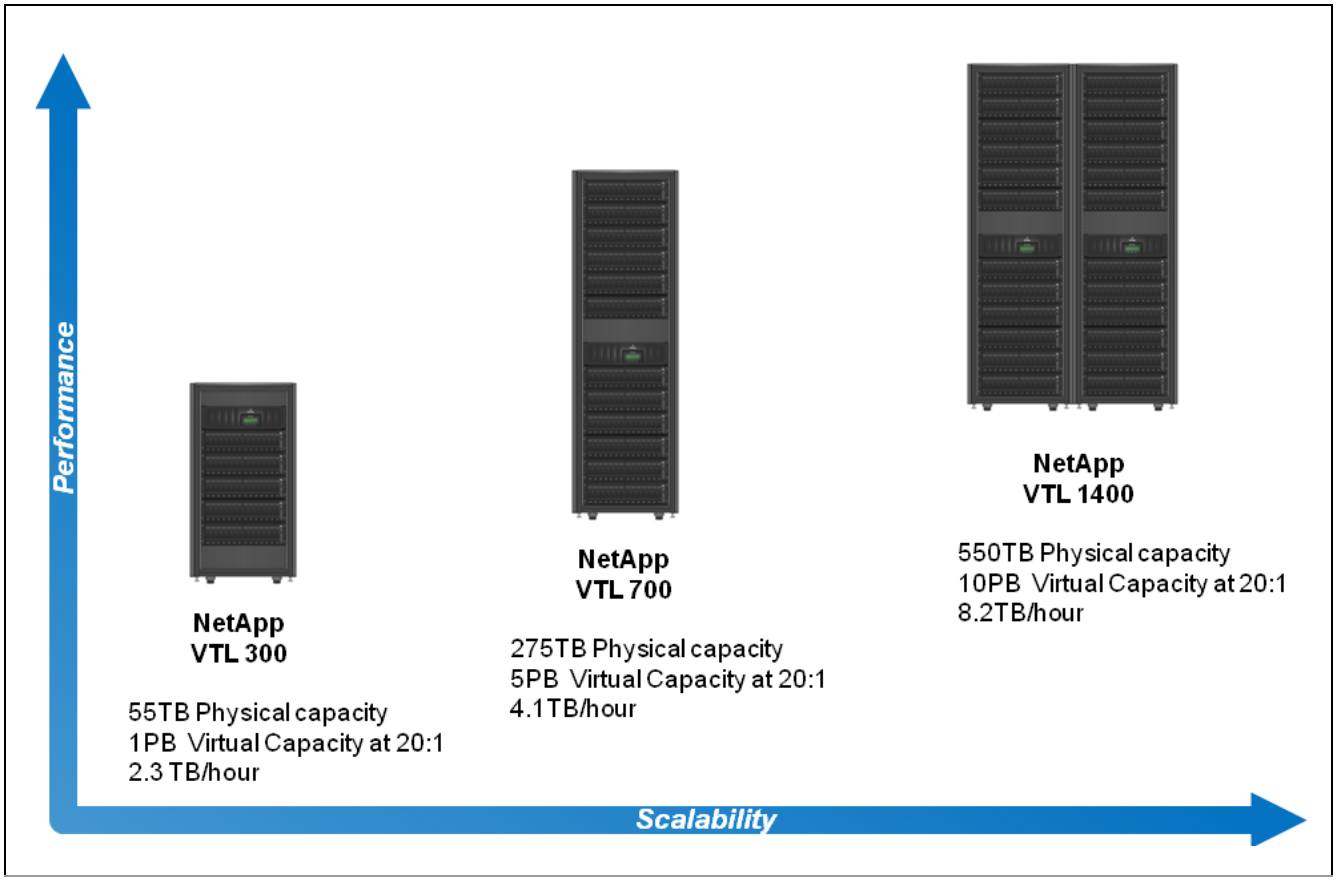


NetApp VTL is a disk based backup appliance that supports virtual tape library emulation and data deduplication. NetApp uses a modular, controller-based architecture to provide scalability and enterprise-class levels of performance and availability. NetApp VTL scales non-disruptively from an entry-level system with 10 TB of usable disk capacity prior to compression and deduplication up to a multiple controller system supporting 48 disk shelves with over half a petabyte of usable disk capacity with 1 TB disks. This could equate to 10PB of virtual capacity assuming a 20:1 deduplication ratio. The NetApp VTL solution is managed as an integrated appliance, but it can emulate hundreds of virtual tape libraries with thousands of virtual tape drives and tens of thousands of virtual tapes. Disks and RAID protection are all managed transparently by the VTL controller. This approach

[1] Source: ESG Research Report, *Data Protection Market Trends*, February, 2008.

provides the modularity needed to address a wide range of capacity and performance requirements while keeping management simple and straightforward. Because the VTL owns the disks, no back-end array configuration or management is necessary. Additionally, NetApp VTL is optimized for large block sequential I/O, which characterizes almost all backup and restore activity.

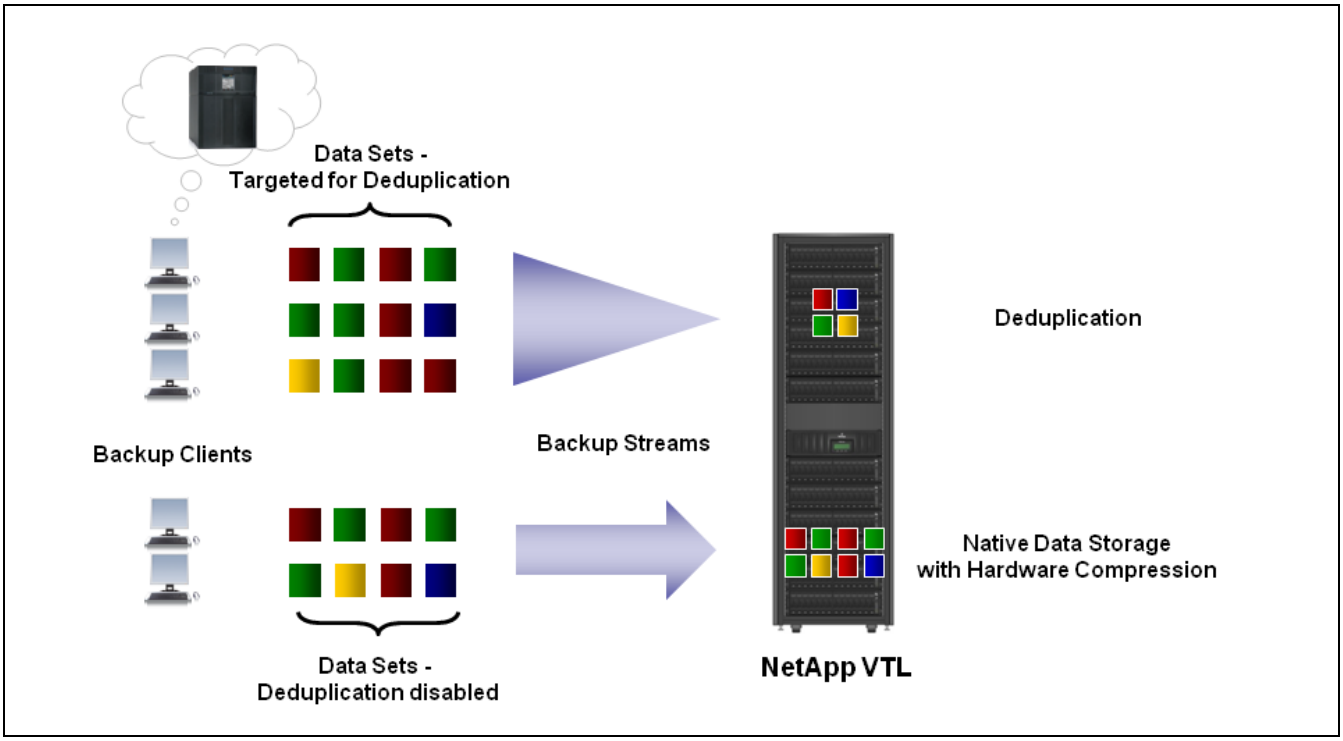
FIGURE 2. ENTERPRISE CLASS SCALABILITY



With the release of NetApp VTL 6.0 in 2008, data deduplication was introduced into the NetApp VTL family. The NetApp VTL allows administrators to enable deduplication on an individual virtual library basis, enabling control over which data sets to deduplicate. This is important as users may want to disable deduplication for data that either does not benefit from deduplication or that will not be retained for any significant period of time. Pre-compressed image files or encrypted data sets do not benefit greatly from deduplication—pre-processing data normally makes it difficult to find repeated patterns and severely reduces deduplication effectiveness.

Figure 3 illustrates how this would work in a real environment. Data sets that can benefit from deduplication such as databases, e-mail and file systems with extended retention periods are backed up to a VTL with deduplication enabled. The NetApp VTL then de-duplicates the data as a post-backup process. Deduplication can be scheduled or performed on-demand. Data sets which would benefit little from deduplication, such as pre-compressed image files, encrypted data sets and database recovery log backups can be sent to a VTL with deduplication disabled, but with hardware compression still enabled. This reduces system overhead and frees up CPU and memory resources that can be used for deduplication of the data sets that can benefit from it.

FIGURE 3. DATA DEDUPLICATION IN THE NETAPP VTL



ESG Research indicates that 87% of organizations using VTL continue to use tape as an integral part of their backup policies.[2] NetApp is certified with all major backup applications and has partnered with several vendors (Symantec, CommVault, BakBone and SyncSort) to more deeply integrate NetApp VTL disk-to-disk-to-tape (D2D2T) technology with their respective backup applications. NetApp was first to market with this level of backup application integration. To date, Symantec (Veritas NetBackup), CommVault (Galaxy) and BakBone have already delivered new releases that take advantage of this capability, and other vendors are in development. NetApp has also written a script-based D2D2T integration for use with EMC Legato NetWorker. It's important to note that even without this integration NetApp VTL can still be used with all supported backup applications.

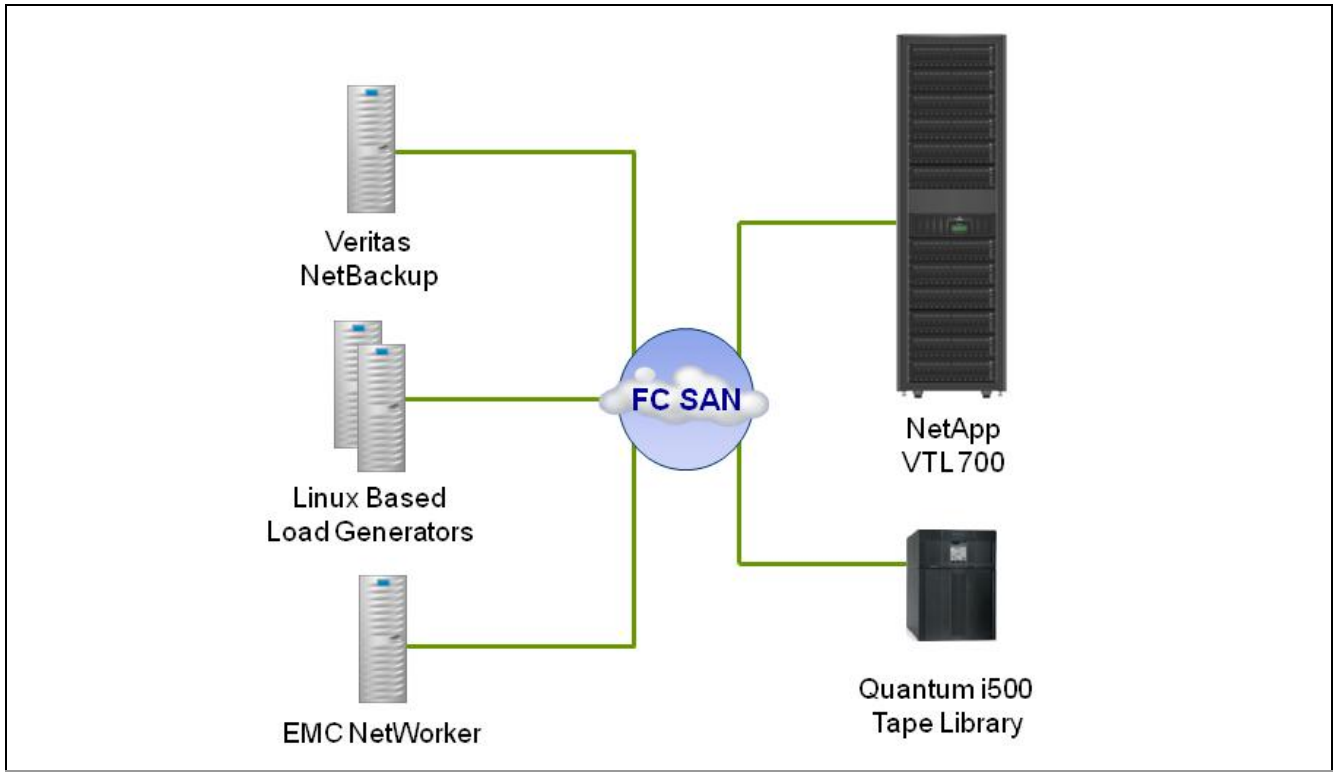
NetApp positions the NetApp VTL family as an enterprise-class VTL solution that is easy to integrate into an existing tape-based environment, simple to manage and scalable in both performance and capacity with native data deduplication. Installation of a new system, backups and restores with leading backup applications, data reduction with deduplication, integration with physical tape, performance, scalability and availability are all explored in this report.

[2] Source: ESG Research Report, *VTL Adoption and Market Trends* March 2007

ESG Lab Validation

ESG Lab performed hands-on evaluation and testing of the NetApp VTL at the NetApp Sunnyvale, California headquarters. Testing was designed to examine ease of implementation and management, test performance and scalability with and without deduplication and validate integration with existing physical tape environments. As seen in Figure 4, the test bed consisted of Veritas NetBackup and EMC Legato NetWorker Windows-based backup servers, along with two Linux-based load-generation servers running a script that utilized data created by the hpcreatedata utility from Hewlett Packard. A NetApp VTL 700 with two fully populated disk shelves was used for the initial installation tests and as a base platform for the online upgrade tests. A Quantum Scalar i500 tape library was utilized for the physical tape integration testing.

FIGURE 4. THE ESG LAB TEST BED



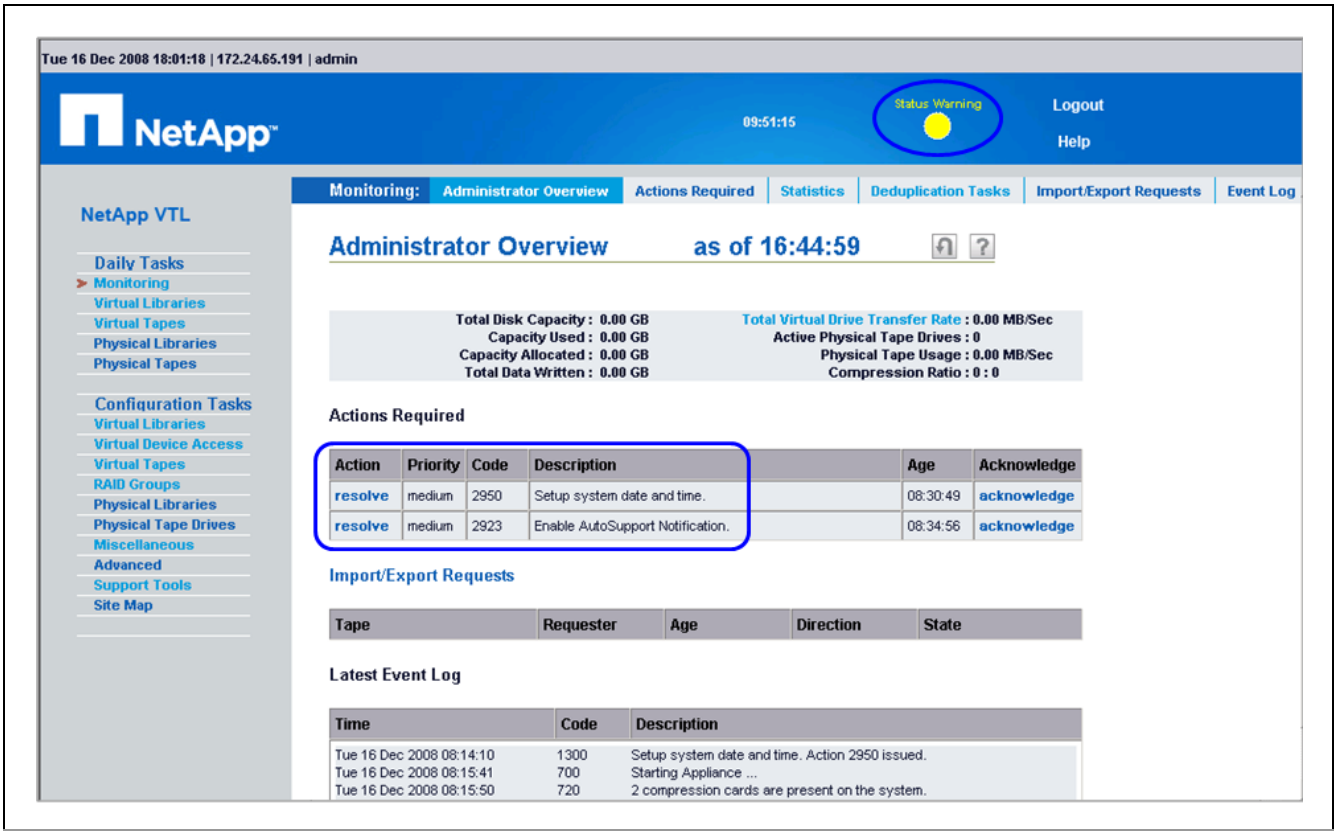
Ease of Implementation and Management

NetApp products have become famous for their streamlined, simple installation and management experience. This phase of ESG Lab testing was performed to evaluate how the NetApp VTL solution's ease of installation stacks up.

ESG Lab Testing

ESG Lab began with a NetApp VTL700 appliance that had been racked, prewired, powered up and connected to the administrative LAN prior to our arrival on site. The first step required was to configure the system with an IP address so the administrator could access the HTML-based management GUI. A full-featured CLI is also available to allow scripted activities. The IP address was set directly from the front panel in one step. Once a valid IP address was configured, the administration GUI was launched from a browser. The opening screen is shown in Figure 5.

FIGURE 5. NETAPP VTL ADMINISTRATOR OVERVIEW

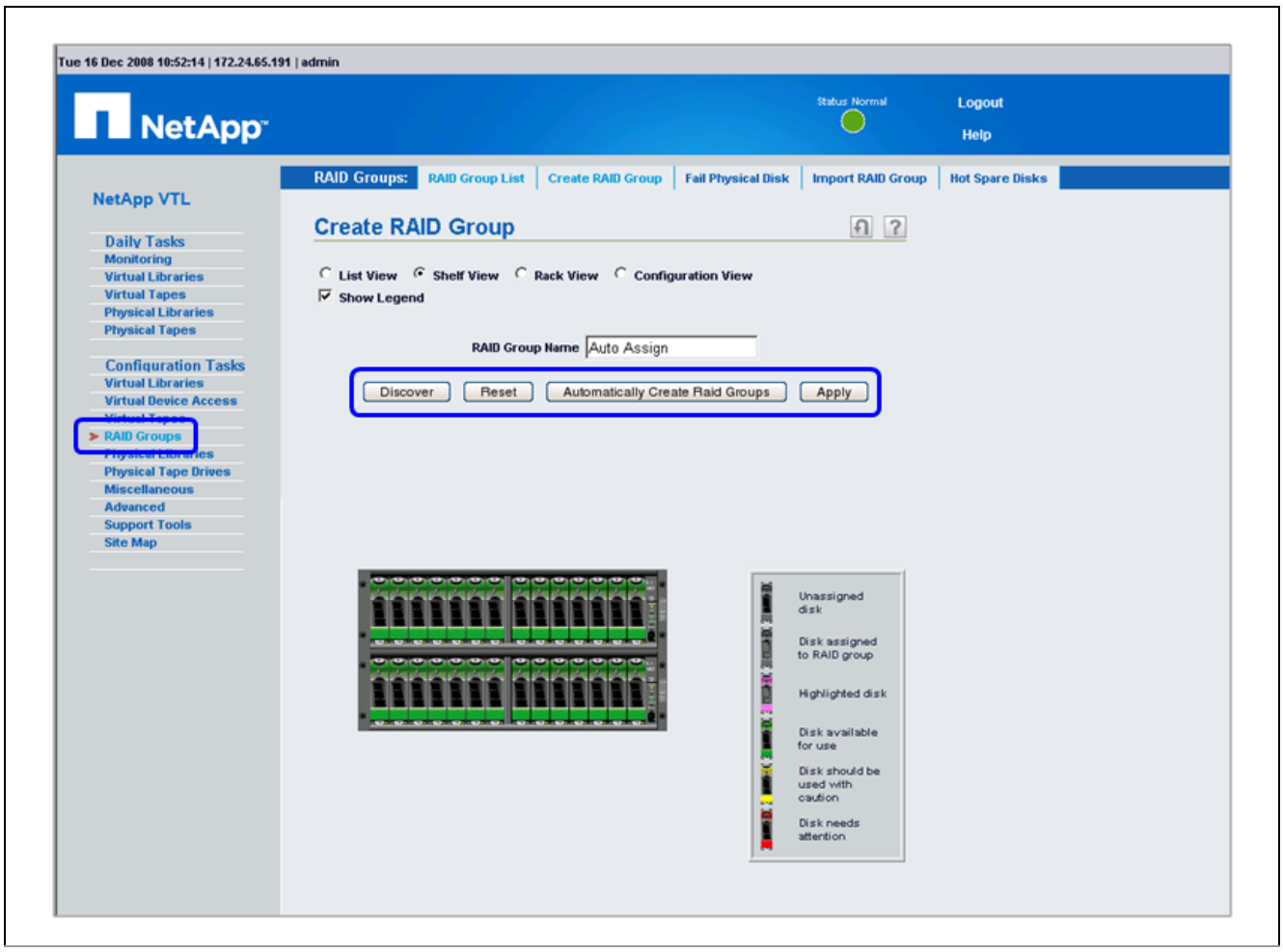


The Administration GUI is laid out logically, with main tasks listed vertically on the left side of the screen and subtasks arrayed horizontally across the top of the active pane. Upon logging in for the first time, the administrator is introduced to the Actions Required list, outlined in blue in the center of the screenshot. Note that the system provides not only a list of actions needed to resolve known issues; it also provides a resolve link next to each one. This takes the administrator directly to the appropriate screen to address the listed issue. At the top of the screen, the status indicator light in the upper right is glowing yellow, indicating a warning. The status indicator is tied to the first digit of the code attached to each action: 1 = Good (green status), 2 = Warning (yellow status) and 3 = Error (red status).

Next, ESG Lab configured AutoSupport. This feature enables the VTL to alert NetApp technical support of fault conditions. AutoSupport is the mechanism used by all NetApp systems to "phone home" and provides NetApp with a powerful tool that can proactively monitor and track systems remotely. Administrators can choose HTTP or e-mail notification. Using the e-mail option, the local administrator can be copied on all messages sent to NetApp.

Before creating any virtual libraries, virtual tape drives or virtual tapes, the administrator must create RAID groups on the raw disks by using the RAID Groups Configuration interface, as shown in Figure 6. The view presented shows a physical representation of all disk shelves that is brightly color coded to indicate the status of each disk.

FIGURE 6. THE CREATE RAID GROUP SCREEN



ESG Lab selected 'Automatically Create RAID Group' to let the NetApp VTL lay out the disks. NetApp VTL RAID management algorithms choose the best layout for performance and availability based on the number of disks and shelves installed. Because the NetApp VTL controller owns the disks directly and acts as the RAID controller, initializing RAID groups is nearly instantaneous. The RAID groups are immediately available for configuration and writing, without the need for any background zeroing or initialization.

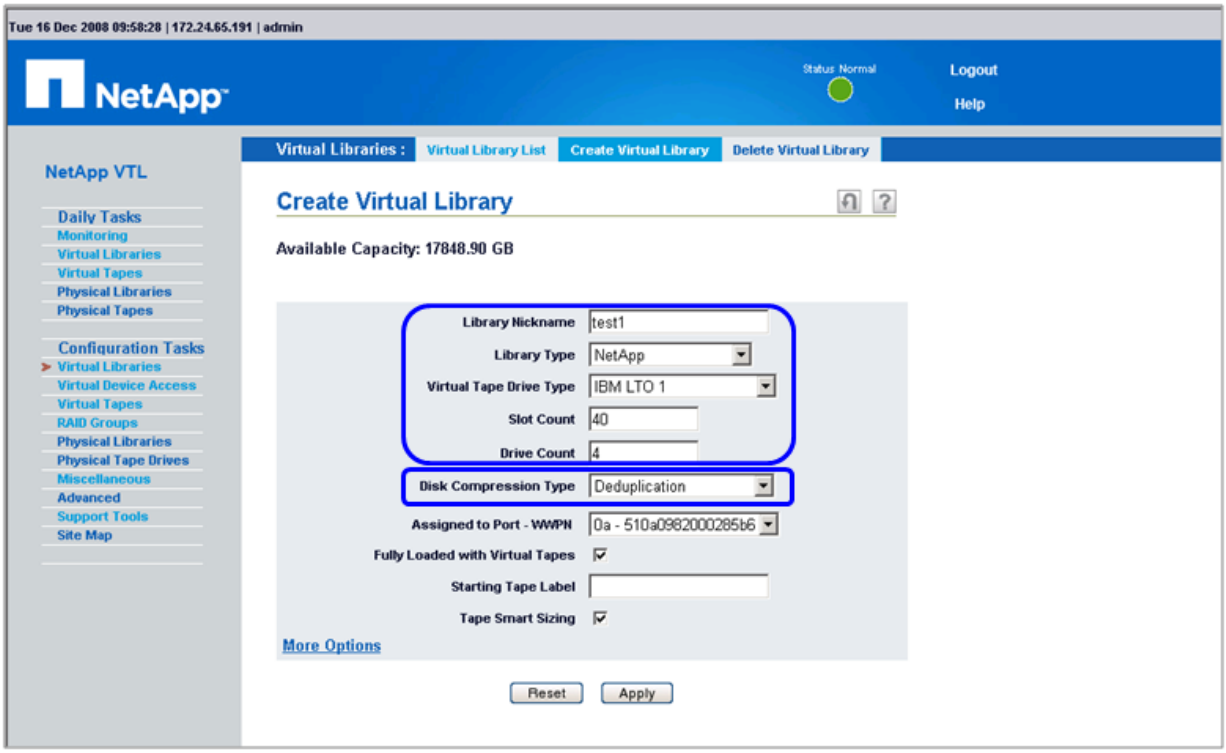
Another important distinction is seen when a previously used RAID group is destroyed and reinitialized. This also occurs nearly instantaneously, with the RAID group immediately available for writing. This capability is critical in systems with high-capacity hard drives. Because NetApp VTL RAID creation occurs so quickly, users can immediately back up their data. A non-NetApp VTL system with 1 TB SATA drives based on traditional RAID 5 controllers could easily take 20 hours to complete this basic operation.

With the RAID group creation completed, the next step was to create a virtual library with virtual tape drives and tapes. Those steps are all accomplished on one screen, illustrated in Figure 7. We named our test library test1 and selected a 'Library Type' of NetApp. This alerts backup applications that the library is a VTL from NetApp and in the case of Veritas NetBackup (and many other backup applications) allows NetBackup to set timings appropriately so that the backup application does not wait for tapes to mount before starting backups and restores.

ESG Lab elected to create four virtual tape drives and forty slots, fully loaded with virtual tapes. A Fibre Channel port for SAN presentation was selected from a pull-down list. An arbitrary tape labeling scheme was chosen for

our testing. An administrator would normally choose a scheme consistent with existing tape labeling standards or allow the NetApp VTL to automatically create virtual tapes that match physical tapes in a library connected to the VTL. New in the NetApp VTL is the 'Disk Compression Type' option, used to enable deduplication for a VTL when it is created.

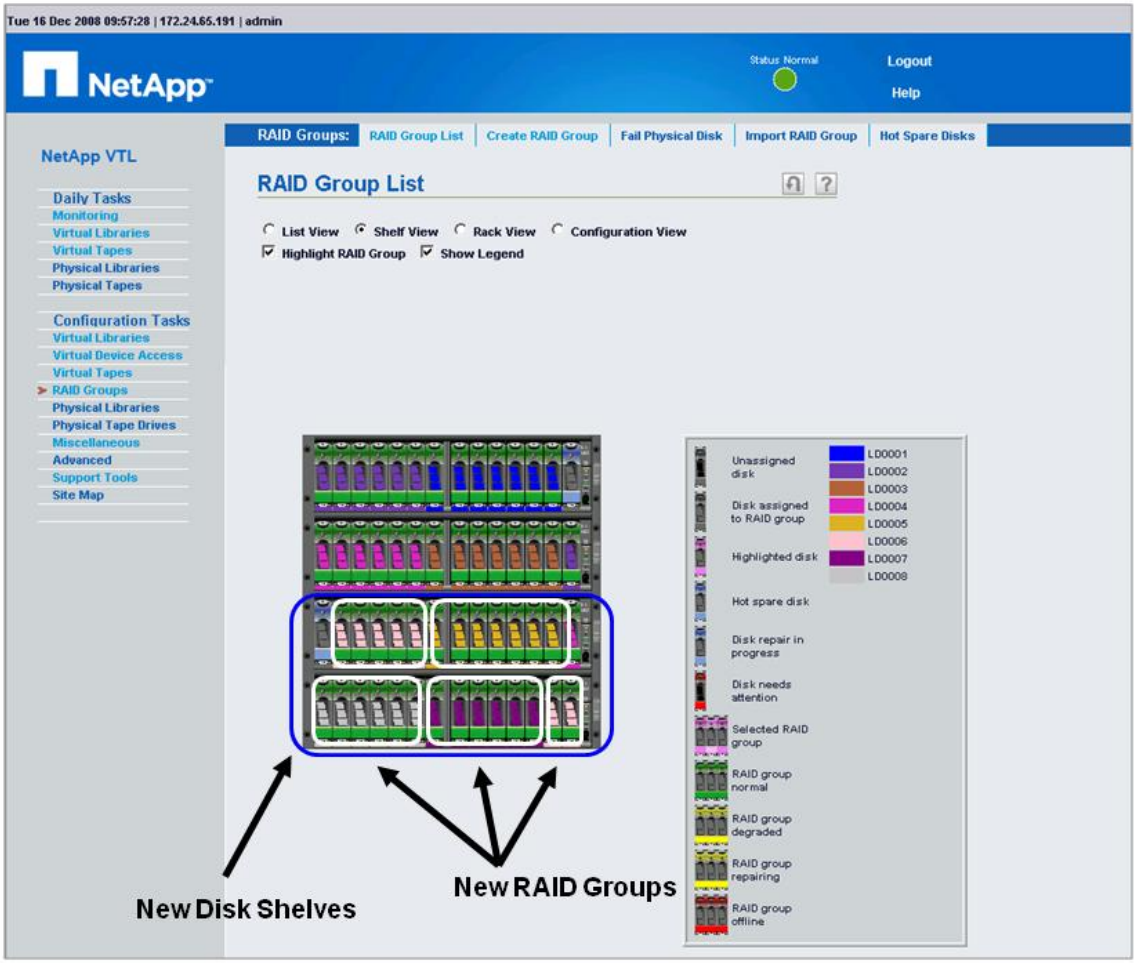
FIGURE 7. CREATING A VIRTUAL LIBRARY



The final step was to discover the new VTL in NetBackup, which was performed in exactly the same manner as discovering a physical tape library and was completed in less than a minute. In less than 10 minutes, the entire system was configured from start to finish and the first backup was kicked off.

Finally, ESG Lab connected two additional shelves to the existing two shelves to validate the ease of capacity expansion. Upon powering up, the shelves were recognized as new, un-configured storage by the NetApp VTL. The only step necessary was to go back to the 'Create RAID Group' screen and select 'Automatically Create RAID Groups,' as was done in the initial installation. The new RAID groups were immediately available for use, as shown in Figure 8.

FIGURE 8. CAPACITY UPGRADE



Why This Matters

ESG Research found that 60% of VTL adopters indicated ease of deployment as the single most important factor in purchasing a VTL solution.[3] This is especially important for enterprise-class VTL appliances deployed in large, complex environments where backup policies span hundreds of servers and dozens of applications—stretching resources to the limit. ESG Lab has confirmed that a NetApp VTL is extremely easy to configure and just as easy to upgrade. The system can be dropped into an existing tape environment and performing backups in less than 10 minutes, while upgrades can occur just as easily on the fly, with new disks dynamically allocated as needed.

[3] Source: ESG Research Report, *VTL Adoption and Market Trends* March 2007

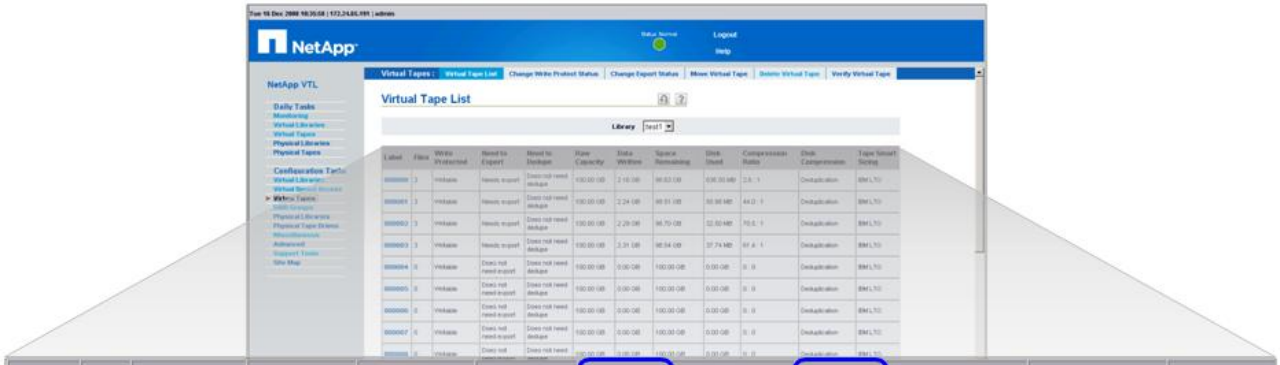
Storage Efficiency

NetApp VTL deduplication reduces the capacity required to store backed up data. Data that has been backed up to a VTL with deduplication enabled is scanned and the NetApp VTL examines the contents of the backup images, recognizing and replacing redundant blocks with pointers to the original blocks.

ESG Lab Testing

To evaluate deduplication in the NetApp VTL, ESG Lab ran multiple full backups of an Oracle database and validated the results. ESG Lab also audited additional in-house testing by NetApp for file system and Exchange data sets. The database was harvested from a corporate data center—full backups performed over ten days were evaluated. The Virtual Tape List screen of the NetApp VTL console, shown in Figure 9, was used to monitor the benefits of data reduction after each NetBackup job had run.

FIGURE 9. THE VIRTUAL TAPE LIST



Label	Files	Write Protected	Need to Export	Need to Dedupe	Raw Capacity	Data Written	Space Remaining	Disk Used	Compression Ratio	Disk Compression	Tape Smart Sizing
000000	3	Writable	Needs export	Does not need dedupe	100.00 GB	2.16 GB	98.63 GB	836.50 MB	2.6 : 1	Deduplication	IBM LTO
000001	3	Writable	Needs export	Does not need dedupe	100.00 GB	2.24 GB	98.51 GB	50.98 MB	44.0 : 1	Deduplication	IBM LTO
000002	3	Writable	Needs export	Does not need dedupe	100.00 GB	2.29 GB	98.70 GB	32.50 MB	70.5 : 1	Deduplication	IBM LTO
000003	3	Writable	Needs export	Does not need dedupe	100.00 GB	2.31 GB	98.54 GB	37.74 MB	61.4 : 1	Deduplication	IBM LTO

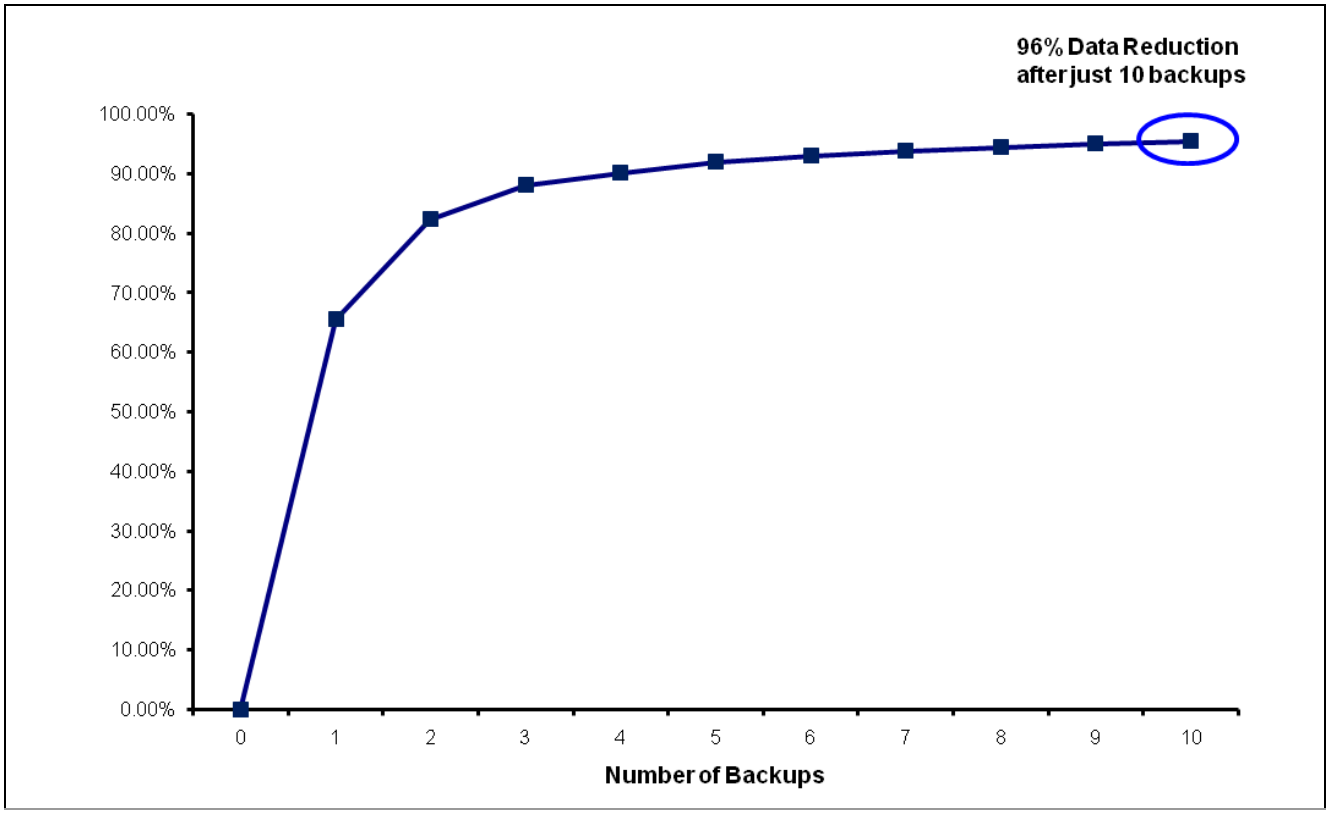
The 'Data Written' and 'Disk Used' fields were collected and analyzed after the tenth full backup was completed. The data collected is detailed in Table 1.

TABLE 1. DEDUPLICATION OVER TIME – DETAILED RESULTS

	Per Individual Event			Cumulative			
Backup	Data Written (GB)	Disk Consumed (GB)	Data Reduction	Data Written	Disk Consumed (GB)	Data Reduction	Cumulative Dedupe Ratio
1	2548	878	65%	2548	878	66%	2.9
2	2562	21	99%	5109	899	82%	5.7
3	2561	4	99%	7670	903	88%	8.5
4	2919	131	96%	10589	1034	90%	10.2
5	2954	45	98%	13543	1080	92%	12.5
6	3000	53	98%	16543	1132	93%	14.6
7	3034	48	98%	19576	1181	94%	16.6
8	3062	48	98%	22638	1228	95%	18.4
9	3091	29	99%	25729	1258	95%	20.5
10	3137	36	99%	28866	1293	96%	22.3

ESG Lab calculated the data reduction over the course of ten backups and plotted the curve of storage savings, as seen in Figure 10. The NetApp VTL achieved 90% data reduction for a production Oracle database after just four full backups and 96% data reduction after ten backups. Said another way, after just ten full backups, a deduplication ratio of 22:1 was achieved.

FIGURE 10. DEDUPLICATION OVER TIME



What the Numbers Mean

- The total amount of data that the NetApp VTL would have stored on disk over the course of ten days without using deduplication was calculated to be 28.87 TB.
- NetApp VTL deduplication reduced the required disk capacity for ten full Oracle DB backups from 28.87 TB to only 1.29 TB. This would enable organizations to keep backups on disk for MUCH longer, reducing the need to go back to tape for recoveries.
- The amount of capacity reduction that can be achieved with the NetApp VTL will vary according to the backup policy in effect, the number of backups retained on disk and the type of data being stored. In this scenario, capacity was reduced by 96% over just ten days.
- ESG Lab saw similar data reduction/deduplication ratios achieved for file system and Microsoft Exchange data sets.

Why This Matters

ESG research[4] indicates that cost and impact to backup and restore performance are two of the leading obstacles to disk-based data deduplicated backup deployments. The ability to deduplicate selective data sets after the backup has completed addresses both of these issues by reducing the amount of data retained on disk while providing optimal performance.

ESG Lab has validated that data deduplication in the NetApp VTL can be used to reduce disk capacity significantly while applying deduplication policy appropriate to data type and retention needs. NetApp VTL administrators can effectively provide high performance backup services, plus fast and reliable restores using greatly reduced disk capacity. This lowers the cost per GB for backup data and enables companies to retain data exponentially longer for recovery purposes while minimizing the impact of deduplication on backup windows and recovery SLAs.

[4] Source: ESG Research Report, *Data Protection Market Trends*, February, 2008.

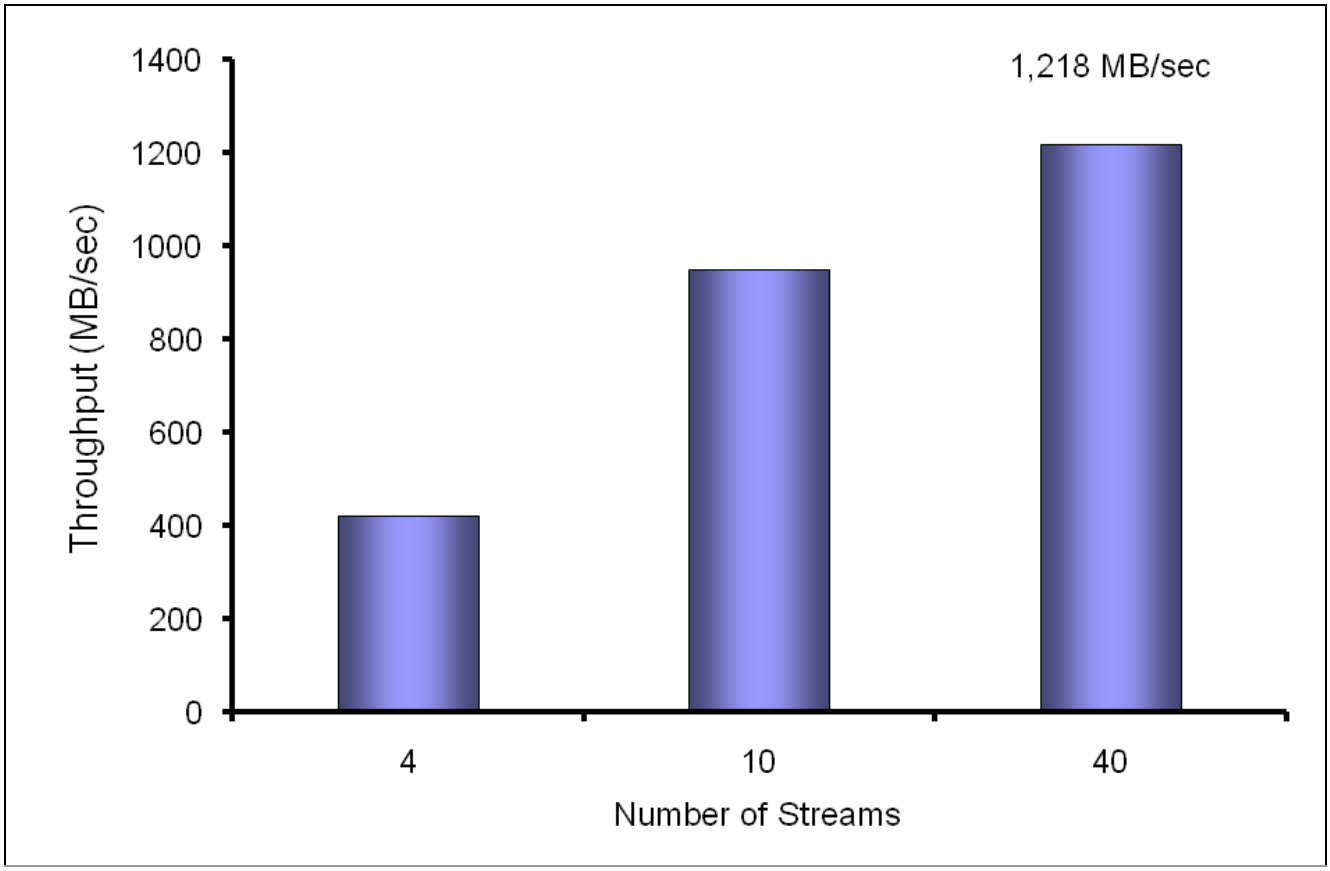
Scalability and Performance

NetApp uses a variety of methods to maximize performance and scalability. The latest server and I/O interconnect technology (4Gbps FC) provides the basic building blocks. A VTL controller with fully integrated RAID and I/O functionality is optimized for large-block, sequential I/O. NetApp VTL optimally distributes deduplicated data across all available disks to maximize VTL performance, Selective data deduplication enables administrators to only apply deduplication, a resource-intensive operation, to data sets that will benefit from it.

ESG Lab Testing

ESG Lab simulated multiple backup streams to measure the backup performance capabilities of the NetApp VTL700 with standard hardware compression and data deduplication enabled. Two Linux servers with two Fibre Channel connections were connected to a NetApp VTL700, as shown previously in Figure 4. A PERL script was used in conjunction with the hpcreatedata utility during this phase of testing to avoid file system overhead in the system being backed up, maximize the stress on the VTL700 and minimize the amount of equipment required to obtain valid results. The hpcreatedata utility creates files on disk with variable levels of compressibility. The script was reviewed by ESG Lab to ensure that it issued the same low-level SCSI commands used by common backup applications.[5] The performance results were independently verified against the statistics reported by the Fibre Channel switch providing connectivity between the clients and the VTL700.

FIGURE 11. WRITE THROUGHPUT: 2:1 COMPRESSIBLE DATA WITH FOUR DISK SHELVES



[5] The script was configured to use a 2MB buffer sending data that was 2:1 compressible.

Figure 11 shows results obtained for simulated backups using four, ten and forty simultaneous streams of 2:1 compressible data with hardware compression enabled and deduplication disabled. It is important to note that performance scales linearly as streams are added and that the results achieved with forty simultaneous streams exceed the performance numbers that NetApp posts for the VTL700 on its public website. Also of interest is the fact that these numbers were produced by a factory default configuration with only four shelves of disks that had not been tuned for performance.

Figure 12 compares the results of testing with data deduplication in three different configurations. Forty backup streams to forty virtual drives in different Virtual Libraries in a VTL700 were used in each test. Deduplication was enabled for all virtual libraries in the VTL in the first pass, half of the virtual libraries in the second pass (twenty deduplication enabled streams), and zero in the third pass. The NetApp VTL performed very well here also; The Mixed Mode configuration (twenty streams to deduplication enabled VTL's, twenty streams with hardware compression only) provided more than 80% of the performance of a system with deduplication completely disabled. Again, it is important to note that these numbers were produced by an un-tuned, factory default configuration with a small number of disk shelves.

FIGURE 12. BACKUP PERFORMANCE ACROSS DEDUPLICATION POLICIES

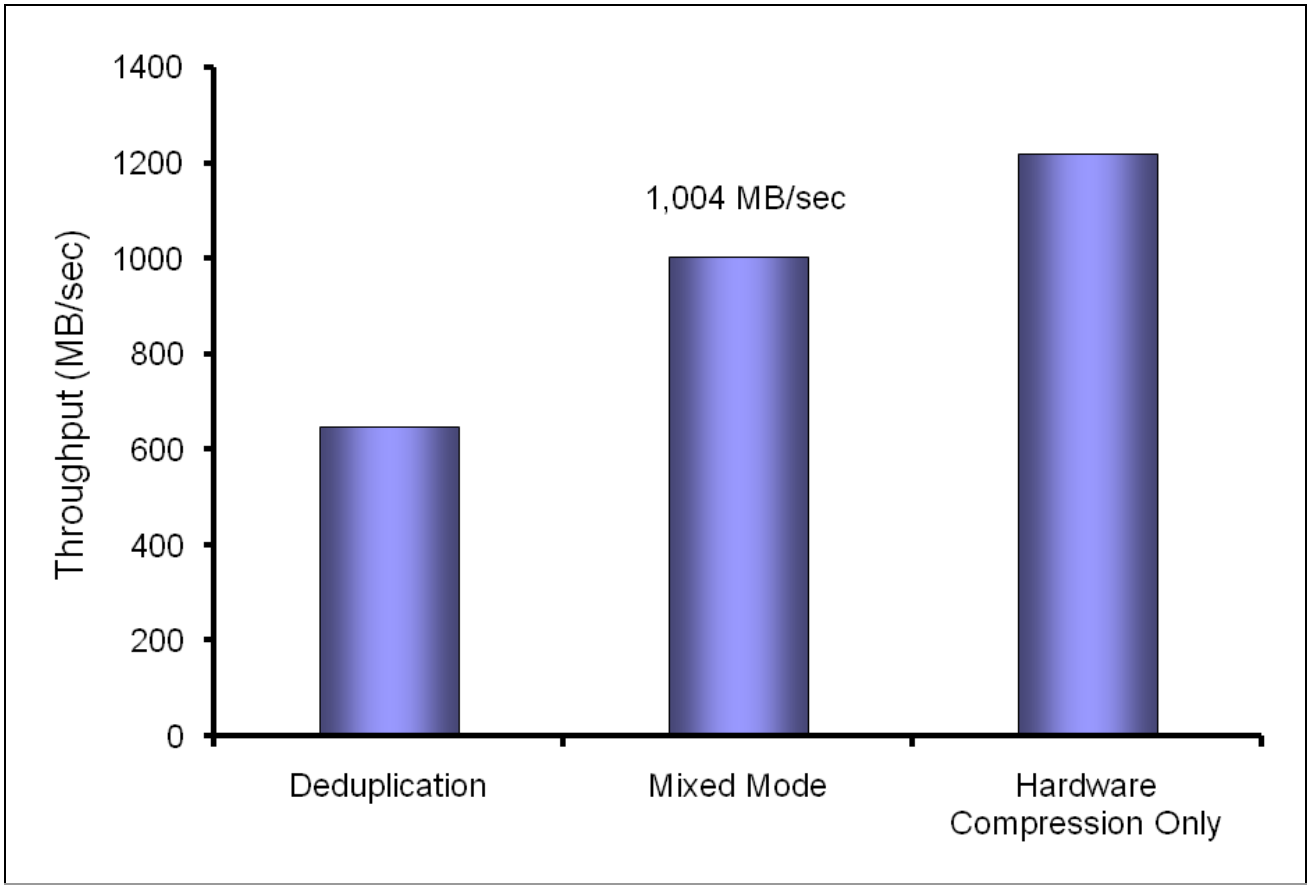


TABLE 2. BACKUP AND RESTORE PERFORMANCE RESULTS

Number of Streams	Deduplicated Write Throughput MB/sec	Mixed Mode Write Throughput MB/sec	2:1 Compressed Write Throughput MB/sec
4	382	400	421
10	568	720	949
40	647	1,004	1,218

What the Numbers Mean

- The NetApp VTL provided maximum performance straight out of the box, with no tuning and a relatively small number of disks.
- Deduplication in a mixed mode environment, (which will be used in virtually all real-world backup environments) ran nearly as fast as backups with deduplication disabled. This increases capacity utilization and improves ROI by requiring less physical disk with a low performance impact.
- Backing up data in a mixed mode environment provided sustained throughput of 1,004 MB/sec. This enables users to back up more than 27 TB in a single 8-hour backup window using a single controller.
- ESG Lab ran a restore test and observed aggregate restore throughput sustained at 562 MB/sec.

Why This Matters

For years, backup administrators have been struggling to complete nightly backups before business resumes in the morning. In the world of tape, this means purchasing more tape drives for more parallelism. Quicker recoveries are needed to increase user productivity and meet service-level agreements and no amount of tape hardware can solve this problem.

As backup windows continue to shrink, IT managers are increasingly adopting backup to disk and virtual tape technologies to get nightly backups done quicker. ESG Research confirmed that VTL early adopters rank performance as one of the top three reasons they chose VTL.[6] ESG Lab has confirmed that a NetApp VTL solution can sustain more than 1000 MB/sec of sustained backup throughput with an out-of-the-box configuration and a relatively small number of disk drives. Based on ESG Lab's experience in testing a number of VTL solutions in recent years, this is an outstanding level of performance and should be attainable for backup environments that can drive this level of I/O.

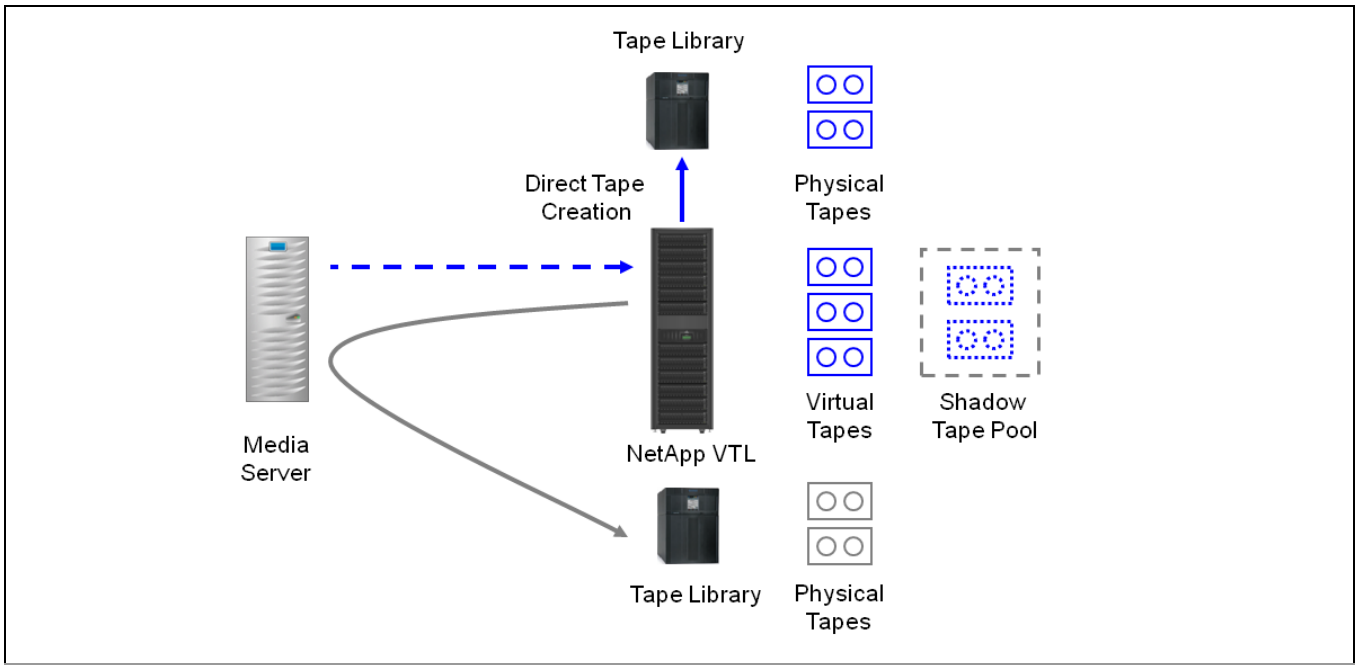
[6] Source: ESG Research Report, *Data Protection Market Trends*, February, 2008.

Integration with Physical Tape

In most backup environments, physical tape is still used for long-term and off-site archiving. VTL users need to export virtual tapes to physical tapes for archiving and import them into a Virtual Tape Library for restores of archived data. To make this work, the virtual library must be configured to emulate drives that are of the same type and capacity as the physical drives in the physical library. The NetApp VTL appliance emulates physical tape libraries and stores large amounts of backup data on disk, automatically copying data directly to and from physical tape.

NetApp VTL uses two features to enhance integration with backup applications: Shadow Tape copies and Direct Tape Creation. Shadow Tape technology automatically caches backups onto disk, enabling fast restores from disk for backups that have already been exported to physical tape. Shadow Tape backups enable storage administrators to meet two often conflicting requirements in a way that is fully compatible with their existing backup software: first, to meet service-level agreements for creating physical copies that must be moved offsite for safe storage within a specified time frame and second, to meet user demands by quickly restoring backup data without having to recall a tape from off-site storage. Direct Tape Creation is used to make copies of virtual tapes without involving the media server in data movement. The net result is the ability to have both a disk-resident and an off-site backup copy both under control of the backup application.

FIGURE 13. SHADOW TAPE BACKUPS AND DIRECT TAPE CREATION



The traditional method for exporting a virtual tape to a physical tape has been to use the backup application and make the copy by using the media server, illustrated by the gray arrow in Figure 13. This consumes significant server CPU, memory and I/O resources that could otherwise be utilized in performing backups or restores. Direct Tape Creation, represented by the blue arrows, utilizes the VTL to perform the actual tape copy. While some VTLs offer this type of functionality, NetApp's integration with backup applications sets it apart. NetApp VTL enables backup software applications to keep track of physical tape cartridges created directly by the Virtual Tape Library

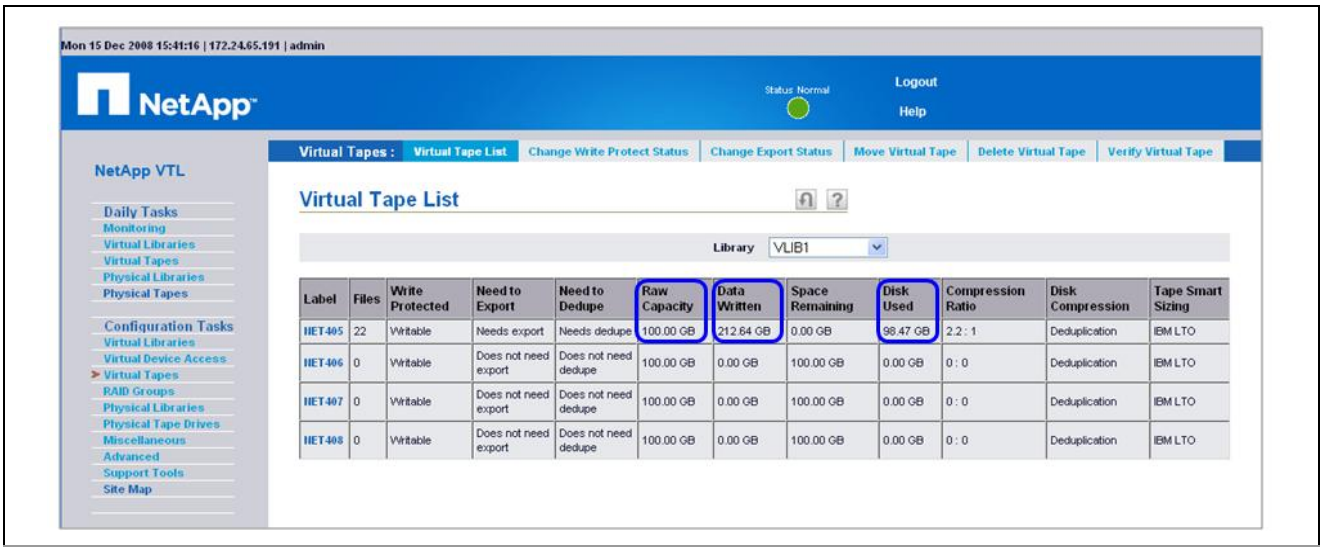
The NetApp VTL retains the native format of the backup application when it writes data to disk. This means that when the NetApp VTL copies data to physical tape, the backup application can restore data from these tapes directly. ESG Lab's goal for this portion of testing was to validate NetApp integration with backup applications and the ability of those applications to keep track of physical and virtual tapes.

ESG Lab Testing

ESG Lab first looked at tape creation procedures. Both virtual tapes and physical tapes created by NetApp VTL are written in the native format of the backup application. This allows the NetApp VTL to import and export tapes without conversion. It also means that the tapes can be restored directly by the backup application without the VTL. In order for this to work, it is crucial that the amount of data on a virtual tape fit on the physical tape without exceeding its capacity, maintaining a one-to-one correlation between a virtual tape and a physical tape. If the data stored on a virtual tape exceeds the capacity of a physical tape, it cannot be exported or cloned.

The NetApp VTL appliance addresses this issue by using a feature called Tape Smart Sizing to control the amount of data written to each virtual tape. Tape Smart Sizing is based on the premise that all data is compressible at differing ratios and that different tape drive types will compress the same data differently. Smart Sizing is designed to allow the VTL to determine the correct size of a virtual tape to ensure that it will fit on a physical tape correctly. Figure 14 shows an example of Smart Sizing in operation. Virtual tape NET-105 was created with a raw capacity of 100 GB and a synthetic data set composed of data with a known compressibility of approximately 2:1 was backed up.

FIGURE 14. SMART SIZING OF VIRTUAL TAPES

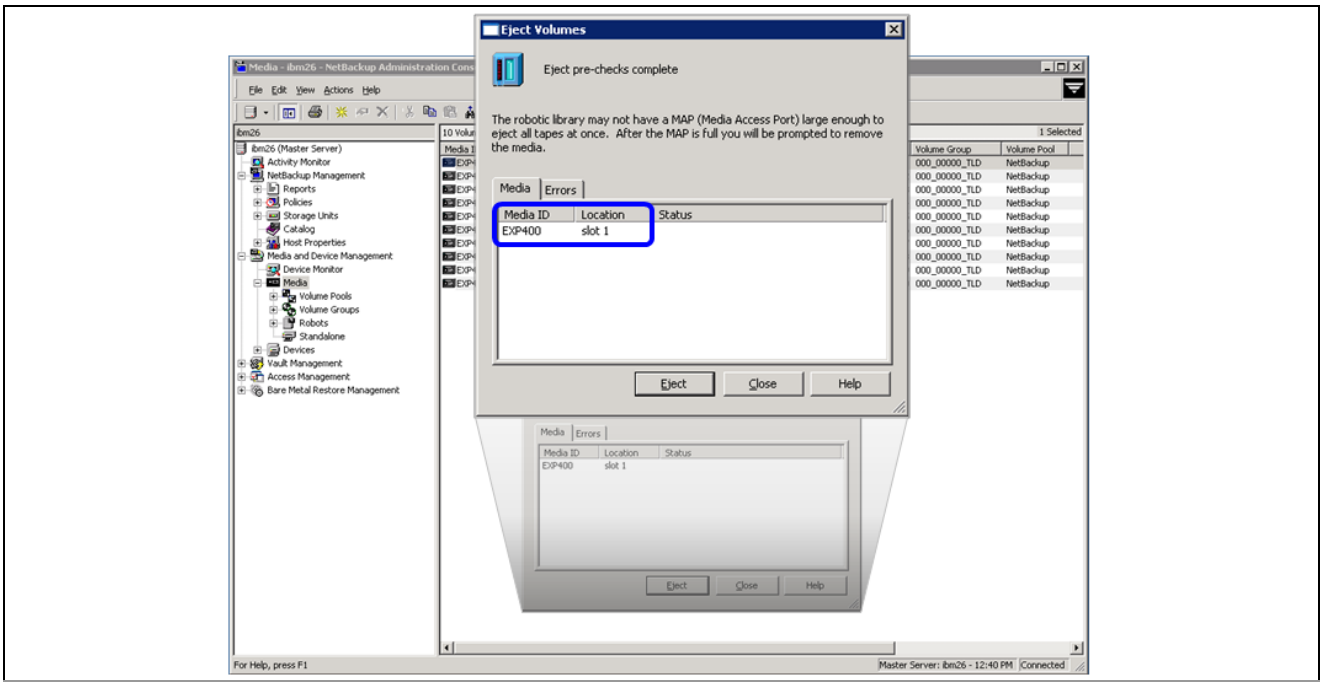


With Tape Smart Sizing, the data stream is monitored to determine the compressibility of the data based on the compression algorithms used by the tape hardware. This allows the NetApp VTL appliance to determine the optimal size for the virtual tape so that it fits correctly on its corresponding physical tape. In the preceding example, 212.84 GB was backed up and stored in one 100GB virtual tape. ESG Lab observed that the virtual tape was exported successfully to a 100 GB physical tape (IBM LTO-1) with no issues.

ESG Lab next tested NetApp VTL Shadow Tape functionality. When enabled, whenever a virtual tape is exported to a physical tape directly from the VTL, the virtual tape is moved to the VTL Shadow Tape Pool. A virtual tape in the Shadow Tape Pool is invisible to the backup application and is not listed as part of a virtual library but is still available for quick access if needed for restore.

Testing began with a backup to a number of virtual tape drives using Veritas NetBackup on a Windows server. Once the virtual tapes were filled, an Eject command was executed from NetBackup for tape volumes EXP400 and EXP401. Figure 15 shows tape volume EXP400 ready to be exported.

FIGURE 15. VIRTUAL TAPE EXPORT



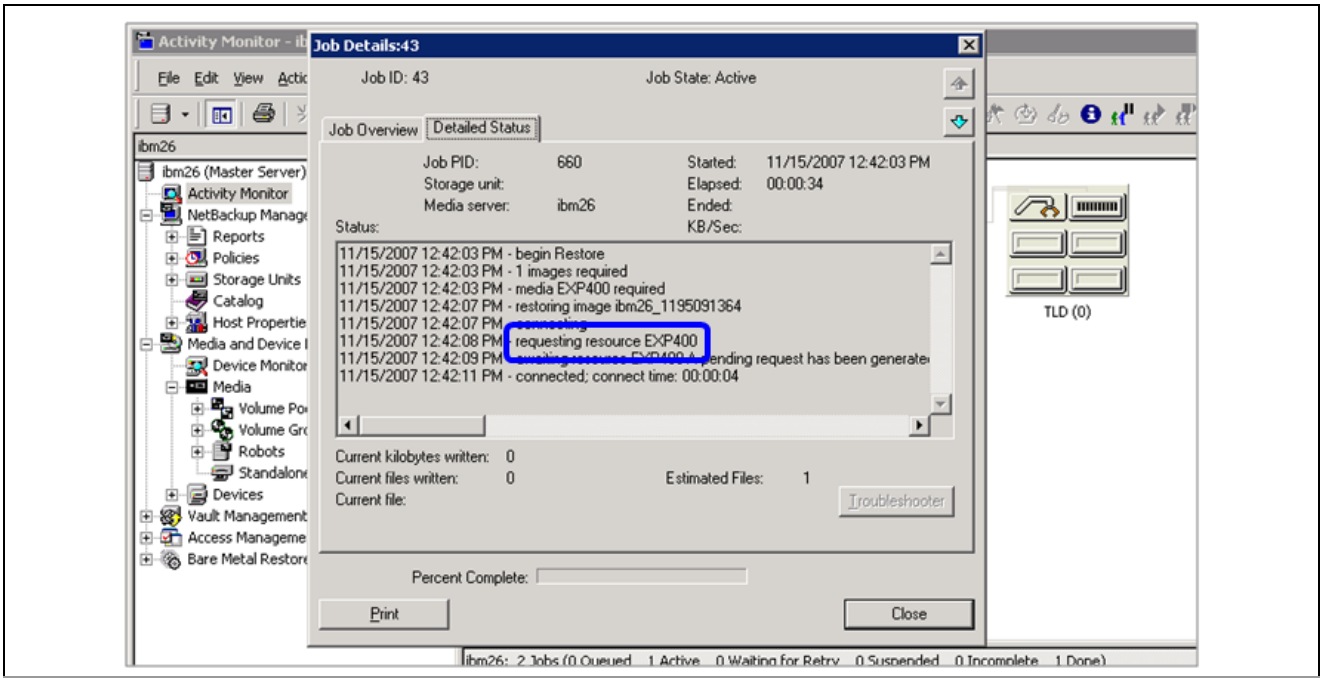
Once the Eject command was completed, ESG Lab used the NetApp VTL GUI interface to examine the tape now in the Shadow Tape Pool. It was noted that the operation to move the tape into the Shadow Tape Pool was, in fact, automatic and occurred nearly instantaneously. Figure 16 shows the Shadow Tape Pool with tape volume EXP400, which had been exported using NetBackup. It is also important to note that this 100GB Shadow Tape was deduplicated and consumed only 1.04 MB on disk.

FIGURE 16. THE SHADOW TAPE POOL

At this point, ESG Lab confirmed that once a virtual tape has been exported to a physical tape, the copy residing in the Shadow Tape Pool can be immediately and automatically available for a Restore operation, even though it

is not visible to the backup application. Testing began by performing a Restore operation from NetBackup. Once the Restore command was issued, NetBackup indicated, as seen in Figure 17, that it had a Pending Request for the tape to be mounted.

FIGURE 17. NETBACKUP TAPE MOUNT PENDING REQUEST



Within seconds of the command execution, the virtual tape residing in the Shadow Tape Pool was automatically moved back into the Virtual Tape Library being used for this test and NetBackup began transferring the data to the backup host. Once the restore was completed, the tape was moved back into the Shadow Tape Pool per the user-defined shadow pool policy on the NetApp VTL. The entire process was completed without manual intervention.

Why This Matters

ESG Research shows that 87% of organizations using VTL still incorporate physical tape in their backup policies.[7] Tape copies are made for DR, archival and compliance requirements. NetApp VTL integration with physical tape enables IT managers to meet their physical tape creation requirements with existing legacy tape systems and to significantly downsize their tape infrastructure as they move an increasing amount of operational backup and restore operations onto disk-based systems such as VTLs. Hard cost savings are realized from reduced expenditures on tape drives, tape libraries and media servers, requiring fewer licenses, and fewer physical media (tape) purchases. Soft cost savings flow from increased reliability, decreased physical tape management and reduced off-site handling requirements, all of which free up dedicated backup administrators for other productive tasks.

NetApp VTL Direct Tape Creation can create physical tapes significantly faster than using the backup server while reducing media server loads and SAN traffic. ESG Lab validated that NetApp VTL Direct Tape Creation, Tape Smart Sizing and Shadow Tape functions are fully compatible and integrated with the backup applications' built-in media management functions, while leveraging the enhanced performance and productivity that a disk-based solution offers

[7] Source: ESG Research Report, *Data Protection Market Trends*, February, 2008.

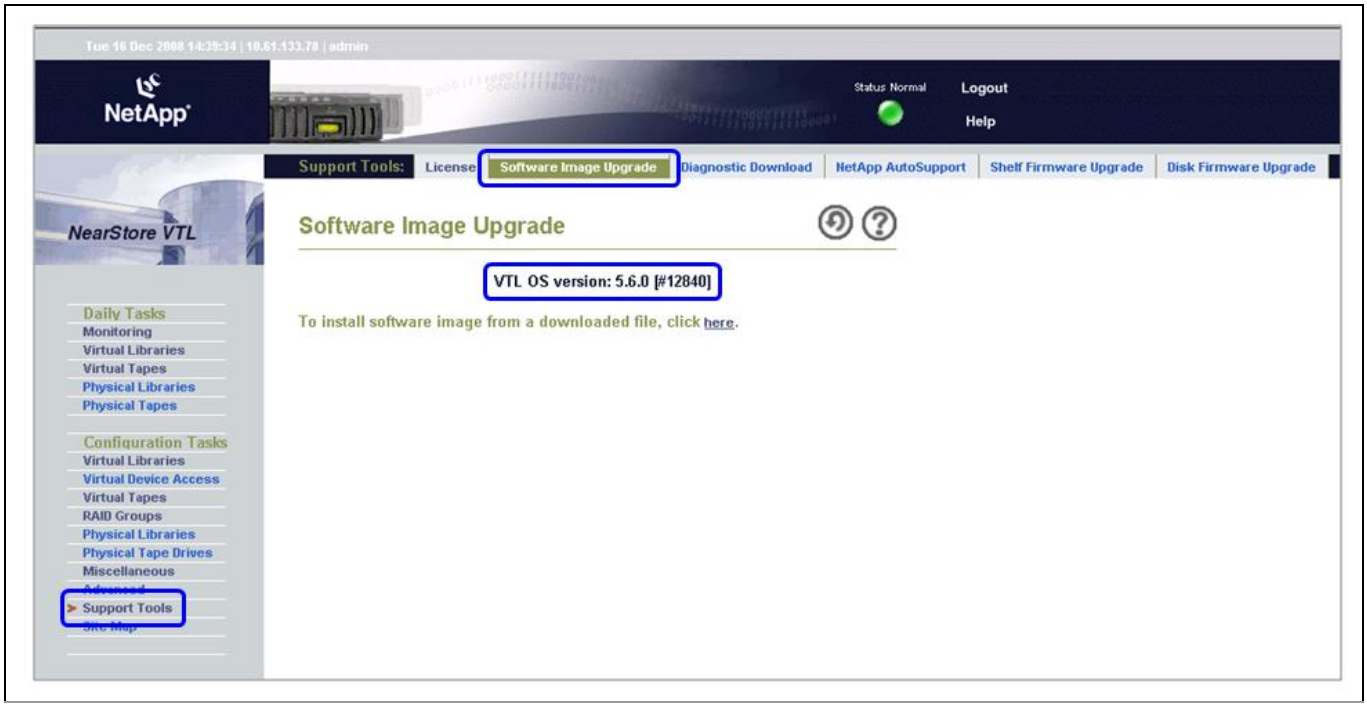
Reliability and High Availability

High availability (HA) and fault tolerance are major concerns for today's backup administrators, who have come to expect the same enterprise-class reliability in a VTL solution that they have with their primary storage systems. The NetApp VTL is architected as an integrated appliance, meaning that disks and RAID protection are managed directly by the NetApp VTL software. This allows tighter control and better response to hardware failures. Additionally, the VTL configuration is stored internally within the appliance on disk, facilitating simple controller replacement and upgrades. VTL software can be updated online, with only a short reboot required.

ESG Lab Testing

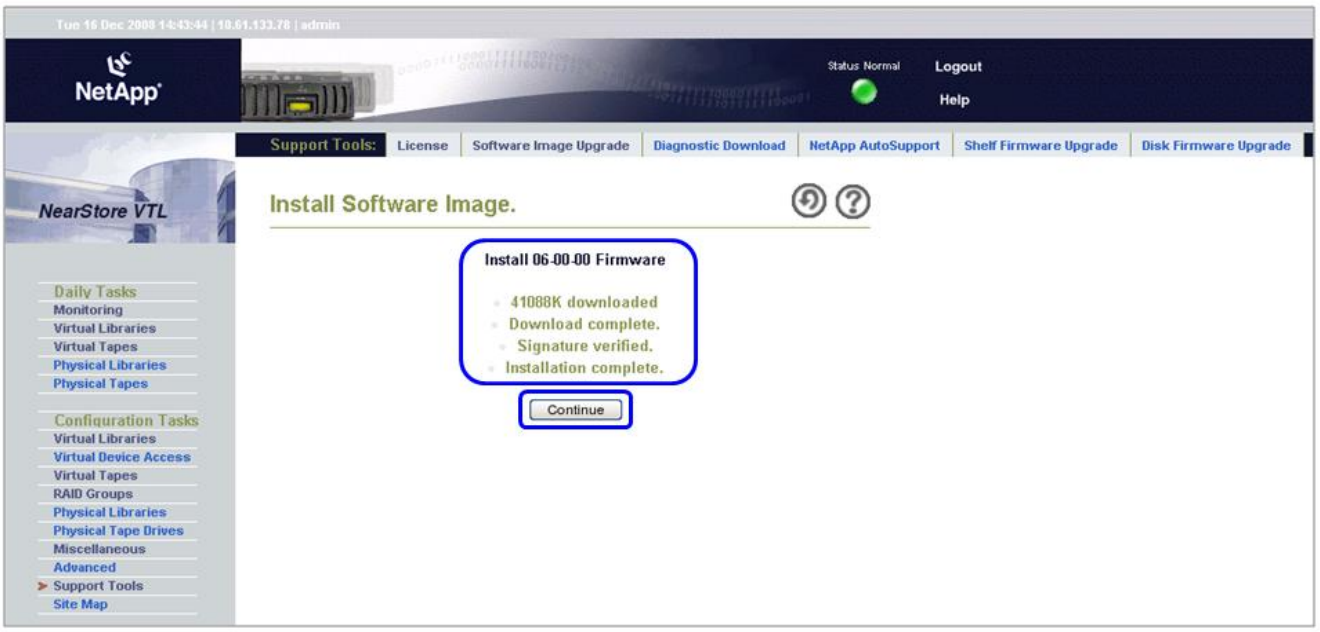
ESG Lab performed a code upgrade, then simulated SAN and controller failures during this phase of testing. A NetApp VTL System running Software version 5.6.0 was used for this series of tests, as seen in Figure 18. After clicking Support Tools on the right side of the screen, ESG Lab clicked Software Image Upgrade to kick off the upgrade.

FIGURE 18. ONLINE CODE UPGRADE



As shown in Figure 19, the NetApp VTL displays progress as the code is downloaded, verified and installed. Once installation is completed, the administrator is prompted to click Continue, which confirms a system reboot is required to load the new firmware.

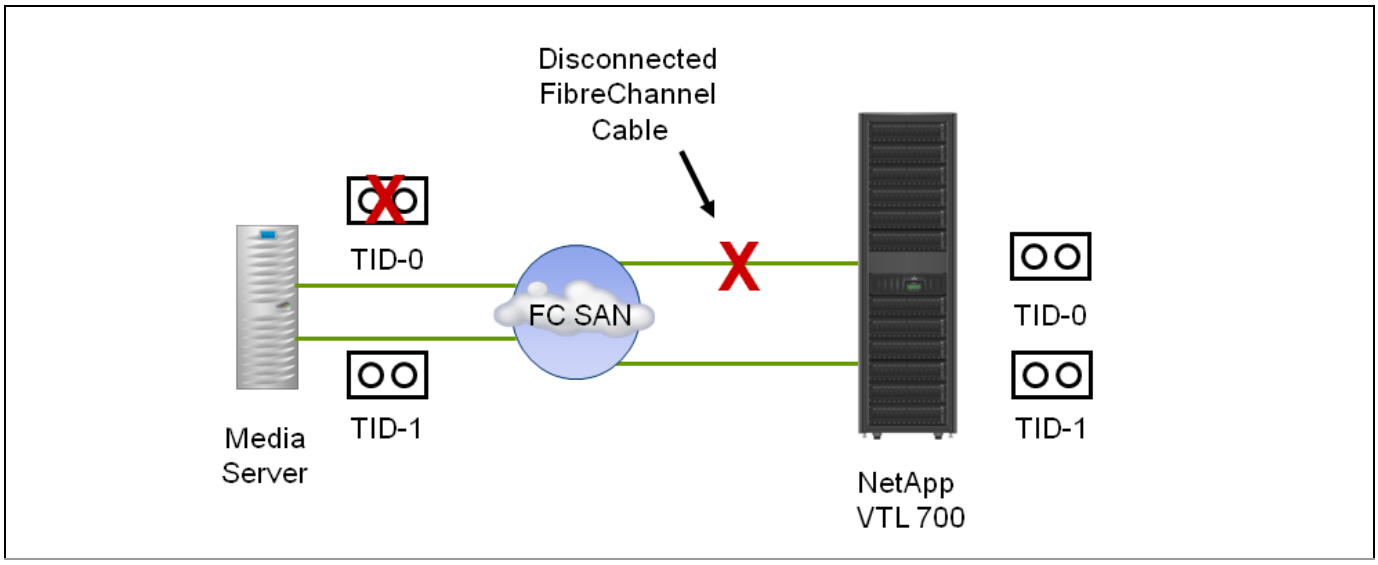
FIGURE 19. ONLINE CODE INSTALL PROGRESS



The entire code upgrade, including the reboot, took just four minutes for a system with two disk shelves. NetApp advises that upgrades to systems with more disk shelves would take slightly longer.

ESG Lab next injected a fault by unplugging one redundant FC (Fibre Channel) connection to simulate a problem in the SAN while data was being backed up by NetBackup. For this test, NetBackup was configured in failover mode with two virtual tape drives in a Storage Unit Group. This configuration lets the backup application continue to operate when SAN connectivity to the VTL is compromised. ESG Lab confirmed that the backup application was able to continue to write data to the Virtual Tape Library via the surviving path, with no intervention required. As shown in Figure 20, when the Fibre Channel cable to TID-0 was pulled, the server could no longer see the tape drive. Visibility to TID-1 was unaffected and NetBackup simply started writing the backup to TID-1.

FIGURE 20. SAN FAULT SIMULATION



Finally, ESG Lab simulated a VTL head replacement by resetting the VTL server head to factory defaults, thereby erasing all configuration data. The net effect of this procedure was exactly the same as removing a failed controller (or a controller scheduled for upgrade) and replacing it with a new controller received from the factory. The NetApp VTL provides the ability to save a configuration file, which can be loaded at this point to recreate all tape libraries. If a user does not have this file, the libraries can be recreated manually, because the VTL uses self-describing data on disk. The final steps are to import the RAID groups, assign tapes to libraries, assign libraries to Fibre Channel ports and adjust the SAN zoning. ESG Lab was able to restore all previously configured tape libraries and virtual tape drive and media configurations in less than 10 minutes. ESG Lab estimates that even with SAN reconfiguration (which was not performed here), the entire process could have been accomplished in 15 minutes.

Why This Matters

Nearly Half of IT professionals surveyed by ESG reported reliability as a challenge with their current data protection processes and technologies.[8] Missing a backup window or a restore SLA can be extremely costly in both lost productivity and hard dollars. Most backup hardware failures happen overnight (when backups are most active) and often aren't discovered until the following morning. A missed backup puts the company at risk until the problem is resolved and the next backup completes successfully. As a result, backup administrators and IT managers frequently dial in to check backup jobs in the middle of the night.

ESG Lab has validated that the NetApp VTL provides online code updates, continues to operate through SAN faults and can be restored to full operation easily even after controller failure. These enterprise-class features ensure that backups complete on schedule and that data is ready to restore when it is needed.

[8] Source: ESG Research Report, *Data Protection Market Trends*, February, 2008.

ESG Lab Validation Highlights

- ☑ ESG Lab confirmed that a NetApp VTL can be deployed in an existing tape environment without changes to policies or procedures and users can be performing backups in less than 10 minutes.
- ☑ ESG Lab has validated that data deduplication in the NetApp VTL can be used to reduce disk capacity significantly while applying deduplication policies appropriate to data type and retention needs.
- ☑ ESG Lab observed block level deduplication in the NetApp VTL reduce the required disk capacity for ten full Oracle DB backups from 28.87 TB to only 1.29 TB, a data reduction factor of 95% or 22:1. ESG lab audited similar backup tests with file system and Exchange data and saw similar reduction ratios.
- ☑ NetApp VTL delivered performance that met or exceeded the claims made on the NetApp website using an "out-of-the-box" configuration and only one-sixth of its potential total disk spindles. Backing up a mixed environment including deduplicated and non-deduplicated data sets drove a sustained throughput of 1,004 MB/sec, enabling organizations to protect more than 3 TB per hour, or 27 TB in a single 8-hour backup window, using the configuration tested by ESG.
- ☑ Tape Smart Sizing eliminated a huge management headache by automatically ensuring that each virtual tape created will fit on a single, matching physical tape, regardless of the compressibility or deduplication of data.
- ☑ ESG Lab was impressed with NetApp Shadow Tape functionality, which enables physical tapes to be sent off-site while maintaining a disk copy for fast restores, as well as the integration achieved with both Veritas NetBackup via API and EMC Networker via PERL scripts.
- ☑ New disk shelves were added nearly instantaneously and immediately put into service.

Issues to Consider

- ☑ The Management Console is crisp and intuitive, but in ESG's opinion could use two adjustments. Users would benefit from the ability to view multiple VTL appliances from a single browser window and select the VTL they would like to manage. In addition, the "daily tasks" and "configuration tasks" menus should be collapsed into one master menu for simplicity.
- ☑ Although detailed statistics are collected and saved internally by the NetApp VTL, detailed reports are currently not available in the GUI. Customers can create their own reports via scripts, but an enhancement to extract that data either in direct report form or in a format that could be imported into a third-party program for reporting would be extremely useful.
- ☑ When customers upgrade from previous versions of VTL code, backups can be converted to the new format, which supports deduplication, on a tape by tape or library by library basis. Conversion is a background task and can happen during the scheduled window or on demand.

ESG Lab's View

ESG has found that interest in—and adoption of—VTL solutions is both strong and growing, with the vast majority of organizations, both small and large, relying on a mix of disk and tape to meet data protection requirements. A small—and shrinking—percentage of users are still using tape exclusively for data protection. Organizations of all sizes are being asked—and in many cases told—that they've got to have secondary data available and quickly accessible in downtime situations. Not doing so can be costly, if not financially devastating.

ESG Lab has validated that the NetApp VTL can be deployed, configured and performing backups in less than ten minutes. The demonstrated capacity savings of NetApp VTL deduplication change the economics of data protection—making backups more cost-effective as more data is stored for longer periods of time on fast and reliable disk. Using a real-world Oracle database, ESG Lab observed 96% data reduction after just ten full backups, a 22:1 deduplication ratio. NetApp deduplication for VTL is a no-cost feature and can be added to all existing VTL 300, VTL 700 and VTL 1400 systems. The NetApp VTL also delivered enterprise class performance, providing 1,004 MB/sec of sustained throughput for a combination of deduplicated and compressed backups through a single controller, enough to protect 27 TB of data in an 8 hour backup window.

ESG Research has confirmed that Virtual Tape Library (VTL) and deduplication technology can no longer be classified as emerging technology and has moved into the early mainstream. Early adopters have confirmed that VTL slips easily into existing environments, while significantly improving the speed and reliability of backup and, more importantly, recovery. Deduplication transparently and dramatically reduces storage requirements, enabling deep archive of backups on disk. NetApp direct-to-tape technology can reduce the number of media servers needed to support a solution. This can reduce hardware, software and management costs by utilizing the VTL to perform tape exports. ESG Lab believes that the enterprise-class performance and scalability of NetApp VTL solutions, combined with deduplication and the long-standing NetApp tradition of easy implementation and integration, provide a compelling approach to a disk-based backup strategy.

Appendix

TABLE 3. TEST CONFIGURATION

NetApp VTL 700 – 4 Disk Shelves	Version 6.0.0
Veritas NetBackup Server IBM x336, 1x 3.0GHz XEON CPU, 2GB RAM 2x HBA -QLA2340	Windows 2003 SP1 Veritas NetBackup: 6.0 MP5
EMC NetWorker Server IBM x336, 1x 3.0GHz XEON CPU, 2GB RAM 2x HBA -QLA2340	Windows 2003 SP1 EMC Legato NetWorker: 7.3.1
2x I/O Load Generators IBM x336, 1x 3.0GHz XEON CPU, 2GB RAM 2x HBA -QLA2340	Linux RHEL 4.0 hpcreatedata ver. 1.2.3
Cisco MDS 9506 FC Switch 3x2GB Blade (96 ports) 1x4GB Blade (40 ports)	
ADIC i500 TLU IBM LTO-3	Library firmware: 300g.gf015 Tape drive firmware: 59D2 F/W
NetApp VTL 700 – Software Upgrade	Version 5.6.0

20 Asylum Street
Milford, MA 01757
Tel: 508-482-0188
Fax: 508-482-0218

www.enterprisestrategygroup.com