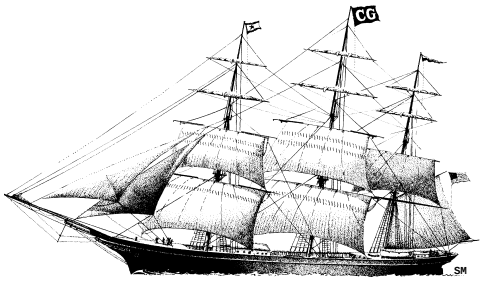
THE CLIPPER GROUP
Navigator™



Navigating Information Technology Horizons℠

Published Since 1993 Report #TCG2008058R October 21, 2008

Virtualizing the High-End Tape Environment — IBM Enhances VTL Offering

Analyst: David Reine

Management Summary

Technology has invaded my personal domain and I am not complaining. In fact, technology has taken over to such an extent where it is now intrinsic to the way we entertain ourselves in our homes. It used to be that if you wanted to watch a TV show and you were not going to be home, you could pop a tape into the VCR, make a few settings (OK, more settings than anyone over 55 could handle), and you could record the program for later viewing. That is, as long as the show was not longer than two-to-four hours. If you wanted to watch a baseball game, or any untimed sporting event, good luck if it went longer than expected. Today, VCRs are history, for the most part, replaced by DVRs with enough capacity to record an entire season of *Desperate Housewives* or any other multi-episode program, and the Internet, where you can view just about any movie ever made by downloading it from a variety of web sites that rent movies on-demand from a virtual library of choices – all for your home-viewing enjoyment. This is not only convenient, but it is economical as well, since no baby-sitter is required while you go to the video store and there is no driving to the megaplex cinema. Thus, you can save money and gas while having an evening of entertainment.

This same technological revolution has also invaded the data center of every enterprise, as well, and just in time. The rapid proliferation of storage in the enterprise data center is having a tremendous, albeit negative, impact on the storage budget of every CIO. With information capture doubling in volume every 12-to-18 months, the IT staff is having a difficult time finding the floor space, cooling, and electrical energy necessary to deploy additional disk arrays to retain the primary copy of this data, let alone multiple backup copies mandated by corporate policy and government regulation, or remote disaster recovery copies required by business continuation policy. *Yesterday*, the data center could neatly file away all of the *megabytes* of data on tape reels or cartridges and store them in automated tape libraries for nearline retrieval, or in a rack in a remote location for disaster recovery. *Today*, few data centers have the required capacity to keep copies of *petabytes* of mission- and business- critical data available on an online basis, even though a number of enterprise SLAs require the IT staff be able to retrieve lost data instantly in order to respond to application requirements. Technology, however, has found a way to meet these data center challenges. By implementing a Virtual Tape Library (VTL), the data center can store backup copies of vital information on spinning media, ready for instantaneous retrieval, while at the same time, migrating less urgent data to a tape library for disaster recovery or archive purposes. Unfortunately, many enterprises still are having a problem keeping up with capacity requirements that continue to double.

IBM has a solution for them. With the latest iterations of the *TS7700 Virtualization Engine*, IBM has dramatically increased the disk cache available to the data center and upgraded the ability to transition less critical information to high-speed, high-capacity tape cartridges. To learn more about IBM's new high-end VTL, please read on.

IN THIS ISSUE

Enterprise Backup Requirements

The seemingly unlimited growth of storage in the data center continues to put a heavy burden on an overworked IT staff responsible for maintaining backups in a mission-critical enterprise data center. Corporate policies and government regulations define the type of information that must be retained, even defining the length of time that it must be retained, varying from one week to one year, or longer, depending upon the application. Many organizations have SLAs that define recovery point objectives (RPOs) and recovery time objectives (RTOs) to ensure a speedy return to operational efficiency in the event of an application or operational failure. **These SLAs become increasingly difficult to manage, however, when the backup window is *limited*.**

Tape has always been a primary medium for backup, disaster recovery, and archive in the data center. It is an inexpensive and portable medium with high-streaming throughput, high capacity, and low energy requirements, to ensure the continuation of data processing in the event of a data loss, or, as we have seen in recent years, the need to recreate an entire storage environment for disaster recovery.[1] Unfortunately, the portability of tape leaves the enterprise open to the possibility of loss or theft. In either case, the IT staff must protect the information contained on tape by encrypting all private information that leaves the data center, whether it is going to another enterprise location, to a partner, or to a vault.

Tape does not always satisfy the RTO goals in a mission-critical environment. In some instances, the data center has replicated data on a disk-to-disk, or D2D, basis in order to accelerate the recovery process. Many businesses, recognizing the need to improve the utilization of tape resources *and* to store backups for very long periods, employ a D2D2T strategy, taking advantage of the economies of tape in a multi-tier ILM environment[2]. The need to improve recovery time is no excuse, however, to lose sight of the overall enterprise objectives to simplify the storage infrastructure, reduce management costs, and control the total cost of ownership (TCO) of the storage environment. The data center staff must balance the requirement to improve overall performance with the need to lower the cost per gigabyte, including the costs for electricity and cooling[3].

One way to improve enterprise storage performance is to automate the backup and recovery processes with a virtual tape library (VTL). A VTL can perform as an endpoint in a D2D architecture, or act as an intermediate depository on the way to physical tape to help enable faster recovery. It employs high-capacity SATA drives for second-tier storage, complementing an ILM infrastructure. This enables more frequent, automated recovery points and, therefore, a faster recovery time without requiring tape interaction.

The IT staff can complement existing physical tape facilities by taking advantage of a multi-tier hierarchy with D2D backup, especially when some applications cannot take advantage of the high-speed streaming characteristics found in new enterprise drives, such as IBM's *TS1130* at 160MB per second. With standard backup software such as IBM's *Tivoli Storage Manager*, Veritas' *NetBackup*, or EMC Legato's *NetWorker*, the data center can stage the initial backup to a VTL, **improving the backup and restore processes.** A VTL can **improve the RTO,** eliminating a slower recall from tape, utilizing higher disk throughput to reduce bottlenecks. A VTL can complement the existing business continuance infrastructure by optimizing IT resources and utilizing physical tape resources more efficiently.

IBM invented the virtual tape system introduced in 1997, and, in 2006, IBM introduced its *TS7700 Virtualization Engine* to provide enhanced tape virtualization for the *System z* environment. IBM designed the TS7700 to provide improved performance and capacity and to help lower the total cost of ownership for tape processing. It introduced a modular, scalable, high-performing architecture for mainframe tape virtualization. The TS7700 integrated the advanced performance, capacity, and data integrity design of the *3592* tape drive, IBM's industry-leading tape technology, with high-performance disk and

[1] See the issue of Clipper Notes dated February 1, 2007, entitled *The Evolving Role of Tape in the Data Center*, available at http://www.clipper.com/research/TCG2007013.pdf.

[2] See the issue of Clipper Notes dated May 28, 2008, entitled *Making Use of Virtual Tape in a D2D2T environment*, available at http://www.clipper.com/research/TCG2008028.pdf.

[3] See the issue of Clipper Notes dated February 13, 2008, entitled *Disk and Tape Square Off Again – Tape Remains King of the Hill with LTO-4*, available at http://www.clipper.com/research/TCG2008009.pdf.

an advanced *System p* server to form a storage hierarchy managed by robust storage management software and an extensive self-management capability.

With the continuation of unprecedented growth in storage requirements, and increased federal oversight in the financial arena, the need for more backup and archive capacity and higher throughput has led IBM to introduce two new versions of their high-end, mainframe virtualization engine: the *TS7720* and the *TS7740*. These additions are part of IBM's availability and retention initiatives within its *Information Infrastructure* strategy, the formal umbrella under which IBM focuses on the solutions required by enterprises, rather than the individual hardware, software, and services components.[4] This report, however, focuses on the new components.

The TS7700 Solutions

The TS*7700 Virtualization Engine Family* represents a family of mainframe virtual tape solutions available to the data center to optimize the processing of backup and archive information, while reducing costs such as power, maintenance, operations, and support that all contribute to increased TCO. With the *TS7740 Virtualization Engine*, the data center can deploy a fully-integrated tiered storage solution, leveraging the best features of both disk and tape to enhance backup performance and provide the tape processing capacity required in today's data center. The *TS7720 Virtualization Engine* provides the high disk capacity needed by mission-critical applications that can take advantage of a large cache to satisfy rapid recall requirements (without requiring physical tape connectivity). Available in a standalone configuration, the data center can also deploy theses TS7700s as a two- or three-cluster GRID to increase capacity or provide high availability service with automated replication to support business continuity.

Both the TS7720 and the TS7740 include functions such as advanced policy management to help automate and simplify IT operations by controlling physical volume pooling, cache management, dual copy, dual copy across a grid network, and copy mode control. They offer the same standards-based management interface and enhanced statistical reporting, available in the previous iteration of the TS7700.

IBM's TS7700 Virtualization Engines are feature rich. Some of the major points that differentiate the TS7700 from competitive offerings include the ability to provide volume consistency at a remote site during an outage. That is, if there were two TS7700s in a grid and one became unavailable the second TS7700 could take ownership of all volumes from associated servers. After the first TS7700 becomes available, all volumes that were affected are brought current. This allows for outstanding business continuity and virtually 100% data availability.

The TS7700 also has flexible replication policies that allow it to send virtual cartridge volumes asynchronously or synchronously over the grid to different TS7700 targets to support data availability and business continuity strategies. In addition, the TS7700 has a relatively low cost to add incremental capacity and performance. An enterprise can add performance throughput or add capacity in scalable increments as granular as 100 MB per second for performance, or a cartridge at a time for capacity, while competitive offerings for this segment are not as flexible or granular.

TS7740 Features

IBM designed the TS7740 to help accelerate the backup and recovery process by using a tiered hierarchy of disk and tape in order to take advantage of the performance of disk and the increased capacity and low cost of tape. Identical in functionality to the original TS7700, the newest release (1.5) of the TS7740 comes with a number of enhancements which:

- Enables a higher capacity disk cache;
- Integrates the Enterprise Library Controller for a simplified infrastructure with reduced cost and floor space requirements;
- Supports the newest high-performance IBM tape drive, the TS1130, with data encryption and automatic reformatting of existing cartridges on the next physical volume usage;
- Supports the TS3500 Tape Library with a high-density frame, providing up to 1PB of capacity, in a single frame; and
- Supports new 1Gb/s Ethernet adapters for expanded high-speed grid connections.

[4] See *The Clipper Group Captain's Log* dated September 8, 2008, entitled *Is IBM is Serious about Managing Information Instead of Storage Infrastructure?*, which is available at http://www.clipper.com/research/TCG2008042.pdf.

Based upon IBM's POWER architecture, the new RAID Array Cache Controller, 3956 Model CC7, comes with two-socket, dual-core, 64-bit Power5+ microprocessors running under the AIX operating system, taking advantage of the outstanding performance of IBM's System p environment. Taking advantage of IBM's investment in disk architecture, the TS7740 borrows from IBM's *DS4700* disk arrays to expand the native tape volume cache from 6TB per node up to over 13TB per node, for a maximum cache capacity of over 40TB uncompressed (over 120TB with 3:1 data compression) per three-node cluster. The data center staff may use a multi-node cluster to increase the capacity of the TS7740 configuration to appear as a single library or used to create a high-availability scenario in a geographically distributed environment.

The TS7740 also comes with the same functionality of the prior model with support for 256 virtual drives and 1 million virtual volumes per node, up to 16 TS1130 tape drives per node, and 4 *FICON* channels per node.

TS7720 Features

The TS7720 provides a cache-centric solution to accelerate the processing of frequently accessed data, with the same benefits and functionality as the TS7740, but without the attachment to physical tape(s). The TS7720 enables the data center to capture, over FICON, up to 70TB of uncompressed backup or archive data per node with a maximum capacity of 210TB using a standard 3:1 compression ratio for mainframe data. A three-node TS7720 configuration can support up to 630TB of information with a throughput of up to 1200MB/sec, or can be used, like the TS7740, in a distributed grid mode. This is a significant advantage for any application that benefits from having its data disk resident. This is especially true for image applications that previously used optical disk as a medium or applications such as *DFSMS/OAM* that manage large objects, such as scanned images or coded data. In the past, these objects were stored on disk, tape, and optical platters.

Conclusion

With no end in sight to rapid data growth, the CIO has to find a means to both capture a limitless stream of mission- and business-critical data and be able to retrieve that information at a moment's notice to in order to retain business continuity, while at all times conforming to rigid budgetary constraints. With version 1.5 of the TS7700 Virtualization Engine, IBM addresses these CIO concerns, especially the critical nature of the TCO of the enterprise data center, adding additional functionality and innovation to an already rich solution.

With the TS7740, the enterprise can deploy a totally integrated and automated platform to support a traditional D2D2T environment with a beefed up disk cache and support for the TS1130 tape drive with its native 1TB capacity, sustained throughput of 160 MB/s, and encryption capability so necessary to protect the resources and brand of the enterprise. With the TS7720, the data center can take advantage of the highest cache capacity of any IBM VTL. With a compressed capacity of up to 630TB, The TS7720 can support the D2D requirements of just about any enterprise.

If your enterprise is battling with increasing backup data and the never-ending war with the ever-shrinking backup window, you need to look at the advantages that IBM is offering with the newest members of the TS7700 family. Whether your requirements are for a D2D or a D2D2T environment, IBM has a well thought-out Information Infrastructure solution that employs the latest technology for your enterprise.

About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

David Reine **is Director, Enterprise Systems for The Clipper Group.** Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. He joined The Clipper Group after three decades in server and storage product marketing and program management for Groupe Bull, Zenith Data Systems, and Honeywell Information Systems. Mr. Reine earned a Bachelor of Arts degree from Tufts University, and an MBA from Northeastern University.

- ***Reach David Reine via e-mail at dave.reine@clipper.com or at 781-235-0085 Ext. 123. (Please dial "123" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.