TANEJA GROUP
TECHNOLOGY ANALYSTS

P R O D U C T   P R O F I L E

Evaluating Enterprise-Class VTLs: The IBM System Storage TS7650G ProtecTIER De-duplication Gateway

September 2008

THE NET NET

Increasingly stringent service level agreements (SLAs) are putting significant pressure on large enterprises to address backup window, recovery point objective (RPO), recovery time objective (RTO), and recovery reliability issues. While the use of disk storage technology offers clear functional advantages for resolving these issues, disk's high cost has been an impediment to widescale deployment in the data protection domain of the enterprise data center. Now that storage capacity optimization (SCO) technologies like single instancing, data de-duplication, and compression are available to reduce the amount of raw storage capacity required to store a given amount of data, the $/GB costs for disk-based secondary storage can be reduced by 10 to 20 times. Virtual tape technology, disk-based storage subsystems that appear to backup software as tape drives or libraries, are one of the most popular ways to integrate disk into a pre-existing data protection infrastructure because they require very little change to existing backup and restore processes. While virtual tape libraries (VTLs) are interesting, SCO VTLs that leverage data de-duplication and other related technologies are compelling.

Given high data growth rates, stringent SLAs for data protection, and the need to contain spending, enterprise customers really need to take a look at SCO technologies. Taneja Group predicts that large enterprises will rapidly move to SCO VTLs over the next 1-2 years while the market for non-SCO VTLs (VTLs that do not have integrated SCO technologies) dwindles rapidly. Data growth rates in the 50% - 60% range will be pushing this transition as much as will the clear cost advantages that SCO VTLs offer over non-SCO VTLs. While SCO is a key requirement, performance remains the number one need of the enterprise data protection environment. After all, if the SLA for completing the day's backup cannot be met, all other criteria are moot. This has significant implications for vendors of SCO VTLs. Their solutions must provide the capacity optimization that the enterprise customer demands, while enabling enterprise-class performance. Vendors that can provide both efficient SCO technology and enterprise class performance offer a very compelling value proposition.

In this Product Profile, we discuss the criteria we recommend be used to compare and contrast enterprise-class SCO VTL solutions from different vendors, and then evaluate how the IBM System Storage TS7650G ProtecTIER De-duplication Gateway performs against these criteria. The TS7650G, IBM's first offering based on technology from the April 2008 acquisition of Diligent Technologies, supports very high single system throughput, multiple PBs of usable capacity, and optional clustering with support for a global de-duplication repository - all important considerations for enterprise SCO VTL prospects.

87 Elm Street, Suite 900   Hopkinton, MA 01748   Tel: 508-435-5040   Fax: 508-435-1530   www.tanejagroup.com

The Inevitability of Disk-Based Data Protection

Disk is in widespread use as a part of the data protection infrastructure of many large enterprises. Evolving business and regulatory mandates are imposing stringent SLAs on these organizations, pushing them to address backup window, RPO, RTO, and recovery reliability issues, and disk has a lot to offer in these areas. Technologies such as VTLs have made the integration of disk into existing data protection environments a very operationally viable option.

Cost has historically been the single biggest obstacle to integrating disk into existing data protection infrastructures in a widespread fashion, but the availability of SCO technologies such as single instancing, data de-duplication, and compression have brought the $/GB costs for usable disk capacity down significantly. SCO-based solutions first became available in 2004, and the SCO market hit $237M in revenue in 2007. Over the next five years, we expect revenue in the SCO space to surpass $2.2B, with the largest single market sub-segment being SCO VTLs (source: Taneja Group Next Generation Data Protection Emerging Markets Forecast September 2008). If you are not using disk for data protection purposes today, and you are feeling some pressure around backup window, RPO, RTO, or recovery reliability, you need to take another look at SCO VTLs. It is our opinion that within 1-2 years, SCO VTLs will be in widespread use throughout the enterprise. With data expected to continue to grow at 50% - 60% a year, the economics of SCO technology are just too compelling to ignore.

A Brief Primer on SCO

Taneja Group has chosen the term SCO to apply to the range of technologies that are used today to minimize the amount of raw storage capacity required to store a given amount of data. Data de-duplication is a common term in use by vendors, but this term really only describes one set of algorithms used to capacity optimize storage. And many vendors of de-duplication use it along with other technologies, such as compression, in a multi-step process used to achieve the end result. That said, de-duplication is the primary technology that enables solutions to reach dramatic capacity optimized ratios such as 20:1 or more. Given the focus and attention on de-duplication - as well as the fact that it is at the heart of IBM's TS7650G - let's take a closer look.

At their most basic level, data de-duplication technologies break data down into smaller recognizable pieces (ie. elements) and then look for redundancy. As elements come into the system, they are compared against an index which holds a list of elements that are already stored in the system. When an incoming element is found to be a copy of an element that is already stored in the system, the new element is eliminated and replaced by a pointer to the reference element. In secondary storage environments like backup where backed up data may only change 3-5% or less per day, there is a significant amount of redundancy that can be identified and removed (a 5% change rate implies a 95% data redundancy rate!). De-duplication algorithms can operate at the file level (this is also referred to as single instancing) or at the sub-file level. Sub-file level de-duplication

tends to produce higher data reduction ratios. Looking across vendor offerings in the market today, it is not unreasonable to achieve data reduction ratios against secondary storage like backup data sets of 10:1 to 20:1 or greater over time.

To provide an example of how de-duplication performs in practice in backup applications, let's take an example. Assume a new data set that has never before been backed up. On day 1, it is backed up to disk and de-duplicated (this may occur during the backup or after the backup, but more on that later). On day 2, the data set is once again backed up to disk, but as de-duplication is applied, it can now look at both backups to find common elements. The data reduction ratio achieved on day 2 is very likely to be higher than that achieved on day 1, particularly if the backed up data has not changed much in the 24 hour period. If we assume that 30 days of backups are retained on disk, then it is very likely that there is a lot of redundant data that can be removed and replaced with pointers. The factors affecting data reduction ratios in backup include the change rate of data (day to day), the number of days of retained backups, and the specific SCO technology in use.

SCO Approaches and Architectures

SCO can be deployed either at the source (backup client) or at the target (backup target). Performing the capacity optimization work requires CPU cycles, so where it is performed may have a performance impact that needs to be evaluated. Source-based SCO typically leverages resources on the backup client to perform the work, which may impact backup and/or application performance, but it does minimize the amount of data that has to be sent across a network to complete the backup. Source-based SCO may offer certain advantages in remote office back office (ROBO) backup environments, but tends to be targeted at environments where each backup client does not have a lot of data.

Target-based SCO presents a backup target, often through a VTL interface, and leverages resources on an appliance or a storage subsystem to perform the work. Target-based SCO supports much greater throughput than source-based, and tends to be targeted for use in enterprise environments to handle large backup volumes per client. Target-based SCO can offer the opportunity to much more efficiently leverage a global data de-duplication repository during the capacity optimization process than source-based SCO can. Vendors that support a global repository can often offer higher data reduction ratios than those that do not since they can perform the redundancy identification and elimination across a much larger number of backup clients.

Capacity optimization can be performed through either an in-line or a post-processing approach. In-line processing performs the capacity optimization work as it is writing data to the backup target. Post-processing allows the data to be first written to the backup target, and then through a separate process picks this data back up and runs it through the capacity optimization process. The operative metric for an end user, assuming that you want your backups

in capacity optimized form, is the amount of time it takes to both ingest the backup and to perform the capacity optimization, not just the time it takes to ingest the backup.

This dichotomy (in-line vs post processing) has some key implications on overall system performance that may not be entirely evident. When an in-line vendor quotes a throughput number, that is the single number necessary to evaluate how long it takes to complete the backup and process the data into capacity optimized form, at which point it is ready for any further processing (e.g. 600MB/sec can process roughly 2.16TB/hour). When a post-processing vendor quotes throughput, that generally refers to how long it takes to ingest the data and does not include the post-processing time necessary to capacity optimize it (e.g. 600MB/sec can ingest 2.16TB/hr but additional time will be required to perform post-processing). To truly understand if a post-processing approach can meet your backup windows, you need to evaluate the total time required to both ingest the backup and to perform the post-processing. Post-processing vendors may argue that since the post-processing is de-coupled from the backup, it doesn't matter how long it takes. In some environments, that may be true but if you have an 8 hour window to complete your backups and capacity optimize them before you clone data to tapes, or replicate your backup sets to a remote site for DR purposes, and you cannot complete the backup ingest and the post-processing within that 8 hour window, then the post-processing approach will impact your DR RPO.

Without a doubt, in-line approaches require less overall physical storage capacity than post-process approaches. For a given environment exhibiting a 10:1 capacity optimization ratio, the system will write 100GB of data for every 1TB it backs up. A post-process method will need to write that 1TB to disk first, then cycle it through post-processing, eventually shrinking the storage required to store that backup to 100GB. Thus, post-processing systems must maintain spare capacity to allow for the initial ingest of data prior to the de-duplication process. Post-processing products clearly require more capacity for a given environment than in-line solutions to allow for this buffer, but the actual amount will vary based on the specific post-processing approach being used.

Post-processing approaches introduce additional time before a capacity optimized backup is ready for further processing, such as cloning to tape, distributing electronically to a DR site, etc. If additional time and capacity are available, then you may be indifferent between the two approaches, but if they are not, then this is something to consider when evaluating solutions. Note that some post-processing vendors allow the post-processing to be started against a particular backup job before it completes, thereby reducing both the capacity and time requirements that would otherwise be associated with approaches which perform these operations sequentially. In-line approaches, however, will generally complete the overall backup processing (ingestion + capacity optimization) faster than post-processing approaches since they complete their work in a single pass.

What To Look For In A SCO VTL

First, you need to understand what your backup issues are and how you prioritize them. If you're like most enterprises, they will be most of the following: backup window, RPO, RTO, recovery reliability, solution cost, and offsite data storage requirements (whether by tape transport or replication). Other considerations include integration issues with your existing data protection infrastructure, whether you're targeting ROBO or data center environments, and what the quantity of data is that you will be dealing with over the lifetime of the solution. Once these issues have been understood, it's time to take a look at the technology options. Over the last several years, we have talked with hundreds of end users that have deployed SCO VTL technology, and that input, combined with our take on the developing trends in data protection, has led us to define the following criteria for evaluating SCO VTL solutions:

Performance. Assuming you want the data in capacity optimized form, the operative issue here is how fast you will be able to complete the backups and get the data into its capacity optimized form so that it is ready to be used for any additional processing, such as tape cloning and/or replication to a remote site. Whether you choose an in-line or a post-process approach may impact backup ingest time, but you still need to understand the total time required to ingest and capacity optimize the backup to ensure that you will have sufficient time to meet any further backup processing requirements.

If your target is to complete daily backup activities within 8 hours, and you have roughly 26TB of data that will have to be transferred each day to perform the backups, then an in-line solution would need to process data at about 900MB/sec on a sustained basis to meet this requirement. With a post-process solution, you would need to be able to ingest the backup and complete the separate SCO processing within that same 8 hour period - a difficult challenge. To make this calculation, you'll need to ask the vendor about the rate at which data is capacity optimized during post-processing.

Scalability. There are several issues to consider here. First, understand what the base capacity of the system is. Capacity optimization ratios generally vary across workloads, but the more base capacity is supported, the more usable capacity will be supported. Let's define some terms here. *Base capacity* is the amount of raw capacity supported after any RAID-based data protection schemes have been taken into account. *Usable capacity* refers to the amount of storage capacity represented after any applicable SCO technologies have been applied against base storage capacity. For example, a system with 50TB of base capacity, when used with a workload that can be capacity optimized at a rate of 10:1, can store up to 500TB of raw data.

Next, understand what kind of capacity optimization ratios you can expect to achieve. If vendors offer a capacity planning tool that can be run against a target workload to provide an estimate, then take advantage of this. If at all possible, test several of the

technologies that look most promising in your environment, and don't just run them against a single backup. The throughput performance of various SCO algorithms may change over time as the indexes grow; conventional hashing and content-aware algorithms may actually suffer decreased throughput once their index has outgrown main memory capacity (something that often happens around 20TB of base capacity with conventional indexing algorithms). In environments that do weekly full and daily incremental backups, ratios will generally improve over time, approaching a steady state. The daily change rate of your data is a critical determinant of the ratios you'll achieve over time, and if you're like most shops your daily rate will vary somewhat.

Finally, understand if the solution you've chosen supports what is called a "global" repository. Earlier, we stated that some sort of index is generally referenced as each element comes into the system. Architectures that allow multiple SCO VTLs to reference a single, global repository that includes all the elements that have been seen before tends to offer better ratios than systems that have a single, separately developed index for each SCO VTL. Architectures that support global repositories tend to offer a better growth path as well, since when the performance capabilities of a single SCO VTL are outgrown, a new one can be added and can immediately take advantage of the index that is already there.

High availability. In today's 24x7 environments, even secondary data has to be highly available so that stringent SLAs can be met. SCO VTLs cannot compromise that

high availability as they are integrated into existing data protection infrastructures. Once data is converted into a capacity optimized form, it is not usable by applications until it can be re-converted back into its original form. If there is a failure, either within a SCO VTL or at the level of the entire SCO VTL, the data may not be available. For that reason, it is important to support high availability solutions that can ride through single points of failure. High availability architectures allow maintenance to be performed on-line as well, further improving the overall availability of the environment. Clustered architectures are a good way to meet this need, and can contribute to higher overall throughput as well if a global repository is supported. Look for support also for various RAID options on the back end storage to protect against disk failures.

Reliability. Because SCO VTLs effectively convert data into an abbreviated form prior to storing it, there is some conversion risk that must be evaluated. How does the system perform the conversion, and what is the risk of false positives (two elements that are not exactly alike being identified as such)? In SCO VTLs that use conventional hashing methodologies, this risk is called out as the "hash collision rate." While nominal hash collision rates may appear to be low with conventional systems, if they are going to be used in enterprise environments that may be dealing with petabytes of usable capacity, they need to be evaluated in light of that level of scale.

When data is read back, it's important to verify the accuracy of the conversion process.

Does the SCO VTL perform data verification to ensure that any retrieved data, after it is converted back into its original form, exactly matches the data that was originally written by the application? How is this done? Any system being evaluated for use in an enterprise environment must offer independent data verification to ensure conversion accuracy.

Solution maturity. With a technology like SCO, there is a learning curve for vendors. Being further down on the learning curve can translate directly into better performance, higher scalability, and improved data reliability. Look for vendors that have at least hundreds of systems deployed in production and can point to a number of references whose environments look similar to your own. Large enterprises often look for very broad support coverage which can address locations they may have on a worldwide basis. Larger, more mature vendors tend to offer better geographical support coverage than smaller vendors.

IBM's TS7650G: An Enterprise-Class SCO VTL Solution

In April 2008, IBM announced the acquisition of Diligent Technologies. With their in-line SCO VTL gateway, Diligent had already achieved considerable success, having established themselves as a leading SCO VTL vendor to large enterprises. The IBM acquisition puts the muscle of a trusted storage supplier behind Diligent's unique and innovative ProtecTIER technologies.

IBM's announcement of the TS7650G ProtecTIER De-Duplication Gateway in

September 2008 represents the integration of Diligent's technology into IBM's Tape Systems product portfolio and includes important new functionality for large enterprises. With this release, IBM offers clustering for high availability, supports a global repository across cluster nodes, and doubles the sustained single system throughput of their SCO VTL to almost 1GB/sec – a number that clearly marks them as the industry leader for in-line, single system SCO VTL performance today. This is a familiar position for them however, since the previous version of the ProtecTIER technology had the industry's highest in-line, single node throughput before it was superseded by the TS7650G.

The ProtecTIER Technology

The TS7650G is a SCO VTL gateway based on an IBM System x with 3 GHz, quad core Intel processors and 32GB RAM, running Red Hat Linux. Available in two models – a single node or a dual node cluster – it supports FC on both the front and back ends and dedicated Ethernet connections for the cluster communications. While the gateway supports heterogeneous storage on the back end, IBM has specifically qualified their own storage subsystems, including the DS4000, DS8000 and IBM XIV storage platforms, as well as storage subsystems from EMC and HDS.

HyperFactor is the patent pending de-duplication technology that is used to perform the capacity optimization. What is so unique about this technology is that it is based on an extremely efficient indexing design that can map up to 1PB of base

storage in a scant 4GB of RAM. This supports the TS7650G's industry leading in-line, single node throughput because element identification and referencing is all performed in main memory – no accesses to disk are required. Competitive indexing technologies such as hashing and content-aware approaches have much less efficient mapping algorithms, forcing them to reference a disk-based index during the capacity optimization process to map more than around 20TB of base capacity. This explains why alternative capacity optimization technologies generally suffer decreased throughput as the repository grows; they run very fast when all the index references can be handled in main memory, but once they outgrow the available memory and must touch disk, reference times can slow down by two orders of magnitude. This efficient index mapping design sets HyperFactor apart, allowing it to scale linearly for repositories up to 1PB in base capacity. After HyperFactor completes the de-duplication process, it then compresses elements before they are stored.

The Importance of SCO VTL Clustering

With this announcement, IBM is unveiling gateway clustering along with support for a global repository. Although today they are supporting two node configurations, the architecture is designed to support up to 16 nodes over time, providing a very scalable growth path for high end customers. Clustered TS7650Gs present a single VTL image to backup servers across which single system throughput can be scaled. Based on data from ProtecTIER's installed base, many of their customers are seeing single node

sustained throughput in the 450MB/sec range, with peak throughputs topping 600MB/sec. In adding a second node and supporting a global repository, IBM is pushing the sustained throughput rate into the 900MB/sec range, with peak throughputs even higher. Because the entire index is mapped into the main memory of each node, it doesn't matter which node a backup stream hits: it will enjoy the same high level of performance.

When it comes to throughput in clustered environments, there is an important distinction between single system and aggregated throughput. Single system throughput identifies a throughput number against a *single repository*, access to which may be spread across multiple VTLs and multiple processing nodes. In the TS7650G's case, multiple gateways leverage a global repository, which makes the single node throughput number additive as nodes are added to scale the system. For example, a single node TS7650G can sustain speeds of 450MB/sec, while a two-node cluster can sustain 900MB/sec, all while accessing a single large repository. Other competitors talk about aggregate throughput numbers for their clusters, which implies that they do not support a global repository. In these products, there is a separate repository for each "node" so the performance numbers for each node are not additive. Such products lead to independent islands of storage, which limits the capacity optimization ratios to those achievable by a single node. Enterprises that are looking to consolidate their backup sets to improve efficiencies and reduce management points, necessarily prefer solutions with high single system

throughput as opposed to throughput that is aggregated across several independent systems.

The introduction of clustering technology has important implications in the areas of performance and high availability. As mentioned above, it allows IBM to increase their in-line, single node performance lead in the industry even further. Very high single system throughput is most important when customers have newer, higher performance FC interfaces between the backup servers and the VTL – just what you'd expect in the large enterprise environments at which IBM is targeting the TS7650G.

Availability is another extremely important consideration in these types of environments. In two node configurations, a single node can fail and the remaining node will immediately begin servicing the entire workload, although the overall throughput of the configuration will drop to that of a single node. The failed node can be replaced on-line and re-integrated into the cluster without having to disrupt the backup applications that are writing to the VTL. Clustering also gives customers additional flexibility in performing maintenance and upgrades to cluster nodes, as well as gracefully expanding cluster size in the future as larger node counts are supported. The TS7650G clustering technology supports *both* improved performance and availability, not just improved availability.

Evaluating the IBM TS7650G

How well does the TS7650G perform against the criteria we identified earlier for evaluating SCO VTL solutions (performance, scalability, availability, reliability, solution maturity)?

Performance. We've already reviewed the TS7650G's industry-leading in-line, single node and single system performance numbers, showing how that is directly related to IBM's patent-pending HyperFactor de-duplication technology. The highly efficient index design of HyperFactor allows it to scale up to 1PB of base capacity without impacting indexing performance, a considerable problem for competitive alternatives that are based on hashing or content-aware algorithms. IBM's roadmap includes expanding the solution to a higher number of nodes over time, which will offer large enterprises a non-disruptive, long-term growth path to higher performance. Competing vendors may offer higher aggregate throughput today, but single system throughput is the operative number for the enterprise data center. What is clear is that the TS7650G supports the industry's highest in-line, single system throughput performance for a SCO VTL today by a wide margin.

Scalability. The data growth rates that most large enterprises are experiencing today mean that most will be managing at least hundreds of terabytes of secondary data in the near future. With ProtecTIER's ability to support up to 1PB of raw capacity, the TS7650G can support multiple petabytes of usable capacity, depending on the achieved capacity optimization ratios across the relevant workloads. Hash-based and content-aware de-duplication algorithms do not even come close to the scalability of

HyperFactor, whose ability to map 1PB of base capacity in main memory supports multiple petabytes of usable capacity. The fact that IBM can scale to this level against a single, global de-duplication repository is key: all other things being equal, they will achieve higher data reduction ratios by using a global repository than vendors scaling to the same usable capacity but that spread that capacity over multiple repositories (one associated with each SCO VTL appliance). And the TS7650G's single node performance and scalability mean that you can build out these large configurations with less hardware, creating simpler, less expensive configurations. Whether you're consolidating multiple existing backup targets or creating a single backup target that can scale to petabytes of capacity, the TS7650G lets you do this very cost-effectively.

Availability. The introduction of clustering not only doubles single system performance but also addresses the enterprise requirement for higher availability. IBM's clustering technology provides a highly available environment that can tolerate the failure of a VTL node while maintaining access to all the data within the repository.

To provide the necessary levels of high availability, enterprise SCO VTL solutions also need to be able to ride through single disk failures. The TS7650G supports heterogeneous storage on the back end, and IBM recommends the use of RAID capabilities supported by this back end disk to provide high data availability. If higher levels of resiliency are desired, users can flexibly configure storage subsystems with the required levels of resiliency. IBM's Best Practices provide tools that recommend certain RAID configurations for it's repository (metadata and user data) for optimal performance and resiliency.

Reliability. Two basic issues were identified earlier in this area: the risk of false positives and the verification of retrieved data. HyperFactor uses a unique approach to identify and confirm redundant elements. At a high level, HyperFactor does a very low latency "fly by" looking for elements that look similar to what it has already seen. A more in-depth analysis is then performed only on the elements identified as "similar" whereas the "new" elements go immediately into the index before they are stored on the back end storage. Competitive approaches execute their full "chunk evaluation algorithm" on each and every element, which in the end generally means they end up doing a lot more work (at very high latency cost since a large percentage of references may require reads from disk) for *every element*. HyperFactor's approach not only handles higher throughput but also more reliably identifies each element.

ProtecTIER retains metadata about each element, one piece of which is a cyclic redundancy check (CRC or checksum). On reads, ProtecTIER assembles the required elements, performing checksums on each element once they have been converted back into their original form to verify that the data element read out of the repository is the exact same data element originally stored there.

The RAID capabilities of the underlying storage subsystems provide yet another level

of data reliability. Most RAID technologies from major storage suppliers offer options to repair bit error rates on the fly if they occur, and will reliably retain data even as disks occasionally fail. ProtecTIER solutions take on this additional level of data reliability as they support these storage subsystems.

Solution maturity. The ProtecTIER technology has an installed base of production deployments that numbers in the hundreds, most all of whom are in the Fortune 1000. Representative industries include financial services, healthcare, telecommunications, oil & gas, retail, media and entertainment, manufacturing, and government. There are customers managing storage capacities in the hundreds of terabytes of usable capacity across multiple ProtecTIER gateways. In its third year of availability, the base HyperFactor technology is mature and proven, with most customers seeing data reduction ratios between 10:1 and 25:1 for backup data.

The IBM acquisition is an important milestone in ProtecTIER's technology life cycle since its target customers care deeply about issues like reliable long-term technology sources, keeping the number of vendor relationships low, worldwide support coverage, and integration with other key data protection technologies. IBM is a company that large enterprises trust to address these issues.

Taneja Group Opinion

Given the amount of data with which large enterprises have to deal, and the expected growth rates for this data over the next five years, non-SCO VTLs are generally not going to provide a sufficiently scalable, cost-effective solution as a backup target. For this reason, we see the VTL market rapidly transitioning to SCO VTLs over the next 1-2 years. How a SCO VTL compares in its overall performance and scalability are the two critical issues which set enterprise-class solutions apart.

With this new announcement, the TS7650G offers the features required in an enterprise-class SCO VTL: industry leading in-line, single system throughput, expandability to multiple petabytes of usable capacity, high availability with a clustering approach that supports a global repository, built-in features to ensure that data is reliably identified, stored, and retrieved, and a mature solution that has been reliably deployed in hundreds of customers across multiple verticals. If you're looking for a SCO VTL solution to handle the kinds of backup workloads common in large enterprises, IBM's pedigree in this area is solid. Just ask their ProtecTIER customers.