

IBM Launches Newest ProtecTIER Appliance — Positioning Data Dedupe for Mid-Market Users

Analyst: David Reine

Management Summary

With millions of barrels of oil having spewed from the floor of the Gulf, attention returns once again to energy consumption, and the manner in which we waste the natural resources that are so essential to our economy and the very basic requirements of our daily life. As we drive around the average city in the U.S., we see thousands of gas-guzzling cars idling at red lights and driving in circles looking for a parking space in over-crowded parking lots. The requirement for innovation in energy use has led the automotive industry to more expensive alternative engines, hybrid and electric, in order to reduce the consumption of gas, while at the same time reducing incremental costs per mile. The size of today's cars is also continuing to shrink around the world, and beginning in the U.S., as we see the popularity of "The Smart Car" and other reduced-in-size vehicles, such as the new *MINI Cooper* from BMW, with better MPG ratings and a smaller footprint in the parking garage. Smaller cars with improved efficiency are becoming the way of the world. Bigger is not always better, however, rightsizing to meet real requirements is something that we have recognized in the datacenter for quite some time.

The server and storage sprawl that has added significantly to the cost and complexity of running a modern enterprise datacenter mandated more innovation. Consolidation and virtualization have reduced that complexity by enabling the IT staff to reduce the size of the infrastructure footprint and energy bill, as well as the acquisition cost for hardware and software. In terms of server applications, new x86 processors have doubled the number of cores per CPU from two to four, and now to eight, all in the same server footprint and energy envelope, enabling more applications, and more users, per server. In terms of storage, we have seen an increase in the capacity of a Tier-2 disk device, now at 2TBs, and a comparable increase in the capacity of open systems tape media, with LTO-5 now sitting at 3TBs on a compressed tape cartridge. Unfortunately, these increases in media capacity have not kept pace with the rapid explosion of information capture in the datacenter or the multiple copies of information that go into a data protection plan. In order to stay ahead of the next disaster, the IT staff has had to deploy new innovative techniques, such as data deduplication to reduce the size of the storage envelope, as well as the duration of the backup window. Data deduplication has been proving very effective in these large environments, but what is being done to enable mid-sized datacenters (and smaller) to enjoy the same technological advantages.

One company that is paying heed to the mid-market customer is IBM. In February, 2009, IBM introduced the *ProtecTIER TS7650* as a down-sized version of their ProtecTIER Gateway. This reduced the entry point for data dedupe to 7TBs, and lowered the entry cost as well. For many smaller datacenters, however, this was still too big and too costly for the mid-market. Now, IBM has announced the *TS7610 ProtecTIER Deduplication Appliance Express* with an entry configuration of 4TBs. To learn more about the TS7610, please read on.

IN THIS ISSUE

- **Smaller Datacenter Storage Dilemma... 2**
- **IBM's ProtecTIER Technology..... 2**
- **IBM TS7610 to the Rescue 3**
- **Conclusion 4**

Smaller Datacenter Storage Dilemma

Through consolidation, the IT staff in all datacenters has simplified the physical management of existing enterprise compute resources throughout the datacenter, replacing under-utilized, under-performing servers with new, multi-core processors in multi-socket systems. This has enabled the datacenter to reduce the server count by a factor of up to 10:1, lowering the TCO of the datacenter. The flexibility of these servers enables the improvement of server utilization through the virtualization of the server environment with virtualizers, such as *Hyper-V*, *VMware*, and *XenServer*. **By virtualizing multiple mission-critical, and business-critical, applications within a single physical server, the datacenter not only improves server utilization but also increases the demand for efficient backup technology from the virtualized environment to the storage network.** This results in the need for an affordable storage array with higher throughput, higher capacity, and the energy efficiency necessary to meet today's unique scalability demands of a growing enterprise, but keeping in mind the higher performance requirements of tomorrow. Specifically, in order to reduce the TCO of storage, most datacenters will need to implement some form of information management program to control the cost of storage.

Complexity and waste cannot help but contribute to an increase in the TCO of the datacenter, due to excess costs for maintenance, systems management, floor space, and energy. The IT staff has to change the architectural infrastructure of the datacenter. They have to change the IT paradigm by removing complexity and duplication from the datacenter through the deployment of scalable, virtualized storage and, especially, the simplification of the backup and recovery scenario, reducing the number of duplicate files that are being preserved.

As data ages, the urgency to access it decreases. *Mission-critical* applications typically have a requirement for instantaneous access to storage. Therefore, the IT staff tends to store this information on the fastest, most reliable media available, typically 15K RPM Fibre Channel (FC) drives. As you might suspect, this is also the most expensive media available in open systems storage arrays. The data for *business-critical* applications, such as email and data warehouse, does not necessarily have to be on Tier-1 drives. This secondary data can be on Tier-2

drives, such as FC drives at 10K RPM. Tier-3 data, such as backups, can clearly utilize the high capacity, low-cost characteristics of SATA drives. However, due to the long-term nature of this data, there is an overwhelming need to control the amount of duplicate data being replicated and the energy required to maintain a responsive backup architecture. Furthermore, there is an absolute necessity to deploy storage with the reliability, availability, and serviceability (RAS) features demanded by the SLA policies of the enterprise.

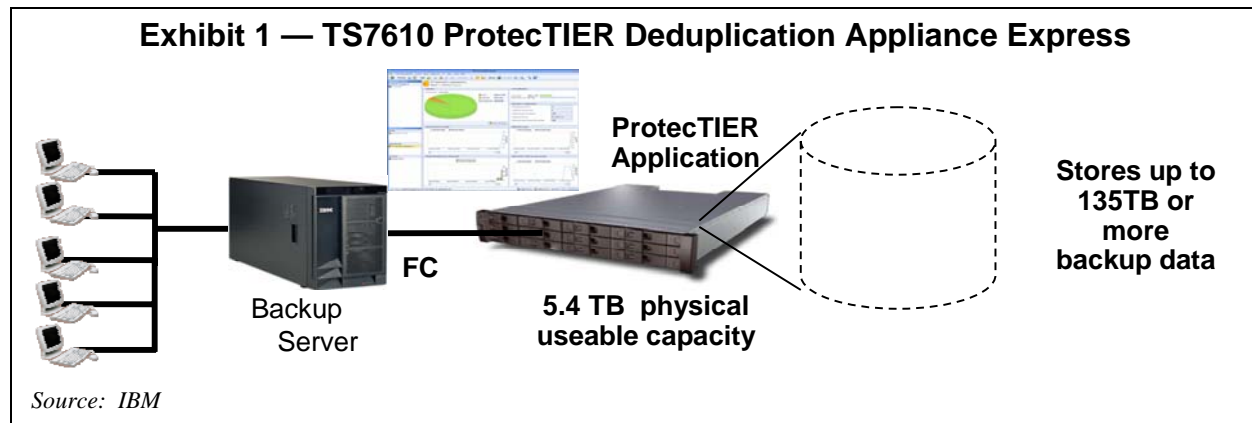
The datacenter of any mid-sized enterprise faces the same issues as those of the largest enterprise, simply on a different scale. Where the largest enterprise is supporting thousands or tens of thousands of users with hundreds of terabytes of data, *and growing*, many mid-sized enterprises may *only* be supporting 500-to-1000 users, on possibly more than 100 servers, with 100s of terabytes of backup data, and growing. There is a third size of datacenter, however, with requirements that pale in comparison. It may have less than 100TBs of backup data and still be running out of Tier-2 storage. Unfortunately, this smaller datacenter largely has been ignored while data deduplication vendors search out the largest enterprises to capitalize on those with the biggest pains.

In order to satisfy the demands for these larger datacenters, IBM delivered the *ProtectTIER TS7650G Gateway*¹, with capacity up to 1PB, and the *ProtectTIER TS7650 Appliance*, with a capacity from 7TBs to 36TBs for Tier-2 data. These are high-performance solutions that combine IBM's energy efficient storage technology with their unique data deduplication technology acquired from the acquisition of Diligent Technologies. IBM has now taken their ProtectTIER architecture one step further, and downward, to address the needs of smaller environments with the same innovative technology used in the largest enterprises. IBM has announced the *ProtectTIER TS7610 Appliance Express*, aimed at businesses with smaller, and less expensive, storage needs.

IBM's ProtectTIER Technology

In order to reduce data proliferation, many enterprises are now deploying data deduplica-

¹ See [The Clipper Group Navigator](http://www.clipper.com/research/TCG2009008.pdf) entitled *Reversing the Requirement for Storage Growth – IBM Consolidates and Simplifies Tier-2 Storage* dated February 25, 2009, at <http://www.clipper.com/research/TCG2009008.pdf>.



tion software to eliminate redundant copies of the same data, reducing it to a single instance. In the deduplication process, duplicate copies of the same data are deleted, leaving only one instance to be stored. However, indexing of all data is still retained in the event that data needs to be recovered. Deduplication can reduce the required storage capacity significantly, since only unique data is stored. For instance, if an email system containing 100 instances of the same 1MB attachment is being backed up, all 100 instances would be saved, requiring 100 MBs of disk capacity. With data deduplication installed, only one instance of the attachment is actually stored. Now, each additional occurrence just references the one saved copy. In this case, a 100MB storage requirement would be reduced to just 1MB. Now, imagine the amount of storage required if the datacenter is doing a full backup weekly. **The 100MB storage requirement would quickly expand to 5,200 MBs!** With data deduplication deployed, only 1MB of capacity would be consumed. This degree of savings cannot be assumed for all storage environments. The IT staff needs to examine their stored data closely to determine if data deduplication can provide them with similar savings.

If your backups do contain duplicate data, then it is easy to see the benefits that ProtecTIER data deduplication can provide in terms of reduced storage requirements, improved SLAs, and reduced TCO. With reduced acquisition costs – and reduced power, space, and cooling requirements – a D2D environment becomes more suitable for primary backup and restore and for retention that can easily extend to months, or years. With backup data on disk, the recovery SLA is higher, tape media handling errors are eliminated, and mul-

iple recovery points become more attractive. Data deduplication also reduces the amount of data that must be transmitted across the network for remote backups, replication, and disaster recovery.

ProtecTIER Data deduplication is an innovative technology that can deliver the storage consolidation and, therefore, the economic advantages of reducing the TCO of the storage infrastructure with industry leading performance, capacity, data integrity, and scalability. For enterprises backing up large amounts (more than 20TBs nightly), the IT staff can choose the ProtecTIER Gateway or acquire data deduplication software from an ISV and integrate it, along with an open systems server and storage management software, themselves. However, now, IBM is making enterprise-class data deduplication affordable for smaller environments, with the IBM *TS7610 ProtecTIER Deduplication Appliance Express*, enabling customers to hit the ground running.

ProtecTIER TS7610 to the Rescue

With the TS7650G ProtecTIER Gateway addressing the largest enterprise datacenters with up to 1PB of usable capacity, and the TS7650 ProtecTIER Appliance serving the mid-sized datacenter with between 7TBs and 36TBs of usable capacity, IBM has now turned its attention to the smaller business or agency and branch offices and departmental requirements of larger enterprises. With the TS7610 ProtecTIER Appliance Express, IBM is well positioned to provide these smaller environments with the same high-end data deduplication technology that has made its big brothers so successful in larger enterprise accounts around the globe. (See Exhibit 1, above.) Clients needing to do full backups of up to 3TBs are prime candidates for the

TS7610 ProtecTIER, the only enterprise-proven deduplication solution available for the mid-market customer.

The TS7610 is a complete plug-n-play solution to enable a rapid deployment by the IT staff. It comes with an easy-to-use graphical user interface (GUI) to simplify a set-up in less than two hours and to facilitate management processes. With an entry configuration at 4TBs of usable capacity, the TS7610 can handle the backup requirements of an enterprise storing 40TBs of data with a deduplication ratio of 10:1. Obviously, if your backup data has more duplication and the IT staff does more frequent backups, you could achieve a ratio of 25:1, reclaiming up to 100TBs of storage. Moreover, with a throughput of 80MB/second for backup performance (in real-world conditions, with no caveats), you can shrink your backup window, returning your systems to mission- and business-critical applications use quicker, while lowering your TCO. In fact, restoring data is even faster than backup. In addition, the TS7610 is easy for the IT staff to upgrade, on-site, scalable up to 5.4TBs of usable capacity, with the transmission of an upgrade key, storing up to 135TBs or more of backup data. With a street price less than \$53K, the 5.4TB model is significantly less than its competition. Moreover, with an entry configuration of 4TBs, the TS7610 is available for only about \$42K, putting it into the right ballpark for the mid-market user.

The IBM TS7610 deduplication appliance uses the ProtecTIER inline deduplication technology, so there is no post-processing burden to consider. ProtecTIER is designed for 100% data integrity, using a bit-for-bit comparison, eliminating the possibility of a hash collision. The TS7610 uses all quality IBM components with a high reliability, availability, and serviceability (RAS) record to ensure that your backup server is ready when needed.

The TS7610 building block is a 3U enclosure that fits in a standard 19" rack, using a standard quad-core 2.33GHz Intel *Xeon* processor. It supports twelve RAID protected 1TB SATA drives formatted to RAID-6. Not only does the TS7610 make an ideal deduplication engine for smaller environments, it also is well positioned for the branch office to leverage its many-to-one replication (for remote office data protection), connecting back to a TS7650 Appliance or TS760G Gateway in the main datacenter in a hub-to-spoke manner back.

Conclusion

With data deduplication, the smaller environments or remote offices of a larger enterprise can store up to 25 times more data on disk; however, for all of the obvious reasons, your mileage may vary! With deduplication in place, the IT staff can increase the data-retention period by storing more backup data on disk at a lower unit cost. The IT staff can reduce backup and restore times with inline data deduplication while at the same time improving the reliability of backup operations by minimizing failures. The IT staff can also drive down the TCO of disk-based backup in terms of energy, cooling, and floor space.

With the TS7610 ProtecTIER Appliance Express, IBM has enabled enterprise-quality data deduplication for the mid-market. With better reliability and a faster recovery time than tape, the TS7610 provides affordable data backup and recovery for both the mid-market and remote offices. It is significantly better than non-deduplicated disk, storing more data on less disk, while consuming fewer energy, cooling, and space resources. With its simplified GUI, the TS7610 enables an improved management for backup without having to implement radical changes. Moreover, IBM provides all of this at a very affordable price. If you are concerned about the time it takes you to backup and recover with your existing infrastructure, looking to update your backup and recovery practices for your remote offices, or if you're dealing with uncontrollable data growth, you should look into IBM's TS7610 ProtecTIER solutions.



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 a Senior Contributing Analyst for The Clipper Group. Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. In 2002, 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

The Clipper Group Navigator, The Clipper Group Explorer, The Clipper Group Observer, The Clipper Group Captain's Log, The Clipper Group Voyager, Clipper Notes, and “*clipper.com*” are trademarks of The Clipper Group, Inc., and the clipper ship drawings, “*Navigating Information Technology Horizons*”, and “*teraproductivity*” are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosures

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.

After publication of a bulletin on *clipper.com*, The Clipper Group offers all vendors and users the opportunity to license its publications for a fee, since linking to Clipper's web pages, posting of Clipper documents on other's websites, and printing of hard-copy reprints is not allowed without payment of related fee(s). Less than half of our publications are licensed in this way. In addition, analysts regularly receive briefings from many vendors. Occasionally, Clipper analysts' travel and/or lodging expenses and/or conference fees have been subsidized by a vendor, in order to participate in briefings. The Clipper Group does not charge any professional fees to participate in these information-gathering events. In addition, some vendors sometime provide binders, USB drives containing presentations, and other conference-related paraphernalia to Clipper's analysts.

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.