THE CLIPPER GROUP
Navigator ™

Navigating Information Technology Horizons^{SM}

IBM's Data Retention Appliance —
The Case for *Just-in-Case*

Analyst: David Reine

Management Summary

When the average consumer goes looking to purchase a new stereo system for their teen or upgrade their home audio system, there are a myriad of options available. A young teenager receiving their first system might be overjoyed with one of the integrated micro-stereo systems from RCA or Panasonic, inexpensive and available at a local department store. Perhaps the doting parent will step up a level and purchase a rack system available from companies such as JVC or Sony, a little more expensive, but available from a variety of specialty stores, such as Tweeter or Circuit City, who can provide all of the guidance that you need. The self-professed audiophile, however, is not content with pre-packaged systems. This consumer will do all the research, check out consumer reports, and select the very best components: amplifier, CD player, graphic equalizer, receiver, speakers, etc. The audiophile will accept responsibility of putting all of the pieces together in order to get the sound quality that he desires, regardless of the expense.

Data center managers, however, are not willing to accept the responsibility, and the risk, of putting together their own mission-critical IT solution. The cost of selecting the best components, integrating and testing them, and maintaining them, is too high. Limitations on staffing and budgets rule. In 2006, lowering the total cost of ownership of any IT solution is a priority. With the explosion in data creation that we have seen recently, some from growth but a lot from compliance with regulations and legislation, backup and retrieval and disaster recovery has moved from being a problem to being a … disaster! CIOs in all enterprises are looking for the integrated solution that will slide into their IT architecture with maximum performance and minimum impact on operations procedures. A cottage industry in data retention appliances has sprouted, with vendors looking to integrate state-of-the-art components into a seamless solution. Unfortunately, when you put together a heterogeneous mix of a commodity x 86 servers with a commodity disk array with a commodity tape library, costs for the solution rise as development and testing processes mount. One company, however, has assembled a set of components from their own catalog, delivering a high-performance, low-cost data retention solution.

With access to servers, disk arrays and tape libraries from divisions within their own company, IBM has upgraded their existing data retention offering, the IBM *System Storage DR550*, with the latest in server architecture to provide the highest performance and reliability, and the latest in disk array technology, to provide scalability levels to satisfy the growth demands of the largest enterprise. To see if the DR550 can satisfy your needs, and budget, please read on.

The Data Retention Quagmire

For many enterprises, stored data has been doubling in size annually and there is no end in sight. Enterprises have always been storing, backing-up, and archiving their mission-critical data. This is the lifeblood of any business. If a disk drive fails, or human error deletes a vital file, the data center staff must be able to recover that information, re-trieving it quickly. Today, however, there is a new category of information that enter-prises of all sizes are required to save: com-pliance data. This is the information that corporate officers send and receive – email and instant messages, financial reports, con-tracts, etc. – and must be saved for years, or even decades. While initially this informa-tion may be considered part of the mission-critical database, after some variable length of time, it must be saved for no other reason than to keep the CxOs out of trouble. **Some of this information will be written and rarely, or never, read**[1]. This includes half of the client data that is recorded. It is pre-served *just-in-case* (JIC) a court or legal authority decides that they need to see it. It must be backed-up along with all of the mission-critical data that is used every day; however, it does not require a mission-criti-cal storage tier. The enterprise must guaran-tee, and prove, that all of this JIC data has been saved to a non-erasable, non-rewritable media, to ensure that there are no "acci-dents". Quite clearly, tape qualifies in this category, and in addition, will reduce the total cost of ownership (TCO) for IT as it will outlive disk storage. **We used to say that if it is worth saving, it is worth pro-tecting. Today, enterprise executives and board members are seriously involved with protecting themselves, as well**.

Recent innovations in disk architecture have lowered the acquisition cost per giga-byte for storage. In response, many data centers have adopted a disk-to-disk (D2D) strategy for snapshots or the *short-term* back-up and retrieval of mission-critical data. This

has been a boon to those data centers where the length of time to do a backup exceeds the size of the backup window, causing a serious dilemma for the CIO when mission-critical applications must be delayed until the backup is completed and additional shifts must be added to the data center staffing woes. Unfortunately, there are other costs involved besides simply acquisition and staffing. The cost to keep the disks spinning and the cost to cool the data center contribute to the TCO for any data retention solution. With com-pliance data, we are talking about spinning that data for, possibly, 7 to 10 years, or longer.[2] In this environment, we see archive to tape as an ideal solution. Reports of the demise of tape appear to be premature. The requirement for the retention of WORR or WORN data has given new life to tape as we can see with the recent investments being made by IBM, and other companies, in the development of higher performance, higher capacity tape drives, in both the mainframe and commodity arenas[3]. The requirement today is for cartridge capacities matching the capacity of the largest disks, and compati-bility with even higher capacity cartridges in the years to come. Furthermore, there is a need for any new tape solution to be introduced with WORM and encryption capabilities to ensure the authenticity of the data and to secure the contents from un-authorized access.

Unfortunately, the long-term retention requirements for compliance data mean that it needs to be managed, from both a media and software standpoint. Retention data will absolutely outlive the media to which it is initially written. Any successful data reten-tion solution must be able to migrate the data and manage the milestone dates on which the compliance data may, or has to be, destroyed. These milestones need to be managed to the data retention policies instituted for any particular enterprise or industry.

---

[1] Rather than WORM (write once, read many) this data is WORR (write once, read rarely) or WORN (write once, read never).

[2] In the case of health records, some states have amended HIPAA regulations to require retention for life plus 6 years, which could extend the period to 50-60 years.

[3] See **The Clipper Group Navigator** dated October 11, 2003, entitled *Tape Virtualization in the Enterprise – Reducing Data Center TCO*, available at http://www.clipper.com/research/TCG2005062R.pdf.

Finding a single appliance that meets all of these criteria is not easy. Most companies with data retention solutions have acquired and integrated commodity platforms and management software from different vendors and attempted to integrate them into an appliance. The costs associated with acquisition, test, integration, and support tend to drive up the acquisition price and complexity of the solution. Few companies, however, have the breadth of product set to be able to do it internally. IBM is one company that has proven it can.

IBM's Data Retention Commitment

IBM has been improving the information lifecycle management process (ILM) for enterprises of all sizes, delivering multi-tiered data retention platforms for the past three years, starting with the announcement of the *DR450* in 2004. Designed to employ a best practices philosophy to protect data, the DR450 integrated existing state-of-the-art hardware and software components, the *POWER4*-based *pSeries 615[4], TotalStorage FAStT600* disk array, and the *Tivoli Storage Manager for Data Retention,* to introduce compliance storage in a box. With disk capacity of up 56TB, IBM provided an appliance capable of managing data for an extended period of time, especially with the availability of *LTO* and *3592* tape drives and libraries.

IBM followed this announcement with a pair of additional products, the *DR550* later in 2004 and *DR550 Express* in 2005, based on IBM's newest *POWER5* architecture, capable of dual-nodes, along with a *Total-Storage DS4100* array[5]. In addition to improving performance, the DR550 extended the storage capacity by 60%, to 89.6TB. This does not even address the improvements in third tier storage provided by the introduction of the *TS1120* tape drive in the Fall of 2005. With a native capacity of 500 GB and a throughput of 100MB/s, the

---

[4] IBM chose to use the high performance POWER-based platform over the commodity based Intel or AMD platforms available in their xSeries, in order to meet the performance requirements of their base.

[5] Based on the Engenio architecture.

TS1120 contributes to a further simplification of the data retention process through cartridge consolidation. The cost/GB for tape continues to be less expensive than disk.

DR550 V2 reduced the entry price for data retention from $140K to $100K for a 3TB system. Apparently, someone at IBM realized that, while competitive and less costly than other solutions, this was still pretty steep for many small and medium-sized enterprises (SMEs), so last summer they announced a single-node version, the *DR550 Express*. Still using the same basic components, IBM restricted the scalability of the solution to 1.1 TB of internal storage, and priced an entry Express solution at $33K. The Express version did not support real-time synchronization with *Metro Mirror[6]*; however, this is a function that few SMEs require.

It is significant to note that IBM should have a price advantage in the data retention space as all of the components come out of their product catalog. While the IBM pServer group has to fund POWER R&D, the tape storage unit funds development of both commodity and mainframe tape architecture, and the disk storage team funds the work with Engenio, the DR550 team can reap the fruits of those investments, riding the coattails of an already in place support team. Those investments have now led IBM to the third generation of data retention solutions.

The DR550 V3.0 Solution

The *DR550 V3.0* is the third generation data retention architecture to come out of IBM, with improvements made to hardware, software, and both implementation and professional services.

Hardware

Once again taking advantage of investments made in sister divisions, the DR550 server component has been upgraded, this time to include the *POWER5+*. With the *p5 Model 52A* server as the core of the new DR550, the data center can utilize the single

---

[6] Metro Mirror requires a DS4000 storage array.

or dual server configurations of the enterprise model. With dual 1.9GHz processors and 36MB of L3 cache in each, the enterprise can take advantage of the ultra-high frequency, multi-core design to improve the overall performance of the appliance for less than $100K[7]. With storage requirements doubling, the DR550 Express, with a 1.5 GHz CPU, has been enabled to expand from 1.1TB of internal Ultra 320 SCSI storage, at $27.6K, to 3.9 or 6.7TB, with the addition of SATA drawers[8]. In order to improve both the logical and physical security of the installation, both models have been made tape-ready for an optional WORM drive, either TS1120 or LTO-3, and provided with a lockable cabinet, front and rear.

In addition, the DR550 has been made flexible enough to attach a wide range of tape libraries to enable a "lights-out" environment, both IBM and commodity, from any number of vendors, from ADIC to Sun (formerly STK), in order to provide a complete archive solution. This capability distinguishes IBM from competitors and enables the data center to lower the platform TCO through the following.

- Backup of the DR550 operational files;
- Migration path for DR550 controlled data; and
- Scalability to petabytes.

Software

The DR550 brings a wide range of mission-critical data retention capabilities to the data center with the ability to establish a full set of data retention policies, including:

- Default and custom policies;
- Terminate based on a predetermined expiration date;
- Terminate based on a retention event;
- Terminate on demand
- Protect forever; and
- Use of a deletion hold and release for audits and legal events.

---

[7] A single server version is available for less than $75K.

[8] A fully configured 6.7TB Express model is priced at $69.1K.

Exhibit 1 – DR550 Features

- **Event and Policy-based data retention**;
- **Data encryption**;
- **Disaster recovery**;
- **Tiered storage management**;
- **IBM System Storage Archive Clien**t;
- **Management of attached heterogeneous storage devices**;
- **Metro Mirror** for synchronous data replication between sites (N/A on Express models); and
- **Global Mirror–Preserved Write Order**, to match I/O completion order on the local subsystem (N/A on Express models).

Source: IBM

A second major feature of the DR550 is data encryption. All models support encryption of the data by the application along with key management. In addition, the DR550 also provides the capability for the *System Storage Archive Manager (SSAM)[9]* client to do the key management, with the DR550 client encrypting the data "in-flight" according to 128 AES/56 DES encryption standards. The DR550 client also decrypts the data upon read when the key is provided. In addition, the SSAM client permits users to archive files automatically from their workstations or file servers to data retention protected storage, and to retrieve archived files to their workstation or file server. A full set of DR550 features are listed in Exhibit 1.

Services

The third category is services, and here IBM shines with an established set of professional and remedial services and a worldwide force ready to implement. IBM refers to the DR550 optional services as "carefree" as they remove the worries that keep most data center managers up at night. The services bundle includes the following.

---

[9] SSAM is a Data Retention variant of the *Tivoli Storage Manager.*

- Three year warranty maintenance service for all DR550 components, including IBM onsite repair[10];
- Onsite software/firmware upgrades for up to two onsite visits per year, including firmware and software component upgrades and patches to provide proactive maintenance to ensure DR550 system concurrency;
- IBM RAID conversion services to provide conversion assistance from the default RAID 5 parity protection configuration to an enhanced disaster recovery protection configuration with mirrored RAID 10;
- IBM implementation services to provide assistance in planning, installing, and configuring new data retention systems or capacity upgrades;
- IBM implementation services for Metro or Global Mirroring; and
- IBM MES services to provide options or upgrades for existing installations.

Conclusion

So, how does IBM's offering differ from any or all of the commodity, integrated solutions that are available from every VAR or ISV knocking on the CIO's door? IBM has attempted to respond to the business requirements for which their customers, and others, have been clamoring. **First and foremost among these is reducing the TCO of the data center.** By basing their data retention family on the POWER server family, IBM has established a solid foundation for price/performance with an architecture that leads the industry in many of the commodity benchmark tests that are used to rate server performance. Using the p5 server today guarantees upward mobility for the data center and protection of the investment that they make in today's architecture, a low investment at that, due to the reuse of existing platforms. The implementation of an Express solution provides an answer for the vast SME market, previously largely ignored. In addition, by using a multi-tiered solution,

integrating disk and tape, IBM reduces the long-term costs of preserving compliance data. **Further, the CIO must remember that the TCO for data retention is not over three years or five years; it is for the life of the data, and that data will outlive the media on which it resides.** It is vital that the enterprise solution provider be a partner for the long haul.

Secondly, **IBM has addressed a major end-user concern of information retrieval performance**. By combining high-performance microprocessor technology with high-performance and high-capacity tape technology, IBM can retrieve more information faster from fewer cartridges than other leading vendors. Not only does this allow for a faster end-user response, but it also minimizes the staffing requirements in the data center and reduces the number of cartridges required, thus reducing expensive data center floor space.

Third, **IBM's DR550 effectively addresses the data center issues surrounding management of the retention data**. Using the Systems Storage Archive Manager solution, the data center can easily manage the defined retention periods and also manage date with undefined retention periods. Along with the DR550's replication capabilities through Metro and Global Mirroring, the DR550 can also protect the security of the information through data encryption, saving the enterprise from possible embarrassment and expensive corrective actions.

Whether you are dealing with files, databases, or images, the enterprise is saving data for a very long time. It is important that you have a solution, and a partner, that you can trust just-in-case you need to retrieve that data. If your enterprise is looking for an affordable solution from a trusted vendor, you should check out the DR550 and the DR550 Express. One, or both, may be the right solution for your enterprise.

---

[10] Three-year warranty maintenance is also available as a separate offering.

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

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group** *Captain's Log*, **The Clipper Group Voyager**, 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.

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.