Find Relief on the Treadmill of Information Storage – Solutions for SMEs

Analyst: Michael Fisch

Management Summary

Information storage is a treadmill of ever-increasing speed. Enterprise data continues to grow with no signs of abating. This important asset fuels the engines of commerce, and enterprises must ensure their data is highly available, protected from loss or corruption, and secure from unscrupulous persons who might want to exploit it. Good corporate governance and even laws demand that it also be archived according to policies and readily retrieved. Meanwhile, budgets are limited. Keeping up with these storage requirements is like running on a treadmill that gets incrementally faster. If you do not keep pace, you are liable to tumble off the back!

Small- and mid-sized enterprises (SMEs) feel this pressure to keep up and adapt every bit as much as large enterprises, perhaps more so since they traditionally have fewer resources to work with. A challenge for SMEs is that the *latest and greatest* storage solutions are generally available at the high end first – because it takes time for new technologies to mature, disperse into the market, and eventually be redesigned and repackaged appropriately for SMEs. Without the big IT budgets and specialized staff of large enterprises, SMEs have even greater requirements for affordability, ease of use, and the ability to scale solutions incrementally as their environments evolve. Last, but not least, they need to be able to leverage existing investments in servers, networking, and storage as they adapt to more demanding requirements. Wholesale "rip and replace" is neither feasible nor realistic.

The good news is that several important storage technologies continue to be enhanced and enriched for the midrange market:

- **Consolidated, networked storage**
- **Storage management**
- **Data protection and disaster recovery**
- **Data archiving**

Their benefits include easier management, lower storage TCO, faster disaster recovery, and intelligent, systematic information lifecycle management (ILM).

Technology adoption is a matter of risk versus reward: While the increasing requirements have raised the risk of not taking action, the maturity and enhancement of these technologies have sweetened the reward. **Therefore, if you are an SME gasping for air on the treadmill of information storage, consider how these storage technologies might benefit your business.** Read on for details.

Assess the Situation

Your objective is to keep up with the information storage requirements. (No one really has a choice here – business depends on it.) The first step is to assess your enterprise's situation, to survey the storage environment. Walk through the data center and note the computer, network, and storage equipment. Map out the infrastructure. Talk to the busy IT administrators about their roles, procedures, and challenges. Ask users about their experiences using IT applications on PCs and workstations. **In other words, consider the whole operation from end to end – how the enterprise stores, manages, and protects its vital information**. This will help answer some key questions:

- Which storage technologies does the enterprise currently employ?

- How well are they meeting business requirements, in terms of performance, availability, recoverability, and cost?

- Which storage challenges does the enterprise most need to address?

If your IT operation is of any size, you will find many processes and technologies involved in keeping it going. Backups and restores, provisioning capacity, growth planning, data migration, adding applications and servers, etc. are all part of the routine. You will also find problems to solve. No matter how sophisticated the operation, skilled the administrators, or large the budget, it is just the nature of the information treadmill. Your enterprise will likely relate to one or more of these common storage challenges.

- Keeping up with data growth without overspending on hardware
- Containing IT staffing levels and operating costs
- Disaster and operational recovery
- Insufficient application performance
- Archiving data for corporate governance and regulatory compliance
- Migrating data without disrupting users

Develop a Plan

Equipped with a sense of your enterprise's state of affairs and a prioritized list of storage needs/challenges, the next step is to develop a plan to address them, including which technology solutions to implement, when, and how. Due to budget and time restrictions, a phased rollout often makes the most sense. Start with projects that address the highest-priority challenges – the low-hanging fruit.

As you build this vision and plan, keep your existing IT assets and capabilities in mind. You may be able to leverage them to simplify and accelerate the deployment of new technology solutions. For instance:

- An IP network could also be used for networked storage, such as a NAS platform or an IP SAN.

- Direct-attached storage may be turned into networked storage via SAN attachment or a NAS gateway.

- Be realistic about technology upgrades; you probably would not want to replace an asset that was recently deployed.

Furthermore, you may choose to enlist outside help for assessment and planning, such as a third-party provider of professional services.

Storage Solutions for SMEs

Below we describe various storage technology options and benefits. Use this as an overview and starting point for researching solution area(s) to satisfy your storage challenges.

Consolidated, Networked Storage

The foremost solution for cost-effectively handling data growth is storage consolidation. Rather than discrete islands, consolidation separates storage from individual servers and establishes a centralized, shared pool. This is more efficient for similar reasons that carpooling or public transportation is more efficient for commuting – utilization rates are higher and operating costs are lower, which means lower total cost of ownership (TCO). Consolidated storage is easier to scale and manage because it is unified and more flexible. Each administrator can effectively handle more capacity, so staffing levels do not have to rise in proportion with data growth. Better utilization and less overhead capacity also contribute to lower environmental costs (i.e., floor space, power, and cooling).

There are three primary architectures to consider:

- Direct attached Storage (DAS)
- Storage area networks (SAN)
- Network attached storage (NAS)

DAS is the most-familiar legacy environment. When one bought a server (host), they received embedded or directly-attached storage. It was simple to acquire but difficult to optimize, as Murphy's Law would attribute – you either had too much or too little storage available to the associated applications. In addition, what made matters worse, you could not share or tap into excess storage on a neighboring platform.

For that reason, most enterprises choose SAN and/or NAS for storage consolidation because of their scalability, extended connectivity, and technological maturity. Networked storage solutions of all sizes are available now, not just ones for larger IT operations, as was the case several years ago. Furthermore, SAN and NAS are complementary rather than competitive, and many operations deploy both to meet a broad spectrum of requirements (i.e., file and block services).[1]

Storage Management

While consolidation simplifies the storage itself, storage management is a tool for simplifying the management process, which is in fact the largest component of storage TCO. It scales the impact of IT administrators by allowing them to do more with less effort. A storage infrastructure has many components that must interface and work together: files systems, volumes, host bus adapters, cables, switches, routers, storage arrays, RAID groups, LUNs, tape drives and libraries, backup servers, and, of course, the data itself. Add in heterogeneous devices, interoperability issues, geographic complexities – and it is a lot to coordinate. Using individual device managers for every piece of the puzzle can be dizzying and costly. To address this challenge, storage management software provides a central point for monitoring, reporting, configuring, mapping, automating, and optimizing. By *automation*, we mean taking action based on user-

defined policies as well as wizards that walk through a setup or management activity.

Though product features vary, general categories include the following.

Storage Resource Management (SRM)

Manages and reports on data and storage capacity resources. Storage resource management (SRM) software looks at data from the perspective of servers and clients and answers the question, "What data is out there and where is it?" It scans file systems and databases, tracks data by type and amount, and often correlates it to specific storage devices. It gives a mountaintop view of the data in the enterprise.

This is very useful for capacity planning, asset management, quota management, and charge-back to users and departments. For instance, SRM finds inactive or duplicate data for deletion or migration to archives, which frees up space without additional storage purchases. It finds unused and forgotten capacity for redeployment. It also identifies storage arrays, volumes, and file systems that are approaching full capacity and should be expanded to avoid downtime due to an out-of-space condition. It works in SAN, NAS, and DAS environments.

SAN management

SAN management tools monitor, map, and configure devices in a storage area network, such as switches, host bus adapters, and storage arrays. They provide LUN masking and zoning. Tools that are more sophisticated can automate the end-to-end process of provisioning capacity, from the storage array through the network to the server and application. Other tools can also analyze performance and help administrators find and correct bottlenecks in the data path that affect application performance.

Data Protection and Disaster Recovery

When data corruption, a system failure, or disaster occurs, the only question that matters is how soon and how completely can the business recover? If the answer is unsatisfactory, it is better to correct the situation sooner than risk unacceptable downtime or, eventually, data loss. The answer lays in effective data protection and disaster recovery measures.

---

[1] See **The Clipper Group Explorer** dated November 18, 2005, entitled *Networked Storage – Evaluation Criteria*, at http://www.clipper.com/research/TCG2005074.pdf.

Backup

Backup is the most common means of protecting data, though it is not without its challenges. Increasing amounts of data reside on heterogeneous servers, PCs, and laptops that number in the tens, hundreds, or thousands – which backup applications and procedures must accommodate. Improved support for heterogeneous environments, automatic agent "pushing", and options for desktop/laptop backup are quite helpful here.

With backup in mind, a major trend is the use of low-cost disk arrays as backup targets because of their advantages over tape in speed and media reliability. The high bandwidth of a disk array and its ability to read and write random, intermittent streams of data make it favorable for fast backups and restores. This helps operations cope with strained backup windows and increasingly stringent recoverability objectives. While both tape and disk media can fail, a RAID array inherently stores data redundantly and knows immediately when a failure occurs. With tapes, one does not necessarily know a media failure exists until the tape is needed for restore, which is not a good time to find out. It also requires physical handling of the media. However, tape still delivers the lowest cost per unit of capacity, especially when storing very large quantities of data. Therefore, many advocate a combination of disk and tape for backup. Initial backups are stored on disk for quick recovery, and tape is used for long-term and, possibly, remote storage.

A technology gaining popularity for its robustness and ease of deployment is virtual tape libraries. These are systems that employ disk-based storage, but present themselves to popular backup applications as if they were tape drives. The benefit of this approach is no need to change backup software or procedures and no retraining of personnel. At the same time, you gain the benefits of using disk-based media for very fast recovery, and the ability to do it on a very granular basis. These products also facilitate the final production of tape for offsite and archive purposes.

Snapshot integration

Snapshot copies are useful for non-disruptive backup as well as fast restores from a recent point in time. A snapshot or point-in-time copy takes a "snapshot" of data on disk at an instance in time. It can be a full copy or a space-saving differential copy. Making a copy is fast and, if properly integrated, non-disruptive to the application. The backup system can then use it as the source for a backup job without disrupting user access to the primary data. This means that backups may run during business hours, providing more flexibility for backup windows. Snapshots are also useful for fast recoveries from logical faults, such as data corruption or accidentally deleted files, and data repurposing.

Remote replication

Remote replication maintains a complete physical copy at a remote site in real time. As the source data changes, so does the target. Like a spare tire, if the source data becomes unavailable for some reason, a completely current (synchronous) or nearly current (asynchronous) copy is available to resume operations. It protects from local system failures or disasters, such as fires, floods, or electricity outages. It can also be used for data copying and migration without disrupting users and applications.

Data Archiving

Data retention or archiving policies are a necessary aspect of corporate governance and regulatory requirements. Sarbanes-Oxley, USA Patriot Act, HIPPA, SEC 17a-4, and other regulations have brought it to the forefront of the corporate conscious. They require archiving data for specified periods, on penalty of serious fines and legal liability. This applies to electronic records - such as e-mail, instant messages, documents, transactions, and images. However, compliance is not the only reason to archive. Databases and file systems are like cargo ships in that they slow and become less stable when overloaded (i.e., too large). Archiving prunes data from primary storage and migrates it secondary tier (disk or tape for long-term archive), which improves performance (maybe enough to avoid a server or storage upgrade) and lowers total storage cost. A streamlined database or file system also speeds up backups, restores, replication, and data migration. Timely search and retrieval is an important aspect of archiving, so rudimentary backup is insufficient for this purpose.

Archiving solutions are now widely available to manage the data migration, retention, and retrieval process. They use descriptive information about data (i.e., metadata) to categorize and move it between storage tiers based on policy. For instance, a policy might require that files not accessed in six months move to secondary storage for long-term archiving. These solutions maintain a link or association between the application and migrated data, so users can still search and access it. Some can even delete data after a specified period, thereby managing the cradle-to-grave lifecycle. Archiving solutions today apply to specific applications, such as e-mail, databases, or file systems. A universal archiving solution that encompasses all forms of data is not (yet) available, so enterprises have to prioritize and decide which data to manage in this way.

Conclusion

Don't let the treadmill of information storage wear you out. Advanced storage technologies that can get small- and mid-sized enterprise up to speed are available. Consolidated, networked storage, storage management, data protection and disaster recovery, and data archiving for SMEs have matured over the last few years. There are now solutions with lower prices, easier setup and management, and modular scalability.

So, what now? Spend some time assessing your enterprise's situation: Identify your strengths, assets/resources, and challenges. Leverage whatever you can from your current storage infrastructure, and lay out an executable plan to upgrade your environment to keep pace with the increasing pressures of growing your business.

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

Michael Fisch is Director of Storage and Networking for The Clipper Group. He brings over ten years of experience in the computer industry working in sales, market analysis and positioning, and engineering. Mr. Fisch worked at EMC Corporation as a marketing program manager focused on service providers and as a competitive market analyst. Before that, he worked in international channel development, manufacturing, and technical support at Extended Systems, Inc. Mr. Fisch earned an MBA from Babson College and a Bachelor's degree in electrical engineering from the University of Idaho.

➢ *Reach Michael Fisch via e-mail at mike.fisch@clipper.com or at 781-235-0085 Ext. 211. (Please dial "211" 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.