Clipper Notes™

VENDOR-AGNOSTIC EXPLANATIONS AND ADVICE
FOR THE INFORMATION TECHNOLOGY BUYER

Navigating Information Technology Horizons℠

Networked Storage – Evaluation Criteria

Analyst: Michael Fisch

Management Summary

Fast data growth and increasing reliance on IT have left many businesses with an over-burdened storage infrastructure. It may function, but not optimally, and a business assumes a number of risks and costs by continuing with the status quo. Consider the following.

- *Do you often worry that your data is not sufficiently protected from disaster or data corruption? Could you recover? In a complete and timely fashion?*
- *Do you know the status of your backup operations? Are backup jobs inconsistent? Is the backup window too short?*
- *Is data growth becoming harder to manage?*
- *Does storage management consume too much of your administrators' valuable time?*
- *Do new applications take too long to roll out?*
- *Is it difficult to add storage capacity or migrate data between servers?*
- *Does storage downtime interfere with the business – causing lost opportunities, frustrated users, and irate managers?*
- *Do you know how well your storage is utilized? If so, is it low (e.g., below 70%)?*

In short, your current storage infrastructure may be a liability to your organization – by costing too much, by putting essential information at risk, or by negatively affecting productivity. The good news is that the problem is solvable. Networked storage combined with advanced data protection and management are an effective solution to many of the information storage challenges that businesses face.

Networked storage represents a different approach than traditional direct-attach configurations. It separates storage from servers and consolidates it on a network with broad accessibility. Storage area networks (SAN) provide block-level access for dedicated applications over Fibre Channel or IP (iSCSI) networks. Content-addressed storage (CAS) is useful for centralized, long-term, online storage of archived information. Networked-attached storage (NAS) provides file sharing over IP networks. While these technologies are optimized for different purposes, all enhance the value and impact of capabilities like data replication, backup, and centralized management. **Networked storage delivers a simpler, more cost-effective, and highly available storage infrastructure.**

While most large enterprises already enjoy the benefits of networked storage, small- and mid-sized enterprises are still making the leap. **The technology is mature, and now may be a good time to consider it, especially if you are feeling the pains of data growth.** Read on for a closer look.

IN THIS ISSUE

The Pains of DAS

Data growth in the Information Age is a constant, like gravity or the changing seasons. For businesses, this growth comes in many forms: database records, e-mail, files, images, scientific measurements, financial data, and any other type of digital object. Storage capacity requirements will be greater next year than today – the only question is how much more. More robust storage service levels are also demanded, driven by the increasing reliance on IT for business operations and communications, the fast pace and global reach of the Internet economy, and corporate governance issues, such as regulatory compliance.

Businesses have to keep up because they run on this information. It is raw fuel for commerce. Imagine what would happen to a call center if the computer systems went down because data was inaccessible. It would not field many calls! The same dynamic applies in a hospital, R&D lab, or procurement department. Productivity and access to data are inextricably linked. No matter how big or small, enterprises need to manage storage effectively.

The traditional approach is to dedicate storage to each server, whether internal or external, and scale it by adding drives or upgrading servers or storage arrays. This direct-attach storage or DAS is sufficient for a single server, or perhaps a handful of servers sharing an array, but its limitations become apparent when IT infrastructure scales. As applications and servers proliferate, they create many disconnected islands of storage. The environment becomes fragmented and complex. It requires increasingly more time for skilled administrators to perform backups, restores, capacity expansion and reallocation, data migration, monitoring, troubleshooting, and so forth. Capacity utilization is poor because of the difficulty in sharing or reallocating it. As a result, the costs of storage management (the largest component of TCO[1]) and equipment are unnecessarily high. The unused capacity still consumes floor space, cooling, and electricity – adding more cost.

To compound the problem, consistency and service levels suffer in a fragmented environment. Backup jobs may fail or not be performed frequently enough, risking data loss. System recovery in the event of a disaster, or even just restoring a file, may take hours or days. Tasks like capacity expansion and data migration are disruptive to applications. All of this has a negative effect on productivity. **In short, the traditional approach to storage is costly, complex, and disruptive, especially at large scale.**

Networked Storage Provides Relief

The industry invented *networked storage* to address the limitations of DAS. It replaces point-to-point cables with a network and unshackles storage from captivity to individual servers. In other words, this architecture separates storage from servers, consolidates it on a network (physically and/or logically), and makes it broadly accessible to multiple applications and users. It is more flexible, dynamic, and efficient. Management and data protection are simpler, utilization is higher, and TCO is lower. (*More on this later.*)

The principal categories of networked storage are SAN, CAS, and NAS, each optimized for specific purposes.

SAN

A storage area network (SAN) is a dedicated network for connecting host servers to storage devices. It provides access to a shared pool of block storage, and possibly other devices like a tape library. A SAN facilitates and speeds up functions like data replication, data migration, centralized backup and recovery, and redundant paths to storage for high availability.[2]

In a SAN, capacity from one or more storage arrays is partitioned and securely mapped to individual servers or server clusters. Applications running on the servers may be anything that traditionally uses block storage: database management systems, e-mail, ERP, CRM, business intelligence, file servers, NAS gateways, and so forth. SANs are especially well suited to applications that favor reliability

---

[1] Total cost of ownership (TCO) = acquisition cost + operating costs – residual value at end of life.

[2] See **Clipper Group Explorer**, dated January 25, 2002, entitled *Business Continuity Goes Better With SANs – The 3 R's of Resilience*, available at http://www.clipper.com/research/TCG2002003.pdf.

and speed, like transaction processing. Compared to other forms of networked storage, the advantages of a SAN include fast, tunable performance, secure partitioning, and virtually infinite scalability for storage consolidation.

The two SAN connectivity options are *Fibre Channel* and *iSCSI*. Fibre Channel is the original network protocol for SANs. It has high performance[3], low latency, and robust transmission. The emerging alternative is *iSCSI*, a protocol for accessing block storage over IP networks. Though Fibre Channel is faster, IP is deployed everywhere, and enterprises can use existing IT skills instead of learning a new technology. IP equipment also tends to cost less, especially at the entry level, though Fibre Channel has come down in price. In short: Fibre Channel for performance, iSCSI for ease of use and cost.

CAS

Content-addressed storage (CAS), or object-based storage, is a fast-emerging category for centralized, long-term, online storage of archived information. CAS addresses *fixed content* or *reference information*, the largest and fastest-growing category of enterprise data. This data plays an indispensable role in business operations and corporate governance, especially regulatory compliance. Enterprises are challenged to store it in an accessible yet cost-effective manner.

Think of CAS as a consolidated, network storage tier optimized for archiving, where the object being sought is a document, form, image, video, etc., accessed by its content rather than its stored location (as with files in a traditional file system). Each object is associated with descriptive metadata and a unique digital identifier generated by a mathematical algorithm. The identifier ensures each object remains authentic and unchanged – important for regulatory compliance. It also eliminates redundant data storage, minimizing capacity requirements. CAS overcomes the scalability limitations of traditional file systems through a flat directory structure. It also does away with the application dependencies, like mounting file systems and tracking directory structures, which interfere with broad and long-term data usability. CAS solutions can incorporate other

useful data management features like policy-based retention, WORM, and full-text searching and indexing.

NAS

Network-attached storage (NAS) provides shared file access over a network. These systems are streamlined file servers with integrated, advanced features for data and storage management. They allow multiple, heterogeneous clients (i.e., users and host servers) to access files concurrently over an IP network using standard *CIFS* (*Windows*) and *NFS* (*UNIX, Linux*) protocols.

NAS consolidates storage at the file system level, provides shared, concurrent access, and is known for ease of installation and use. NAS systems are faster than general-purpose file servers and offer advanced features, which could include clustering for high availability, point-in-time copy, remote mirroring, and backup integration. Businesses use NAS for file sharing and collaboration applications like CAD/CAM, software development, Web serving, and general file storage for users. Some also host applications on it.

The majority of NAS systems have integrated disk arrays, which combine file access and storage components in one platform. Their advantage is "plug-and-play" – ease of installation and management. An increasingly popular alternative is the NAS gateway, which provides only the file access component and connects to SAN storage on the back end. It offers a greater degree of storage consolidation through a converged SAN and NAS infrastructure. Furthermore, some businesses are turning to IP-based, all-in-one storage platforms with file access (NAS), block access (iSCSI SAN), and basic archiving capabilities.

Storage Complements

SAN, CAS, and NAS are complementary technologies, each serving different but important roles in the arena of consolidated, networked storage: SAN for fast, dedicated block-level storage; CAS for long-term, online archival; and NAS for collaboration and file sharing. **In any given situation, application and data requirements are the primary determinant of the most appropriate technology.** There is overlap between them, which leaves room for a particular solution's merits and an IT department's particular

---

[3] Fibre Channel bandwidth is 4 Gbit/s, with a lot of 2 Gbit/s still installed.

preferences. However, their respective roles are clear, by and large, and many businesses use multiple forms to satisfy their spectrum of storage requirements.

Advanced Data and Storage Management

Networked storage is also a superior foundation for deploying advanced data and storage management. These capabilities improve service levels in terms of storage performance, availability, recoverability, and cost. **Networked storage functions as an amplifier for them, magnifying their reach and performance.** Consider the following.

- **Mirroring** – Maintains a real-time copy of data on disk at a remote site for failover purposes. As the source data changes, so does the target. It functions like the spare tire in a car trunk: If the source data becomes unavailable for some reason, a current or nearly-current secondary copy is available to resume operations. A dedicated storage network can facilitate and increase the performance of the link between the two sites, enhancing the full solution.

- **Point-in-time (PIT) copy** – Creates a copy of data on disk at a specific point in time. If a remote mirror is like a window with a real-time view, a PIT copy is like a photograph that captures a scene at a specific point. This feature works in conjunction with a storage network to enable non-disruptive backup, data warehouse loading, and application testing. It is also useful for quick recovery from data corruption or accidental deletion.

- **Backup and recovery** – Protects data by periodically backing it up and keeping a copy offsite to recover from local disasters. By moving data over a dedicated, high-speed network, a SAN improves this process in many ways. Backup and restore jobs are faster and more consistent, which relieves constrained backup windows and gets systems back online more quickly. It takes backup traffic off the LAN, so it does not interfere with application performance and user access. It helps move data offsite for electronic vaulting. Moreover, a SAN makes it easier to share backup targets, such as tape libraries or ATA disk arrays for fast backups and restores, and to centrally manage the backup and restore process. In other words, a SAN lets IT administrators sleep better knowing that mission-critical data is more secure.

- **Data migration** – Moves data between systems in support of routine operations like load balancing and equipment upgrades, as well as more sophisticated information lifecycle management (ILM)[4] strategies. The traditional approach is to back up a server to tape and then restore the data to another one – a slow and disruptive process. However, a storage network speeds up this process by moving data directly between storage systems. Technologies are available that perform data migration without disrupting production applications.

- **Multi-pathing** – A network offers redundant connections between servers and storage. Load balancing and failover products use these paths to increase the performance and reliability of storage access.

- **Archiving** – Manages the process of data migration, retention, and retrieval for specific applications. Archiving solutions categorize and move data between storage tiers based on policy and maintain a links so users can still search and access the migrated data. Networked storage is ideal for establishing the tiers and migrating data between them.

- **Storage resource management (SRM)** – Gives a mountaintop view of data in the enterprise. "*What data is out there and where is it?*" SRM software scans file systems and databases, tracks data by type and amount, and often correlates it to specific storage devices. Businesses use this information to perform a variety of useful tasks, like capacity planning and expansion, archiving inactive data, load balancing, quota management, and charge-back. Networked storage facilitates the activities involving data movement and capacity expansion and reallocation.

- **Networked storage management and provisioning** – Provides a central point of management point and simplifies the end-to-end process through monitoring, map-

---

[4] See **The Clipper Group Explorer** dated May 11, 2004, entitled *Top Ten Things You Should Know About Information Lifecycle Management*, available at http://www.clipper.com/research/TCG2004041R.pdf.

ping, configuring, and automation. In a SAN infrastructure, many components must interface and work together: files systems, volumes, host bus adapters, cables, switches, routers, storage arrays, RAID groups, LUNs, and the data itself. Add in heterogeneity, interoperability issues, and distributed computing – and it is a lot to coordinate!

A Mainstream Market

The market for networked storage has grown quickly and undergone widespread adoption. In fact, it is the dominant form of storage sold. SAN and NAS in particular have been around for many years and are proven, mature technologies. New, lower-cost IP technology has emerged, and traditional Fibre Channel prices have come down. Deployment and management are easier.

If you are struggling with outdated DAS configurations, it may be time to have a close look at the benefits that networked storage can deliver. The benefits of storage consolidation are real and immediate.

Benefits to the Business

Though specific benefits vary according to a solution's architecture and features, networked storage in general offers these important categories of business benefits.

- **Minimize storage acquisition costs** – A consolidated pool of storage allows capacity to be readily shared and reallocated among servers, minimizing the amount of unused or "overhead" capacity and improving utilization. It also allows servers and storage to scale separately, so adding server capacity does not require upgrading storage, and vice-versa.

- **Protection and availability of enterprise data** – By turbo-charging the backup, protection, and recovery of enterprise data, networked storage helps keep it available to support business processes. For instance, an administrator can replace a failed application server by pointing it to the same networked volume or file system the original server used – no need for lengthy tape backups and restores. Downtime is reduced; workers keep working; and the business keeps humming.

- **Lower storage operating costs through**

simplified management – This is a major benefit because the cost of managing storage over its useful life can be several-to-many times its acquisition cost. With networked storage, common tasks like scaling capacity, migrating data, and performing backups are faster and easier, especially if using centralized management tools. This empowers each administrator to handle a much higher quantity of storage, with time-consuming tasks now automated. Furthermore, storage capabilities that are based in the array or network can be applied to multiple, even heterogeneous servers.

Vendor Considerations

When considering networked storage solutions, you will want to find a vendor who can meet your requirements – today and tomorrow. **Below are vendor characteristics to consider. Some may be more important than others, depending on what you need.**

- **Reliability and a proven track record** – A vendor's successful history in networked storage products is a positive indicator of its ability to be a long-term storage partner. It may go without saying, but established vendors are less risky because they are more likely to be around in the future and to be supported by third-party hardware and software providers. Ongoing investment in R&D and innovation is also a positive indicator of future competitiveness.

- **A broad storage line that can meet a range of requirements** – A business may require all forms of networked storage: block and file and online archiving. It may need one today and another tomorrow. Ideally, a vendor offers a range of solutions that includes SAN (Fibre Channel and iSCSI), CAS, NAS, and converged networked storage platforms. Products should support multiple technologies or easily upgrade between them.

  Performance and availability to meet application requirements – These service level characteristics are a function of the storage platforms themselves as well as any advanced data and storage management capabilities. Again, even if you do not need these features now, you may want to add them later.

Case in Point: *Blue Rhino*

Storage Blues

Blue Rhino, acquired in 2004 by Ferrellgas, is the leading brand propane cylinder exchange services – that is, tanks for barbecue grills plus related items. It distributes through thousands of retail locations, including the retail giant Wal-Mart. Blue Rhino relies heavily on electronic supply-chain management and uses a Web portal as the main point of contact for its distributors.

Eight years ago, its data center was typical for a mid-sized enterprise. It had 10 servers running a variety of applications on Windows NT/2000, including Microsoft Exchange, SQL, General Ledger, Office, an imaging solution, and proprietary applications. The servers had internal and direct-attached SCSI storage.

Then growth happened. Blue Rhino's revenues grew from $70 million in 2000 to $260 million in 2003, prior to its acquisition in 2004. Corporate data surged even faster during this period, from 200 GB to over 2 TB. This presented major challenges for its small IT staff. Tracking and managing storage on each server was labor-intensive. Adding capacity was disruptive, and much of it remained idle and unused because it was not transferable between servers. Application response times suffered. The process to refresh a database for application development took a full 8 to 10 hours and included stopping the production application and backing it up to tape. Backup windows and restore times were a concern. The final straw was a disk array failure that caused a couple days of downtime. **Blue Rhino's business success created a storage problem that needed fixing.**

Networked Storage Delivers

The IT department considered SAN solutions from several major vendors. They wanted a fast, highly-available storage array that could start small, in terms of capacity and entry price, and then scale over time to 10+ TB without a forklift upgrade. A track record of product evolution and upgradeability was important. The small staff made it necessary to have centralized management. Finally, they looked for a solid line of software features that they could "grow into" over time.

After narrowing it down to a single vendor, they convinced the CFO and purchased a midrange, SAN-attached storage array containing 1.2 TB of capacity. It included a pair of FC switches connected to 8 servers, snapshot copy, and GUI-based software management. The snapshot capability was especially a hit because it made copies for non-disruptive backup, refreshing a development database, month-end reporting, and fast recovery. The reseller Dimension Data performed the assessment and installation.

Blue Rhino's new SAN delivered major improvements. It virtually eliminated storage outages. Storage management activities were cut by 75%, saving the expense of hiring additional head-count. The time to refresh a database fell from 10 hours to only minutes, accelerating the application development cycle. This was especially useful when Blue Rhino had to revise its IT applications and processes for Sarbanes-Oxley compliance. Moreover, daily snapshots ensured clean backups and allowed fast recovery from disk. They began backing up to disk as well as tape. Storage consolidation on a SAN also facilitated server consolidation, saving costs on the processing side.

Over time, Blue Rhino made additional SAN investments. They connected six more servers to the SAN and installed dual host connections with failover for high availability. They equipped the staff with a storage resource management tool, so they could rely less on outsourced management. Of course, they purchased capacity to keep up with data growth. Future plans include better staff training and possibly a second SAN at remote data center for redundancy.

In retrospect, installing a SAN was a shrewd business and technology decision. It brought more flexibility to operations. "One thing that helped our growth is that ability to turn on a dime," said Brad Hinshaw, then Blue Rhino's Manager of Technical Operations. Greater reliability also brought peace of mind. The storage vendor replaced a disk drive at one point that was predicted to fail, before it might cause downtime for business users. "Nobody knew, and that's the beauty of it," he said. **With a scalable SAN in place, they felt confident about meeting the company's data requirements.**

- **Solutions that start small in capacity and price and scale up in a non-disruptive, pay-as-you-grow fashion** – This especially relevant to small- and mid-sized enterprises (SMEs) and branch offices that do not have the budget to invest in a major solution upfront.

- **Ability to deliver a complete solution** – More than pieces and parts, integrated solutions and professional services are important to SMEs who may lack the time and resources for in-house integration and testing.

- **Interoperability with and support for the existing IT environment** – This may include cooperative support agreements with others vendors and/or a single point of accountability.

- **Easy implementation and management** – It should not be too difficult or take too long to deploy. Pre-integration, deployment wizards, and centralized management software are helpful here.

- **Reasonable pricing** – You get what you pay for, of course, but the value should be there.

Conclusion

Storage is a balancing act between meeting/exceeding business requirements and keeping costs of ownership at a reasonable level. Due to the relative expense and importance of storage in business, it is worth taking the time to do it right.

Networked storage, especially in conjunction with advanced data and storage management capabilities, offers a better and more cost-effective way to store data than traditional DAS. After many years of evolution and maturity, it is a proven solution for many of the storage aches and pains that SMEs and branch offices experience today.

In addition, the reality of never-ending data growth means that storage infrastructure has to evolve and scale over time. It is certainly not a static entity. **Deploying a solid, though flexible, foundation of networked storage would be a smart move for meeting your businesses' storage requirements – today and tomorrow.**

Fibre Channel will live on today, and tomorrow, in the data center for high performance, mission-critical applications, especially where significant investments in FC architecture have already been made. There are hundreds of thousands of SANs installed in the enterprise today. The vast majority of these are Fibre Channel, even though they all do not have to be. iSCSI has been accepted. It has made a significant penetration in thousands of high-end data centers that are deeply committed to ILM and have added iSCSI to their portfolio of tools being used to reduce the TCO of the storage network.

Every enterprise needs to consider a primary and secondary storage network to implement an ILM environment. For some enterprises, iSCSI can be their single storage networking solution.

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,** **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.

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.