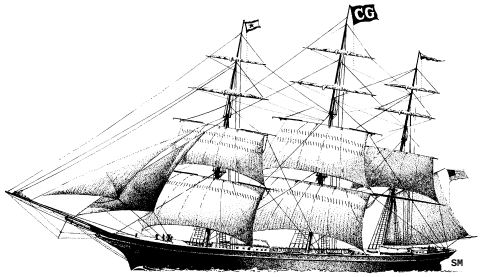
THE CLIPPER GROUP
Explorer™



Navigating Information Technology Horizons℠
Published Since 1993 February 22, 2002

Intelligent Storage Networks - Creating A More Cost-Effective Storage Infrastructure

Analyst: Michael Fisch

Management Summary

Once again, the creaks and groans of a disruptive shift are beginning in the storage industry. **The world is just getting used to the idea of networked storage.** Storage area networks (SANs) are an increasingly common way to connect multiple servers to a centralized pool of storage. Network-attached storage (NAS) is used for sharing files over the corporate network. Market adoption of these technologies is still brisk, even with the current economic downturn. But now the siren song of *the next thing* is being heard again.

The upcoming shift involves placing intelligence in the storage network. Whereas storage software or intelligence used to be the exclusive domain of servers and storage arrays, it has begun to move into the network that joins the two together. Intelligent switches and dedicated networking platforms will take an expanded role for managing, moving, and protecting data. The reason for this shift is to drive greater efficiencies and economies of scale into the storage infrastructure. The network provides a natural common ground from which to manage and manipulate all storage data and create a truly unified infrastructure. Enterprises will benefit from:

- **Lower administration costs through greatly simplified management**
- **Lower acquisition costs by establishing classes of storage**
- **Enhanced value and extended life of existing storage assets**

These benefits are especially attractive as enterprise data continues to grow and storage, especially the management of storage, consumes more and more of the IT budget.

In light of this, customers must ask some important questions. How inevitable is this technology shift? And when is the best time to buy? **Adopting *the next thing* entails trade-offs and risks.** If you buy now, will you be stuck driving a station wagon when everyone else wants to drive an SUV? If you wait to adopt, will you be late to drive your enterprise with better technology?

These are difficult questions, and unfortunately, there are no simple formulas or easy answers. **However, it is time to start thinking about, and even planning for, intelligent storage networks.** The value of this architecture is such that it is only a matter of time before it is widely adopted. A number of products are already on the market, and many more will be released this year. Perceptive enterprises should start to think about how they can gain a competitive advantage by making the leap, or at least sticking their toe in the water. **Read on to find out why intelligent storage networks are a compelling approach and how they will enable a more functional and cost-effective storage infrastructure.**

IN THIS ISSUE

Advent of Storage Networks

Today, the storage industry is in the midst of widespread adoption of networked storage. Previously, each server had its own storage – either internal or directly attached with a SCSI[1] cable. However, this led to disconnected islands of information distributed throughout organizations, and it required a large and costly cadre of hard-to-find IT workers to manage them. It was also difficult to share, protect, and fully utilize the information with this approach. It was a situation in search of a better solution, more so as the amount of information grew.

Storage area networks, or SANs, provided the quantum leap in connectivity that allowed storage to be separated from servers, consolidated, and accessible by the whole organization. These high-performance networks based on Fibre Channel (FC) technology[2] offered longer cabling distances, faster transfer rates, and practically infinite scalability. Storage could be centrally administered and allocated, even more so through virtualization[3]. The result was dramatic economies of scale in management, higher utilization of storage capacity, and significantly-lower total cost of ownership (TCO).

At the same time, network-attached storage, or NAS, provided a means to readily share information in file format over a local-area network (LAN). **Special-purpose file servers, or NAS appliances, provided high-speed access to files from clients running a variety of operating systems like Unix/Linux and Windows NT/2000.** These devices were relatively inexpensive, easy to install, and solved a problem of getting information to users who needed it.

Moving Intelligence into the Network

Looking forward, the next evolutionary step is to make storage networks smarter. **This "plumbing" between servers and storage will start to do more than just move data.** It will take on higher-level roles of management, protection, and routing of information.

Intelligence is fundamental to a robust storage infrastructure. It enables fast, continuous access to information in the midst of rapid growth, equipment failures, geographically-dispersed systems, environmental disasters, and multiple systems accessing data simultaneously. Intelligence in the form of software exists to address these issues, but it is currently distributed among a multitude of servers and storage arrays. **In the next generation of storage, much of this functionality will migrate to central locations in the network.**

The specific features to migrate are open for debate. Some functionality seems clearly to belong on storage arrays because it is tightly coupled to disks, such as processing RAID[4] algorithms and re-building failed drives. Other features – like dynamic load balancing – are best placed on servers because they are closely linked to I/O processing from an application perspective. Nevertheless, a number of important storage functions could fit well in the storage network:

- **Virtualization** – Presentation of heterogeneous storage arrays as a single, logical entity and ensuring only authorized servers can access each volume.
- **Remote replication** – Copying data to a remote site for general content distribution or for failover in the event of a disaster.
- **Point-in-time copy** – Creation of a "snapshot" of production data for other purposes such as backup.
- **File services** – Open systems file sharing traditionally performed by NAS appliances.
- **Security** – Enforcement of storage network-wide security policies and ensuring only authenticated users and servers have access.
- **Storage area management** – Configuration, monitoring, and control of multiple or all devices in a SAN. Ideally, this includes automatic provisioning and policy-based management.

[1] Small Computer Systems Interface

[2] See *Fibre Channel: The Defending Champion Has Staying Power* in **The Clipper Group Explorer** dated December 14, 2001, at www.clipper.com/publications.htm.

[3] See *Storage Virtualization in 2001: A Space Odyssey* in **The Clipper Group Explorer** dated April 9, 2001, at www.clipper.com/publications.htm.

[4] Redundant Array of Independent Disks – describes various techniques for protecting data by placing it on more than one disk in case of failure.

- **Quality of service** – Prioritization of data traffic to ensure appropriate service levels.
- **Consolidated tape backup** – Centralized backup that ensures consistency, fast restoration, and, ideally, does not interfere with servers or the LAN.
- **Protocol conversion** – Conversion between various network protocols and transports, such as FC, TCP/IP, InfiniBand, iSCSI, etc., to allow multiple networks to work together, leverage each other's capabilities in a synergistic way.

Though this list is neither exhaustive nor conclusive, it does provide a reference point for considering next-generation solutions.

Another factor in moving intelligence into the network is the platform in which it resides. The platform must have "traffic cop" capabilities – the ability to open up data frames[5] and process the data in real time – or the ability to interact with and potentially control other devices on the network. Most existing SAN equipment is not "smart enough" in this sense, but many vendors are developing or already offer products with some of these capabilities:

- **Intelligent switches** now are coming to the market. These purpose-built networking devices are becoming multi-lingual and able to converge multiple protocols and transports to create a unified storage network. Some will also offer advanced features like storage virtualization, file services, etc.
- **In-band storage networking servers and platforms** are also available. (See box at right.) These provide a front-end for multiple storage arrays, and then connect to servers either directly or through the network. The key function of these devices is storage virtualization, though all offer additional features such as remote mirroring, point-in-time copy, and file services.

In-Band vs. Out-of-Band

There are two basic architectures for storage networking platforms: in-band and out-of-band. Both have unique strengths.

In-band means the device is in the data path between servers and storage arrays, and information (files, data, etc.) must pass through it. This affords the opportunity to examine and process the data in real-time and independent of servers and storage arrays – well-suited for implementing features like virtualization and remote replication. The downside is the potential for slowing data transfer and becoming a point of failure. Vendors have worked to overcome these problems with high-performance chip sets, cache memory, greater queue depth, and redundant, failover configurations.

Out-of-band means the platform is outside of the data path and interacts with servers, switches, storage arrays, etc. via device-specific APIs (application programming interfaces), MIBs (management information bases), or agents. This architecture is well suited for applications like storage area management that need to monitor, manage, and configure all devices in the network. It also eliminates the in-band weaknesses of slowing data transfer and becoming a point of failure. The downside is that it is one-step removed from the action, so to speak, and relies more heavily on intelligence distributed among devices in the data path. The challenge of this approach is to ensure consistency and interoperability among a wide variety of devices, makes, and models.

- **Out-of-band storage networking platforms** can encompass a broad range of features such as storage area management and virtualization. (See box above.) This approach is well suited for monitoring, managing, and even controlling all devices in the network.

Casting Out Inefficiencies

Let's look back in time to understand why intelligent storage network devices are clearly the direction that storage networking is headed. After all, history really does tend to repeat itself.

There was a day when craftsmen worked in small shops using basic tools to produce goods.

[5] Packets of data, as well as information describing the data, such as its origin and destination.

A blacksmith, for example, may have melted and refined iron, poured it into a mold, and shaped and tempered it into whatever was needed that day – say a set of horseshoes, an axe, and a knife. Then came the Industrial Age. Large factories filled with machines churned out goods in greater quantities and at lower costs. The manufacturing process became more specialized or *disaggregated*[6], *mechanized*, and *consolidated* in large facilities. This resulted in economies of scale that endowed the masses with inexpensive goods. The solo craftsman simply could not compete. And thanks to this revolution, we now enjoy modern wonders like Wal-Mart and the choice of 50+ different models of SUVs. Civilization marches forward!

Similar concepts are now being applied to information technology. When one takes a broad look at IT, as if looking down from a satellite, two paradoxical patterns emerge. **The first is a breaking apart or *disaggregation* of IT systems.** When once there were only self-sufficient mainframes with dumb terminals attached, now there are application servers, database servers, Web servers, storage platforms, and so forth, with networks gluing them together. A computational task involves all of them working together to generate an "answer". The other, seemingly contradictory, trend is *consolidation*. **Within each area that is broken apart, assets are being centralized.** For instance, servers can access a common pool of storage rather than each having a dedicated storage array. IT systems are being rearranged.

As in previous ages, the reason for this rearrangement was efficiency. **Each component that is broken apart and consolidated can be managed as a single entity.** It requires fewer people, and it is much easier to make sure each asset is utilized to the fullest (i.e., no unused and wasted resources). There is more flexibility, allowing the system to be reconfigured and changed on the fly. It is also much easier to scale a discrete entity – critical in light of ever-faster processors, ever-fatter network pipes, ever-larger storage capacities, and ever-more-powerful applications to consume it all. **The overall result is an IT system that can grow and change with the times, minimize TCO, and provide greater value to the enterprise.**

This disaggregation and consolidation has already been applied to storage with the advent of storage networking. **However, a close look reveals that inefficiencies still exist.**

For one, SAN and NAS storage are typically separate. There are RAID arrays connected to FC SANs and there are NAS appliances connected to IP LANs – with no link between the two. It would be better if all storage was consolidated and clients/servers could access it in either block or file format. Then storage would be truly consolidated and easier to manage.

Even more importantly, the software for protecting, managing, and moving information is not consolidated. It comes in a variety of forms and is spread among the servers and storage arrays, creating a number of limitations and management complexities.

For example, storage software that runs on application servers slows production applications and leads to administration headaches. An instance of software must run on each server, including typical corporate environments where multiple operating systems are deployed, such as Windows NT/2000 and the many flavors of Unix/Linux. Even if a software package supports them all, an enterprise must find a way to procure, install, and manage potentially hundreds of software instances. Furthermore, software consumes server CPU cycles and detracts from application performance on *every server*. This can negatively affect worker productivity and even detract from enterprise objectives such as profitability and customer satisfaction. (Have you ever called a customer support center and had to wait several minutes while the agent's computer processed your transaction?) There has to be a better way!

One solution is to move intelligence into the storage arrays and thereby offload application servers, though this brings another set of drawbacks. **Today, array-based software is proprietary – each vendor has their own version that only runs on their arrays.** Customers are discouraged from buying hardware from multiple vendors because of the complexity of learning and managing multiple software packages. Features like remote replication also require customers to purchase multiple units from the same vendor. This creates a sort of vendor lock-in, with the potential for higher prices.

In addition, some software functionality is only wholly effective when applied to all

[6] For example, one factory produced steel, another used it to make tools, and a third party sold it, etc.

servers and storage arrays simultaneously, including heterogeneous makes and models. Storage virtualization and storage resource management fit into this category, for example. Array-based solutions for these features could never be optimal, and server-based solutions possess the previously-described shortcomings.

A More Efficient Approach

The solution is to disaggregate intelligence from both servers and storage arrays and consolidate it in the storage network. All storage traffic must pass through the storage network, so it is a natural common ground from which to manage all of the data. The scope of the solution is maximized; it can encompass all servers and storage arrays. The inefficiencies associated with array-based and server-based software disappear. There are relatively few devices where intelligence must reside, and management is greatly simplified. The framework can incorporate all existing storage assets as well as future acquisitions. And customers are able to increase the functionality and efficiency of their storage.

Benefits to the Bottom Line

The net effect of moving intelligence into the storage network is a more economical storage infrastructure, especially as it scales. The amount of information corporations must handle is rapidly growing, sometimes doubling annually, and storage is becoming an increasingly large portion of the IT budget. Greater economies of scale can relieve some of this pressure.

The benefits of centralized storage administration only increase as the infrastructure scales. **SAN and NAS as well as heterogeneous storage can be combined into a single entity, resulting in simpler management, higher utilization of capacity, and lower TCO.**

Intelligent storage networks also give customers the flexibility and freedom to deploy storage from multiple vendors without hindrance, according to the price/performance requirements of each application. **This ability to match "classes of storage" with applications is a major cost saver.** High-end (and more costly) storage can be used with mission-critical applications, while other applications can use mid-range or low-end storage.

And perhaps the best news is that this approach does not mean scrapping investments in existing storage. On the contrary, by making advanced functionality available to even JBOD[7] arrays, it extends its value and useful life. **In short, putting intelligence in the network offers a more functional and cost-effective storage infrastructure.**

Conclusion

For all of these benefits, and in light of the trend toward disaggregation and consolidation, **the deployment of intelligent storage networks is only a matter of time for many enterprises. It offers a "quantum leap" in value.** The market is already moving in this direction, and products are now available. Hundreds of millions of venture capital dollars have been poured into startups such developing products. And even some established vendors are moving in this direction, despite vested interests in keeping intelligence on servers or storage arrays.

Since the writing is on the wall**, it makes sense for forward-looking enterprise to start considering intelligent storage networking devices.** This is a strategic issue to be resolved in light of both current and future storage needs. Though there is a risk in being one of the first to jump on a bandwagon, it is mitigated by the assurance that the market is moving in that direction. **Furthermore, there is a real opportunity to garner a competitive advantage by creating a more capable storage infrastructure and, more importantly, dramatically lowering its cost.** In an era of rapid information growth, and in an economy that competes on the basis of information more than any other asset, having the best storage infrastructure would be, well, smart.

[7] Just a Bunch of Disks – an array of disks without a RAID controller

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 a Senior Contributing Analyst with the Clipper Group.** He brings over six years of experience in marketing and engineering at computer hardware and software manufacturers. Mr. Fisch worked at EMC Corporation as a marketing program manager focused on service providers and as a market analyst. Prior to 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 MFisch@clipper.com or at (781) 235-0085 Ext. 36.***

Regarding Trademarks and Service Marks

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