THE CLIPPER GROUP
Navigator™
Navigating Information Technology Horizons℠
Published Since 1993 April 26, 2002

The Future of Big Blue Storage – IBM Explains It All

Analyst: Michael Fisch

Management Summary

Big changes are afoot in the storage industry. Storage intelligence – the software for managing, moving, and sharing data – is migrating from servers and storage arrays into the network that joins the two together. The distinct and sometimes contentious camps of NAS and SAN are on a collision course to form a more efficient, merged entity. Furthermore, individual device managers are giving way to SAN-wide management platforms that look after the effectiveness of the entire storage environment. **The purpose of these changes is to enable greater efficiencies and ultimately lower storage total cost of ownership (TCO).** This is what enterprises need most in light of the rapid, unabated growth of information.

IBM has recognized these industry shifts and announced a storage product strategy that plants itself firmly on high ground as the trends play out. This comprehensive plan addresses SAN and NAS, block and file, hardware and software, multiple protocols, and open standards. **IBM intends to deliver in 2003 a unified, networked storage environment that serves all needs of the enterprise at the lowest possible TCO.**

The product strategy has three key components:

- **Virtualization Engine** – a SAN-wide virtualization platform
- **Storage Tank** – a SAN-wide and eventually enterprise-wide file system
- **Open management** – based on the Common Information Model (CIM)

The *Virtualization Engine* will reside in the network and present multiple, heterogeneous disk arrays as a single resource to be allocated as and where needed. It will also support advanced features like data replication and security. *Storage Tank* will deliver a common, high-performance file system that enables heterogeneous servers to access files via the SAN. In effect, it converges NAS and SAN by consolidating file and block data in the SAN. Finally, IBM will support *CIM*, the management interface standard developed by the Storage Networking Industry Association (SNIA), for storage management software and devices.

With this announcement, IBM has set its course for navigating into the next generation of enterprise storage. Read on for details.

IN THIS ISSUE

The Next Generation of Storage

The world is becoming accustomed 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. Compared to direct-attach storage, both SAN and NAS are much more efficient and cost-effective. The market recognizes this, and these technologies are enjoying widespread adoption and fast growth. Even so, the storage industry is feeling the rumblings of another disruptive shift on the horizon.

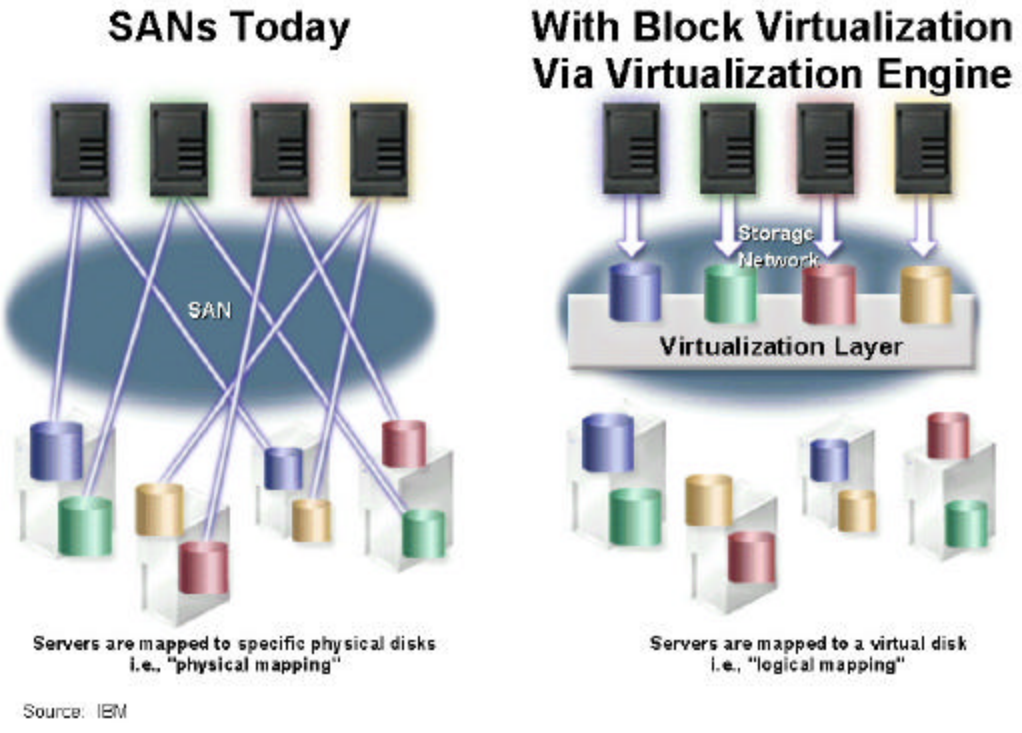
The rapid growth of information has continued nonstop for years, and there is no end in sight. While the Information Age has connected the world together and made it a smaller place, it has also brought ever-increasing stockpiles of data that must be managed, protected, and made easily accessible. **The benefits of SAN and NAS have helped enterprises to cope, but unrelenting data growth is driving the need for a "next generation" of storage networking technologies that deliver even greater efficiencies and economies of scale.**

IBM's storage product plans reflect this need. Its new hardware and software architectures are intended to deliver the best possible set of benefits in all relevant dimensions – performance, availability, flexibility, and cost-effectiveness.

IBM's Virtualization Engine

The first component of the strategy is a *Virtualization Engine*, a SAN-wide virtualization platform that resides in the data path and joins together heterogeneous servers and storage arrays. Unlike SANs today which map servers to specific physical disks, the **Virtualization Engine will abstract all attached disks and present them as a single, block-level storage resource to servers, allowing them to be securely shared, reallocated, and scaled as necessary and without disruption.** It will be able to process data real-time and perform advanced functions like point-in-time copy, remote replication,

Exhibit 1: Block-Level Virtualization in the Network



Source: IBM

and data migration across all connected devices. **Developed by IBM Labs, the Virtualization Engine is intended to be the cornerstone of a highly cost-effective and unified storage infrastructure.** (See *Exhibit 1*.)

The Virtualization Engine is designed for high availability and centralized management. Each network-resident node will be an IBM *xSeries eServer* with up to 4 GB of cache. The nodes will be paired for redundancy, with mirrored cache, and LUNs shared between each pair. **There will be no single point of failure.** Furthermore, up to four pairs will be managed as a single pool of virtual storage – with additional pairs added over time.

The Virtualization Engine will deliver several important benefits to the enterprise:

- **Lower administration costs** through simplified management.
- **Lower acquisition costs** by improved resource utilization and smarter storage procurement.
- **Extended value and useful life of existing storage assets.**

With a Virtualization Engine, storage administrators will be able to manage all attached arrays centrally, as one giant disk, and significantly decrease the administrative workload. This can have a major impact since administrative costs are by far the largest part of storage TCO.

A consolidated view of all storage also enables a more efficient use of capacity. Enterprises can then defer incremental storage purchases and effectively lower acquisition costs. Furthermore, **the Virtualization Engine will allow enterprises to establish classes of storage and rationalize storage procurement.** By assigning highly-robust but more costly storage to mission-critical applications and less expensive storage to less-critical applications, enterprises can optimize procurement and save money. Since all connected devices can take advantage of the product's advanced functionality, there will be less need to purchase licenses for array-based or server-based storage software functionality, such as replication. Finally, customers do not need to scrap investments in existing storage assets. On the contrary, connecting them to Virtualization Engine will extend their value and useful life as well as increase the level of utilization. **In short, lower management costs, lower acquisition costs, and longer useful life all mean a greatly improved TCO.**

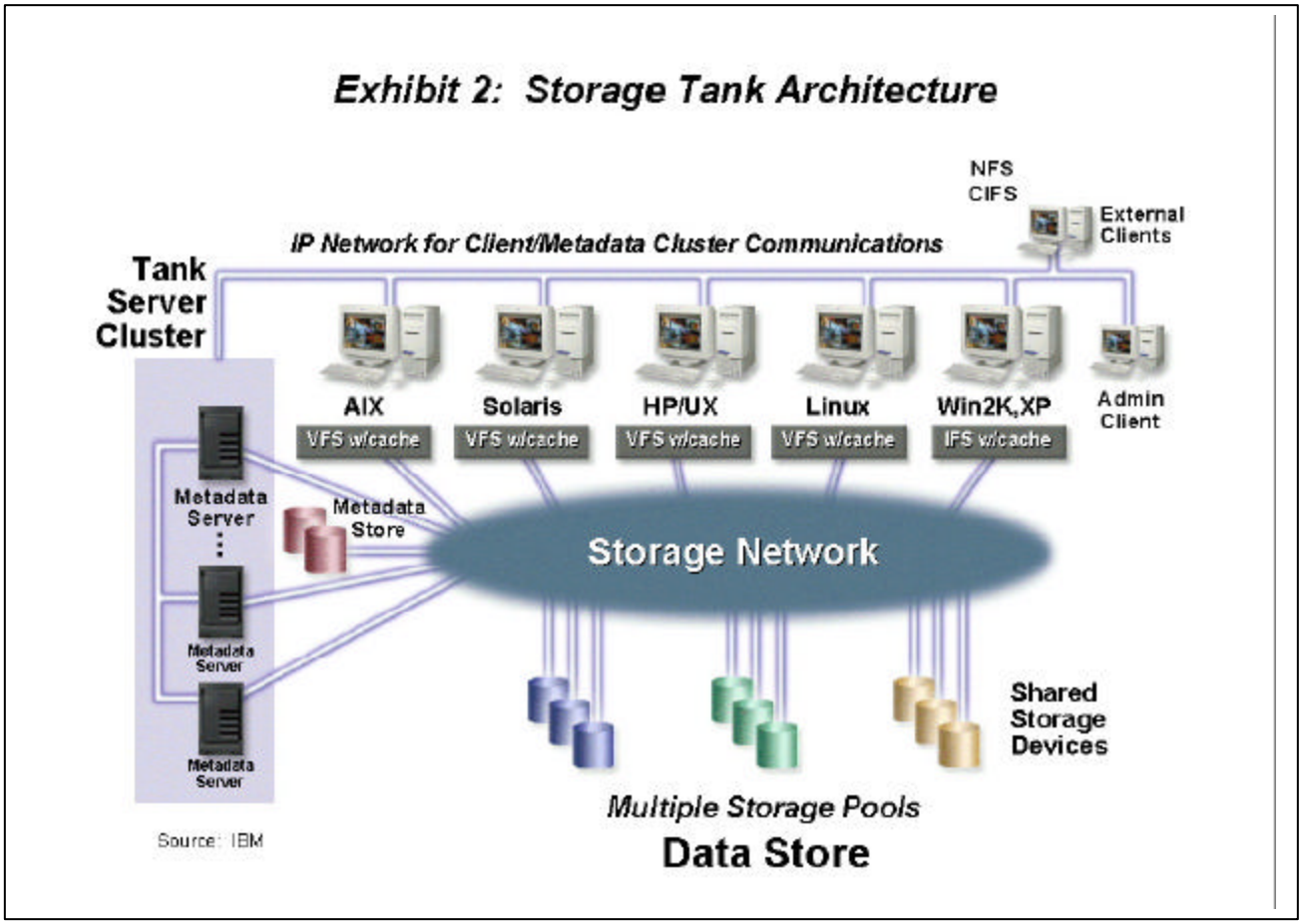
IBM's Storage Tank

The second part of IBM's strategy is to develop a common, enterprise file system called *Storage Tank*. Initially, Storage Tank will be a SAN-wide file system that allows heterogeneous servers to access files directly over a SAN. Later versions will support inter-SAN and wide-area configurations. Traditionally, servers access files over the local-area network (LAN) through a file server or special-purpose NAS appliance. With Storage Tank, servers can pull files directly from storage arrays over a high-performance SAN. **In effect, it converges NAS and SAN and allows all data – both block and file – to be consolidated into a single pool of SAN-attached storage.**

As shown in *Exhibit 2*, Storage Tank works by storing file metadata (i.e., the information about files) on metadata servers that are connected to both the LAN and the SAN. These servers are clustered for high availability. A small agent runs on each application server, which allows the Storage Tank file system to appear native to the server.[1] When a file is requested, the server retrieves metadata over the LAN from the metadata servers. With this information, it can access the file directly over the SAN[2] without in-band interference. In effect, many servers running different operating systems will be able to access the same name space. Storage Tank metadata servers will also act as NFS/CIFS file servers for external clients not

[1] Each Unix server has an NFS-like file system stub (labeled "VFS w/cache" in Exhibit 2) that looks like NFS to the application and operating system, but serves as a cached conduit to the Metadata Server and storage network. Each Windows NT and 2000 server has a similar CIFS-like stub (labeled "IFS w/cache").

[2] SAN interconnect can be Fibre Channel or TCP/IP over Ethernet.



connected to the SAN.

Storage Tank will offer a number of features to facilitate policy-based automation and centralized management. It will automatically provision capacity according to pre-defined policies, including the ability to choose from different classes of storage according to quality of service requirements. Data migration will be non-disruptive. As a common file system, Storage Tank will centrally manage all file data and universally apply features such as capacity planning, backup/restore, point-in-time copies, and security.

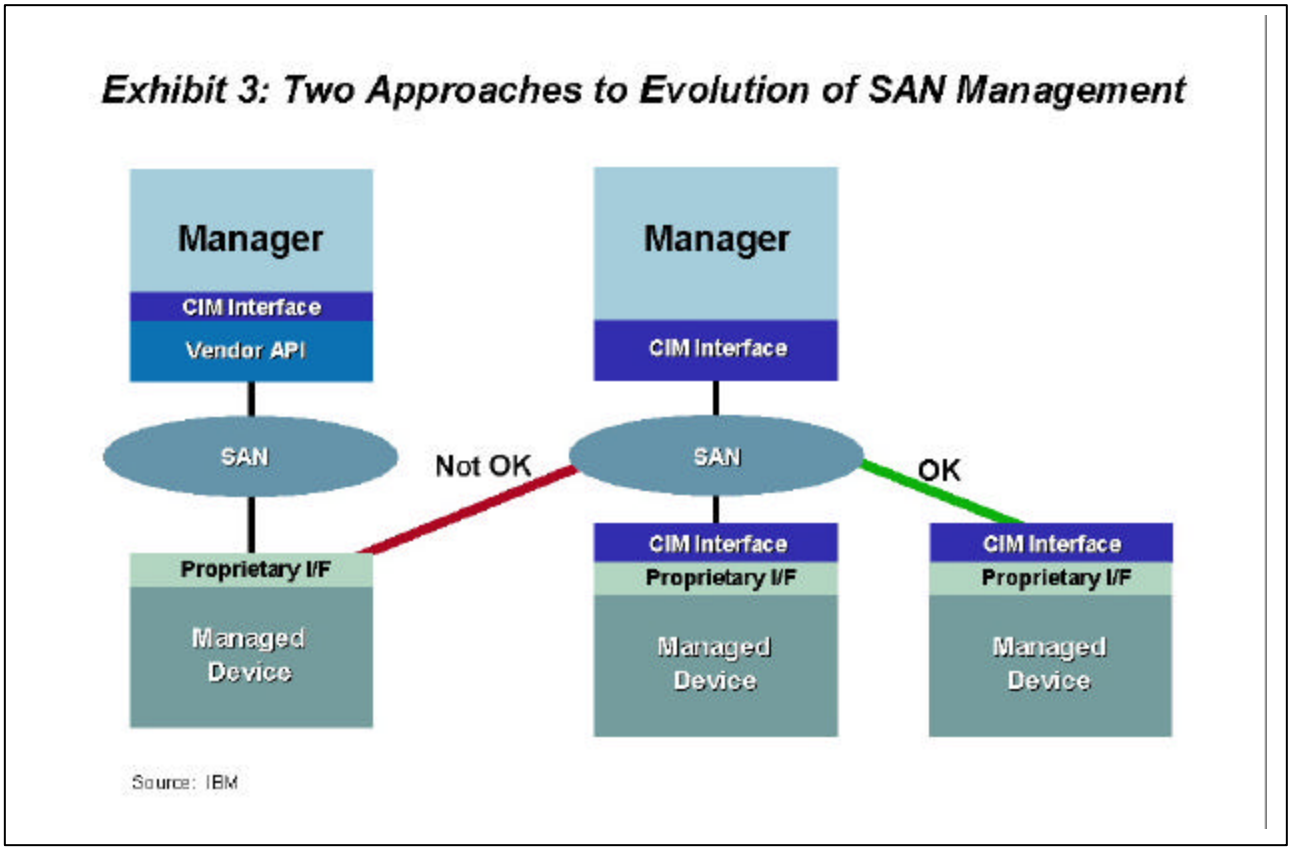
In an effort to promote an industry standard, IBM will license Storage Tank server agents at no cost and openly publish the metadata server protocol. IBM clearly desires Storage Tank to become *the* way for enterprises to access files.

Storage Tank will deliver several benefits to the enterprise:

- **Lower acquisition costs** through improved capacity utilization and smarter storage procurement.
- **Lower administration costs** through centralized management.
- **Increased productivity** by improving the availability and performance of enterprise applications.

Like IBM's Virtualization Engine, Storage Tank improves capacity utilization by centrally managing all file system data and unifying it with block storage in the SAN. It also enables IT departments to rationalize storage procurement by establishing classes of storage. Storage consolidation and creating a single name space for all files greatly simplify management. All of these benefits work to lower TCO.

Furthermore, Storage Tank extends the performance and availability characteristics of SANs to file systems. As dedicated, high-speed storage networks, SANs are designed for robustness and performance. **Applications that rely on Storage Tank's SAN-based file**



system can experience faster response times and higher availability, enabling workers that use the applications to do more. As a result, Storage Tank can improve the productivity of business processes.

CIM: Open Storage Management

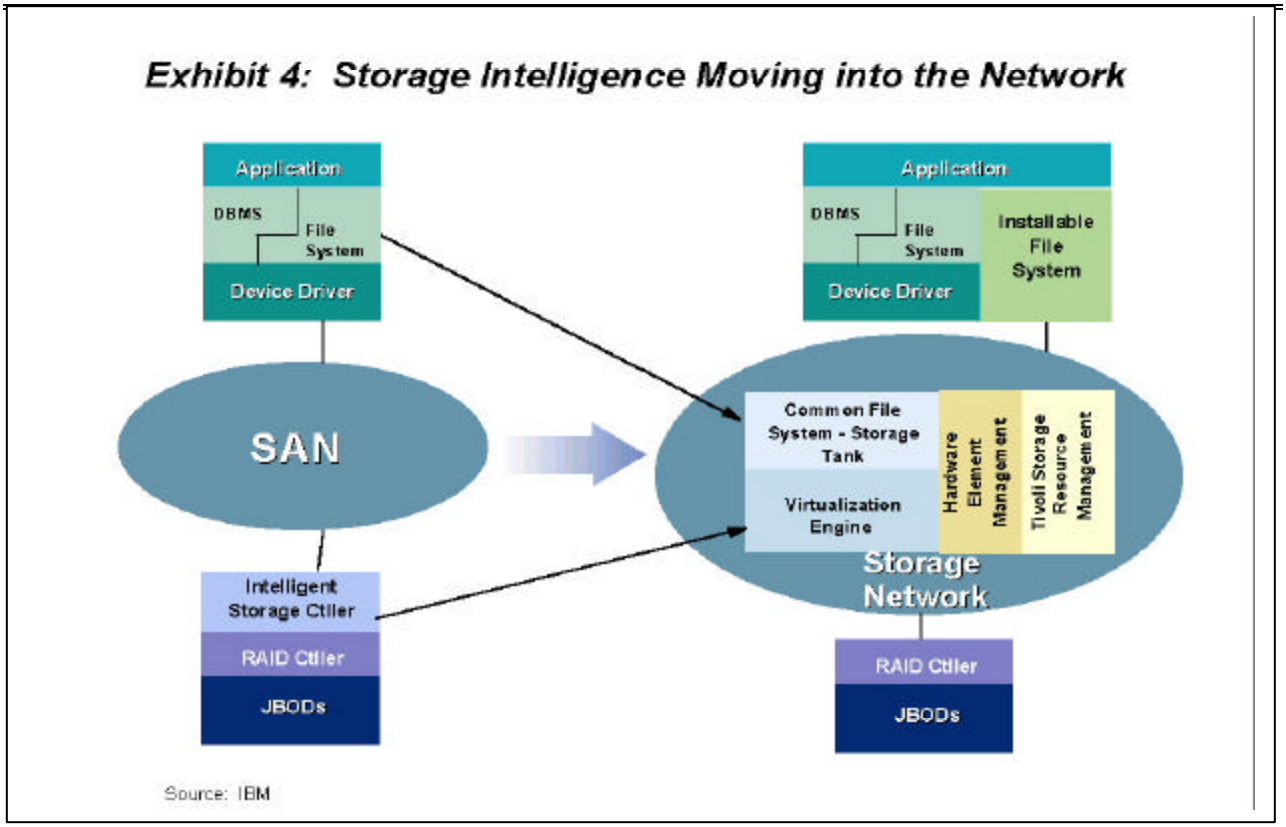
The third component of IBM's strategy is open management based on the *Common Information Model (CIM)*. CIM is a standard interface developed by SNIA[3] for managing storage hardware and software. **It simplifies the integration and management of multi-vendor SANs and gives customers what they have long been clamoring for: interoperability – at least at the management interface level.** By supporting CIM, IBM opens up its storage products for management by third-party applications. Of course, it hopes that enterprises will prefer its own storage resource management (SRM) software being developed under the Tivoli brand.

IBM will support open management by incorporating CIM interfaces in its storage devices as well as in Tivoli SRM (See *Exhibit 3*, right side). **In this way, Tivoli could manage *any* CIM-enabled device, whether from IBM or another vendor, and Tivoli or any CIM-enabled manager could manage IBM storage devices.** This universal inter-operability is in contrast to the alternative method of incorporating proprietary, device-specific interfaces into the management application. This requires vendors to license or reverse-engineer an interface for every device it wishes to support – a costly and never-ending effort. (See *Exhibit 3*, left side.)

Since it will take time to integrate CIM support into its products, IBM will offer an interim solution to speed time to market. It is a proxy server that translates between IBM's proprietary management APIs[4] and CLIs[5] (and potentially those for non-IBM devices) and the CIM interface. This will allow enterprises to take advantage of CIM support more quickly, as well as incorporate legacy products.

[3] Storage Networking Industry Association.

[4] Application Programming Interface.
[5] Command Line Interface.



Meanwhile, IBM plans to deliver native CIM support in the Virtualization Engine and Storage Tank, enabling integrated management of both products.

CIM support means enterprises will have more flexibility and freedom in their storage procurement. They can more easily manage mixed-vendor environments, especially since CIM enjoys broad industry support. CIM will also accelerate the development SRM solutions from IBM (Tivoli) and other vendors. SRM handles the complexity of storage management from a top-down perspective and helps users get the most out of storage assets while minimizing administrative work. CIM will free vendors to focus on developing value-added features rather than deciphering and integrating proprietary interfaces from the plethora of storage products on the market. **All of this is good news because it means more flexibility and openness in the storage industry.**

Conclusion

IBM has put a stake in the ground and signaled that it intends to be a leader in the next generation of storage technologies. As shown in *Exhibit 4*, **IBM's approach will allow it to deliver a common, enterprise file system for all servers and a virtualization platform that pools all storage arrays.** These plans are consistent with, and even accelerate, the trends toward intelligent storage networking, SAN/NAS convergence, and open, SAN-wide management.

For enterprises, it means that even better tools are on the horizon for coping with spiraling storage requirements. **SAN and NAS are available today, and something better – an open, unified, networked, and highly cost-effective storage infrastructure – will be here tomorrow.**

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 Storage and Networking 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. 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 MFisch@clipper.com or at 781-235-0085 Ext. 25***

Regarding Trademarks and Service Marks

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.