THE CLIPPER GROUP
Navigator™

Accelerated Real Time Compression and Automated Three Tiers of Storage — IBM's Enhanced Storwize V7000 and SVC Drive Real Value

Analyst: David Reine

Management Summary

Have you seen those TV commercials for the storage bags that (with the aid of a vacuum cleaner) remove the air out of the clothing or linens so that they take up less room to store? That's a very visible form of compression. Unfortunately, this can be a time-consuming, labor intensive effort. You have to pack the items in the bag efficiently, seal the bag, and then suck out the air. When you want to use the item, you then have to break the seal and let the items inside reflate themselves. None of this is instantaneous. Wouldn't it be nice if you could just put something in an appropriate container and have this done without any involvement or even awareness that it was being compressed/decompressed? Another way to express this is to ask, "Wouldn't it be great if the whole compression operation was invisible and sufficiently speedy, i.e., being done without you ever knowing about it?"

Shifting gears just for a moment, Big Data is a big idea; one whose time clearly has come. Unfortunately, Big Data means a lot of stored data and dealing with this adequately and economically is a big challenge. By "adequately", I mean in a sufficiently speedy manner that doesn't slow the pace of business. By "economically", I mean at the lowest possible cost while still meeting the service delivery requirements. The first is a matter of effectiveness and the second is a matter of efficiency. If you don't have both, your storage solution most likely will come up short. The key to all of this is understanding the speed of storage and retrieval depends on the content and the circumstances. While you could put all of your data on SSDs and thus satisfy the performance requirements all of the time, doing so would be costly and wasteful. Some data (like criminal records) may always need to be delivered at the fastest speeds. However, in the world of Big Data, what may be needed urgently today may not be so clearly anticipatable. What you really want is to be able to treat whatever data now is in hot demand when it becomes hot (and vice versa) and thus be able to keep the overall storage costs down.

None of this is new. The IBM Storwize V7000 and the IBM SVC multiplatform storage controller have had real-time compression and automated tiers for some time.[1] However, IBM has enhanced the Storwize 7000 and SVC in several ways, including (1) enhancing its *Real time Compression (RtC)* capability with the addition of Intel's *QuickAssist Acceleration Technology* to speed the compression and decompression so that it is totally invisible to the user or application

---

[1] The same SVC software functionality runs in the standalone SVC and the IBM Storwize V7000; this includes IBM's Real time Compression capability. For an earlier discussion of this functionality see **The Clipper Group Navigator** entitled *Significantly Improving Storage Efficiency - IBM Delivers Real-time Compression on Unified Storage,* dated October 2, 2012, and available at http://www.clipper.com/research/TCG2012013.pdf.

and (2) increasing the number of tiers of auto-mated storage from two to three (with its new *IBM Easy Tier 3* offering), thus allowing pools of storage at three tiers (SSD, high-performance rotating disks, and high-capacity rotating disks) and automated movement of data to the right tier. Read on to learn more about these and other new capabilities and enhancements.

Data Growth in the Enterprise Data Center

The modern data center has to deliver more information faster than ever before. This requirement has been created by applications with a need for faster data delivery to keep up with the need for faster decision making. Enter-prises are looking for answers to today's prob-lems right now, not tomorrow. At the same time, they are also looking to identify and solve future problems before they become trouble-some.

In the current age of "Big Data", data often has become the most critical asset of the enter-prise data center. Each data center is collecting massive amounts of information, some or much of it created by a variety of external sources, such as video cameras, sensors, and social me-dia, often with much of this data being captured in the cloud. In fact, the storage requirement for both structured and unstructured data for many enterprises may be doubling every 12-to-18 months. There is a tremendous urgency within the Big Data-dependent enterprise to have the flexibility to meet ever-changing business needs with virtualized and efficient storage and, at the same time there is great pressure to lower the total cost of ownership (TCO) of the storage infrastructure. Of course, this sounds like an impossible challenge, similar to that magically deflating and inflating storage bag described in the introduction.

The enterprise that can obtain the most criti-cal insights from that data is sure to have the edge in business preparedness. *Has your enter-prise deployed its own big data initiative, yet? If not, why not?* Perhaps you are concerned about the ROI on any new investments made in data center infrastructure. *However, what if the value returned from higher performance, better stor-age density and lower TCO outweighs the acqui-sition cost of any new infrastructure?*

Today's data center needs to address a multi-tude of issues, none more important than:

- *Reduce physical storage space* to help free up valuable data center floor space and also reduce energy costs.
- *Increase performance* to enable faster analytics and shorten the time for reaching valuable insight.
- *Reduce complexity* to lower operational expenses with easy to use management tools and, thus, enable the data center staff to dedicate more time to focus on valuable strategic priorities.
- *Leverage integrated support* for both file and block data in order to consolidate workloads.
- *Deliver high reliability and business con-tinuity*, in order to satisfy the service level agreements (SLAs) to which the data cen-ter already has committed. This includes the need to improve network utilization for remote mirroring and data replication.

There is a common theme to all of these points: they all contribute to the success of the enterprise by meeting important data storage and delivery goals while, at the same time, lowering the TCO for handling more data, more speedily. In more practical terms, the data center staff cannot allow storage to bust the IT budget. The TCO includes acquisition cost, but also takes into consideration administrative costs, mainte-nance, energy, and even the floor space required to house this massive storage infrastructure. Additionally, with the included virtualization, wasteful practices of over-allocation out of fear of running out of headroom can be eliminated. This significantly improves the ROI for storage investments.

In order to improve storage efficiency, the storage system must be able to support multiple tiers of storage (classes of storage infrastructure with differing performance and costs per unit of storage) with intelligent migration of data among the tiers. Intelligence, whether provided by hu-mans or policy-driven automation, is needed in order to optimize the use of each storage tier. All data does not have the same value to the en-terprise and that value may change over time, both down and up. The most important data and the data in the hottest demand usually need to be stored within the tier that provides the applica-tion with the fastest data delivery – when it is required. It should be migrated to less expensive storage when it loses some of its "heat" or importance.

Virtualization removes the need to worry about where data is stored physically and also

simplifies the deployment of new applications and the creation of multiple storage tiers. Trying to optimize multiple tiers without virtualization is quite burdensome.

The rapid growth of data often is approaching or even exceeding the physical capacity of the modern data center. Now, the data center staff has a new mandate: to store even *more* data in the *same* space, or even harder, in *less* space. The limitations mandated by the decreasing availability of additional floor space in today's data center require the staff to find a way to increase capacity within that same footprint, or run the risk of requiring the construction of a new data center at a possible cost of millions. The data center staff needs to find a way to dynamically fit more data into the same space, whether on HDD or SSD, in order to eliminate the need to build another data center. *How can they do that?*

Traditional solutions, such as in-place compression, have proven to be inefficient, wasting much of the storage space that they were designed to preserve because they often also need to store a pre-processed (uncompressed) copy of the data. Furthermore, many of these same compression algorithms also consume server compute cycles that were intended for the mission- and business-critical applications upon which the enterprise depends. One way to improve storage efficiency is to use innovative, and faster, compression techniques. Some storage vendors have even replaced traditional methods with real-time compression to expedite that process. However, this dynamic compression (and decompression) process still takes more time than some mission-critical applications can afford to wait.

In addition to an improved compression capability, the modern data center also has a need to *automatically* migrate data between the multiple tiers of storage. Having administrators do this manually is a waste of precious staff resources; automation is the answer. "Hot" data needs to be resident on the highest performing drives, usually a solid state drive (SSD) if it is compatible with the budget for the IT infrastructure. "Warm" data should be stored on a high performing disk drive (HDD), usually a 10K or 15K RPM enterprise drive. Finally, "cold" data can reside on a low-cost, high-capacity nearline drive, typically a lower-performing SAS or SATA drive. The storage system must be able to migrate data between any two of these three

tiers, deployed with flexible configurations to meet an ever-changing set of enterprise requirements. This means that the migrations need to be done dynamically to be able to respond rapidly to changes in the information being requested.

As we have often stated before in this forum, *infrastructure really does matter.*[2] The right infrastructure enables the data center to control the TCO of the IT infrastructure, investing more in critical applications that will improve business flexibility, resulting in a better ROI. It all comes down to value. *Are you getting the maximum value for the dollars that you are expending for your infrastructure?* Improving storage value is exactly what IBM has done with the updated IBM *SAN Volume Controller (SVC)* and the newest version of the IBM *Storwize V7000*. IBM has enhanced both of these storage products to create a common foundation of a data architecture that will increase the value of the storage architecture to the enterprise.

IBM SAN Volume Controller and IBM Storwize V7000

The enhancements that IBM has incorporated into both IBM SVC and IBM Storwize V7000 enable the data center with the flexibility and responsiveness they need to meet changing business urgencies and data demands. Through virtualization, IBM has freed data from being tied to any legacy storage architecture. IBM's Storwize V7000 and SVC enable the data center to extend the benefits of Storwize to any block storage attached to it. This improves the efficiency of that storage with the newly announced IBM Easy Tier 3[3] for automated migration between tiers and improved *IBM Real time Compression (RtC)* capability for data compression, reducing the amount of data sprawl running rampant in the data center and transforming the economics of storage, as well.[4] These enhanced features allow the Storwize V7000 and the SVC to maximize storage utilization while minimizing data movement outages.

Storwize also has multi-tenancy support,

---

[2] See **The Clipper Group Navigator** entitled *When Infrastructure Really Matters – A Focus on High-End Storage,* dated October 2, 2012, and available at http://www.clipper.com/research/TCG2012021.pdf.

[3] Much more on this to follow.

[4] For a detailed discussion on RtC, see **The Clipper Group Navigator** entitled *Significantly Improving Storage Efficiency – IBM Delivers Real-time Compression on Unified Storage,* dated October 2, 2012, and available at http://www.clipper.com/research/TCG2012013.pdf.

enabling more file and block consolidation. In addition, *Storwize V7000 Unified*, a file system enhanced version of the Storwize 7000, can be configured with *NFS v4* to improve performance and security capabilities and *SMB 2.1*[5] to enable faster file transfers and better security via access control. IBM's *Active Cloud Engine* allows for file sharing across multiple sites, permits multiple users to update the same file, and enables non-disruptive migration from existing NAS systems.

These features enhance the ability of the data center staff to execute the rapid deployment of any new application and virtually eliminate downtime for data movement, meeting all of the enterprise transformation needs. The SVC and Storwize V7000 simplify and standardize the deployment of storage management with efficiency and consistency, reducing operational costs while, at the same time, improving the performance for big data and analytics workloads.

What is the difference between the IBM SVC and IBM Storwize V7000?

Basically, they are the same in that all of the Storwize functionality found in the latest V7000 is also available in SVC; however, the SVC is a controller without any built-in storage. With the IBM SVC, you need to attach storage arrays to it (from a selection of many brands and models), while the IBM Storwize V7000 is a complete storage solution, including the storage (but you also optionally can attach other arrays behind it). This results in a similar, consistent approach to virtualized storage, allowing the data center to choose which approach best suites their current needs for enterprise storage. The Storwize family has been designed for efficiency, ease-of-use, and reliability. For example, if you need real-time compression or automatic tiering, simply turn them on and forget them. There is nothing else for you to do – and that is the whole idea.

The IBM SVC

IBM has reconfigured SVC as a 2U server to enable flexible, enhanced processing and I/O, with buyer configuration options for RtC, CPU, memory, and I/O. SVC has front-mounted dual system batteries to remove the necessity for an external UPS. When RtC is elected, the SVC requires a cache upgrade and the second CPU is required. SVC can be equipped with optional compression acceleration cards based upon the *Intel QuickAssist* technology, as will be described shortly.

The IBM Storwize V7000

The IBM Storwize storage solutions can be configured for any size data center, beginning with the *IBM Storwize V3700*[6] for the SMB, to the *IBM Storwize V5000* for the mid-sized enterprise, to the newest version of the high-end *IBM Storwize V7000* with accelerated real-time compression. The new V7000 has been designed to double both the IOPS performance[7] and the capacity as compared to the prior V7000 offering, with up to a ten-fold improvement in real-time compression performance.

The new V7000 controller comes with a larger pair of canisters (control modules) than in the previous SVC to support more hardware resources within the same 2U footprint. Each V7000 system comes with eight CPU cores per canister, and up to 64 cores in a clustered system.[8] Each also comes with up to 64GB cache per canister, and up to 512GB of cache per clustered system. With this range of components, the IBM Storwize V7000 can extend your application processing capability as a result of the enhanced memory capability, for example, enabling your analytics applications with faster insight into an ever-increasing proliferation of data. This will allow more applications and users to gain access to data in a meaningful time frame. IBM has doubled the possible number of FC and FCoE ports, giving any rapidly growing enterprise the flexibility required to protect their investment well into the future, by being able to process more data more quickly, when that need arises.

With twice the storage capacity of its predecessor, the V7000 can be configured with up to 20 enclosures per control unit, for almost 4PBs of raw capacity with these enclosures connected via 12 Gb per second SAS ports. Each 2U control enclosure can support up to 24 SFF[9] drives. The V7000 cluster has a maximum of 1056 SFF or 960 LFF drives. In order to protect the investment already made in Storwize infrastructure, IBM has enabled the new V7000 to cluster with existing systems, allowing non-disruptive

---

[6] See **The Clipper Group Navigator** entitled *IBM Enhances Storwize V3700 One More Time - Increasing Storage Capabilities for the Smaller Data Center,* dated May 6, 2014, and available at http://www.clipper.com/research/TCG2013011.pdf.

[7] Without compression.

[8] Each clustered system has a simplified single-system management interface for ease of use.

[9] SFF=Small Form Factor (2.5" wide).

---

[5] SMB=Service Message Block, a storage protocol.

migration and a growth path without discarding assets. *Storwize V7000 Unified* systems can use a mix of old and new V7000 block systems as their back-end storage too.

What is the impact on performance? As depicted in Exhibit 1 at the right, even with real-time compression, the new IBM Storwize V7000 has significantly more IOPS performance (around 40% more) than when stored uncompressed on the prior model.

In addition to improving the hardware infrastructure, IBM has taken steps to also improve the storage efficiency of the entire Storwize family (including SVC) with enhancements to IBM's *Easy Tier* software to enable the dynamic migration of data among three separate tiers[10] – to ensure that your data is stored on the devices that are most appropriate to their urgency. The Storwize family has also been configured to support storage pool balancing to ensure the highest storage efficiency possible with no bottlenecks.

Now, let's take a look at some of the important functional and performance enhancements in detail.
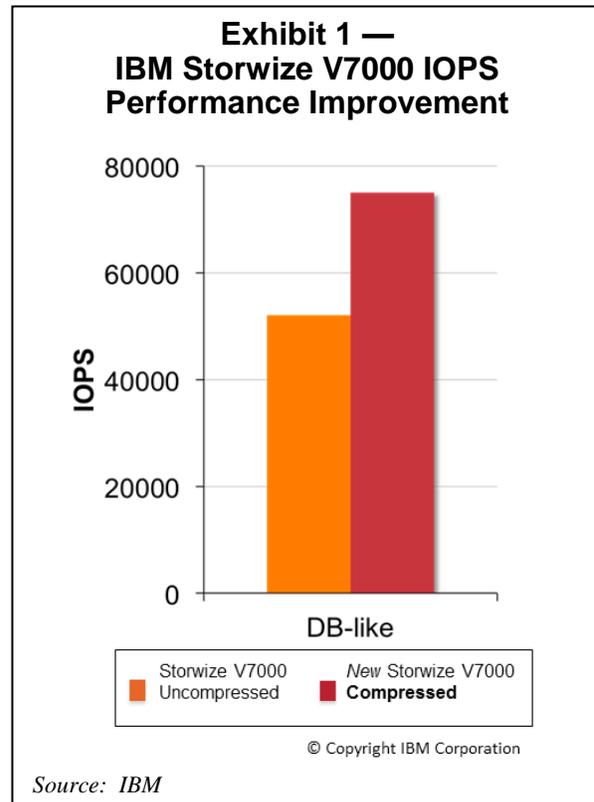
IBM Storwize Enhancements

Real time Compression with Intel Quick-Assist Technology

The IBM SVC, the IBM Storwize V7000, and the Storwize V7000 Unified use a real-time compression architecture designed to enable the data center to store up to five times as much data in the same physical disk space.[11] IBM RtC maximizes the benefits of compression with an increased number of compressed volumes and it can be used with active primary data without the need to reserve space for uncompressed data waiting for post-processing..

With the new hardware platform, IBM SVC and IBM Storwize V7000 utilize a compression accelerator based upon the Intel QuickAssist technology. There even is an option for a second Intel QuickAssist accelerator for the data center with the greatest performance needs, delivering a blistering 10 times improvement in real-time compression performance. The Intel QuickAssist technology was designed to provide additional computational resources for demanding workloads, such as cryptology, compression, and pattern matching, which also have high

**Exhibit 1 —
IBM Storwize V7000 IOPS
Performance Improvement**

Source: IBM

levels of complexity that continues to grow. Intel QuickAssist can do the following:

- *Accelerate performance* for demanding applications with specific hardware acceleration modules.
- *Support migration to advanced designs* using system-on-chip and multi-core processors.
- *Increase business flexibility* with solutions that fit changing business requirements without being tied to a specific-purposed accelerator.
- *Free up valuable cycles on server cores* to perform value-added and differentiating functionality.

The Intel QuickAssist technology makes it easier for vendors, such as IBM, to integrate built-in accelerators into their product designs. In fact, IBM has measured the I/O performance with Intel QuickAssist at about 40% better than the previous IBM Storwize V7000 without compression; this is good evidence that there can be no penalty for using real-time compression.

Easy Tier 3

IBM Easy Tier has been upgraded to support three tiers of storage in V7.3, up from two tiers. Now called IBM Easy Tier 3, it allows automatic and adaptive self management of fractional

---

[10] Up from two in the prior version.

[11] It depends on the nature and mix of the data being stored.

Exhibit 2 — IBM Easy Tier 3 – Automated Storage Tiering

- Constantly analyzes I/O and migrates data automatically to **optimize performance**

- Supports two or three tiers

- Deploy flash and enterprise disk for performance

- Grow capacity with low cost disk

Improve performance up to **3x**

… using as little as **5%** flash storage

Flash Arrays

Less Active Data Migrates Down   Active Data Migrates Up

HDD Arrays
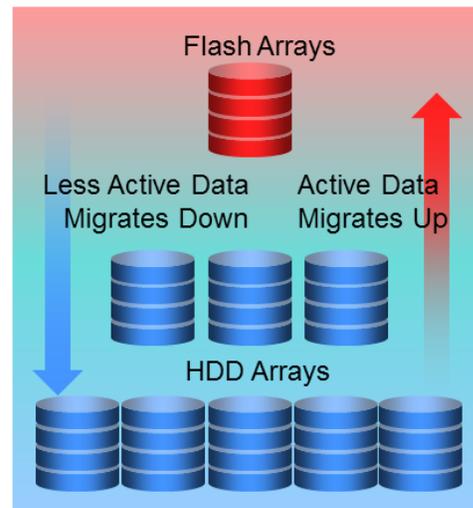
© Copyright IBM Corporation

Source: IBM

LUNs, whether stored on SSDs, high-performance disks, or high-capacity disks. It automatically and intelligently adjusts data allocation in order to achieve the best performance in an existing three-tier storage configuration. (See Exhibit 2, above.)

- **Tier 0** typically is for the applications that require the fastest storage, which usually means SSDs, which tend to be the most expensive per GB (but not per thousand IOPS).

- **Tier 1** is for applications requiring lesser performance than at Tier-0. Today, this usually means higher-performing, SAS-connected hard disk drives, which cost less than SSDs and high-performance disks. These drives spin at 10K-15K RPM.

- **Tier 2** is for less-demanding (in terms of performance) and less urgently needed data – at a much lower cost per GB than SSDs. Typically, this is where the bulk of the enterprise's data should be stored. The focus for Tier 2 is on the bottom line, and the cost per GB is much less than the two higher tiers. Typically, this is provided by SAS hard disk drives rotating at 7200RPM.

With Easy Tier 3, data is moved between the three tiers without administrator involvement or user/application awareness, depending on what data is "hot" and what is not. Easy Tier 3 is a no-charge upgrade for those data centers that have already deployed Easy Tier as part of v7.2.

Storage Pool Balancing

Auto-balancing in V7.3 enables the IBM Storwize family to automatically balance workloads among the arrays in a storage pool (including external virtualized storage), as part of the standard deployment. The balancing is based on the performance capability of the different arrays and is intended to optimize performance by eliminating "hotspots". When new storage capacity is added to a pool, workloads are automatically distributed across all the capacity, new and old. Storage resources remain optimized for both performance and cost objectives, reducing administrative costs and enabling data center staff to focus on strategic initiatives.

Continuing Storage Efficiency

These days, there's a lot more to storage than just hardware. A rich software set is what likely will separate the leader from the less desirable. So let's look at the functionality that IBM also includes in the Storwize offering.

- *Data Migration* enables a non-disruptive migration from existing legacy SAS storage to V7000.

- *Virtualization* of internal and external Storwize storage enables rapid, flexible application deployment, to simplify and standardize application management, and enable configuration changes to meet changing business needs. It also enables the data center to transform the economics of storage with consistent

data efficiency, virtually eliminating application downtime for data movement.

- *Thin Provisioning* enables applications to consume only the space they actually are using, not the total space that has been allocated to them.
- *IP replication* is provided via the integrated *Bridgeworks SANSlide* network optimization. It provides a lower-cost option for deploying remote replication, using artificial intelligence to improve network bandwidth utilization by a factor of three times. It enables the use of Ethernet connections for optional remote mirroring with no separate appliance required, and is transparent to both servers and applications. Faster replication cycles result in better remote data currency and faster recovery. Local replication also is available.
- *FlashCopy* allows the administrator to create copies of data for backup, parallel processing, testing, and development.
- *A simple and easy-to-use GUI* to enable storage to be quickly deployed and easily managed.
- *IBM Storage Mobile Dashboard* to provide monitoring and health check for Storwize systems from mobile devices.
- *VMware 5.5 and VASA block support*, in order to retain currency with the latest VMware capabilities.

Conclusion

Unquestionably, the largest data centers have the greatest demand, not only in terms of required capacity and performance, but also in terms of functionality. With the new releases of the IBM SVC, the IBM Storwize V7000, and the Storwize V7000 Unified, IBM has addressed those concerns with more performance, higher and denser capacity, along with all of the functionality required to run a Big Data data center in 2014 and beyond.

Changes made to the configuration of the new IBM SVC and IBM Storwize V7000 enable double the storage performance over the previous offering, but also, and perhaps more importantly, a tenfold increase in compression performance using the latest IBM RtC techniques and two Intel QuickAssist accelerators. What may have seemed difficult or impossible to do without the QuickAssist acceleration now is possible and practical. Storage with a real-time compression capability now should be the expected norm, with the economic benefits of

reduced physical storage far outweighing the costs of the accelerated compression, as is true with the IBM SVC and IBM Storwize V7000. In addition, the latter now can deliver twice the storage capacity of its predecessor.

IBM Easy Tier 3 with three tiers enables the data center to deploy the most highly efficient storage possible, putting the right data at the right tier at the right time, with the highest value data being stored in the most performant storage devices. With the capability to migrate that data between tiers automatically, the data center can store their data in the most cost-optimized manner – without administrator involvement.

What does this mean for the enterprise? First of all, with IBM SVC or IBM Storwize V7000 the data center can now virtualize even larger configurations than before for even greater consolidation advantages. Compression can be applied more extensively for greater benefit to more applications on fewer disk devices and thus at a lower TCO per GB stored. In addition, current storage can be migrated from existing SVC systems without disruption. If any, or all, of these benefits can improve your enterprise's delivery of data and answers in the requisite short time allowed while also focusing on the bottom line, you need to look into IBMs latest innovations on the IBM SVC and IBM Storwize V7000. IBM has the compression magic that you so badly may need!

About The Clipper Group, Inc.

The Clipper Group, Inc., now in its twenty-second year, is an independent publishing and consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 40 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 a Senior Contributing Analyst for The Clipper Group. Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. In 2002, 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, Clipper Notes, The Clipper Group Calculator, 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.

Disclosures

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.

After publication of a bulletin on *clipper.com*, The Clipper Group offers all vendors and users the opportunity to license its publications for a fee, since linking to Clipper's web pages, posting of Clipper documents on other's websites, and printing of hard-copy reprints is not allowed without payment of related fee(s). Less than half of our publications are licensed in this way. In addition, analysts regularly receive briefings from many vendors. Occasionally, Clipper analysts' travel and/or lodging expenses and/or conference fees have been subsidized by a vendor, in order to participate in briefings. The Clipper Group does not charge any professional fees to participate in these information-gathering events. In addition, some vendors sometime provide binders, USB drives containing presentations, and other conference-related paraphernalia to Clipper's analysts.

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.