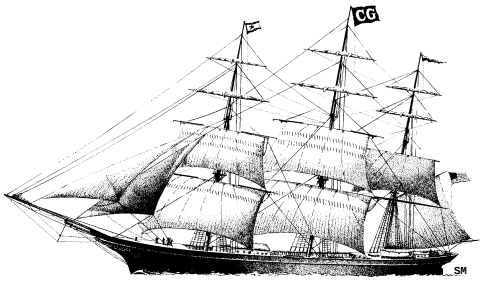
THE CLIPPER GROUP
Navigator™



Navigating Information Technology Horizons℠

Published Since 1993 Report #TCG2009049 November 4, 2009

Accelerating Database Access — Sun Shines with Flash-Based SSD

Analyst: David Reine

Management Summary

You have to admit it: we all get a laugh out of those stories about pampered rock stars. You know the ones where a star's dressing room has to be stocked with 24 bottles of *Evian* water, two jars of *Peter Pan* Peanut Butter (one creamy, one chunky), crackers from *Harrod's*, and, oh yes, a bowl of red *M&Ms*. How does one fill a bowl, a really big bowl, with only red M&Ms? Well, first you have to buy dozens of bags of M&Ms. Then you have to sort through them, putting the red ones in the bowl, and discarding the rest, or waiting for the next act that only wants blue M&Ms! This may sound frivolous, but it really is a waste if you only intend to use a fraction of the contents of each bag. The cost per M&M can become outrageously high and you are wasting a resource. The irony here is that same waste is going on in data centers around the globe where high performance applications demand extremely high access to stored data.

In order to achieve the desired rate of I/Os per second (IOPS), many data centers are forced to spread a relatively small database across hundreds, or even thousands, of enterprise-class disk devices, each with an IOPS rate measured in the tens. By short-stroking the device, i.e., only using a few, adjacent, outside cylinders and tracks, the data can be read at a higher rate by minimizing head movement. Unfortunately, the data center must then discard the rest of the M&Ms – oops, I mean storage capacity – because it cannot be accessed without compromising the high IOPS performance. This is where an old technology, solid-state disks (SSDs), with a new approach, NAND Flash memory, can improve the performance and reduce the total cost of ownership of your IT infrastructure, even though each SSD may carry a higher price tag *per GB* than the mechanical devices (HDDs) they replace. In fact, a single SSD can provide more IOPS than a chassis full of enterprise HDDs. For this reason, more SSDs are finding there way into the enterprise data center for a variety of applications. The key here is to find an integrated SSD solution that can fit into a multi-tiered storage environment, thus providing the most value for the data center dollar.

It should come as no surprise to anyone, then, that Sun Microsystems, with a newly supercharged emphasis on database access, has developed the world's fastest and most power efficient flash array for accelerating database I/O. The *Sun Storage F5100 Flash Array* can cut transaction times in half, while doubling application throughput. With scalability from 20 flash drives up to 80, the F5100 provides the flexible performance and capacity that your enterprise requires to meet growing business needs, while managing critical energy and floor space requirements in order to prolong the life of the enterprise data center facilities. In fact, Sun recently has completed an industry-standard transactional benchmark using its own *SPARC Enterprise Server* and a storage hybrid of Fibre Channel, SATA, and SSD Flash to attain the highest result ever recorded. To learn more about Sun's F5100 Flash Array, please read on.

IN THIS ISSUE

High Performance Data Center Issues

There has always been a requirement for a wide variety of storage devices in the enterprise data center. Mission-critical applications have always demanded high performance, high availability Fibre Channel (FC) devices to ensure rapid response with high availability. These Tier-1 devices are more expensive than the high-capacity Tier-2 SATA drives needed to satisfy the high-scalability storage demands of a rapidly growing business. The typical data center has also maintained a Tier-3 storage requirement to satisfy its archive and long-term storage needs. This Tier-3 requirement has usually been satisfied by tape, due to its low cost structure and low energy demands.

Tier-1 FC requirements can be further subdivided in terms of performance and capacity, with 15K RPM devices carrying a premium over the 10K RPM drives. The same is true for Tier 2, where low-cost 1TB 5400 RPM and 7200 RPM SATA drives often can provide more value to the data center in terms of business-critical application need.

Recently, however, high-performance applications with extremely high IOPS requirements have been placing a tremendous strain on the data center storage budget. In fact, some applications with the need for speed are demanding hundreds of thousands of IOPS to satisfy "C" level management requirements. In order to meet this demand, the data center has been deploying their mission-critical FC resources with a "short-stroke" architecture, only accessing the outermost cylinders of the disks in order to accelerate the response. In order to achieve the required level of IOPS, many enterprises are deploying thousands of expensive, short-stroked FC devices, filling banks of racks with energy-consuming disks, usually loaded at less than 10%. This is a waste of resources and budget dollars, both CapEx and OpEx.

In order to reduce the total cost of ownership (TCO) of the IT infrastructure, many data centers are implementing a new Tier-0 storage class for applications with the highest IOPS performance requirements. This premium Tier-0 level is being deployed with solid state disks (SSDs) configured with Flash memory[1].

The Role of SSDs

Solid-state flash drives can accelerate operational performance by improving application response times. They can increase storage efficiency by doubling server throughput while reducing space and energy requirements. They can help to reduce the TCO of the IT infrastructure by eliminating the necessity to add additional servers and HDDs to boost performance. They will also increase reliability because SSDs contain no moving parts.

Where do Flash SSDs belong? In the server, as direct attached storage (DAS), or in the storage array (SAN)? The answer is both, as the amount of SSD storage required will determine the deployment. While the cost per GB for SSDs is higher than HDDs, the $/IOPS is much lower. Flash, however, is much faster than HDDs but slower than the DRAM often found within the server. While DRAM speed is measured in nanoseconds, SSD speed is only measured in microseconds. SSDs, however, are significantly faster than HDDs, with speed measured in milliseconds. DRAM, unfortunately, is still very expensive, has limited capacity, and is volatile[2]. This makes SSDs more attractive to the data center for applications with very high IOPS requirements.

An SSD can replace a drawer of HDDs; a drawer of SSDs can replace a rack of HDDs. The relatively higher cost per GB of SSDs, however, often leads to the deployment of a hybrid environment with SSDs being used for high performance, and SATA drives deployed for high capacity. Some mission-critical applications can still be satisfied with a modest amount of traditional FC drives, however, leading to the deployment of a multi-tiered architecture of SSD, HDD, and SATA drives.

SSDs bring a new level of performance to the storage array, delivering over 100 times more IOPS to operational performance than HDDs. With no spinning media, SSDs reduce latency, lower energy consumption, and eliminate head movement, improving reliability and performance, while lowering TCO. In addition, with a familiar format, the IT staff can rapidly deploy SSDs into existing environments with the same form factor as traditional HDDs. With data stored on integrated circuits that can with-

[1] See the issue of Clipper Notes dated February 10, 2009, entitled *A New Tier of Storage Appears – Faster, Solid-State Drives State Their Case*, and available at http://www.clipper.com/research/TCG2009006.pdf.

[2] *Volatile* means that data is lost when power is removed, as in a power failure. Non-volatile memory doesn't require constant power to preserve its contents.

stand shock and vibration, SSDs last longer, reducing diagnostic time and minimizing the number of devices that need to be replaced.

There are two types of flash memory – *NOR* and *NAND* – that form the circuitry of flash devices. With long erase and write times and large voltage requirements, NOR flash is well-suited for random access to data that needs to be updated infrequently. NAND flash, on the other hand, provides block access to data and supports greater storage densities, provides greater endurance, and costs less per unit of storage, making NAND flash the more commonly implemented variety.

There are two types of NAND flash technology: *single-level cell* (*SLC*) and *multi-level cell* (*MLC*). SLC devices store a single binary value in each memory cell, while MLC supports four or eight values per cell and may contain two or three storage bits. **Because of its longer lifespan and better performance SLC NAND flash devices are preferred for enterprise applications.** However, there are some very attractive use cases for MLC flash when coupled with new designs in the Flash Memory Controller firmware for enterprise applications. Many of these are read-intensive applications, such as caching and situations that are dominated by reading of databases, like those found in Telco applications. Capacity, as opposed to performance is the main reason to consider MLC. Sun's *F5100 Flash Array* is based upon the high-performance, highly-reliable NAND SLC non-volatile, solid-state technology, designed to deliver the maximum level of I/O performance, with a minimum consumption of power and space.

Sun's F5100 Flash Array

Sun's F5100 Flash Array is a high-performance attached array for I/O intensive workloads. It can redefine how your most frequently accessed databases perform, accelerating application response time, and providing the data center with unprecedented speed, scalability, and TCO savings. The F5100 can enable the IT staff with the capability to replace thousands of power-hungry HDDs with energy-efficient SSDs, consuming a mere fraction of both the power and floor space of racks of HDDs required in a traditional disk deployment. The IT staff can improve database performance in new or existing applications, immediately, without data migration or extensive tuning, in addition to reducing the noise pollution caused by spinning media. The F5100 is optimized for the applications with the heaviest read/write workloads, such as *Oracle OLTP* databases, indexes, hot tables, and data structures. The F5100 delivers improved business productivity where I/O performance, high-transaction throughput, and low latency are most critical. See Exhibit 1, above, for the benefits of Sun's F5100.

Exhibit 1 — F5100 Benefits

- **Faster application response times** – reduces latency, eliminates storage-related I/O bottlenecks with over 1.6 million read IOPS and 1.2 million write IOPS.
- **Outstanding price/performance** – on a cost per IOPS basis – reducing operating costs by up to 90%.
- **Increased efficiency** – with the performance of 3,000 HDDs in 1RU of rack space consuming only 300 watts with a typical read/write mix.
- **Higher reliability** – due to non-volatile enterprise-class solid state Flash technology, with redundant power and cooling.
- **Lower energy costs and noise pollution** – due to the lack of any rotating media.
- **Greater productivity and lower TCO.**

Source: Sun Microsystems

The F5100 is a scalable array with from 20 modules up to 80. Fully configured, it can support up to 2TB of raw storage in 1RU of rack space. It accelerates database application performance by 2 times and cuts transaction time in half. The F5100 will cut energy consumption by up to 100x, depending upon configurability, and supports databases such as Oracle and *MySQL*, as well as working well with *DB2* and *Sybase*. Sun configures 32GB Flash drives in the F5100, with 24GB useable. The remainder is kept in reserve for wear-leveling to prolong the life of each SSD module.

With up to 64 SAS lanes (with 16x4-wide ports), four domains, and SAS zoning, the F5100 assures maximum performance, throughput, and configuration flexibility to facilitate easy deployment and scalability. The F5100 was designed from the ground up for maximum reliability, with advanced wear-leveling techniques, data protection, RAID mirroring, and solid-state robustness, using flash modules with

an MTBF of two million hours. The F5100 also incorporates super capacitor backed *Energy Storage Modules* for write cache and meta-data protection to assure the highest level of data integrity. The F5100 also takes advantage of Sun's *Solaris Fault Management Architecture* to detect and diagnose any problems before they can affect a mission-critical application. With SSDs, there is no need to deploy massive pools of high-cost, high-performance 15K RPM drives; the data center can deploy more economical high-capacity SATA disk drives, instead. In addition, the data center staff can take advantage of features like host-based RAID protection from *Oracle ASM*, or data integrity features in *Solaris ZFS*, such as automatic data integrity checking and correction, to automate data management, enhance data protection, and optimize storage hierarchy for maximum performance at the lowest cost.

High Performance is not simply a marketing buzzword for the F5100. Sun has recently reported the highest rating ever for the Transaction Processing Performance Council TPC-C benchmark – **7,717,510.6 tpmC** throughput at a price/performance level of **$2.34/tpmC**. Using a configuration consisting of Sun *SPARC Enterprise T5440* servers running *Solaris 10*, *Oracle Database 11g*, and hybrid storage, focused around 60 Sun Storage F5100 Flash Arrays, Sun has eclipsed all previous TPC standards with the first new Sun TPC-C listing in over eight years. This rating is over 25% faster than the previous record, based upon IBM's *Power 595* architecture with *Power 6* technology. Sun's vast improvement in performance primarily is attributable to its F5100 Flash Array technology.

Conclusion

With the pending acquisition of Sun by Oracle, it should come as a surprise to absolutely no one that Sun would try to optimize their IT infrastructure in support of accelerating on-line transaction processing. The fact is that we are surprised by the extent to which they have accomplished this goal! This is a testament to the innovative technology employed by Sun with their SPARC Enterprise Servers and their F5100 Flash Array technology and the tuning done in conjunction with Oracle to improve database performance. Sun has broken the previous record established by IBM in June 2008, by more than 1.7M TPM while keeping price/-performance under control.

Fast access to business-critical information is the lifeline that enterprises in every industry rely upon to accomplish their goals. With rapidly expanding databases and new applications being deployed every day into virtualized environments, the ability to handle more IOPS may well be the gating factor between success and failure for every enterprise. With the introduction of a fully-integrated flash array, Sun has enabled the data center to increase not only the number of transactions being processed, but also control the number of devices being deployed and the amount of energy being consumed by the IT infrastructure. The F5100 Flash Array not only helps the data center to improve overall system performance, but it enables the data center to lower the total cost of ownership of the IT infrastructure. Sometimes you have to spend money in order to save money!

If your enterprise is trying to improve system response time while putting a cap on the amount of storage being wasted, you should look into the possibilities provided by solid-state storage and the advantages of redeploying your mission- and business-critical OLTP processes with Sun's F5100 storage.

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

David Reine **is Director, Enterprise Systems for The Clipper Group.** Mr. Reine specializes in enterprise servers, storage, and software, strategic business solutions, and trends in open systems architectures. 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**, 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.