Industry Trends and Technology Perspective White Paper

”Achieving Energy Efficiency using FLASH SSD”

Leveraging SSD as part of a tiered storage building block to support server and storage consolidation to achieve energy efficiency

By Greg Schulz

Founder and Senior Analyst, the StorageIO Group

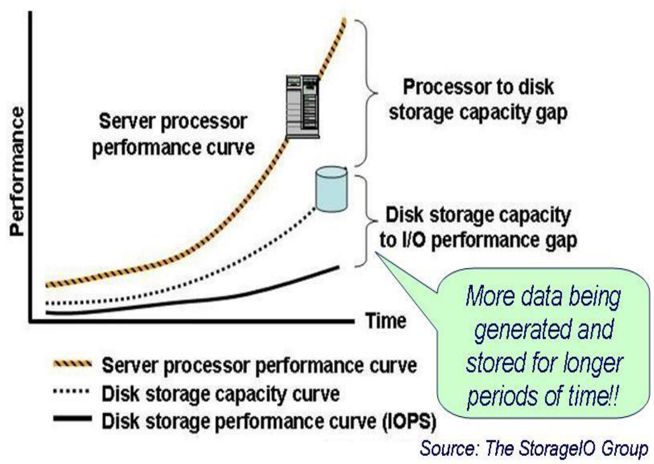
December 7th, 2007

Reducing energy consumption is an important topic for many IT data centers. While there is a lot of discussion about reducing energy by doing less, another form of energy efficiency is doing more with less power. For example, processing more transactions in a given time frame is an effective approach to energy efficiency for time and performance sensitive applications. This paper looks at how solid state disk (SSD) and, in particular, FLASH based SSD can be incorporated into an energy-efficiency strategy as part of a tiered storage architecture to address power and environmental concerns while reducing server and storage I/O performance bottlenecks.

Introduction

IT organizations of all sizes have to do more with their available resources such as power, cooling, servers, networks and storage while supporting growth. Meanwhile, most IT data centers have or will be facing bottlenecks as their IT resources become strained due to consolidation, growth and new applications while maintaining (or improving) service delivery to IT customers and users. At the same time, coverage of green IT, green data center and green storage issues and awareness continues to grow. Just as there are many different facets of being green, there are also several approaches to addressing various needs and requirements. These range from energy avoidance, such as disabling or powering IT equipment down; to using energy more effectively to get more work done while using less power.

Server and Storage I/O performance Gap



The quantity and performance capabilities of servers continue to grow with a corresponding decrease in cost, physical footprint and electrical power required. Similarly, the capacity and reliability of magnetic hard disk drives (HDD) continue to increase while the physical size, power consumption and cost decrease from previous generations.

However, the performance and capacity gap between server performance, storage capacity and performance continues to follow a several decades old trend that is becoming more pronounced. That trend is that I/O performance and, in particular, I/O operations per second (IOPS) and latency have not kept up with the corresponding improvements in server performance, server memory capacity and HDD storage capacity. While the trend[1] is not new, it is becoming more apparent and pronounced with storage device capacities currently at, and soon to be exceeding, 1TByte for 3.5” HDDs.

Countering the storage I/O performance gap

To compensate for lack of I/O performance and the resulting negative impact to IT users of poor response time or increased latency, a common approach is to add more hardware to mask, virtualize, or move the problem elsewhere. A technique that has been used to counter the growing server and storage capacity I/O performance gap has been to deploy storage systems with a large number of HDDs combined with mirrored battery backed-up read and write cache. There are, however, some collateral impacts and effects to the approach of using a large number of HDDs to compensate for a lack of performance including excessive power consumption and cooling requirements, under-utilized storage capacity, and application data hot spots.

Adding more hardware often leads to extra storage capacity being added to make up for a shortfall in I/O performance. By over-configuring to support peak workloads and prevent loss of business revenue, excess storage capacity must be managed throughout the non-peak periods, adding to data center power and cooling demands as well as associated management costs. The resulting ripple affect is that now more storage needs to be managed, including allocating storage network ports, configuring, tuning, and backing

[1] Read about IT data center bottlenecks and the growing server-storage I/O performance gap in the industry trends and perspectives report “Data Center I/O Performance Issues and Impacts” at www.storageio.com/xreports.htm

up of data. This can and does result in environments that have storage utilization well below 50% of their useful storage capacity.

An alternative to HDD is solid state disk, commonly referred to as SSD, which is composed of either a volatile (RAM) or non-volatile (FLASH) memory as the primary storage media. Using this type of storage compared to HDD, the cost per IO is drastically reduced as each SSD can support tens to hundreds of thousands of IOPS as compared to 100-300 IOPS for a HDD. Power consumption is also noticeably reduced per IO when using FLASH which uses significantly less power than HDD as there are no mechanical parts involved in storing data.

Countering under-utilized storage and IT resource capacity

Over time the under-utilized storage capacity can be consumed by application growth and data that needs to be retained for longer periods of time. However, unless the capacity that is going to be consumed by application and data growth is for static or dormant data, any increase in I/O activity would further compound any I/O performance problem being addressed by sparsely allocating storage in the first place. There is a similar activity taking place with servers that have been seen as inexpensive to acquire and, until recently, not as much of an issue to operate, thus leading to the belief that since hardware is in-expensive, using more of it is a workable solution to an application problem.

With the shift in focus and awareness pertaining to rising energy costs and, more importantly, on data center power and cooling bottlenecks, under-utilized servers are being consolidated using virtualization techniques including VMware, Virtual Iron or Xen among others. The result is that power, cooling and floor space resources are being freed up or re-allocated, physical servers whose workload or applications were consolidated are being re-deployed to support growth, and unused servers are being disposed of to make room for new IT equipment.

Power, cooling and facilities bottlenecks

A common issue the StorageIO Group hears from data center IT professionals in organizations of various sizes and geographical locations, especially in the United States, is limited availability of electrical power, cooling and floor space to support application and business growth. IT data centers rely on electrical power and there is a finite supply of electrical power generation (production) and transmission networks (delivery) capabilities. Meanwhile, there is the need for expansion of servers, storage and networking devices to support IT service (demand) delivery enhancements, business growth and new applications. IT data centers must balance IT resources, including power, cooling, floor space, server, network and storage (capacity and performance), with increased demand for IT services delivery.

There is plenty of discussion about the benefits of a green data center, but the reality today is that most IT data center personnel are concerned with addressing limits on electrical power availability and cooling capabilities. For some data centers this means reducing electrical power consumption for cost containment and budget purposes. The more common theme, however, with IT data center professionals is that there is simply not enough available power or cooling capacity to support more servers, storage and networking capacity and thus inhibiting business growth.

Data center power availability issues range a lack of power generating capacity to insufficient transmission or delivery networks between power generation sources in a given geographical area or, even, specific buildings. In addition to limits or bottlenecks on power generation and delivery, cooling capacity in many data centers is tapped out as is backup or standby power capabilities.

Preventing the perfect storage I/O performance and power storm

Ironically, some of the consolidation currently taking place in data centers is to boost the utilization of IT resources that have been configured to achieve a given level of performance vs. higher capacity utilization to compensate for previous technology performance limitations. The solution is to address the problem rather than moving the bottleneck elsewhere (which is rather like sweeping dust under the rug).

For example, if performance is an issue then simply consolidating storage to larger capacity disk drives and fewer storage systems may reduce your energy and cooling requirements; you would, however, also miss your application service and performance objectives. IT data centers need to balance the performance, availability, capacity and energy demands of different applications to meet different classes or tiers of application service level objectives, as seen in Figure-2.

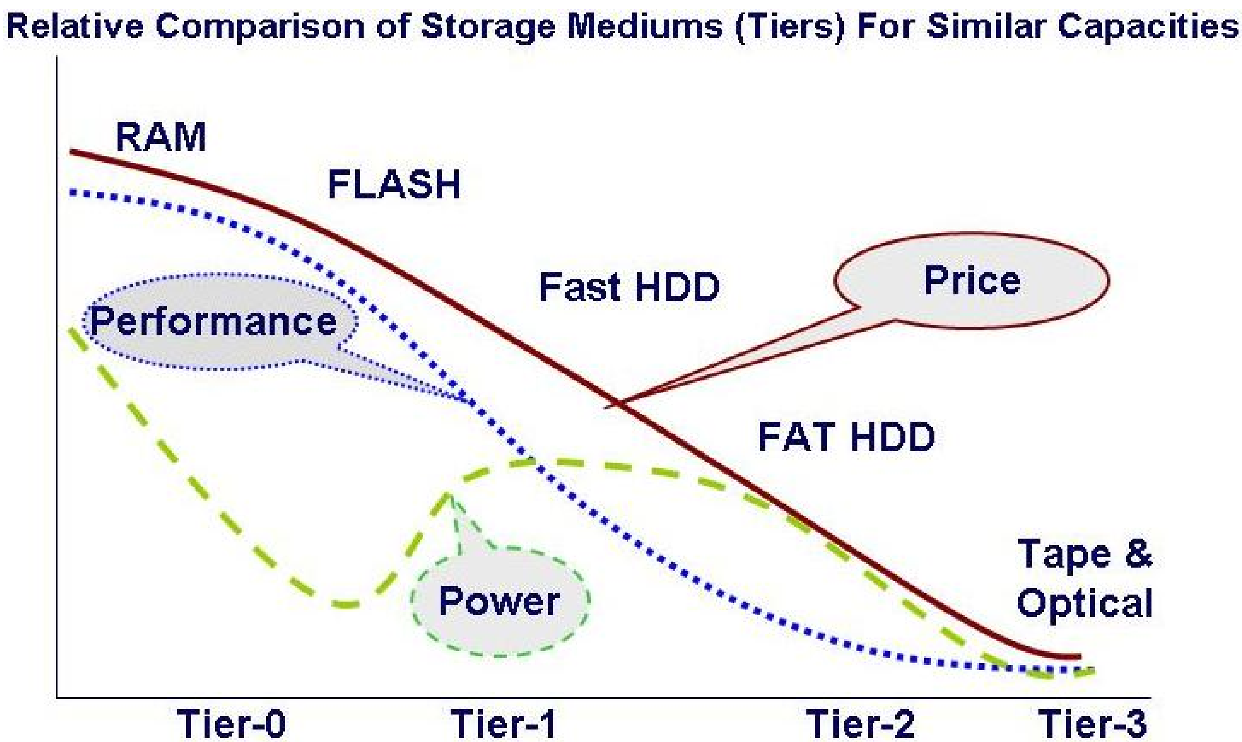


Figure-2: Service characteristic of various classes or tiers of storage

Figure-2 shows different tiers or categories of storage (SAN or NAS) independent of specific architectures (for example cache centric or monolithic frame based, modular, distributed or clustered) contrasting performance, availability, capacity and energy consumption as well as relative cost positioning. Alignment of the most applicable tier of storage to applications needs is an effective technique, along with others such as data footprint reduction, to address data center bottlenecks without continuing the vicious cycle of sparse storage allocation and subsequent consolidation.

From a power and performance perspective, SSD provides very good IOPS per watt or bandwidth per watt of energy used and capacity per watt. There is, however, a trade-off of cost and capacity. As an example, 1 TByte of usable FLASH based SSD can fit into 2 rack units height (2U) in a standard 19" rack consuming 125 watts or .125 kWh of power capable of delivering 100's of MB/sec bandwidth performance. The FLASH SSD would compare with two other storage tiers, those being storage centric high capacity low cost disk drive or high performance medium capacity disk based storage systems. The key for SSD is that no moving parts are involved other than cooling fans, and, thus, less power is required to deliver higher levels of performance.

Solid state disk can be a good fit for I/O intensive and time sensitive applications or workloads where the number of disk drives and subsequent power consumption can be reduced as part of a tiered data storage infrastructure. Historically, the high price of RAM and relative low capacities compared with hard disk drives (HDD) has left SSD to ultra-high performance time sensitive applications. More recently, with the increasing capacities, reliability and decreasing cost of FLASH based memory, SSD combing FLASH and DDR RAM memory have become more practical and affordable. Relatively speaking, FLASH memories are on a faster price per GByte decline and capacity increase basis as compared to other storage mediums. As this trend continues the amount of storage in absolute, and as a percentage of total storage, that is best fit for a SSD will expand due to the performance advantages and power savings associated with SSDs.

Storage Tier Description	Zero Cache or SSD	One Primary	Secondary Near-line or off-line
Value proposition	Time is money, lost time is lost money or life threatening	Balance of disk based performance, capacity, cost, and availability	Low cost, low performance high capacity
Service Characteristics	I/O Intensive, relatively small data sizes that are time sensitive or hot spots for larger application	High performance (IOPS or bandwidth), time sensitive more capacity	Storage capacity centric, infrequent access of static data, tolerance for performance
Footprint	One 19” Rack	One Cabinet	One Cabinet
Devices / Drives	20 x 1TB FLASH with DDR/RAM Cache	165x15.5K 300GB 4GFC	8 x 42TB (336 x 1TB 7.2K SATA)
Capacity	20TB	49.5TB	336TB
Power used[2]	2.5 kWh	4.96 kWh	6 to 2.8 kWh
Annual power	21.9 mWh	43.5 mWh	52.6 to 24.5 mWh
Performance	***Very fast***	Fast	Slow to very slow
Drive Type	DDR RAM or NAND FLASH or caching devices	146GB and 300GB FC & SAS 15,000 RPM HDD	750GB and 1TB SATA HDD, IPM[3] and some MAID and RHDD[4]
Comparison Metrics	Activity per watt including IOPS or Bandwidth per watt, transactions per watt, users or sessions per watt	Activity per watt including IOPS or Bandwidth per watt, transactions per watt, users or sessions per watt	Bandwidth per watt and capacity per watt trading performance for lower energy used

Figure-3 – Different tiers of storage and service characteristics in the same footprint

Even with the continued drop in price of RAM/DDR and FLASH based SSD along with the increase in capacity, for most IT data centers and applications there will continue to be the need to leverage tiered storage, including HDD based storage systems, as seen in Figure-3. What this means, however, is that a balance of SSD for low latency high I/O or performance hotspots along with storage capacity, high

[2] Excludes power for cooling based on normal energy use per vendor spec sheets with 1,000 watts being 1 kWh and 1,000 kWh being 1 mWh
[3] Intelligent power management with variable power consumption to level of service and performance
[4] Removable hard disk drive (RHDD)

performance disk drives in the 146GB and 300GB 15,000 (15K) RPM class are a good fit with 500GB, 750GB and 1TB class disk drives for storage capacity centric workloads.

For active storage environments that do not require the ultra low latency of SSD but need high performance and large amounts of affordable capacity, energy efficient 15K RPM Fibre Channel and SAS disk drives provide a good balance between activity per watt, such as IOPS per watt and bandwidth per watt, and capacity as long as the entire capacity of the drive is used to house active data. If "short-stroking," where only part of the disk is used to improve performance (usually the outer tracks for sequential bandwidth or inner tracks for random IOPS), the HDD is not being used for its storage capacity making such applications good candidates for using SSD. For dormant data and ultra large storage capacity environments with a tolerance for low performance, larger capacity 750GB and 1TB "fat" HDDs that trade I/O performance for higher storage capacity provide a good capacity per watt of energy story.

Applications that can benefit from SSD or some combination of SSD and high performance disk drives include: performance sensitive databases, video and electronic surveillance or security, financial currency exchange and trading, transaction logs and journals, telecom, simulation or modeling for energy exploration, non-linear editing, web and high performance file serving, OLTP, e-commerce as well as others with data and I/O hot spots including Email or text messaging.

What Figure-3 does not show is a relative performance indicator in terms of throughput for sequential bandwidth or IOPS for random workloads. Some storage system vendors will simply quote the HDD manufacturers' rated performance metrics aggregated for the number of HDDs independent of what a storage system can actually perform, providing a false sense of performance capabilities.

Avoid assuming that a storage controller can actually utilize and leverage the full disk performance capability; in many systems this is, surprisingly, not the case. Short of actually testing different storage systems and architectures with your applications under different workload conditions, applicable benchmarks can help provide insight as to how different solutions compare. Exercise care to use benchmarks that are applicable and related to your application needs and even then use those comparisons as only a relative guide to how your applications will perform.

FLASH vs. DDR/RAM

- ✓ FLASH, also referred to as NAND, is relatively low cost and persistent memory that does not lose its content when power is turned off. USB thumb drives are a common example.
- ✓ DDR/RAM is dynamic memory (like what in your computer) that is very fast but is not persistent and data is lost when power is removed. DDR/RAM is also more expensive than FLASH.
- ✓ Hybrid approaches combine FLASH for persistency, high capacity and low cost with DDR/RAM for performance.

Another variable in all of the previous examples is how the storage system is configured in terms of RAID level for performance, availability and capacity. The various RAID levels impact energy consumption based on the number of disk drives or SSD (FLASH or DDR/RAM) modules being used. Ultimately, the right performance, availability, capacity and energy consumption (PACE) balance should be weighed with other decision and design criteria including vendor and technology preferences.

Trends, Tips and Recommendations

While magnetic HDDs continue to decline in price per capacity, FLASH price per GB is declining at a faster rate which makes storage using SSDs a very complementary technology pairing to balance

www.storageio.com P.O. Box 2026 Stillwater, MN 55082 651-275-1563 info@storageio.com

performance, availability, capacity and energy across different application tiers. An industry trend is for more storage systems to be able to intelligently use the built-in capabilities of modern HDDs to reduce power by parking disk heads when not in use, stepping down the RPM or speed of the drives while going into low consumption modes. Many storage vendors are either supporting multiple power settings to vary the power consumption of the disk drive without having to spin disk drives completely down.

General tips and recommendations for using SSD to support server and storage consolidation include:

- ✓ Look to SSDs to alleviate I/O performance bottlenecks by consolidating to improve utilization
- ✓ Balance application service requirements to applicable tier of storage as part of consolidation
- ✓ Energy efficiency can be achieved by using higher capacity disks for static and inactive data
- ✓ Time sensitive data applications can achieve energy affiance by doing more work per watt of energy
- ✓ FLASH and RAM based SSD storage solutions provide good IOPS per watt for energy efficiency
- ✓ FLASH provides good capacity per watt for energy efficiency
- ✓ When comparing either activity or capacity per unit of energy, also look at total energy consumed
- ✓ FLASH provides data persistence and RAM has low latency making for a complimentary solution

Summary and conclusion

From an enterprise and IT perspective, companies such as Texas Memory with the RamSan-500 can now store up to 20 TBytes of data in a standard data center cabinet drawing about 2.5kWh of electrical power. Not only is the RamSan-500 an effective means for reducing electrical power when compared on a capacity per watt of energy basis the FLASH based RamSan-500 also has a good activity per watt value with the ability to support tens of thousands of IOPS or 100s of MB/sec bandwidth without having to occupy several data center floor tiles worth of space or draw 15-30 kWh of power.

Some Applicable Benchmarks Include

- ✓ Storage Performance Council - SPC
- ✓ Microsoft Exchange Simulation - ESRP
- ✓ Transaction Processing Council – TPC
- ✓ The best is your own application!!!

IT organizations are realizing that in addition to power conservation and power avoidance, addressing time sensitive applications with performance enhancements can lead to energy efficacy. By leveraging more energy efficient solutions that are capable of doing more work per unit of energy consumed, for example transaction per watt, IOPS per watt or bandwidth per watt is similar to improving the energy efficiency of an automobile. Keep performance, availability, capacity and energy (PACE) in balance to meet application service requirements and avoid introducing performance bottlenecks in your quest to reduce or maximize your existing IT resources including power and cooling.

About the author

Greg Schulz is founder and senior analyst of the StorageIO Group and author of the book *Resilient Storage Networks — Designing Flexible Scalable Data Infrastructures* (Elsevier Digital Press).