IDC
Analyze the Future

WHITE PAPER

The Need to Standardize Storage Device Performance Metrics

Sponsored by: Seagate Technology

John Rydning David Reinsel
Jeff Janukowicz
September 2008

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

IN THIS WHITE PAPER

In this IDC white paper, sponsored by Seagate, IDC lays the foundation and discusses the importance for storage device vendors to define and coalesce around a set of common performance metrics that reflect real-world performance. This paper explores the need for hard disk drive (HDD), flash-based solid state drive (SSD), and hybrid hard disk drive (H-HDD) vendors to develop standard metrics and testing methodologies in an effort to assist system OEMs and other storage integrators in evaluating and choosing the storage device that best fits their system design, performance metrics, and application goals.

SITUATION OVERVIEW

Hard disk drives are over 50 years old, and the technology has developed over time to become one of the most trusted devices in the computing environment. Today, HDDs are an ideal choice for storing data given their high capacities, good performance, and low prices. HDDs are the de facto standard for storage in most systems available on the market, especially those with large capacity requirements. Given the overall performance and value of HDDs, it is easy to see why they are found in such a wide array of applications from everyday consumer electronics devices and personal computers to servers and storage systems, the backbone of today's digital infrastructure.

Today's computing market is more pervasive than ever before, and datacenters are storing an ever-increasing amount of digital content. As a result, IDC estimates that datacenter storage requirements will increase between 50% and 60% per year over the next several years. Yet, commensurate with the increasing storage requirements is the quest for more performance, better utilization, and faster access to stored data.

What is sometimes overlooked is the fact that HDDs and many of the systems into which they are integrated have evolved together over the past half century. The disk drives and the systems in which they are used have been optimized around the characteristics of rotating magnetic storage to take full advantage of the benefits they provide, as well as to mitigate the challenges they present for any given usage scenario. However, the emergence of more advanced computing power and evolving usage models has opened the door for system OEMs to consider new storage technologies as they bring new products to market and improve the designs of existing solutions.

SSD technology, based on flash memory, and hybrid HDDs, based on a combination of flash memory and rotating magnetic storage, are two storage technologies garnering much interest recently. Both technologies promise to use solid state semiconductor technology to improve performance, lower power consumption, and increase reliability. While these technologies are not new, they do present a new set of challenges to address, as well as a new set of benefits on which to capitalize. Thus, storage system OEMs and storage integrators are now challenged in evaluating a broader array of components and solutions while choosing the storage device that best fits their system design, performance metrics, and application goals. In the end, understanding the real-world performance of storage devices and systems is a key factor in the decision-making process.

Will the Real-World Performance Metric Please Stand Up?

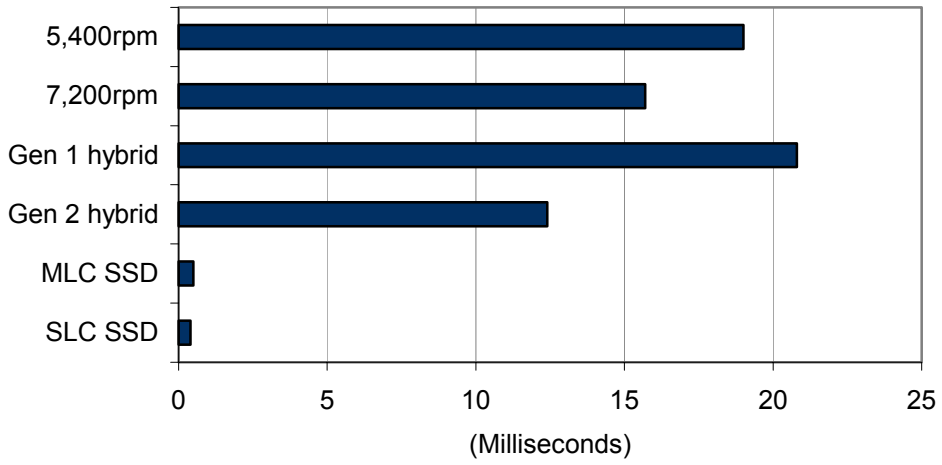
IDC research reveals that the real-world performance of storage devices often differs, sometimes substantially, from the commensurate specifications published for these devices. A number of valid reasons could exist for this disparity, but the disconnect could result in reputation management for a technology or vendor, substandard system performance, or worse yet, a total system failure (and unplanned downtime for a customer), none of which are positive for the industry.

While no set of storage device performance metrics can guarantee complete harmony between specified and real-world performance, it is vital for storage device vendors and their customers to develop standard metrics and testing methodologies to ensure successful integration at the system level.

To highlight the important of real-world metrics, IDC evaluated six storage devices under one computing platform using a common set of tests in a single setting. (See *Benchmarking Storage Options for PCs: The Results Are In — Exposing the Strengths and Weaknesses of HDDs, SSDs, and Hybrids*, IDC #213285.) As shown in Figure 1, each storage device can exhibit significant differences in performance when measuring just at the device level and in a controlled laboratory environment. In this example, a multilevel cell (MLC) SSD has an access time of 0.5ms compared with an access time of 15.7ms for a 7,200rpm drive — a 30x improvement.

FIGURE 1

HD Tach — Access Speeds



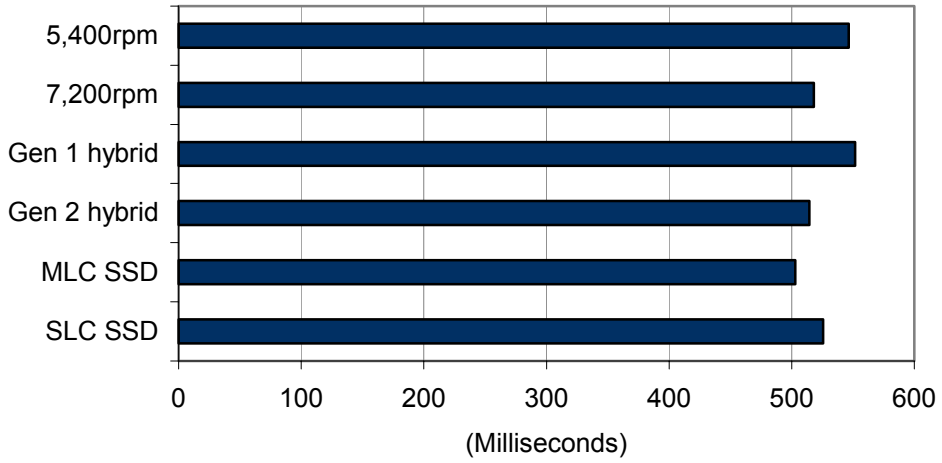
Note: Lower is better.

Source: IDC, 2008

However, once the device is integrated into a system, in this case a notebook PC, the perceived advantages, unless specifically designed around, can be diminished or disappear altogether. Figure 2 highlights this phenomenon when the same storage devices are integrated into a system and real-world tests are performed.

FIGURE 2

Internet Explorer Launch



Note: Lower is better.

Source: IDC, 2008

We see the Internet Explorer launch time for the MLC SSD is approximately 500ms compared with 525ms for the 7,200rpm drive. This result is nowhere near the improvement that might be expected based on the access time test results at the component level. Thus, application-specific benchmarks are needed as opposed to just component-level testing to give customers and systems integrators a fair idea of what kind of performance to expect from the integrated device.

Choosing the Metrics That Matter

Not every metric will impact system performance, and some metrics are less important than others. However, with any given application, there are a number of metrics that system OEMs and system designers will evaluate with a critical eye in order to ensure that their systems will perform as designed and, more important, as expected by their customers.

While power consumption, form factor, capacity, cost, and other metrics are important to consider for enterprise storage platforms, perhaps no other metrics are as important as those related to the speed at which the system reads and writes data. Computing applications are varied in nature — and so are the types of data stored for a given application and the corresponding performance requirements to read and write the data. Consider the three common storage system performance requirements discussed in the following sections.

Small Block Random Performance (Throughput IOPS – IO/s)

High-transaction applications such as those used in the financial services industry, airline reservation systems, or enterprise resource planning (ERP) applications have a substantial volume of random read and writes. The data moving to and from the storage system is often small files disaggregated into blocks of data, typically 4KB or 8KB in size, but the size can vary from system to system. The speed of a storage system to read and write small blocks randomly is measured as inputs and outputs per second (IOPS or IO/s).

A single performance-optimized SSD often provides a more than 1,000 IO/s capability to read and write 4KB or 8KB blocks. However, depending on the storage system design, the speed at which a given storage system writes data may be slightly or significantly slower than the speed at which it reads data. Therefore, the mix of reads and writes required by a given application in combination with the storage system design will be a key determining factor for overall storage system IO/s performance.

Additionally, a flash-based SSD can support output requests at much faster speeds than HDDs due to its superior access times. However, the write speed is typically not as fast as the read speed. Thus, storage system performance gains using flash-based SSDs will vary, depending largely on the mix of read and write input/output requests demanded by the computer application. For example, an early generation SSD that provides more than 1,000 read IOPS may yield lower IOPS in a real-world, mixed read/write application.

Sequential Performance (Transfer Rate – MB/s)

In contrast to applications that generate a high volume of small block random inputs and outputs, data transfer–intensive applications make large I/O requests for blocks of data that tend to reside adjacently to each other on the disk surface. Typical data transfer–intensive applications include streaming video or recording scientific or medical images.

For storage systems, data transfer rates usually are expressed in millions of bytes per second (Mbytes/s or MB/s). Performance-optimized HDDs can achieve a sustained read and write data transfer rate of more than 170MB/s and will be faster as the size of the data block increases. Flash-based SSDs can achieve data transfer rates that are similar to or faster than those of HDDs. Again, the mix of reads and writes for a given application along with the overall design of the storage system will affect the performance of the storage device.

Response Times (to Various Commands – ms)

Inputs and outputs for a storage system commonly are performed through a command/response sequence. The inherent design of an HDD may limit input/output performance since the read and write head inside an HDD must seek a specified data track mechanically. Measured in milliseconds, the seek time for performance-optimized disk drives can range from 3.5ms to 10ms.

Further, an HDD may have to wait for the spinning disk to rotate a specific block address on a given data track under the read and write head. Again, measured in ms, the latency for the block address to rotate under the read and write head for a performance-optimized HDD will vary with the spin speed of the rotating disk, but it is often 4.0ms for a full revolution or 2.0ms for a half revolution.

HDD seek-time and latency delays often can be overcome in storage systems through parallelism and overlapping operations where input data is disaggregated into a number of separate blocks that are distributed over multiple disk drives (aka, overprovisioning drives). In addition, a practice known as "short stroking" is sometimes used to reduce response times. Essentially, short stroking drives is limiting the available write surface of the drives' internal media to a small percentage of the outer tracks, thereby shortening seek times (while reducing the overall capacity of the drive significantly). But even when using these techniques, HDD rotational latency and seek times can impact storage system command response times negatively.

Flash-based SSDs eliminate the mechanical latencies inherent in disk drives and offer the potential for faster command response. But not all data associated with a given application will require quick response times.

Again, performance of HDD or SSD storage devices in a given storage system will vary, depending largely on the mix of read and write input/output requests demanded by the computer application.

Methodologies Matter

It is not enough to provide just the measurement results for any given metric. The context and methodology used in obtaining the measurement must also be understood and controlled.

Hard drive OEMs long have reported various metrics around access times, maximum data rates (both internal and external), and maximum sequential throughput, as well as "typical" or average performance for these measures that system OEMs and other storage integrators can use to evaluate a given storage device.

Integrating and Understanding SSDs

NAND flash-based SSDs are relatively new to enterprise storage, the vast array of and sometimes extreme workloads, as well as the benchmarking activities that support these business critical environments. This is not to say that SSDs are not up to the challenge, but challenges must be addressed to ensure that benchmark results are repeatable and accurate.

These challenges specifically are highlighted by the NAND flash write process that can have a direct impact on write performance. For example, while the read access time of flash typically is 100 µs, the write access time is 200 µs, However, this is only part of the story.

NAND flash is organized into blocks and pages of memory. A memory page may contain eight 512-byte logical block addresses (LBAs) for a total capacity of 4K. A page is the smallest unit that can be read or written (programmed) on many flash chips. For example, a write command for only LBA 0, which is 512 byte, will result in a physical write of 4KB of data to the flash chip. However, the flash chip contains a unit called an erase block that typically consists of 64 pages or a total of 256KB. The erase block is the smallest unit that can be erased. Therefore, if the example of writing to LBA 0 requires an erase to be performed prior to programming the page containing LBA 0 data, then the device must erase 256KB in order to be able to write the 4KB page that contains the 512-byte LBA 0.

To optimize SSD performance, designers use several techniques to minimize these write penalties. One strategy is to cache the write data until more data is received for other LBAs. For example, if the drive receives write commands to LBAs 0, 235, 689, 13, and 1000, the drive will write the data for all of these LBAs to the same 4K page. This strategy reduces the number of writes from a possible 5 to just 1. The result of this technique is that the drive needs to keep track of the locations where it wrote the data in some type of table for future reference. Unfortunately, if a sequential read command is received for LBAs 0 to 50, the drive needs to locate all of the LBAs and reassemble the data in the correct order. This strategy and others deployed by SSD vendors can impact the results of various benchmark tests.

The following dynamics should be understood and addressed in order to achieve accurate benchmark results:

- The data on the drive as it comes from the manufacturer may contain data patterns that are optimal to the device but that are not realistic if used by a given benchmark program.
  - Unlike HDDs, NAND flash-based SSDs can possess technology that can skip the writing of an LBA of data if the contents of the LBA are unchanged. For example, if the benchmark program requests a write of "00h" data pattern to an LBA, and the device determines that the LBA in question already has the "00h" data pattern, the device may not actually perform the write operation, hence artificially improving the write performance.
  - As received from the vendor, the drive may contain flash that has been pre-erased in order to enhance initial write performance. In this case, the initial write performance would be higher than it would be later with actual application usage.
- The device may contain additional memory overhead beyond the stated capacity as a designed method to improve performance and reliability. The device may use this actual pool of additional memory to delay write performance penalties due to erase cycles until a later point in time.

Ensuring Accurate Results

In order to ensure realistic and accurate benchmark results on SSDs, testing the drive only in the initial condition can thus provide unreasonably different results than in a real-world setting. The degree to which some of these dynamics are manifest depends on the design of the SSD and the vendor. Thus, the following points should be considered to ensure accurate results:

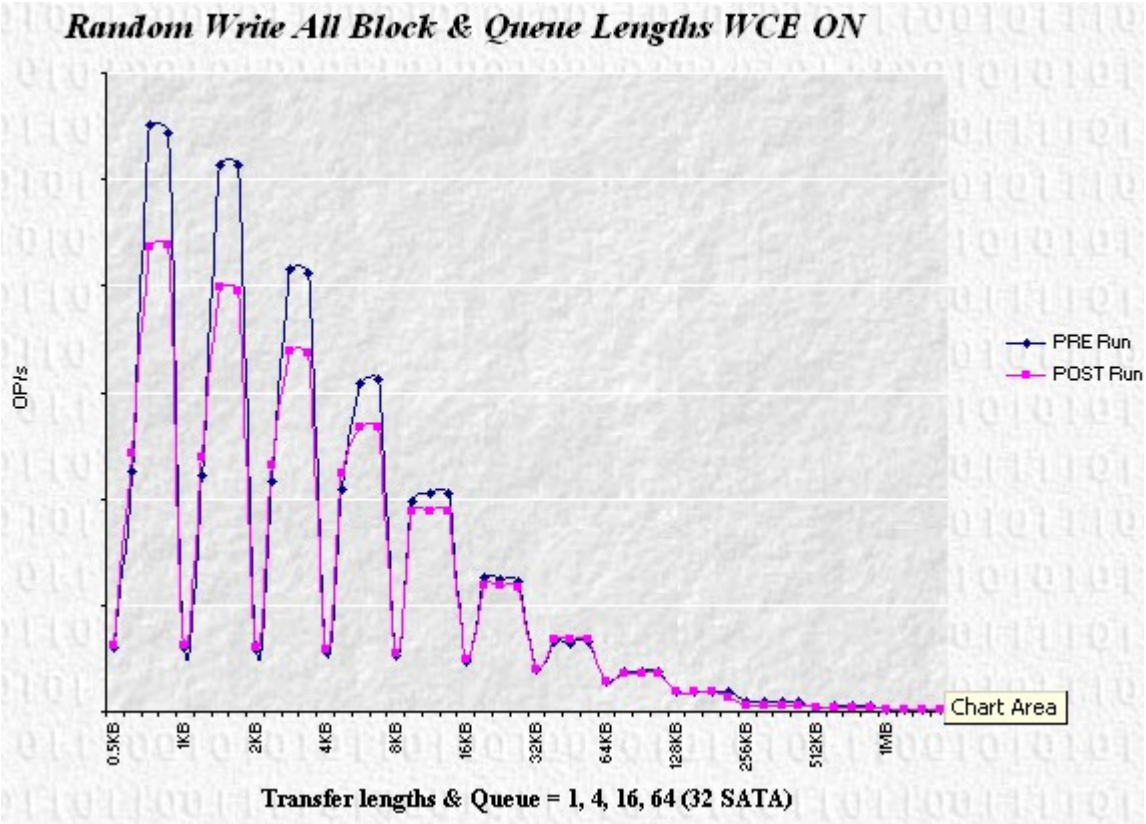
- The entire drive (equal to the drive's stated capacity) should contain typical or representative application data to avoid any unrealistic conditions that might produce unreliable or unrepeatable results.
- Some level of random writes should be performed in order to scatter the LBA locations into a pattern that is realistic for the application.
- Additionally, since some SSDs may implement idle time garbage collection techniques to erase blocks, the benchmark duration needs to be long enough and the time from conditioning the drive to starting the benchmark should be relatively short. If repeating the test multiple times, try to keep the delay between conditioning and benchmark execution as close to the same time as possible.

A Tale of Two States: Before and After

To highlight the importance of understanding SSDs and the underlying technology, Figure 3 compares the performance of one SSD device before and after conditioning the drive with data writes as described earlier. This testing was performed using IOMeter by Seagate.

FIGURE 3

Before and After Testing Results



Source: Seagate, 2008

Again, the degree to which this phenomenon manifests itself depends on the design of the SSD and the vendor. However, as the results show, in order to ensure realistic and accurate benchmark results on SSDs, it is vital to develop standard metrics and testing methodologies in an effort to assist system OEMs and other storage integrators in evaluating storage devices.

The Importance of Standards

After the metrics and preparation methodologies are defined for storage devices, the final step is to choose the appropriate measurement tool. In order to achieve an industry standard to which vendors must adhere when reporting the agreed-upon storage metrics, an industrywide benchmarking or testing tool must be identified or created by an independent, nonbiased body of experts.

A number of new efforts are under way among a variety of standards organizations, such as the Storage Performance Council (SPC) with its set of benchmarks focusing on storage subsystems; Storage Networking Industry Association (SNIA) with its Solid State Storage Initiative (SSSI), the Joint Electron Devices Engineering Council (JEDEC), a standards organization focusing on SSDs; and the International Disk

Drive Equipment and Materials Association (IDEMA), which has worked with the HDD industry for decades.

Perhaps what is most important is to foster the free and open exchange of ideas and information among as many standards organizations as possible. In so doing, the storage industry will put itself in the best position to have a set of common metrics, methodologies, and benchmarks that are used to assist end users and system designers in choosing the right storage device for the right applications and solutions.

CHALLENGES/OPPORTUNITIES

There is little doubt about the importance of storage in today's world. The vast industry that makes up the storage ecosystem has faced and overcome a myriad of challenges over the decades by delivering the right capacity, performance, and reliability that end users demand. However, with new storage devices, such as SSDs and hybrid HDDs, it is imperative to maintain the high standards that are expected in these environments. Thus, storage device vendors and their customers must be aware of the challenges ahead and customers must demand that vendors deliver to the high standards expected within the enterprise:

- System design can enhance, complement, or compromise any given storage device performance advantage/metric. In the end, understanding the real-world performance of storage devices and systems is a key factor in the decision-making process.

- With such a large number of storage device vendors, it is vital to communicate effectively with customers to avoid confusion in the marketplace. A failure to do so could delay adoption of any new storage technology and result in premature consolidation.

- Standardization promotes commoditization and typically accelerates consolidation among the vendor base. This dynamic is bound to happen because system OEMs desire to have at least dual-source strategies of components that exhibit similar performance and integration results.

- Intellectual property is a two-edged sword. While it may give a vendor an advantage over its competition, if it is not shared or licensed at some point, then adoption of the technology could be hindered or, worse, stopped.

- Future applications, technologies, or system designs could prompt the need to define and develop other metrics and methodologies.

Copyright Notice