iJRASET
International Journal For Research in
Applied Science and Engineering Technology


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 4 **Issue:** V **Month of publication:** May 2016

**DOI:**

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

www.ijraset.com
IC Value: 13.98

Volume 4 Issue V, May 2016
ISSN: 2321-9653

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

# Survey on Client SSD and Enterprise SSD Performance

Radhika Makadia[1], Rashmi R[2]
[1]Student, [2]Assistant professor
*Department of Information Science R.V. College of Engineering, Bangalore, India.*

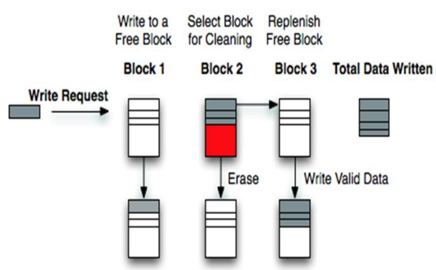
***Abstract: In current Storage system, NAND flash memory based Solid State Disk (SSD) is starting to replace hard disk drive (HDD) in desktop systems. Integrating SSD into enterprise storage systems, however, is much more challenging. Solid-State Drives (SSDs) offer significant performance improvements over hard disk drives (HDD) on a number of workloads, SSDs can exhibit significant performance degradations when garbage collection (GC) conflicts with an ongoing I/O request stream***

***Key Words- SSD, RAID Client Enterprise SSD***

## I. INTRODUCTION

Solid State Devices (SSDs) have emerged in the last few years as a viable secondary storage option for laptops and personal computers. Now, SSDs are poised to replace conventional disks within large-scale data centers, potentially providing massive gains in I/O throughput and power efficiency for high-performance applications. The major barrier to SSD deployment in such settings is cost [1]. This barrier is gradually being lifted through high-capacity Multi-Level Cell (MLC) technology, which has driven down the cost of flash storage significantly.

SSDs, suitably optimized for random read and write performance, could effectively replace whole farms of slow, rotating disks, small SSDs are starting to appear in laptop computers because of their reduced power-profile and reliability in portable environments. As the cost of flash continues to decline, the potential application space for solid-state disks will certainly continue to grow. Despite the promise that SSDs hold, there is little in the literature about the architectural tradeoffs inherent in their design.




Where such knowledge exists, it typically remains the intellectual property of SSD manufacturers.[2] First, SSDs can only write data to clean pages, and so they necessitate erase operations to reset flash blocks back to the clean state. However, each flash block can only tolerate a limited number of erasures before wearing out. Second, bit errors are very common in SSDs due to read disturbs, program disturbs, and retention errors[3,4,5][6].Thus, SSD reliability remains a legitimate concern RAID (redundant array of independent disks) [7] provides an option to improve reliability of SSDs.A Case for Redundant Arrays of Inexpensive Disks (RAID). Using parity-based RAID the original data is encoded into parities, and the data and parities are striped across multiple SSDs to provide storage redundancy against device failures. RAID has been widely used in tolerating hard disk failures. However, traditional RAID introduces a different reliability problem to SSDs Conventional wisdom suggests that parities should be evenly distributed across multiple drives to achieve better load balancing, such as in RAID-5. [8] find that RAID-5 introduces the problem of correlated device failures, in which all SSDs wear at the same rate and fail simultaneously, thereby causing data loss. [8].

## II. BACKGROUND RESEARCH

An SSD is composed of multiple chips, each of which is organized in multiple blocks. Each block typically has 64 or 128 fixed-size pages of size 4KB or 8KB each. Both read and write operations are performed in units of pages. Data can only be written (done by flash level program operations) to clean pages. SSDs perform an erase operation in units of blocks to reset all pages of a block into clean pages. To improve write performance, SSDs use out-of-place writes, reading data from a page may also have a weak

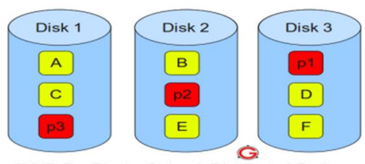
Volume 4 Issue V, May 2016
ISSN: 2321-9653


programming effect on other pages in the same block, which again leads to data corruption. Therefore, both bit-level protection and device-level protection like RAID are necessary for SSDs.

### *A. SSD RAID Basics*

We now describe the organization of an SSD RAID. Figure shows an SSD RAID organization. We consider the device-level RAID organization where the array is composed of N + 1 SSDs numbered from 0 to N RAID [9] Each SSD is divided into multiple non-overlapping chunks, each of which can be mapped to one or multiple physical pages. The array is further divided into stripes, each of which is a collection of N + 1 chunks from the N + 1 SSDs. Within a stripe, there are N data chunks, and one parity chunk encoded from the N data chunks. Since we focus on single-fault tolerance, we require that each stripe contains at most one erroneous chunk without data loss so that it can be recovered from other surviving chunks in the same stripe. Suppose that each SSD contains B blocks, and the array contains S stripes (i.e., S chunks per SSD). For simplicity, we assume that all S stripes are used for data storage, although in practice an SSD is usually overprovisioned [9].

Design Tradeoffs for SSD Performance To generalize our analysis, we organize parity chunks in the array according to some




RAID 5 – Blocks Striped. Distributed Parity.

probability distribution. Let SSD i contain a fraction pi of parity chunks (where $0 \le i \le N$ and $0 \le pi \le 1$). In the special case of RAID-5, parity chunks are evenly placed across all devices, so pi = 1 N+1 for all i if the array consists of N + 1 drives. For Diff-RAID, pi 's can be arbitrarily defined subject to the condition PN i=0 pi = 1. Each block in an SSD can only sustain a limited number of erasures, and it is worn out after the limit[12]

## III. PERFORMANCE ANALYSIS

### *A. Client SSD Performance Considerations*

Client use cases typically involve a single user running various Client software applications over the course of the day. Client SSD demand intensity is low relative to Enterprise class SSDs which have multiple users or concurrent workloads that continuously access the SSD. For Client SSDs, IOPS are often listed as “up to” a certain number of IOPS, at “sustained” IOPS or perhaps (rarely) at “Steady State” IOPS. What is the difference between these claims and what do the different performance states mean? How many IOPS are good? How many IOPS are enough? One should always try to understand the performance state in which any measurements are taken. Each performance state is highly dependent on a variety of factors that can have a significant impact on performance. FOB and “up to x IOPS” statements are transient levels of performance that may never be seen once the user installs and begins to use the SSD. Claims of sustained IOPS can also be problematic (and misleading) if the sustained level is not precisely explained. Which metric is more important? In absolute terms, a higher value for IOPS and Bandwidth (MB/sec) is better (more) while a lower value for Response Times or Latency is better (faster). How many more IOPS will make a difference in my use case? Are the additional IOPS worth the additional price? What if IOPS are better but the Response Time is worse when comparing SSD A to SSD B? Which is more important: higher IOPS or lower Response Time; higher IOPS or higher Bandwidth? And at what cost? For many Client use cases, IOPS and Bandwidth may be more important than Response Times (so long as the response times are not excessively slow). This is because the typical Client user would not typically notice a single IO taking a long time (unless the OS or software application is waiting for a single specific response). However, choosing an SSD based solely on the highest IOPS or Throughput rate can have a point of diminishing returns. The user’s computer may not be able to utilize IOPS rates past a certain point. The Client use case may consist of low end-user SSD demand (“read the SSD, wait while the user processes the data on the screen, maybe then write a little to the SSD”). higher Bandwidth, while always good to have, may be advertised based on a block size or read write (RW) mix that does not match the workload that the user’s computer generates or that the SSD sees. Thus, it is critical to understand what and how the advertised metrics are obtained and to know what kind of workloads will be generated in the intended use case.

### *B. Enterprise SSD Performance Considerations*

The Enterprise market requires SSDs with higher overall performance. Enterprise SSDs are generally measured at Steady State

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

under a full workload for seven days a week for 24 hours per day operation. The constant usage requires a substantially different design, fewer errors, higher general performance and greater data integrity. Enterprise SSDs are often aggregated as RAID, Tiered, Direct Attached Storage (DAS), Network Attached Storage (NAS) and Storage Attached Network (SAN) for more complex storage solutions. SSDs are also packaged as higher performance solid state storage in PCIe connected devices – both as traditional SSD like devices as well as more advanced, higher speed products (such as Persistent Memory, Memory cache and Direct Memory). Enterprise SSD use cases also tend to be more homogenous than Client with SSDs tuned for the access patterns associated with specific applications such as On-line Transaction Processing (OLTP), Virtualized Machines (VM), Virtual Desktop Infrastructure (VDI), Video-on-Demand (VOD) edge servers, Tiering, data center, database and others. This allows the Enterprise user to deploy SSDs with high performance in the areas that relate to the intended Enterprise use case. While IOPS are an important metric in the Enterprise, it is often the management of response times that is paramount. It is often said “IOPS are easy, latencies are hard.” Competing Enterprise SSDs may offer substantially the same number of IOPS, but may differ significantly in response times and latencies. Thus, Enterprise focus on response times can differ from Client. In a Client application, it may be acceptable to have many response time spikes (instances of very high response times for a number of IO operations) since the result may merely be an annoying delay to the user (hourglass or spinning gear for a while). In the Enterprise, the SSD must often meet a Quality of Service (QoS) level - the requirement that a given application complete all requested processes within a specified time limit (and effectively disallowing an application from accepting or processing a given percentage of requests whose response times exceed a fixed threshold).

*C. The Dimensions of Performance Three Performance Metrics*

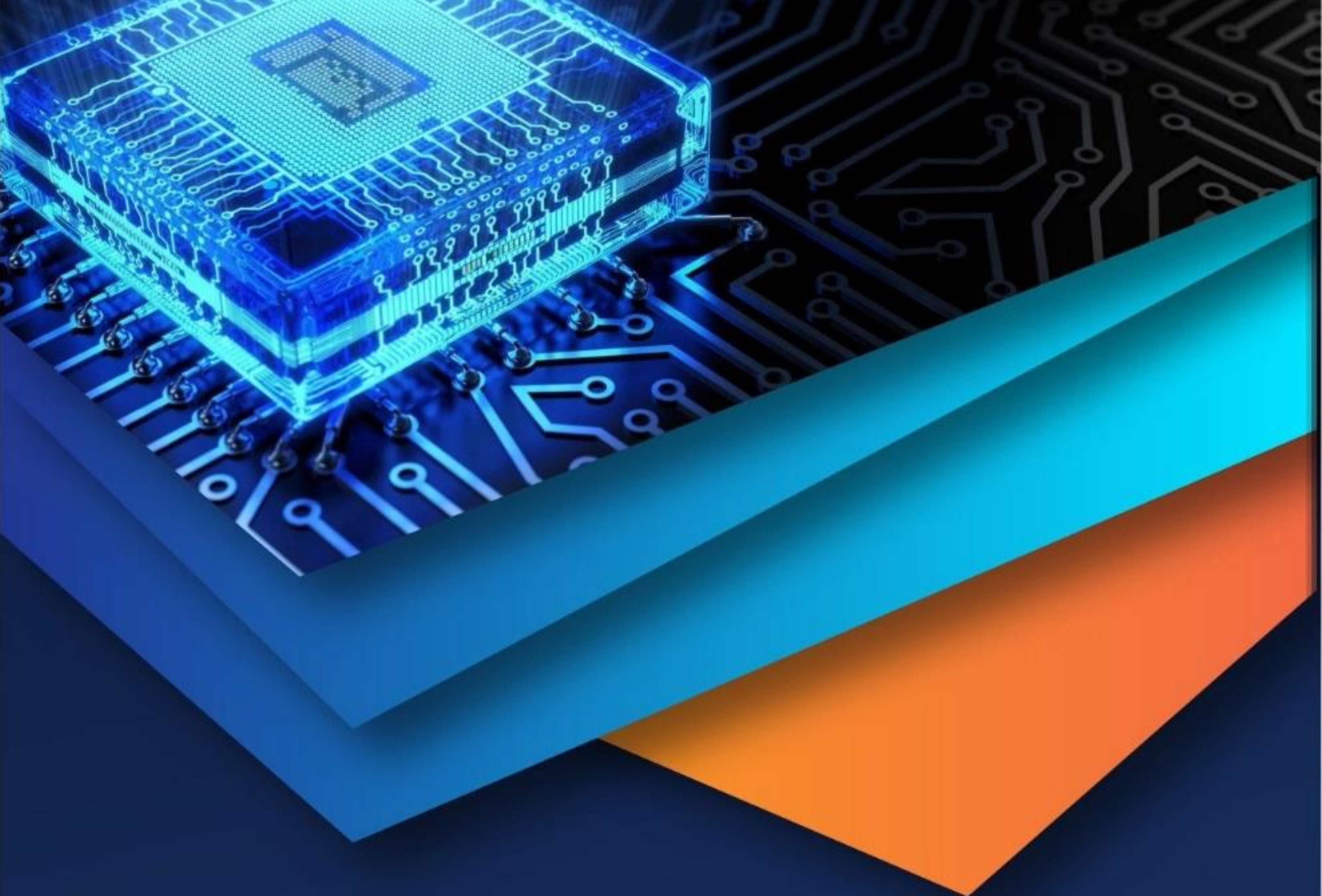
IOPS, Throughput (TP) & Latency (LAT) Each metric represents a different dimension of performance. Evaluation of SSD performance should consider all three dimensions: IO transaction rate (IOPS), sustained data transfer rate (TP) and the latency (LAT) or response times (RT) of the IOs. The optimal values for each metric, and the relationship between them (relative importance of each) is a function of the anticipated use case workloads. It is important to evaluate all three dimensions of performance because the user will want to know: How many IO operations can be completed (IOPS)? • How much data can be transferred (Throughput in MB/s)? • How much delay there is for a given IO operation (Response Time/Latency)? For example, when I boot my system (on a laptop or Enterprise virtual desktop), how fast will the OS load the small block read and write operations (IOPS)? The department manager may look at this as “howquickly are my workers getting productive?” How fast will my system load large graphic files for editing or stream a training video from my storage device (Throughput)? How long will I have to wait to find/load/ save a specific file (Response Time/Latency)? Access Patterns and Test Workloads An access pattern is the type of storage and retrieval operations to and from a storage device. Access patterns are described in three main components: • Random/Sequential – the random or sequential nature of the data address requests • Block Size – the data transfer lengths • Read/Write ratio – the mix of read and write operations Any particular workload or test stimulus is approximated by some combination of access patterns. That is, an access pattern is one component of a synthesized equivalent IO workload. In contrast, an IO operation access is considered to be random when its starting LBA is not contiguous to the ending LBA of the preceding IO operation. Sequential operations can be faster than random operations. When HDD recording heads are on a single track, sequential IOs can be faster than random IOs. In NAND flash SSDs, sequential IOs can be faster than random IOs when SSD mapping tables are less fragmented. NAND SSDs use a virtual mapping scheme whereby LBAs are mapped to Physical Block Addresses (PBAs). SSD virtual mapping is used for several reasons. For example, wear leveling algorithms distribute newly recorded data to new cell locations to promote even wear on the memory cells and thus improve the memory cell life or endurance. Hence, the SSD must keep track of the LBA – PBA associations. The SSD controller keeps track of the LBA to PBA association by using a mapping table. Depending on its design, the SSD mapping table can become highly fragmented (and slower) if comprised of many random access entries whereas the mapping table can be more efficient (faster) when mapping a more organized table of fewer sequential access entries. Block Sizes Block sizes (BSs) can be in varying sizes from small to large Blocks are more efficiently stored in an SSD when they are aligned with the NAND flash memory cell boundaries so that, for example, a 4KiB block will fit exactly in a 4KiB NAND page size. So far, we have talked about two metrics: IOPS and TP. Now let us turn to our third metric: Response Time (also called Latency). Response Time measures how long a particular IO transaction takes to complete (which is a affected by everything that touches the IO as it traverses the “IO Stack”). Knowing how an IO is affected by the IO Stack is important to understand when evaluating Response Time measurements. IO Stack The performance we measure for a particular access pattern can be very different depending on where in the system we do the measurement. Device level (or Block IO) tests typically measure access patterns as close to storage hardware as possible (desirable for SSD drive testing), whereas file system level tests are more often used to test the software application in user space (which measures the overall system level performance of the storage device). Access patterns generated by software applications must traverse the IO Stack (Figure 2) **-** to get from the user space to the SSD and back again. This means that the file system and various drivers will affect the IOs as they pass them up or down the IO Stack. What are workloads? In very general terms, a workload can be described by the access patterns measured during an observation period (e.g. 10 minutes of random 4KiB 100% Writes). Having determined that the key performance metrics (IOPS, TP and LAT) are described in terms of their access patterns, the key questions are: What access patterns are generated by my application, and What access patterns are applied to my SSD. Workload IOs at the user level can be very different from those seen at the device (e.g. SSD) level because the IO stack can alter or change the nature of the access patterns as the IOs traverse the stack. User level workloads are complex, multiple stream access patterns that can change over time. A system can start with a largely small block random read/write (RW) boot workload, shift into a predominantly sequential read (R) workload as software loads, and then generate a mixed stream workload as multiple programs interact – all while the operating system software generates its own mix of access pattern activities in the background. While user workloads are complex mixtures of various access patterns, there are commonly accepted access patterns used by the industry for certain applications.

## IV. CONCLUSION

SSD product design and architectures continue to evolve. Advances in SSD NAND flash technologies, increasingly sophisticated SSD firmware, optimization of the IO software stack and more powerful hardware will lead to increased SSD performance levels into the hundreds of thousands to millions of IOs per second. The use of the SSS Performance Test Specification and RTP can provide a means to accurately and reliably test these high performance SSDs and to create an industry wide baseline for SSD performance test.

## REFERENCES

[1] Narayanan, E. Thereska, A. Donnelly, S. Elnikety, and A. Rowstron. Migrating Server Storage to SSDs: Analysis of Tradeoffs. In EuroSys 2009. ACM, 2009.
[2] D. Roberts, T. Kgil, and T. Mudge. Integrating NAND Flash Devices onto Servers. Commun. ACM, 52(4):98–103, 2009.
[3] Y. Cai, E. F. Haratsch, O. Mutlu, and K. Mai. Error Patterns in MLC NAND Flash Memory: Measurement, Characterization, and Analysis. In DATE, 2012
[4] H. Sun, P. Grayson, and B. Wood. Quantifying Reliability of SolidState Storage from Multiple Aspects. In SNAPI, 2011.
[5] E. Yaakobi, L. Grupp, P. Siegel, S. Swanson, and J. Wolf. Characterization and Error-correcting Codes for TLC Flash Memories. In ICNC, 2012
[6] L. M. Grupp, J. D. Davis, and S. Swanson. The Bleak Future of NAND Flash Memory. In FAST, Feb 2012.
[7] D. A. Patterson, G. Gibson, and R. H. Katz.
[8] M. Balakrishnan, A. Kadav, V. Prabhakaran, and D. Malkhi. Differential RAID: Rethinking RAID for SSD Reliability. ACM ToS, 6(2):4, Jul 2010
[9] N. Agrawal, V. Prabhakaran, T. Wobber, J. D. Davis, M. Manasse, and R. Panigrahy.
[10] N. Mielke, T. Marquart, N. Wu, J. Kessenich, H. Belgal, E. Schares, F. Trivedi, E. Goodness, and L. Nevill. Bit Error Rate in NAND Flash Memories. In IRPS 2008: IEEE International Reliability Physics Symposium, 2008.
[11] www.snia.org/sites/default/files/SNIASSSI.SSDPerformance-APrimer,2013
[12] IEEE TRANSACTIONS ON COMPUTERS 1 Analysis of Reliability Dynamics of SSD RAID Yongkun Li, Member, IEEE Patrick P. C. Lee, Member, IEEE John C. S. Lui, Fellow, IEEE




10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY