# Evolution of Storage trend for Big Data


## Abstract:

Data is a collection of raw facts that has been translated into form that computers or user can understand and convert it into knowledge. The cycle of data explosion runs in following manner information created, stored and used. Tremendous increase in data generation has been seen over the period, i.e. 800,000 PB in 2000 to 35 ZB expected by 2020. The data with big volume, variety and velocity can be considered as Big Data. The increase in data generation has raised plethora demand of storage requirement. The significance and evolution of storage device from HDD to SSD will be explained based on performance, characteristics, mechanism, cost and reliability. Comparison of NVMe drive with respect to SATA SSD against MongoDB and Cassandra database will be present and discussed. Finally, the comparison between primitive data storage and flash storage will be explained.


## Introduction:

The inevitable growth of data generation day by day is indirectly boosting the technology to keep constantly build and invent new storage techniques, platform, databases and so on. The storage technology has evolved tremendously to meet the ever-increasing performance demands. While every new generation of storage technology was able to achieve modest performance improvement over its predecessor. Several innovations were made in storage technology over previous state of art. Solid state drives were the first technology leap which contains no moving part, provide better random access read performance then HDD. Second technology leap was made by Nonvolatile memory express (NVMe), it is a software standard for optimization of SSD connection through PCIe interface. It can also be seen that traditional mainstay of storage technology named HDD is still compared with NVMe drives and SSD regarding the unique characteristics. The basic eye-catching thing in traditional and latest storage device hustle is between cost and performance. HDD is falling to bring cost effective solution, where high capacity and performance is required for cloud, enterprise, web and virtualized application. The latency, interface, protocol, form factor, reliability will be compared for storage technology. The detailed discussion with respect to performance for NoSQL database by SATA HDD, SATA SSD, and NVMe drive will be analyzed.

**Storage technology SSD:** Solid state drive are in big demand because of its performance capacity over HDD. Some of the important features of SSD are wear levelling, error correction and TRIM SSD which leverage SSD to achieve read/write speed in excess of mechanical hard drive. It is also having an effective cost in comparison of HDD, which has again attracted IT sector towards it. Let us go through it performance details:

**Write speeds:** With increase in Data over the years, every organization want the better through put speed from its storage device so is the case with the drive it has been observed that read/write speed has been increase in both mechanical and solid state drive too. Discussing about performance side the major advantage will be available for fast boot time, faster random read/write and faster sequential read for big files, for application user less waiting time at start-up, loading application and in general use. Common read and write speed is attain upto 500MB/s and 300 MB/s respectively. This speed is continuously increasing with the advancement of technology. With the use of PCIe slots available now with increased bandwidth tends to increase read/write speed commendably even further.

| | Typical Solid State Drive | Typical Mechanical Hard drive |
|---|---|---|
| Typical **Access Time** | 0.5ms | 11ms |
| Sustained Read | 90+MB/s | 70MB/s |

***Side-by-side Performance Comparison (21/03/2012)*** [1]

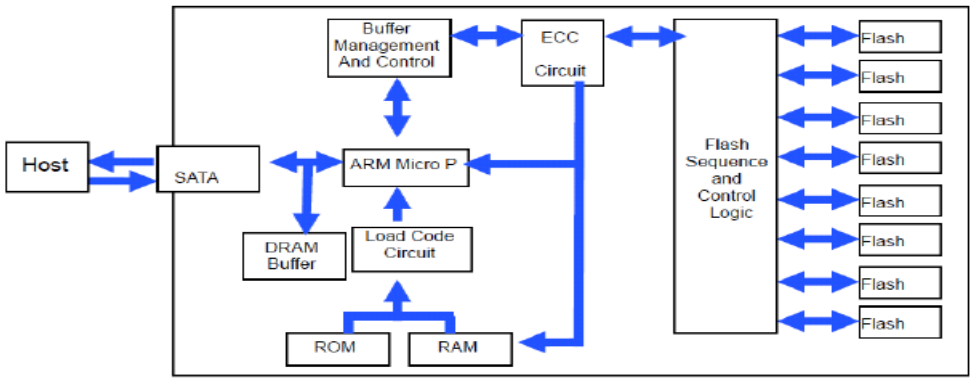
**Shock and Vibration [1]** The capacity of resisting shock and vibration and ability to work in extended temperature range makes SSD a better replacement for applications where mechanical hard drive will fail to work.

| | Typical Solid State Drive | Typical Mechanical Hard drive |
|---|---|---|
| Shock (Operating) | 1500G | 300G |
| Operating Temperature | 0°C - 70°C<br>Higher temperatures are available | 0°C - 60°C |

***Shock and temperature comparison (21/03/12)***

**Components:** Main two components of SSD are Controller and memory where it uses to store data. Since DRAM volatile memory was developed was used as primary memory component of SSD. Now NAND flash non-volatile memory.

The control act as a brain and that work as a bridge between NAND Flash memory to host computer for following functions: Error correction, Wear levelling, Bad Block mapping, Read and write caching, Garbage collection, Encryption




***Inner workings of a SSD*** [1]

**Storage technology HDD vs SSD:** Mechanical hard drive known as HDD is the traditional way of storage used for data. As the demand of data is increasing the expectation of IT sector is getting higher with respect to the performance of storage system. The major difference between HDD vs SSD is performance and cost. Some important performance characteristics comparison of HDD vs SSD is discussed here.

**Latency:** Time taken for access is known as latency. SSD Read access time is 509 times faster than HDD, which is 15.75 ms sounds for HDD ad 0.031 ms for SSD. In HDD write access time is faster than read access. In write access time also SSD win over HDD because it is slower by 102 times with respect to SSD.

**Throughput:** Throughput speed of SSD is 514.28 MB/s as against 149.86 MB/s of HDD. It can be inferred from stats that HDD is 3.4 times slower. Write speed of HDD is 1.22 MB/s is much better than read speed but decreased by 128.65 MB/s in SDD which is 105 times faster.

**Capacity:** HDDs work as workhorses while SSD are only affordable at lower capacities. SSD capacity are extremely high expensive.

**Bandwidth:** it is evident that SDD can provide followings things much better than HDD are high performance for restarting and peek read/write for enabling multitasking capabilities. While booting, loading data and launching can be provide by SSHD.

**Lifetime cost or total cost of ownership (TCO):** The hardware cost of HDD is lower than the SSD but on the other hand cost for bandwidth (in terms of $/IO) is much lesser for SSD. Which is more important for enterprise than capacity per dollar ($/GD). [5]

**Power Consumption:** Normally, Power consumption of SSD is very less in comparison of HDD, which give the advantage of cost saving and less heat generation. [5]

**Reliability:** SSD has a failure rate 0f 0.5% whereas HDD has a failure rate of 2-5%. [3]

**Below table will show the comparison between HDD vs SSD: [2]**

| | HDD | SDD |
|---|---|---|
| Speed | Compared to SSD, HDD has longer read/write times, supports fewer IOPs and higher latency. | Compared to HDD, SSD has faster read/writes, supports more IOPs and lower latency. |
| Heat, Electricity, Noise | To generating heat, rotate the platters and noise, hard disk drives use more electricity. | In solid state drives, no such rotation is needed; they do not generate heat or noise and use less power. |
| Defragmentation | Due to fragmentation, the performance of HDD drives worsens; so that, they require to be defragmented periodically. | SSD is not require defragmentation because drive performance of SSD does not get affected by fragmentation |
| Components | HDD include moving parts – thin layer of magnetic material is covered by one or many platters holds by a motor-driven spindle. On disks, top write head and read head are positioned | A memory chip is nonmoving parts of SSD. With an interface, connector SSD is integrated circuits (ICs), which is interconnected. Controller, cache, and capacitor are the three main components. |
| Weight | HDD's drives are heavier than SSD's. | SSD drives have motor, spindle and rotating disks, hence they are lighter than HDD drives. |
| Dealing with vibration | Due to vibration, HDDs moving part make them more habitual to crashes and damages. | SSD drives have 2000Hz vibration which is more durable than HDD. |

The trend of SSD vs HDD in last five is quite impressive between 2012 and 2017, The revenue of enterprise-class SSD grew from $3.0 billion to $11.4 billion with a CAGR of over 30%, based on study by Gartner, The Register, and Stifel. The rise of SSD vs HDD is commendable and sold value exceeded the value of HDD first time. [5]

**The Rise of NVMe:** From 2005 and 2010 the communication buses for drive was SATA/SAS on which both SSD and HDD both use to communicate, over the time this become inadequate to meet the faster access time and higher data rate typically for SSD. In 2011 new host control interface standard released called Non-Volatile Memory express (NVMe) It was open logical to capitalize low latency and internal parallelism of flash storage drives. It is a specification which allow SSD to make effective use of high speed PCI Express bus in computer. There are variety of form factor. The performance advantage of NVMe over mechanical drive and SATA SSD is present in below table.

HDD vs. SATA vs. NVMe

| Interface | PCIe | | SATA 6 Gb/s |
|---|---|---|---|
| Protocol | NVMe Protocol | AHCI Protocol | |
| | (Optimized for flash SSDs) | (Optimized for mechanical HDDs) | |
| Bandwidth | PCIe x2 or x4 lane | PCIe x2 or x4 lane | SATA 6 Gb/s |
| Form Factor | M.2/PCIe Expansion Card/U.2 | M.2/PCIe Expansion Card | M.2/2.5" SSD |
| Max. Read Performance | >3000MB/s | >2000MB/s | >500MB/s |
| Max. Write Performance | >2000MB/s | >1500MB/s | >500MB/s |

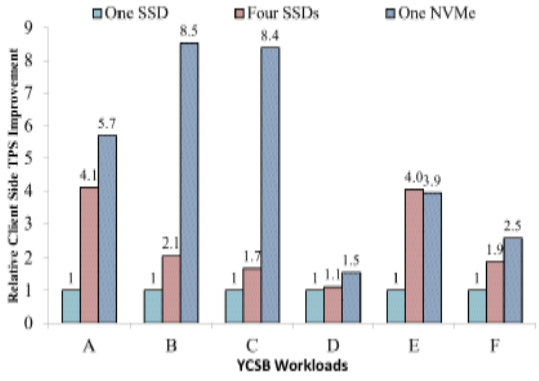
Source: Hyperion Research 2018

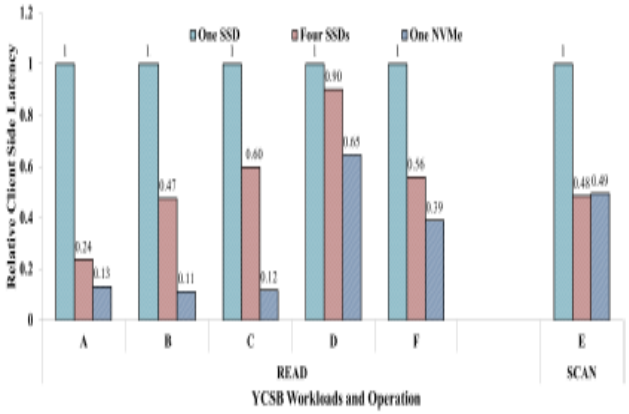
[5]

**Evaluation of performance on Database with storage technology:**

Since DBMS are prime example of I/O intensive application for datacenter it has been expected that NVMe drives will be deployed in datacenters. The discussion will be on NoSQL databases (Cassandra and MongoDB) and compare NVMe performance with SATA SSD in multiple configuration. In real world production databases commonly, serve multiple clients but in this experiment, we are using single server-client setup. We use YCSB to exercise NoSQL databases. Two comparisons will be made as follows one will be baseline characterization on single SATA SSD in addition four SSD in RAID0 configuration. Second both SSD configuration to single NVMe drive to analyze its impact on performance.[4]

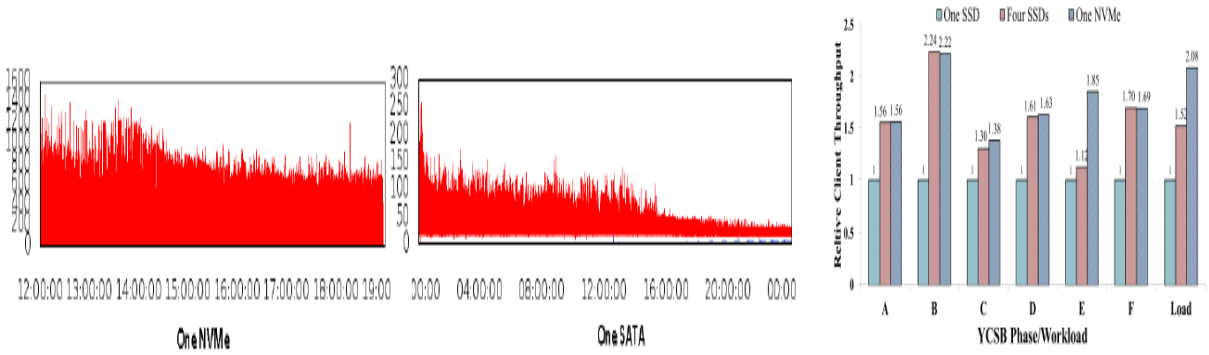
**Cassandra:** One of the trending data store of NoSQL in last few years is Cassandra. It is designed for fast writes and scale-out, it's throughput will scale linearly when new nodes are added to cluster. Any data written in Cassandra is written in parallel to memtable (DRAM resident cache) and commit log (disk resident datastructure) when memtable exceed threshold it flushes the disk and update data will be written in SSTable. The write completion to commit log are not in critical path while in contrast to write read in Cassandra have no longer datapath if they miss memtable. It uses periodic compaction to avoid various lookup and merge SSTable into single SSTable. The performance is compared for 6 YCSB workloads. The performance improvement over SSD range from 55% in workload D to a whopping 8.34 X in workload C has been noticed for NVMe. Write and update happen in DRAM while missing read value are store in SSTable, NVMe speedups are most profound in read-heavy workloads B and C. it is modest in write heavy workload A and many read-modify-write workload F. There is no improvement in workload D is because it consists of latest distribution and great cache memory. Data stripping in RAID0 help in scan read large chunks of data for workload E. [4]



Now comparing the client side read latencies of different storage media. Neglecting the update and insert because it same across all media. It can be seen that throughput been benefits for Cassandra correlate well with corresponding latencies. Workload B and C have lowest relative latency for NVMe compared to SATA SSD (11%-13%) and shows best benefits in throughput. The scan and throughput for workload E is very close to that of four RAID0 SSD. YCSB does not came close to the bandwidth provided by NVMe drive. The bottleneck for operation is either software stack or I/O latency. Peak disk bandwidth for NVMe is 350 MB/s while one single Cassandra instance barely sustain 70 MB/s. a good amount of time was spent for waiting for I/O operation. In contrast, NVMe instance has negligible I/O wait time. Hence, we can say that Cassandra read performance is very sensitive to read latency of storage media. Benefits demonstrated for Cassandra is due its lower read latency is only for NVMe drives. [4]

**MongoDB:** Since MongoDB uses memory mapped I/O, there will be different benefits of using NVMe drive for this. The benefit of NVMe drive is evident in case of disk performance dedicates to system performance which is workload E (scan and load phase). As is it evident that NVMe drive is able to sustain 3X higher write bandwidth compared to SATA SSD, leading to an overall 2X increase in throughput. Now 4 SSD RAID 0 has help but not as much as compared to NVMe. In this case NVMe drive performs 65% and 36% better for load phase and workload E, respectively. For rest of the test owing towards memory mapped I/O the performance of NVMe SSD is like RAID 0. [4]




**Reference:**


https://www.amplicon.com/docs/white-papers/SSD-vs-HDD-white-paper.pdf

https://www.researchgate.net/profile/Anagha_Kadve/publication/309159564_Trade_Of_Between_SSD_and_HDD/links/5801d84808ae1c5148ca7461/Trade-Of-Between-SSD-and-HDD.pdf

http://www.enterprisestorageforum.com/storage-hardware/ssd-vs-hdd.html

https://www.cs.utah.edu/~manua/pubs/systor15.pdf

Flash Storage Trends and Impacts Sponsored by HPE Bob Sorensen, Earl Joseph, Steve Conway, and Alex Norton April 2018