ENTERING THE YOTTABYTE ERA USING ENTERPRISE FLASH DRIVE TECHNOLOGY

EMC Proven Professional Knowledge Sharing 2011

Milan K Mithbaokar
Systems Integration Advisor
Dell Inc.
Milan_Mithbaokar@dell.com
Table of Contents

Abstract .......................................................................................................................... 3
Brief History of SSD ...................................................................................................... 4
What are Enterprise Flash Drives? .............................................................................. 4
What is the need for SSD Drives? .............................................................................. 6
Benefits of using SSD versus Hard Disk Drives ......................................................... 6
Misconception Comparing SSD with Hard Disk Drives ............................................. 6
   Current Trend (Mixing Solid State Drive and HDD) ................................................. 6
Architecture of SSD ..................................................................................................... 7
   Controller .................................................................................................................. 7
      Error correction Code (ECC) .................................................................................. 7
      Wear leveling ........................................................................................................ 7
      Read scrubbing ....................................................................................................... 8
      Read and write caching ......................................................................................... 8
      Garbage collection ............................................................................................... 8
      Encryption ............................................................................................................. 9
Memory .......................................................................................................................... 9
   Single-level cell ........................................................................................................ 9
   Multi-Level cell ......................................................................................................... 9
   Host interface .......................................................................................................... 9
   Form factor .............................................................................................................. 10
      Standard HDD form factors ................................................................................ 10
Challenges Facing SSD Drive Technology ................................................................. 10
   Standards for SSD Drive Technology ..................................................................... 11
   Application Classes ............................................................................................... 11
   Endurance Rating and Verification ........................................................................ 11
   Workload .................................................................................................................. 11
Current State of Digital Universe Storage ................................................................. 12
Conclusion ................................................................................................................... 12
References .................................................................................................................... 13

Disclaimer: The views, processes, or methodologies published in this article are those of the author. They do not necessarily reflect EMC Corporation’s views, processes, or methodologies.
Abstract
Enterprise Flash Drives (EFDs) have entered the expanding Digital Universe market and will soon lead us toward the Yottabyte Era. The yottabyte is a unit of information storage equal to one septillion \(10^{24}\) bytes. The unit symbol for the yottabyte is YB.

An EFD—also known as a Solid State Drive (SSD)—is comprised of integrated circuits and has no moving parts. Since they do not contain moving parts, there is no such parameter as seek time or rotational latency. Solely using integrated circuits without any mechanical parts improves the performance of EFD’s and helps them deliver a tremendous amount of input/output per second (IOPS) with very low response times, as well. The difference between the IOPS produced by traditional Fibre Channel (FC) Hard Disk Drives (HDDs) and EFD’s can be approximately thirty times (i.e. EFD’s can approximately produce 30 times more IOPS than traditional Fibre Channel Drives). Additionally, since EFD’s do not contain any moving parts, they consume less energy compared to traditional FC HDDS.

EFD technology will help organizations make the best decisions in terms of speeds (IOPS) which were previously not possible using the traditional HDD’s. Organizations can now opt to go with Hybrid Storage models (using SSD and FC Drives), a current trend in the storage industry.

The introduction of SSD technology and this Hybrid Storage model into the IT infrastructure business world will offer a lot of options to company leaders who can take the benefits of the speed (IOPS) that this technology has to offer to help them achieve their business goals.

This article explores the challenges facing the SSD technology and how it will lead us to the Yottabyte Era.
Brief History of SSD
SSD is based on flash memory. During the 1970's, the semiconductor industry was not that progressed and the idea of non-volatile memory, which would retain information even when the power was lost, was fairly new. This idea of so-called flash memory occurred in the 1970’s to Dr. Fujio Masuoka who kept working on his idea and finally invented the flash memory and was later also awarded the 1997 IEEE Morris N. Liebmann Memorial Award from the Institute of Electrical and Electronic Engineers.

What are Enterprise Flash Drives?
Applications that require high IOPS, reliability, and energy efficiency use EFD’s. The EFD’s have a higher set of specifications compared to SSDs used in notebook computers. The term was first coined by EMC in January 2008 to define the SSD manufacturers who would provide products meeting these higher specifications and standards. Figure 1 shows the diagram of traditional HDD versus the current SSD.
Figure 1: Diagram showing the Enterprise Flash Comparison of the inside of a hard disk drive (HDD) versus a solid-state drive (SSD), which has no moving parts, enabling quieter and more reliable operation.
What is the need for SSD Drives?
As referenced in my Abstract, the Digital Universe is expanding at an alarming rate. Performance is the key factor in today’s economy and, consequently, enterprise storage is at a crossroad. Legacy mechanical-based storage systems that have dominated the open systems market since the late 1980s have reached the limits of their ability to deliver acceptable, cost-effective, and power-efficient performance for even the most basic enterprise workloads. There is a need for new storage systems based on devices such as flash memory that can surpass the limitations of mechanical storage and meet the demands of the enterprise.

Benefits of using SSD versus Hard Disk Drives
- SSD’s can boot up up to 5 times faster than traditional hard disk drives
- SSD’s have no mechanical parts, meaning they can save more energy and contribute toward a ‘greener’ planet
- SSD’s are very light compared to hard disk drives
- SSD’s have longer write cycles compared to traditional wear and tear of mechanical components used in hard disk drives, resulting in longer life
- SSD’s provide approximately 30 times more IOPS compared to traditional hard disk drives. Fewer drives are needed to meet IOPS needs, directly contributing to energy savings.

Misconception Comparing SSD with Hard Disk Drives
The key factor in comparing the HDD with SSD’s is IOPS even though both are capacity devices. SSD’s produce almost 30 times more IOPS than traditional HDD’s and should be measured in terms of $/IOPS (performance) as opposed to HDD’s which are measured against $/GB (capacity).

Current Trend (Mixing Solid State Drive and HDD)
SSD’s have given momentum to the storage markets and companies are slowly accelerating its adoption. Companies have realized the IOPS benefit this technology has to offer and have started migrating their read-intensive applications on EFD’s and mixing them with FC HDDs for write operations. Slowly becoming a current trend in the storage Industry, this arrangement, apart from meeting the IOPS needs of the organization, is helping them use fewer drives, resulting in lower power consumption.
Architecture of SSD

I hope the benefits I have described and the current trend have heightened your curiosity about this technology. Now, let's take a look at the SSD architecture.

The key components of an SSD are the controller and memory to store the data.

Controller

Every SSD includes a controller that incorporates the electronics that bridge the NAND memory components to the host computer. The controller is an embedded processor that executes firmware-level code and is one of the most important factors of SSD performance. Some of the functions performed by the controller include:

Error correction Code (ECC)

An ECC or forward error correction (FEC) code is a system of adding redundant data, or parity data, to a message, such that it can be recovered by a receiver even when a number of errors (up to the capability of the code being used) were introduced, either during the process of transmission, or on storage.

Wear leveling

There are three basic types of wear leveling mechanisms used in the Flash memory storage devices.

No wear leveling

A Flash memory storage system with no wear leveling will not last very long if it is writing data to the flash. Without wear leveling, the Flash controller must permanently assign the logical addresses from the operating system (OS) to the physical addresses of the Flash memory. This means that every write to a previously written block must first be read, erased, modified, and re-written to the same location. This is very time consuming and highly written locations will wear out quickly while other locations may be unused entirely. Once a few blocks reach their end of life, the drive is no longer operable.

Dynamic wear leveling

Dynamic wear leveling uses a map to link Logical Block Addresses (LBAs) from the OS to the physical Flash memory. Each time the OS writes replacement data, the map is updated so the
original physical block is marked like invalid data, and a new block is linked to that map entry. Each time a block of data is re-written to the Flash memory, it is written to a new location. However, blocks that never get replacement data sit with no additional wear on the Flash memory. The name comes from only the dynamic data being recycled. Although the drive may last longer than those with no wear leveling, there are blocks still remaining as active that will go unused when the drive is no longer operable.

**Static wear leveling**
Static wear leveling also uses a map to link the LBA to physical memory addresses. Static wear leveling works the same as dynamic wear leveling except the static blocks that do not change are periodically moved so that these low usage cells are able to be used by other data. This rotational effect enables the SSD to operate until most of the blocks are near their end of life.

**Read scrubbing**
It is important to check each memory location periodically, and frequently enough, before multiple bit errors within the same word are likely to occur. This is because one bit errors can be corrected, but multiple bit errors are not correctable. In order to not disturb regular memory requests from the CPU—which can decrease performance—scrubbing is usually only done during idle periods. Memory scrubbing increases reliability.

**Read and write caching**
In computer engineering, a cache is a component that transparently stores data so that future requests for that data can be served faster. The data stored within a cache might be values that have been computed earlier or duplicates of original values that are stored elsewhere. If requested data are contained in the cache (cache hit), this request can be served by simply reading the cache, which is comparatively faster. Otherwise, (i.e. cache miss), the data have to be recomputed or fetched from its original storage location, which is comparatively slower. Hence, the more requests can be served from the cache, the faster the overall system performance.

**Garbage collection**
Once every block of a SSD has been written one time, the SSD controller will need to return to some of the initial blocks which no longer have current data (also called stale blocks). The data in these blocks were replaced with newly written blocks, and now they are waiting to be erased so that new data can be written into them. This is a process called garbage collection (GC). All SSDs will include some level of garbage collection.
**Encryption**

Encryption is the process of transforming information (referred to as plaintext) using an algorithm (called cipher) to make it unreadable except by that possessing special knowledge, usually referred to as a key. The result of the process is encrypted information to protect data on a SSD.

**Memory**

Most SSD manufacturers use non-volatile NAND flash memory in the construction of their SSDs. Flash memory-based solutions are typically packaged in standard disk drive form factors (1.8-, 2.5-, and 3.5-inch) or smaller unique and compact layouts due to the compact memory.

There are two types of Memory used in SSD; single-level cell (SLC) or multi-level cell (MLC).

**Single-level cell**

A single-level cell (SLC) is a memory element made of floating-gate transistors that stores one bit of data in each cell. SLC memory has the advantage of faster write speeds, lower power consumption, and higher cell endurance. However, because it stores less data per cell, it costs more per megabyte of storage to manufacture. Due to faster transfer speeds and longer life, SLC flash technology is used in high-performance memory cards.

**Multi-level cell**

A multi-level cell (MLC) is a memory element capable of storing more than a single bit of information. MLC NAND flash is a flash memory technology using multiple levels per cell to allow more bits to be stored as opposed to SLC NAND flash technologies, which use a single level per cell. Currently, most MLC NAND stores four states per cell, so the four states yield two bits of information per cell. Lower priced drives usually use MLC flash memory, which is slower and less reliable than SLC flash memory.

**Host interface**

The host interface is not specifically a component of the SSD, but it is a key part of the drive. The interface is usually incorporated into the controller discussed above. The interface is generally one of the interfaces found in HDDs. They include:

- Serial ATA
- Serial-attached SCSI (generally found on servers)
- PCI Express
- Fibre Channel (almost exclusively found on servers)
- USB
- Parallel ATA/IDE interface (mostly replaced by SATA)
- (Parallel) SCSI (generally found on servers; mostly replaced by SAS)
**Form factor**

The size and shape of any device is largely driven by the size and shape of the components used to make that device. Traditional HDDs and optical drives are designed around the rotating platter or optical disc along with the spindle motor inside. If a SSD is made up of various interconnected integrated circuits (ICs) and an interface connector, then its shape could be virtually anything imaginable because it is no longer limited to the shape of rotating media drives. Some solid state storage solutions come in a larger chassis that may even be a rack-mount form factor with numerous SSDs inside. They would all connect to a common bus inside the chassis and connect outside the box with a single connector.

**Standard HDD form factors**

The benefit of using a current HDD form factor would be to take advantage of the extensive infrastructure already in place to mount and connect the drives to the host system. These traditional form factors are known by the size of the rotating media, e.g. 5.25", 3.5", 2.5", 1.8", not by the dimensions of the drive casing.

**Challenges Facing SSD Drive Technology**

Having discussed SSD architecture helps us understand the technology a little better. The following points list the challenges facing the SSD drive technology:

1. One of the biggest challenges facing the SSD’s is its stability and longevity.
2. Known write cycles limits of NAND flash memory that this technology uses caused people to think about the reliability of this drives.
3. Lack of industry standards for SSD’s.

The technology has matured to the point where there are fewer concerns about this problem. Standards are developing for this technology and standards play a critical role in technology adoption and proliferation.
Standards for SSD Drive Technology
The Joint Electronic Devices Engineering Council (JEDEC) is the global leader in developing open standards for the microelectronics industry, with more than 3,000 volunteers representing nearly 300 member companies.

To achieve the goal of consensus-based industry standards for SSDs, JEDEC’s JC-64.8 Subcommittee for Solid State Drives has taken the lead to provide meaningful, real-life, endurance and reliability metrics to better enable customers to select the right SSD for their expected applications and workloads. In September 2010, JEDEC announced the publication of two widely anticipated standards for solid state drives.

Going forward, JC-64.8 plans to continue work on additional SSD standards. All interested companies are encouraged to join JEDEC and participate in the development effort.

Application Classes
For each class of SSDs defined in the standard, JESD218 Solid-State Drive Requirements and Endurance Test Method define conditions of use and corresponding endurance verification requirements. As SSDs are subject to different levels of demand depending on the applications in use, the standard defines two application classes: Client and Enterprise. It further establishes specific requirements for each, an approach intended to help consumers and enterprise IT managers choose products that are the best fit for their needs.

Endurance Rating and Verification
JESD218 also creates an SSD Endurance Rating that represents the number of terabytes written by a host to the SSD (TBW), which provides a standard comparison for SSDs based on application class. A standard endurance rating will be a welcome change for end users seeking to compare SSDs from different manufacturers. In addition, the standard establishes two approaches—direct verification and extrapolation—for endurance and retention verification.

Workload
Since workloads are expected to change as applications evolve, they are described in a separate, complementary standard: JESD219 Solid-State Drive Endurance Workloads. Because workloads that a SSD is subjected to have a significant impact on the amount of data that may be written to a drive, a standard workload is required to have
comparable results. At the present time, JESD219 defines workloads for enterprise applications only; client workloads will be added in the near future.

**Current State of Digital Universe Storage**

The Digital Universe has already entered the Exabyte Era, and we are quickly moving toward the Yottabyte Era. The following references will help you understand the current state of the Digital Universe:

1. According to an International Data Corporation (IDC) paper sponsored by EMC Corporation, 161 Exabyte’s of data were created in 2006, "3 million times the amount of information contained in all the books ever written".

2. According to the June 2009 update of the Cisco Visual Networking Index IP Traffic Forecast, by 2013, annual global IP traffic will reach two-thirds of a zettabyte or 667 Exabytes. Internet video will generate over 18 Exabytes per month in 2013.

**Conclusion**

The Digital Universe is growing at an alarming rate and is also demanding more performance (IOPS). Traditional HDD’s are unable to keep up with the demands of IOPS. There is a need for the growth of EFD’s which will meet the IOPS needs of organizations and also help them deliver energy-efficient solutions that lower power consumption and reduce space requirements.

Market standards such as JEDEC are also getting in place with the hope that companies work together and quickly adopt these standards and produce high quality, reliable EFD’s.

The best way to move forward is to start with the Hybrid Model. Doing so would increase the demand for this technology and help foster development and promote growth in this technology.

Let’s enter the Yottabyte Era using Enterprise Flash Drive technology.
References


EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.