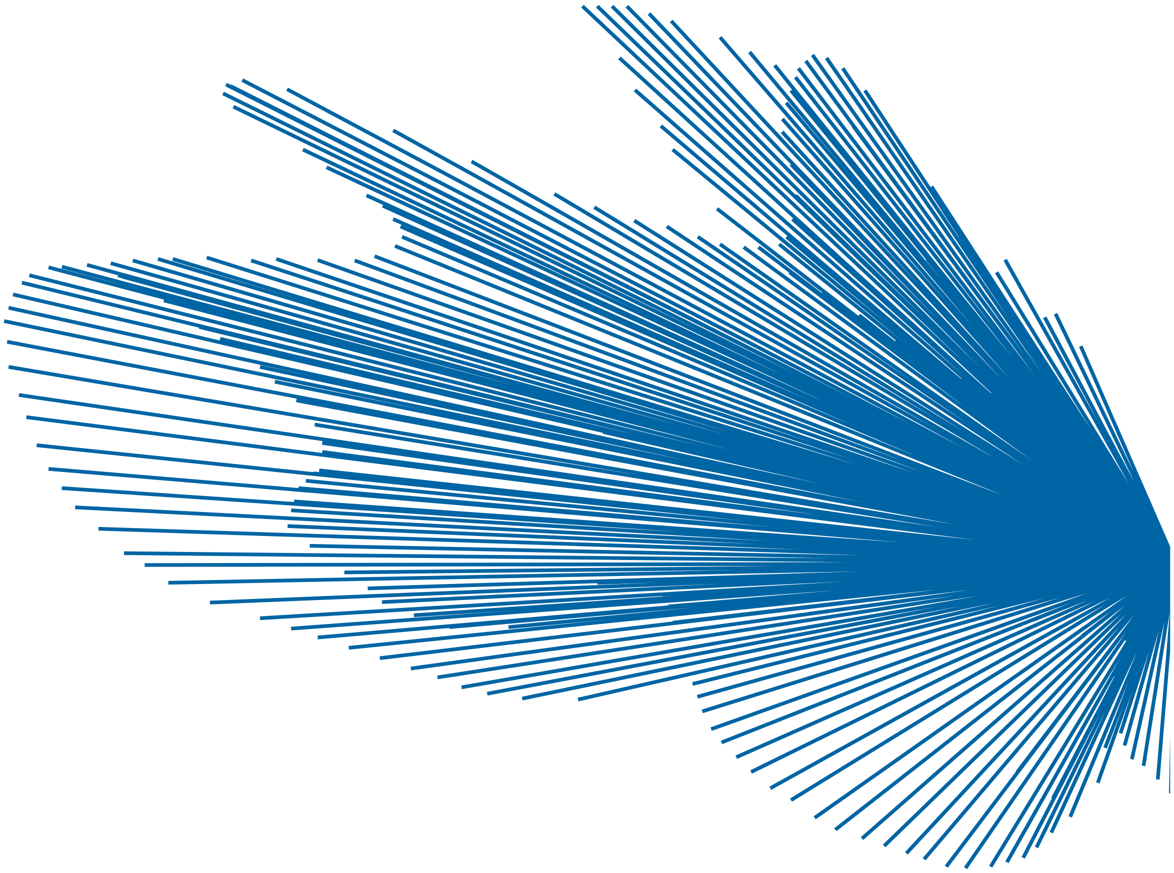
Whitepaper | August 2013



Deciphering SSD Endurance

Which SSD will hold up under your workload?

HGST
a Western Digital company

Table of Contents

Executive Overview

Solid-state drive (SSD) endurance is an important factor in system design. Due to characteristics of flash media, repeated write cycles eventually wear out an SSD. Drive replacement is more frequent under heavy enterprise workloads, which can raise the total cost of ownership (TCO) and periodically cause application downtime. Indeed, once endurance is taken into account, a solution that initially seems less costly may actually, over time, be more expensive. Cost per terabyte written is now required for users to know over the life of their SSDs.

Before purchasing an SSD, it is important to understand your application workload and how long an SSD will last in that environment. Vendors may measure and specify SSD endurance differently, so look closely at how specifications are derived. Finally, select an SSD that supports your workload requirements at the lowest TCO.

HGST's advanced CellCare® Technology extends and maximizes drive lifetime. This unique technology cuts costs and delivers other SSD benefits, such as greater reliability and consistent performance with use.

SSD Endurance Characteristics

SSDs are succeeding HDDs as the storage solution of choice for high-performance enterprise workloads. Built from NAND flash memory, SSDs are similar to hard drives in that they store data persistently. . However, SSDs offer such advantages as:

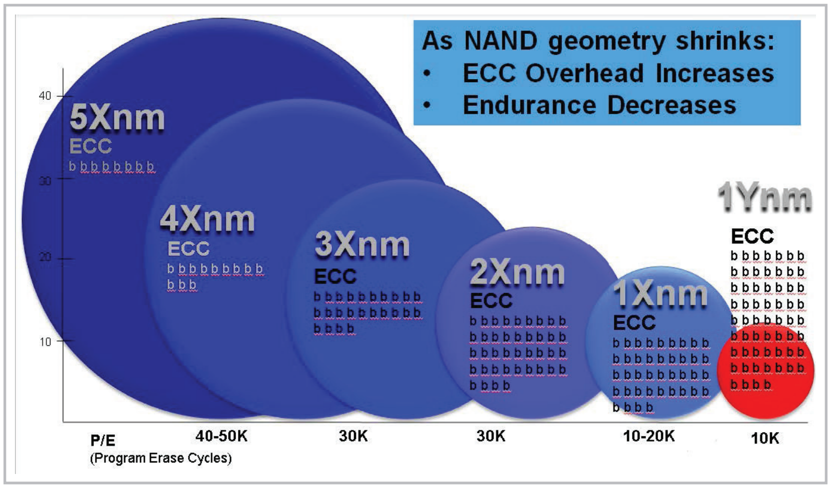
- Faster data access, as much as 1000 times the random I/O performance of HDDs
- Drive greater data center energy efficiency by reducing the overall space and energy footprint
- Lower price on an input/output per second (IOPS) basis
- Cost per terabyte written is a new SSD metric to help evaluate various brands and models

A factor to consider in system design is that SSDs eventually wear out with use due to the nature of flash media. Program erase (P/E) cycles slowly wear out flash media, while reads have no effect and can be sustained indefinitely. Consequently, flash endurance is measured in P/E cycles, with memory cells able to survive anywhere from 3,000 to more than 100,000 cycles. Cell architecture is responsible for some of that difference. For instance, single-level cell (SLC) architectures store one bit per cell. This approach yields memory that is more durable and faster than that of multi-level cells (MLCs), which store multiple bits per cell. MLCs produce denser and more cost-effective storage—advantages that are making MLCs an increasingly popular choice.

Additional endurance impact arises from shrinking process features, with these dropping from around 40 to 19 nanometers—or billionths of a meter—over the past few years. This reduction has cut the cost of flash storage, but the smaller transistors are more prone to bit errors and other endurance-decreasing effects. As NAND flash geometries shrink, P/E cycles decrease, while the error correction code (ECC) overhead required actually increases. As ECC requires drive resources, the more ECC that is needed will have a negative impact on the overall speed and performance of the SSD. (see Figure 1)

In an SSD, flash is written in pages that vary from 1 to 8 kilobytes in size. Changing any bit on a page causes

Figure 1: ECC and NAND geometry scale showing P/E reduction.



it to be erased and rewritten. Due to this and other sources of write amplification, the data written to flash is a multiple of the data actually written by the host, with the exact ratio varying by workload. More specifically, changing a bit on a page is done by writing a new copy of the page elsewhere, and marking the old one for deletion. When the "garbage collector" (GC) comes around, it tosses the marked-for-deletion pages, but has to copy those pages that are still valid to a new erase block before it can clean up the old one.

A variety of techniques are used to extend flash memory endurance. The memory cells themselves can be made more robust. For instance, enterprise-suitable eMLC provides greater endurance and speed than consumer-targeted cMLCs. Other techniques involve the controller and the way data is written to the SSD. In wear leveling, writes are adjusted so that they are spread evenly. Error correction code uses mathematical algorithms—and extra parity bits—to catch and fix errors. Compression and deduplication squeeze data down so that less is written. Memory can be reserved, or overprovisioned, so that failed blocks can be replaced as needed. Overprovisioning ensures that failed blocks don't subtract from the device capacity. More important, overprovisioning reduces write amplification by effectively giving the GC algorithm more space to play with.

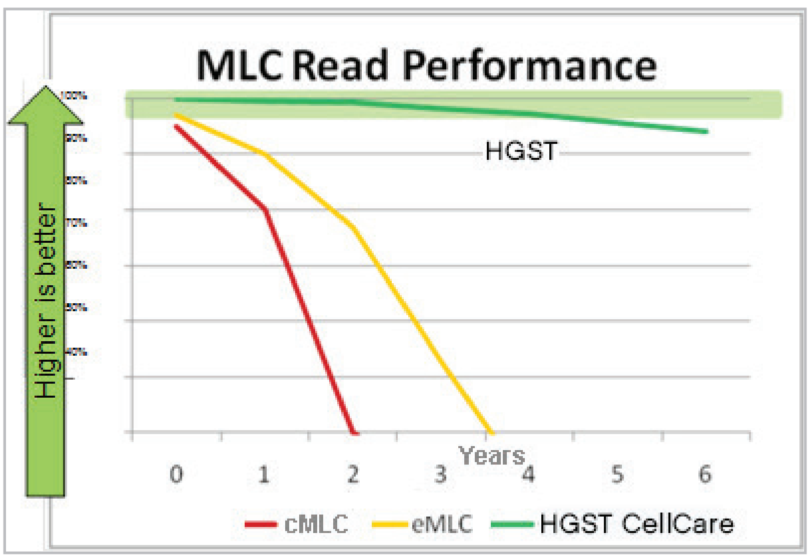
All of these methods have drawbacks. Overprovisioning adds to the cost of flash media. Compression and deduplication can be computationally intensive, and their effectiveness depends on the data pattern. Not all enterprise workloads are compressible. Wear leveling and error correction do not directly address underlying endurance issues. They only try to make best use of the current flash characteristics, but don't actually improve the flash P/E cycles—or endurance—of SSDs.

HGST SSD Technology Delivers Unmatched True Endurance

The long lifetime of HGST solutions is the result of careful attention to, and correction of, endurance-robbing factors of flash media. For instance, there is CellCare, a proprietary flash controller technology that dynamically measures the health of cells as they wear. From this, the technology applies predictive techniques to optimize reads and writes, thereby creating an adaptive feedback loop between flash memory and controller.

One benefit is that HGST solutions require much less overprovisioning and achieve higher levels of endurance for longer SSD life. Whereas some SSDs set aside as much as 37% for overprovisioning, HGST technology accomplishes the same result at only 23%.

Figure 2: HGST CellCare sustains performance over the life of the drive.



Another advantage is that P/E cycles for MLC flash are improved to 70K. With CellCare Technology, HGST SSDs can sustain more than 10 continuous full-capacity drive writes per day for three or more years without sacrificing performance (see Figure 2). It also sustains the health of the drive, and ensures that data stored is secure, reliable and available. Additionally, HGST has created several models with unlimited write options. For those solutions, endurance-related costs are completely removed from the TCO equation.

Not All Endurance Specifications Are the Same

In looking at endurance, it is important to remember that not all specifications—or warranties—are the same. For instance, with regard to the latter, some vendors have stated that their products are good for either 5 years or 1 petabyte written (PBW)—whichever comes first. This is sometimes shown as the number of drive writes per day. However, the total number of petabytes the SSD can write over its lifetime is the ideal way to measure and compare SSDs, and allows for matching to a given use case.

Manufacturers use a variety of methods to specify endurance. Some opt for a 50/50 mix of random and sequential writes, which is not a real-world scenario. Others write random blocks of data but do not reveal the block size, and sometimes the workload is not published at all. Thus, it can be difficult to judge how stringent or reliable an endurance rating is.

In contrast, HGST uses the worst-case workload of 100% random writes of 4K blocks to provide customers with an endurance rating that can be easily relied upon. The sample endurance ratings for SAS/SATA SSDs in Table 1 show that HGST offers the greatest endurance even under the most stringent workloads—nearly twice as durable as rival SSDs.

Table 1: Sample SAS/SATA SSD endurance ratings from manufacturers.

Brand	Interface	Capacity	Endurance	Workload	Notes
Vendor A	SATA	400GB	7.3 PBW	4K Random, 100% Write	Based on vendor documentation
Vendor B	SAS	400GB	3.3 PBW	Random, 100% Write	Estimated from vendor documentation in which vendor stated “7.9PBW 50% seq./50% random.”
Vendor C	SAS	400GB	10.2 PBW	4K Random, 100% Write	Previous model endurance was listed at 1 PBW
HGST s842	SAS	400GB	**13.9 PBW**	4K Random, 100% Write	HGST delivers the best endurance even when using worst-case workload
Vendor D	PCIe	1.2 TB	16.3 PBW	4K Random, 100% Write	Estimated from vendor documentation
HGST s1122	PCIe	1.6 TB	**34.0 PBW**	4K Random, 100% Write	HGST delivers the best endurance even when using worst-case workload

When considering endurance specifications, remember that many data centers and enterprise applications use SSDs for heavy, continuous workloads. Hence, it is very important to understand how an SSD vendor measures endurance and if it applies to a given situation. After all, price, capacity and performance become meaningless if a drive fails due to endurance not keeping up with the use case.

SSD Endurance Affects TCO and Application Availability

Greater endurance pays off in the bottom line. The TCO for an SSD over a typical 3 to 5-year lifespan begins with the initial cost. To that is added:

- Cost of replacement SSDs
- Cost of application downtime or slowdown during the replacement procedure
- Cost of labor to install and configure replacement SSDs

The price per PBW for the s1122 SSD is significantly lower then the competiion because it is far more durable. The total write workload that an s1122 SSD can sustain over a 5-year lifetime could require up to three or four of the competitive products—the initial card plus a series of replacements. As a result, the TCO of the HGST solution could be as low as 25% to 35% of the price of competitors' SSDs when looking over the life of the drive. These figures only consider hardware costs alone. Additional costs could include server downtime, lost productivity and hardware replacement labor.

In another example comparing enterprise SAS SSDs, the s842 SSD from HGST is not only priced lower per drive than two major competitors, but lasts longer and delivers greater random write throughput. The list price of the S842 SSD is 14 to 20% lower, but when lifetime endurance is factored in, the savings is as much as 75%.

To the cost of hardware must be added the cost of application downtime. For instance, an HGST customer in the Oil and Gas industry, GeoEnergy, found that replacing a failed SSD led to four hours of downtime, during which jobs were not running, deadlines were not being met, and revenue was not being generated. Warranties are great in protecting you from having to buy another device, but they don't address costs associated with downtime. For example a 5-year warranty on a drive where you have to replace sooner will still cause downtime, labor and maintenance costs.

Another component of TCO is the administrative overhead to monitor, procure and replace SSDs.

All of these figures and calculations are informed estimates, not exact science. However, it is important to think through your endurance requirements and understand an SSD's capabilities to make an optimal purchasing decision. The last thing you want is to select a drive that seems inexpensive but costs more in the long run and periodically wreaks havoc on application availability for your users.

SSD Endurance Checklist

1) Understand SSD endurance specifications:

- √ What is the rated total drive-writes per day or petabytes written (PBW)?
- √ Is the workload used to measure endurance worst-case (4K random,100% write), or something easier?
- √ What is the warranty period? Is warranty subject to a media wear-out limitation?

2) Estimate your application workload:

- √ What is the block size and read/write mix?
- √ Random or sequential? Continuous or periodic?
- √ What is the IOPS throughput requirement?
- √ Are there latency or caching software requirements to consider?

3) Calculate SSD endurance for your workload:

- √ SSD lifespan = PBW / Daily write workload

4) Estimate SSD TCO over expected system lifespan by adding:

- √ Cost of SSD and replacements, if any
- √ Cost of application downtime or slowdown during replacements √ Cost of labor for replacements

Conclusion

In summation, SSD endurance is an important but often overlooked characteristic. HGST SSDs maximize endurance using advanced CellCare Technology, and the endurance is reported at the worst-case workload. Thus, an HGST drive is more durable and will perform the same after years of use as it did fresh from the factory. When evaluating vendors, ask if they can make the same claim.

The TCO for HGST drives is lower over time due to the advantages of great endurance. Longer SSD life equals fewer or perhaps no replacement drives. That eliminates the associated downtime and labor costs, which is not the case for products of inferior endurance.

Finally, HGST has created several models with unlimited endurance options. Such products eliminate lifetime-related replacement expenses from the TCO equation altogether.

HGST
a Western Digital company

WP017-EN-US-0813-05