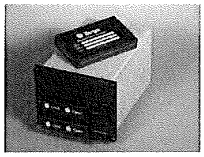
Increasing Flash SSD Reliability

classic article by SiliconSystems .... April 2005



Removable Military Solid State Disks from **Targa Systems**

SSD reliability papers
After SSDs... What Next?
Can you trust flash SSD specs?
Bad block management in flash SSDs
Are MLC SSDs Safe in Enterprise Apps?
SSD Myths and Legends - "write endurance"
Can you believe the word "reliability" in a 2.5" SSD ad?

SILI
SYSTE
The Future of Storage

WD Western Digital Solid State Storage
How long will this SSD last?
Forecast SSD life. No more guesswork.

Editor's intro:- Solid state disks, based on flash technology, have greatly improved in performance in recent years and now compete head to head with RAM based accelerator systems. Flash also has significant advatanges in servers compared to RAM SSDs due to low power consumption. But if you think that all solid state disks which use flash are equally reliable and enduring then think again. That's a bit like saying that a Mercedes 300SL sports coupe is as tough as a Tiger tank because both were made in Germany and both are built out of metal. But as Oddball (Donald Sutherland) says in the movie Kelly's Heroes "I ain't messing with no tigers." This article by SiliconSystems, shows how their patented architecture cleverly manages the wear out mechanisms inherent in all flash media to deliver a disk lifetime that is about 4x greater than of other enterprise flash products and upto **100x** greater than intrinsic flash memory.

SiliconSystems' SiliconDrive technology is specifically designed to meet the high performance, high reliability and multi-year product lifecycle requirements of Enterprise System OEMs in the netcom, military, industrial, interactive kiosk and medical markets. One of the measures of storage reliability in Enterprise System OEM applications is endurance. Endurance is defined as the number of write/erase cycles that can be performed before the storage product "wears out."

It is important to note that endurance is not just a function of the storage media. Rather, it is the combination of the storage media and the controller technology that determines the endurance. For example, magnetic media is an order of magnitude less reliable than NAND flash, yet the controller technology employed by rotating hard drives can compensate for this deficiency.

{NOTE: This is just an example of how a controller, if it is well designed, can compensate for the deficiencies of the media. It is a completely different discussion to compare the mechanical reliability of rotating hard drives to solidstate storage that has no moving parts.}.

Write/erase cycle endurance for solid-state storage is specified in many ways by many different vendors. Some specify the endurance at the physical block level, while others specify at the logical block level. Still others specify it at the card or drive level. Since endurance is also related to data retention, endurance can be specified at a higher level if the data retention specification is lower. For these reasons, it is often difficult to make an "apples to apples" comparison of write/erase endurance by solely relying on these numbers in a datasheet. A better way to judge endurance is to break the specification down into the main components that affect the endurance calculation

- Storage Media
- Wear-Leveling Algorithm
- Error Correction Capabilities

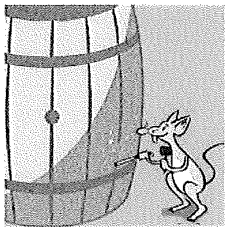
Other factors that affect endurance include the amount of spare sectors available and whether or not the write is done using a file system or direct logical block addressing. While these issues can contribute to the overall endurance calculation, their effects on the resulting number are much lower than the three parameters listed above. Each of these factors will be examined individually, assuming ten-year data retention.

The final section of this white paper provides a calculator to assist in the understanding effects of each of these parameters on the overall endurance in an application.

Storage Media

The scope of this white paper is confined to non-volatile storage – systems that do not lose their data when the power is turned off. The dominant technology for non-volatile solid-state storage is NAND flash. While NOR flash is also a possible solution, implementation of NOR technology is generally confined to cell phone and other chip-on-board applications. For these applications, NOR provides execute-in-place, boot and data storage functionality in a single chip. The economies of scale and component densities of NAND relative to NOR make NAND the ideal solution for non-volatile solid-state storage systems.

The two dominant NAND technologies available today are SLC (single-level cell, sometimes called binary) and MLC (multi-level cell). SLC technology stores one bit per cell and MLC stores two bits. A comparison of SLC and MLC is shown in Figure 1.



SSDs on StorageSearch.com

Over **130** SSD manufacturers listed and profiled on our classic SSD page.

2.5" SiliconDrives from **Western Digital**

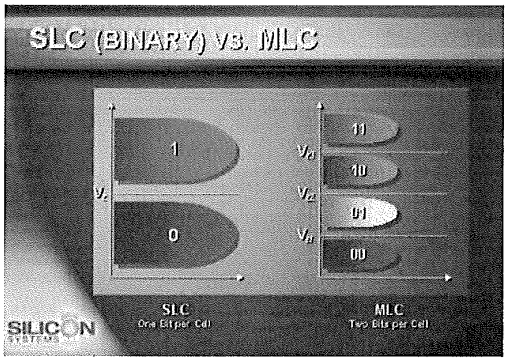
...Later:- in January 2006 - SiliconSystems published more information about how they were engineering increased reliability into flash disks. Below is the text of their press release.

SiliconSystems Introduces the Industry's First Self-Monitoring Solid-State Drive

ALISO VIEJO, Calif., January 30, 2006 - **SiliconSystems, Inc.** today announced a breakthrough storage system monitoring and usage technology called SiSMART.

This new, patent-pending technology accurately monitors storage system usage to predict useable life, and is incorporated into the company's entire SiliconDrive product line. By monitoring read/write activity, SiSMART technology provides users of solid-state storage systems a level of confidence and accuracy about the viability of their storage solutions previously unobtainable.

SiSMART technology constantly monitors and reports the exact amount of storage system useable life available allowing users to make any necessary adjustments or schedule preventative maintenance to ensure system availability and data integrity.



SLC NAND is generally specified at 100,000 write/erase cycles per block with 1-bit ECC (ECC is explained in greater detail in this white paper). MLC is generally specified at 10,000 cycles with ECC. While the datasheet for the MLC device does not specify the level of ECC required, the MLC manufacturers recommend 4-bit ECC when using this technology. Therefore, when using the same controller, a storage device using SLC will have an endurance value roughly 10 times that of a similar MLC-based product. A more thorough discussion of SLC versus MLC components can be found on the respective websites of various NAND flash component manufacturers on their respective websites.

Wear Leveling

Wear leveling allows data writes to be evenly distributed over the storage media. More precisely, wear leveling is an algorithm by which the controller in the storage device re-maps logical block addresses to different physical block addresses in the solid-state memory array. The frequency of this re-map, the algorithm to find the "least worn" area to which to write and any data swapping capabilities are generally considered proprietary intellectual property of the controller vendor.

It is important to note that wear leveling is done in the solid-state memory controller and is independent of the host system. The host system performs its reads and writes to **logical** block addresses only. So as far as the host is concerned, the data does not move.

To illustrate the effects of wear leveling on overall endurance, assume three different storage devices with the following characteristics:

- Flash Card with no wear leveling
- Flash Card with dynamic wear leveling
- SiliconDrive with static wear leveling

In addition, assume that all three storage devices use the same solid-state storage technology (SLC or MLC – for purposes of this discussion, it doesn't matter). All three devices will have 75% of their capacity as static data, which is defined as any data on a solid-state storage device that does not change. Examples of static data include operating system files, look-up tables and executable files.

Finally, the same type of write is performed to all three systems. The host system single block of data to the same logical block address over and over again.

No Wear Leveling

Figure 2 (below) shows a normalized distribution of writes to a flash card that does not use wear leveling. In this instance, the data gets written to the same physical block. Once that physical block wears out and all spare blocks are exhausted, the device ceases to operate, even though only a small percentage of the card was used.

In this instance, the endurance of the card is only dependent on the type of flash used and any error correction capabilities in excess of one byte per sector. Early flash cards did not use wear leveling and thus failed in write-intensive applications. For this reason, flash cards with no wear leveling are not recommended for Enterprise System OEM applications.

SiSMART technology is ideal for enterprise system OEM applications in the netcom, military, industrial, interactive kiosk and medical markets.

"Solid-state storage technology offers major benefits over rotating disk drives, such as added security and unmatched ruggedness, but until now there were valid concerns about the inability to accurately predict storage system lifespan," said Michael Hajeck, CEO at SiliconSystems. "After receiving overwhelmingly positive feedback from some of our most demanding tier-one customers regarding the accuracy and dependability of our SiSMART technology, we decided to include this breakthrough technology in our complete line of SiliconDrive products."

Solid-state drive lifespan is becoming a topic of concern among analysts, according to Gartner Senior Analyst Joseph Unsworth. "There is a strong need in the market for a means to track drive usage and make more accurate predictions concerning lifespan. Technology that enables customers to have the ability to set their own parameters and anticipate when problems will arise with their drives will be attractive in order to manage risk."

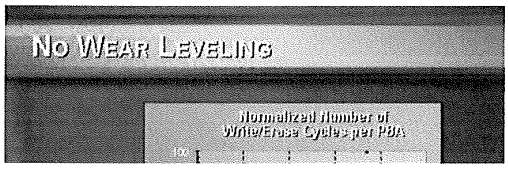
Achieving What SMART Technology Cannot

Rotating hard disk drives employ Self Monitoring and Reporting Technology (SMART), which was designed to act as an early warning system for pending problems with mechanical media. Though this technology is useful for monitoring wear on rotating hard disk drives, it cannot be used to monitor the useful life of a solid-state drive. Since solid-state storage products have no moving parts many of the parameters monitored by the SMART function are not applicable.

Solid-state storage components, which are the fundamental building blocks of every solid-state drive, can lose the ability to retain programmed data after hundreds of thousands to millions of write/erase cycles. With no method to determine or predict when write/erase cycle endurance will be exceeded, a solid-state storage product is typically allowed to operate until it ultimately fails, leading to unscheduled system down times and significant data loss.

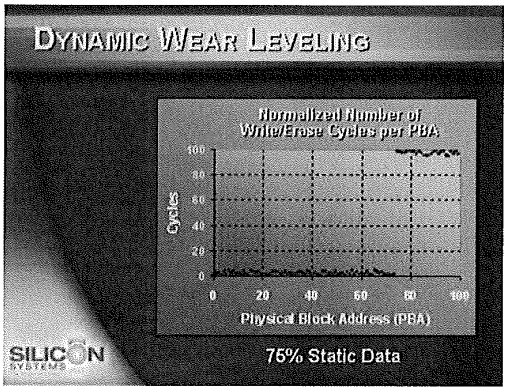
In contrast to SMART, SiSMART monitors how many write/erase cycles have occurred on a solid-state storage system -- the only real failure mechanism present in solid-state storage. By incorporating a patent-pending algorithm that tracks all data transactions internally in the SiliconDrive, SiSMART is able to accurately monitor and report storage system usage to the host system. This enables users to model future usage, set thresholds to perform maintenance and adjust data collection requirements to match the required life of the deployed equipment.

Beginning in February 2006, SiliconSystems' entire SiliconDrive product offering will come equipped with both SiSMART technology and the company's patented PowerArmor technology. PowerArmor is another innovative technology from SiliconSystems that was developed to eliminate storage system field failures by virtually eliminating drive corruption and data loss in the event of unexpected power disturbances. SiliconDrives equipped with SiSMART and PowerArmor will provide a level of data integrity and data reliability never previously available.



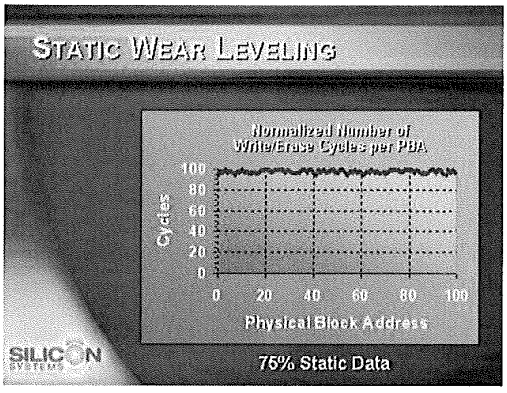
Dynamic Wear Leveling

Figure 3 (below) shows a normalized distribution of writes to a flash card that employs dynamic wear leveling. This algorithm only wear levels over "free" or "dynamic" data areas. That is to say, if there is static data as defined above, this area is never involved in the wear leveling process. In the current example, since 75% of the flash card is used for static data, only 25% of the card is available for wear leveling. The endurance of the card is calculated to be 25 times greater than the card with no wear leveling, but only one-fourth that of static wear leveling.



Static Wear Leveling

Figure 4 (below) shows a normalized distribution of writes to a SiliconDrive that employs static wear leveling. This algorithm evenly distributes the data over the entire SiliconDrive. The algorithm searches for the least-used physical blocks and writes the data to those locations. If these locations are empty, the write occurs normally. If they contain static data, the static data is moved to a more heavily-used location prior to the new data being written. The endurance of the SiliconDrive is calculated to be **100 times better** than the card with no wear leveling and, in the example discussed here, four times the endurance of the card that uses dynamic wear leveling.



Error Correction

Part of a solid-state memory component specification is related to error correction. For example, SLC NAND components are specified at 100,000 write/erase cycles with one-bit ECC. It goes to reason that the specification increases with a better error correction algorithm. Most flash cards employ error correction algorithms ranging from two-bit to four-bit correction. SiliconSystems' SiliconDrive technology is based on the Company's industry-leading six-bit correction.

The term six-bit correction may be slightly confusing. Six-bit correction defines the capability of **correcting up to six bytes in a 512-byte sector**. Since a byte is eight bits, this really means the SiliconDrive can correct 48 bits as long as those bits are confined to six bytes in the sector. The same definition holds true for two-bit and four-bit correction.

The relationship between the number of bytes per sector the controller can correct is not directly proportional to the overall endurance, since the bit error rate of NAND flash is not linear. To

state it another way, six-bit error correction is more than three times better than two-bit ECC since the probability of getting a three-bit error is significantly greater than the probability of a seven-bit error.

Summary of Media,, Wear Leveling and ECC

There is much confusion about the definition of "industrial grade." Many companies are seeking to only define industrial grade in terms of the solid-state memory components in the storage device – namely SLC vs. MLC NAND. While this is an important issue, the capability of the controller to compensate for the media is even more significant. Use of wear leveling and error correction technologies can dramatically affect the reliability and enhance the usable life of the storage device in an Enterprise System OEM application.

The matrix below summarizes the effects of the different items discussed throughout this white paper.

In the table (below), a **"1" indicates the best** possible endurance scenario, and a "10" indicates the least desirable configuration. Values 2-9 are a bit more subjective, but their relative positioning makes sense in the context of most types of data transfers.

N = No Wear Leveling; D = Dynamic Wear Leveling; S = Static Wear Leveling

ECC	SLC NAND			MLC NAND		
	N	D	S	N	D	S
2-bit	6	5	4	10	9	7
4-bit	5	4	2	9	8	6
6-bit	4	3	1*	8	7	5

* = *SiliconSystems' SiliconDrive Configuration*

Wear leveling is important as it allows data writes to be evenly distributed over the entire storage device. A device with no wear leveling wears out faster because data is written to the same physical block. Flash cards that use dynamic wear leveling algorithm only write across dynamic or free data areas. By far the best endurance is provided by static wear leveling, where the data is written equally to all blocks of the storage device.

Equally important is the error correction capability. Most flash cards use error correction algorithms ranging from two-bit to four-bit correction. Industrial grade solutions should in general use more robust algorithms. SiliconSystems has designed an industry-leading six-bit error correction into its entire product family of SiliconDrives.

SiliconSystems' SiliconDrive technology provides the optimum mix of controller and storage component technology to maximize endurance. SiliconDrives use the powerful combination of the most reliable solid-state memory components currently available, static wear leveling and industry-leading six-bit ECC to deliver highly reliable industrial-grade solid-state storage solutions for Enterprise Systems OEMs.

Endurance Calculations

To get an idea of how long a solid-state storage device will last in an application, the following calculations can be used.

Note: These calculations are valid only for products that use either dynamic or static wear leveling. Use the solid-state memory component specifications for products that do not use wear leveling. To calculate the expected life in years a product will last:

$$\text{Years} = \frac{(\alpha - \beta) \times \lambda \times (1 - \varphi)}{(\omega \times \xi) \times k}$$

Where:

α = Capacity in MB (when converting from MB to GB, MB = GB x 1,024)
β = Amount of Static Data in MB (this value should be 0 for static wear leveling)
λ = Block Level Endurance Specification
φ = Safety Margin
ω = File Size in MB (when converting from KB to MB, KB = MB x 1,024)
ξ = Number of Writes of file size ω per minute
k = Number of minutes per year = 525,600

To calculate the number of data transactions:

$$\text{Transactions} = \frac{(\alpha - \beta) \times \lambda \times (1 - \varphi)}{\omega}$$

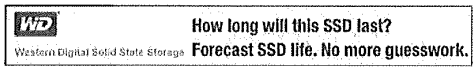
Where:

α = Capacity in MB (when converting from MB to GB, MB = GB x 1,024)

...SiliconSystems profile

Here are some more storage articles you may be interested to read

- the **Fastest** SSDs
  the SSD Buyers Guide
  the Top 10 SSD Companies



Editor:- July 18, 2011 - a new blog on StorageSearch.com looks at PCIe SSDs They may all look alike (bunch of chips on a card) but this is no clone market. What they do and how they do it is diverging more than ever....

STORAGEsearch is published by ACSL