Flash Solid-State Disk Reliability

By Woody Hutsell, Jamon Bowen, and Neal Ekker
Texas Memory Systems

November 2008

Contents

Introduction **1**
Problems at the Chip Level **2**
Evolution of Flash *2*
Problems with Endurance *3*
SLC vs. MLC Flash *4*
Flash Reliability Solutions at the Board Level **5**
Wear Leveling *5*
Adding Flash Capacity for Write Endurance and Write Performance *6*
ECC *7*
Flash Reliability Issues at the System Level **8**
RAID *8*
Caching *9*
Flash Reliability Solutions at the Data Center Architecture Level **10**
Host-side Reliability *10*
Storage Network Reliability *11*
Storage System Reliability *12*
Data Center Reliability *12*
Still Don't Trust Flash…Try These Low Risk Architectures **13**
NAS Read Cache *13*
SAN Read Accelerator with Preferred Reads *13*
Conclusion **15**
Checklist for Enterprise Flash SSD Reliability **16**

Figure 1: 2GB Single Layer Cell NAND Flash Chip 3

Section 1

Introduction

Though SLC NAND Flash has many valuable attributes that make it attractive for use in enterprise storage, its tendency to wear out after a relatively few number of erase/write cycles has until recently hindered its adoption into mission critical enterprise environments. Within the past two years, leaders in solid state drive (SSD) technology engineering have successfully addressed the Flash endurance and reliability issues with a number of innovative solutions. Reliable, high performance Flash SSD products are currently available to enterprise customers.

By taking a narrative journey from the chip point of view all the way up to the enterprise information system architecture perspective, this paper will explore why reliability problems occur in NAND Flash and describe some of the ways these issues are being addressed.

Section 2

Problems at the Chip Level

Evolution of Flash

The truth is, at the chip level Flash does present some challenges for SSD engineers. Flash memory is a 1980s invention that offered a revolutionary property to the integrated circuit community: it could be both programmed and erased electronically and retain its state without power being applied. RAM (Random Access Memory) chips have been available for decades allowing electronic reading and rewriting, but power has to be constantly applied for the RAM to maintain its state. Alternatively, EPROM (Electronically Programmable Read Only Memory) chips allowed manufacturers to electronically program chips that would retain their state without power, but they could only be programmed once and never erased. Then, EEPROM (Electronically Erasable Programmable Read Only Memory) was invented, becoming the first chip that could be both programmed and erased electronically and maintain the programmed value without power.

Despite EEPROM's ability to be rewritten, it was considered, for all practical purposes, to be read-only memory. At first, EEPROM was primarily used to store programs for embedded processors and thus only needed to change when programmers fixed bugs or added new features. This happened fairly infrequently, so EEPROM was ideally suited for this purpose. Programming the EEPROM took several seconds to several minutes but was not considered a time critical event. Additionally, programming and erasing the EEPROM required much higher voltages than the circuits that used the EEPROM, so external tools with higher voltages were used to program or erase the chip. There was a limit to the number of times a chip could be erased and programmed, ranging from several hundred up to one thousand. Production embedded devices rarely see more than several dozen program changes, so this limitation was not seen as a major issue.

The characteristics of EEPROM did not lend themselves well to storage outside of fairly static code. However, the technology evolved. Semiconductor process advances have allowed dramatic miniaturization of electronic circuits: a new chip was designed to encapsulate many EEPROM and programming circuits, at which point Flash was born.

Problems with Endurance

As noted in the diagram, a 2GB SLC (Single Layer Cell) NAND Flash chip is composed of 4 dies, 8 planes, 16,384 blocks, and 1,048,576 2K pages. Each plane has its own block program/erase circuit and pool of blocks. To increase the capacity of a Flash chip without dramatically increasing the cost, four silicon Flash dies are packaged together and share a set of pins coming out of the chip. The four dies are divided into two chip enable (CE) groups; each group operates like a separate chip with two dies.

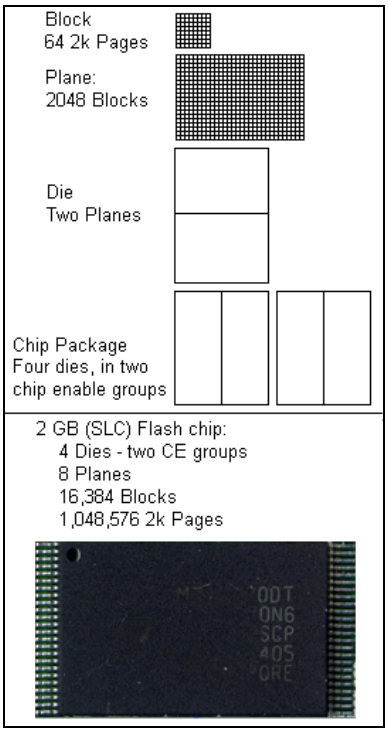

Figure 1: 2GB Single Layer Cell NAND Flash Chip

This organization affects how read and write operations occur in a Flash chip. There are three main operations: read, program, and erase. Each set of dies in a CE group operate independently, as long as the shared data pins are not busy. However, within a CE group, only one of these operations may take place on a plane at a time. Separate operations may take place on separate planes in parallel as long as they are submitted to the CE group at once. The operations are summarized below:

- *Reads* – Reads can occur randomly anywhere within a Flash chip. A read has an access time of <0.1 milliseconds (ms).
- *Erase* – Erasing can only be performed on an entire block. It sets all of the bits in the block to 1. Blocks may be erased randomly within a chip. An erase operation takes approximately 1.5 ms. the erase operation stresses the block and will eventually cause it to fail. For SLC chips, 98% of the blocks will last at least 100,000 write/erase cycles.

- *Program* – Programming should only occur on blocks that are erased, because the program operation can only change a bit from a 1 to a 0. Pages within a block must be programmed sequentially; however, an entire block does not need to be programmed at once. Blocks within a chip can be programmed randomly. A program operation requires a minimum of approximately 0.3 ms.

Flash drives targeted toward the consumer market as either thumb drives or laptop data storage accept the limited endurance or write cycles of Flash and its poor write performance because writes occur fairly infrequently and cost is the overriding concern. In the enterprise, however, the erase stress and endurance issues, coupled with the slow write performance, explain why until recently Flash drives were seldom deployed in mission critical enterprise environments

SLC vs. MLC Flash

Up to this point, the discussion has concerned only SLC Flash. There is a second type of Flash called MLC (Multi-Layer Cell). The two types of Flash are very similar; in fact, their manufacturing processes are nearly identical. The faster, more reliable type is SLC. The difference between the two is fairly straightforward; each holds a voltage to indicate a value stored in physically identical Flash cells. But SLC stores only two values, 1 or 0 (with either a high or low voltage level) while MLC can store four values (high, medium high, medium low, low) representing two bits per cell (00, 01, 10, 11). MLC has a much smaller voltage tolerance than SLC and therefore fails ten times more often because each program erase cycle tends to increase the variance in the voltage that is stored. MLC has a cost advantage because each cell holds two bits instead of one, giving blocks, pages, and chips each twice the capacity. However, this also causes chip operations to take twice as long. This performance penalty, coupled with the 10x greater failure rate than SLC, has ruled out MLC Flash from being suitable for enterprise applications.

Section 3

Flash Reliability Solutions at the Board Level

At the chip level, Flash presented some problems, but it also offers a number of important advantages:

- Nonvolatile
- Fast random reads
- Fairly fast sequential writes
- Low power
- Ruggedness
- Higher density than RAM
- Lower cost per capacity than RAM.

The advantages offered by Flash motivated enterprise SSD vendors to find solutions for the chip endurance and write performance issues. Engineers first explored solutions at the board level, beginning with a concept called wear leveling.

Wear Leveling

The first step to improve Flash write endurance is to use SLC NAND Flash chips, because they have a 10x write endurance per cell advantage over MLC NAND Flash. Chip manufacturers are beginning to offer higher write endurance chips, but key questions remain concerning the impact of these chips on performance and cost.

Because the NAND Flash memory chips wear-out with writes, it is important that frequent data accesses do not occur on a single or limited number of Flash chips. For this reason, every Flash SSD manufacturer incorporates engineering solutions that abstract the physical storage layer from the logical. This means that though the computer system writes data to the same logical address every time, the Flash controller moves the underlying physical address around to the least written Flash memory chips or least written blocks within the chip. This process is what most SSD vendors refer to as wear leveling.

Even this simplest of wear leveling techniques implemented at the circuit board level can be very powerful, For example, using only basic wear leveling the Texas Memory Systems RamSan-500 can operate at peak write loads for years without experiencing reliability issues. The RamSan-500 holds 2TB of usable storage and can sustain a 2GB/s write rate. At 2GB/s it takes 1000 seconds to write across the entire 2TB capacity. To reach the write/erase limit this process would have to be performed 100,000 times (because wear leveling would spread the writes evenly over all the available blocks), requiring 100,000,000 seconds. Dividing 100,000,000 by the number of seconds in a year (31,536,000) gives a Flash write endurance of 3.25 years. This endurance level is achieved with none of the Flash blocks exceeding their endurance specification and using only basic wear leveling, but assuming absolutely no help from the large RAM cache engineered into the RamSan-500 and no help from a large pool of hot spare Flash blocks storage available in the RamSan-500.

A unique characteristic of Flash blocks is that if they are bad, a write to the block will fail, which is easily detected by the Flash controller that then writes the data to a different Flash block. In this way, and other ways as well, Flash SSD controllers monitor the health of individual blocks and chips and when a particular area of a chip has passed a set threshold, its data can be moved and that particular chip or block removed from service with no loss of overall storage capacity, thanks to the extra storage normally included. The controller continues to identify and remove bad blocks, regions of chips, or even entire chips so that ordinary Flash chip wear-out does not result in any data loss or corruption.

While it seems logical to compare Flash SSDs based on their endurance times, this may not be a good way to make a Flash SSD purchase decision. For example, some of the worst Flash SSDs on the market have fantastic write endurance but terrible write performance (because limiting the number of writes improves endurance by decreasing the total number of writes performed). High quality enterprise Flash SSD products solve this problem with more sophisticated wear leveling and added Flash capacity. The discerning buyer will assess their write requirements and attempt to find an SSD that can handle the required number of writes without slowing down their applications.

Adding Flash Capacity for Write Endurance and Write Performance

As mentioned earlier, some Flash SSD products, such as the RamSan-500, have more Flash memory than is reported as usable. This architecture decision is made to increase write performance and plays a role in increasing write endurance by spreading writes across additional Flash chips. Each Flash module in the RamSan-500 contains over 20% more Flash memory than reported.

Why does extra Flash memory increase write endurance? The write endurance for Flash specifies that no more than 2% of blocks will fail after 100,000 write/erase cycles. But what happens to the other 98% of the blocks? The average number of write/erase cycles that the Flash block can sustain is significantly higher than 100,000; by designing a Flash module to tolerate a higher block failure rate a much higher number of write/erase cycles can occur prior to wear out. The RamSan-500 is designed to sustain up to a 10% block failure rate before a module is flagged as needing to be replaced. This increases the number of total system write/erase cycles that can occur to 500,000. This increases the worst case wear-out in a 24x7 2 GB/s write environment from 3.25 years to over 15 years – well beyond the typical life of IT equipment!

ECC

Unfortunately for end users, Flash reliability problems are not limited to write endurance. Flash chips experience unique types of errors that are caused by activity within a chip in nearby cells. These errors are referred to as read-disturbed and write (or erase)-disturbed errors. A disturbed cell is one whose contents have been changed by activity in a nearby cell. These problems are a function of small scale physics at the media level and cannot be prevented by the Flash board designer. The key for the Flash board designer is to prevent these read or write disturbs from impacting data to the server and application. Texas Memory Systems has implemented error correcting codes (ECC) on each Flash module in the RamSan-500. The ECC detects and corrects single bit errors automatically. ECC resolves the majority of problems caused by read or write disturb errors. If the ECC function detects a multi-bit error, then the problem is handled at the RAID controller layer where proper data is reconstructed from the other functioning Flash memory modules.

Combining the various wear leveling strategies with ECC greatly increases the overall reliability of Flash SSD. But some Flash SSD vendors have gone a step further by adding additional system level reliability solutions.

Section 4

Flash Reliability Issues at the System Level

While much can be done to promote the reliability of Flash boards, modules, and drives, it is also the case that some vendors who take a systems view may engineer additional reliability into their SSD solutions. A practical example in the hard drive world is the use of RAID protection for reliability and performance or the use of cache to buffer hard disk drives from small block random inputs and outputs. It turns out that these exact features can be applied to Flash-based solid state storage as well.

RAID

In a traditional HDD storage array, RAID (redundant array of independent disks or redundant array of inexpensive disks) is a way of storing the same data in different places (thus, redundantly) on multiple hard disks. Placing data on multiple disks increases the mean time between failures and fault tolerance, thus increasing reliability. RAID basics can be applied to an array of Flash boards (modules or disks), resulting in similar increases in reliability.

While module level wear leveling and ECC handle the majority of Flash reliability problems, a few other concerns need to be addressed by the system designer, such as multi-bit ECC errors and failed boards due to some other single point of failure on the board itself (or modules or disks). Texas Memory Systems' RamSan-500 resolves these errors at the system level by implementing RAID across the Flash boards in the system. A multi-bit ECC error, a very rare side-effect of read/write disturbs, would normally result in data corruption. But a system such as the RamSan-500 with RAIDed Flash boards is able to reconstruct the corrupted data from multi-bit ECC failures.

A Flash module is much more than just the Flash chips used to store data; it includes some sort of interface connector, interface controllers, Flash controllers, etc. A failure in one of these other components is extremely rare, and yet system designers design solutions that anticipate these potential problems. In this case, RAID protection is the answer; thanks to RAID protection, a failure of a non-Flash component on the module does not result in data loss.

Caching

It is somewhat unusual to think of caching as a mechanism to improve system reliability, but in the case of a large RAM cache in front of Flash RAID, it is an excellent way to improve system write endurance as well as system write performance. The Texas Memory Systems RamSan-500 includes as much as 64GB of RAM cache and a caching algorithm that is explicitly designed to optimize the writes to the Flash RAID subsystem. The large RAM cache isolates the Flash RAID from small block random I/O. This serves to improve write performance and also gives caching algorithm designers a place to aggregate writes into patterns that are Flash friendly. A Flash friendly write is one that makes the most efficient possible use of every Flash erase cycle by writing the most data possible with each write. In this way, caching can decrease the number of write/erase cycles that limit Flash endurance.

If a Flash system is using RAM caching, it is good to be aware of the reliability and persistence of the RAM cache. It is easy enough for a designer to include a small amount of RAM in order to improve write performance, but what happens if a RAM chip experiences a bad bit or bad chip? Well designed RAM caching systems, like the one in the RamSan-500, include ECC (to protect against single bit errors) and Chipkill to protect data from even chip failure. Also of concern is what happens to the data in RAM if external power is lost. Enterprise Flash SSDs need to be able to store data from cache to the Flash subsystem in the event external power is lost. The RamSan-500 solves this problem by including two redundant internal UPS devices inside the RamSan-500. They power the system long enough to flush the RAM cache to the Flash RAID. Texas Memory Systems has developed RAM-based solid state disks for more years than any company in the market and has decades of experience developing solutions for RAM reliability and persistence.

Yet another way RAM cache can be used to protect a Flash system from high write volumes is to actually lock a logical unit of storage (LUN) in the RAM cache. A locked LUN looks like another disk drive to the operating system but happens to be operating entirely out of RAM (with asynchronous backup to the internal Flash modules). It makes the most sense to lock small high write LUNs into cache. In this way, the Flash subsystem is isolated from small block random writes. The RamSan-500 offers this feature, called Turbo, as a way to provide exceptional write performance for enterprise applications and also as a way to promote still greater Flash endurance.

As noted above, SSD engineers are incorporating a wide range of reliability solutions at the board and system levels into Flash SSD products currently available on the market. But reliability strategies don't stop at the product level; data center architectures themselves are optimized for reliability.

Section 5

Flash Reliability Solutions at the Data Center Architecture Level

In the data center, many hardware and software components can come into play to ensure that reliability is not an issue when deploying SSD storage solutions. Normally, Flash SSDs are implemented using the same data protection and performance regimes already established for existing storage. But it's worth noting and discussing a few of the particular wrinkles Flash SSD is bringing to the reliability conversation at the data center architecture level. One of the first and most critical decisions is where to implement SSD, within the servers themselves (internal) or external to the servers.

Host-side Reliability

As the move to blade and cluster computing has shown, data centers are increasingly moving to highly redundant and scalable computing architectures. Blade and cluster servers operate with a processor, RAM, local storage for the operating system, and an interface to external storage. The modern server has little need for large capacities of internal solid state storage because the storage can be more easily protected and accessed on the storage network. Having said that, it is clear that small capacities (< 128GB) of Flash SSD, such as those offered by Intel and Samsung, can be used to store the operating system and applications for these compute nodes. Flash SSD, when used for booting, should be a safe, low power, and effective way to improve boot performance. Vendors determined to have extra reliability can use internal RAID protection (such as mirroring) to get additional reliability. Space and power are a premium in these environments so solutions that optimize reliability without requiring a doubling of components are beneficial.

The main reason that blade computing and cluster computing has taken hold so strongly is that end-users do not want the availability and performance of their most important application dependent on just a single really expensive server. Individual servers are vulnerable to internal failures that can take the server down. When a server goes down everything internal to the server goes down as well. For this very reason, enterprise deployments of solid state storage to accelerate applications will be on the storage network where the storage media can be shared across multiple servers and applications.

Storage Network Reliability

Reliability for external SSD systems is achieved differently from internal SSD devices. External storage, in general, was devised in order to provide storage that was not vulnerable to a host level failure and to allow creation of redundant storage networks with access to highly reliable storage. Enterprises requiring 24x7 uptime use multiple external storage units and replicate/mirror the data between them and external sites to ensure availability.

In the typical data center, either a Fibre Channel storage area network (SAN) or an Ethernet based Network Attached Storage (NAS) architecture is already in place. For solid state disks, attaching to existing NAS filers also involves using Fibre Channel (while the filer connects to the server infrastructure with Ethernet). Fibre Channel attached solid state disks are the best choice for deployment into existing data center and application architectures. Fibre Channel offers very low latency, very high performance, and world-class reliability characteristics that can only be earned after years of real-world deployments.

When evaluating Fibre Channel solid state storage, it is important to consider compatibility with various operating systems, host bus adapters, multipathing drivers, OS or application clustering methods, switches, appliances, and NAS filers. Texas Memory Systems RamSan solid state disks benefit from a fourth generation 4Gbit Fibre Channel controller (up to eight ports are available on a RamSan-500). The system is interoperable with every major open systems operating system (Windows, Solaris, Linux, OpenVMS, Tru64, HP-UX, AIX, Tandem, and others). The RamSan is certified with every major HBA and SAN switch product (many of these manufacturers actually use our products to benchmark their own products). Without interoperability, reliability is not even a consideration.

The RamSan-500 supports the various multipathing drivers used in these environments, including Windows MPIO drivers, AIX MPIO, Sun MPxIO, Linux DMP, Symantec Veritas DMP, etc. Multipathing from servers to storage is essential for protecting data availability in the event a host or storage controller fails; it can also be used to improve performance. The RamSan-500 supports active:active multipathing in most of these environments.

In order to promote application availability, most data centers use clustering of servers. Clustering is available in many operating systems and even at the application layer (for example Oracle RAC). Clustering provides availability even if an entire server is lost. Solid state disks selected for this environment must be able to support clustered computing. The RamSan-500 is a SCSI-III compliant system and is widely deployed with application clustering solutions.

Organizations that have deployed NAS have developed architectures that promote availability between the servers and the NAS filer. Additionally, the NAS filer manufacturers build in features like multipathing in order to promote the reliability of the storage network supported by the filer. The Texas Memory Systems RamSan-500 has been certified with the BlueArc Titan and OnStor NAS solutions.

Storage System Reliability

As reliable as enterprise storage systems have become, many end users need levels of protection beyond those inherent in the devices (i.e. all of the layers of protection described in this document). In this situation, it is not uncommon for enterprise end users to implement synchronous mirroring of two similar storage systems. All operating systems can implement host-based mirroring. A host-based mirror sends writes to two or more storage devices simultaneously and then has the ability to read from one or more of these devices to retrieve data. In a host-based mirror environment, the failure of an entire solid state storage system does not impact application availability. RamSan-500 systems can be deployed in mirrored pairs to provide application and performance availability. Mirroring has very little impact on performance (<2%) but can offer tremendous peace of mind for data center architects.

Data Center Reliability

For data centers worried about disaster recovery, some sort of replication needs to be set-up between a main data center and one to many fail-over data centers. There are a variety of methods to handle this problem, ranging from the application to the host to the array. The RamSan-500 has been deployed in each of these environments. As an example, one method that Oracle database users apply to protect data is Oracle's Data Guard. It does redo log shipping, sending redo logs to remote sites so that a database can be restarted in the event the main data center is lost. Similarly, RamSan systems have been deployed with host-resident clustering and replication tools. Companies that use array-based mirroring from companies like EMC or HDS might prematurely write-off RamSan solutions due to the fact that our systems are not managed by the existing array. There are some solid alternative architectures discussed below that are available to help companies gain from solid state performance without any changes to existing disaster recovery operations.

Section 6

Still Don't Trust Flash...Try These Low Risk Architectures

NAS Read Cache

Pools of fast RAM or Flash memory can be used as read cache in network attached storage architectures. NAS vendors have developed read caching features, such as BlueArc's Dynamic Read Caching, where a Flash solid state disk like the RamSan-500 can be attached as the storage for the read caching environments. Such read caching solutions dynamically migrate frequently read data to the cache, with an added benefit of isolating the impact of SSD failures. With these products, writes are still delivered to the same back-end storage that users have grown comfortable with, while reads are accelerated because they are cached on the RamSan-500 system. The beauty of this architecture is that in high read environments raw latency for data accesses decreases from 5+ milliseconds to sub-millisecond and yet system reliability is not affected. Any snapshots or other disaster recovery (DR) processes still occur on the NAS filers. A failure of the read cache will not result in any data loss. This architecture provides a risk free way to benefit from solid state disk.

SAN Read Accelerator with Preferred Reads

Preferred read architectures are essentially read acceleration solutions for SAN environments. With a preferred read implementation the host server mirrors a solid state disk system to the incumbent storage array (such as an EMC DMX). All writes go to both systems, thus write performance will be no faster than the current implementation. All reads, on the other hand, come from the solid state storage device, resulting in a big performance boost for read intensive environments. Many generations of RamSan systems have been deployed in preferred read environments. Preferred reads options (sometimes referred to as preferred-plex) are built into some operating systems and are available as part of Symantec Veritas Foundation Suite.

The preferred read architecture is especially attractive for environments that have array based replication. In these scenarios, all writes continue to go to the array conducting the replication, therefore consistency groups are maintained and the application can be recovered. If the Flash SSD volume fails, there is no data loss.

Many enterprise data centers have adopted DR solutions based on array-based replication. These array-based replication schemes are central to the business, so much so that the business becomes locked into the products from the array manufacturer. Businesses with array-based replication are reluctant to invest in alternative storage technology because introducing new products would lead to a forklift upgrade at multiple data centers. Preferred reads are a method to integrate solid state disks into these data centers that takes advantage of existing array-based replication mechanisms, delivers extreme read performance benefits, and maintains high reliability and availability.

Section 7

Conclusion

Some design compromises at the Flash chip level lead to erase/write limits that would otherwise prevent Flash-based storage from serious consideration and adoption in mission critical enterprise environments. But Flash offers so many benefits that SSD engineers have been motivated to overcome chip endurance issues. At the circuit board level, solutions such as wear leveling and ECC have been widely implemented. At the system or whole product level additional reliability features are being incorporated, including cache regimes and RAID. Finally, at the level of data center architectures Flash SSDs are already becoming commonplace, and so are the strategies, software, and hardware utilities that allow Flash to function reliably in the most demanding environments.

We have reached the point where Flash SSD products now exist that offer as much enterprise-grade reliability as hard disk storage, if not more.

Appendix A

Checklist for Enterprise Flash SSD Reliability

- Chip Level
  - Uses SLC NAND Flash
- Board Level
  - Implements ECC
  - Implements Wear Leveling
  - Extra Flash Chips for wear leveling (and faster write performance)
- System Level
  - Implements RAID
  - Has a large RAM cache to buffer Flash subsystem from random writes
  - RAM cache is flushed to Flash automatically if external power fails
- Storage Network Level
  - OS interoperability
  - Network interoperability
  - Supports multipathing
  - Supports clustering
  - Can be synchronously mirrored
  - Certified with NAS filers
- Risk Free Implementation Options
  - Certified as read cache for NAS filers
  - Can be synchronously mirrored to third party array with preferred reads from SSD.