E-MLC vs. MLC NAND Flash

Application Note AN002

February 2012

Corporate Headquarters: 39870 Eureka Dr., Newark, CA 94560, USA ♦ Tel:(510) 623-1231 ♦ Fax:(510) 623-1434 ♦ E-mail: info@smartstoragesys.com
Flash Design Center: 2 Robbins Road, Westford, MA 01886, USA ♦ Tel:(978) 303-8500 ♦ Fax:(978) 303-8757
Flash Design Center: 2600 W. Geronimo, Chandler, AZ 85244, USA ♦ Tel:(480) 792-8900 ♦ Fax:(480) 792-8901
Asia: Plot 18, Lrg Jelawat 4, Kawasan Perinudstrian Seberang Jaya 13700, Prai, Penang, Malaysia ♦ Tel:+604-3992909 ♦ Fax:+604-3992903

Table of Contents

1 Overview 1

2 E-MLC vs. MLC NAND Flash 1

2.1 Flash Endurance: Program/Erase Cycles 2

2.2 Manufacturing Process 2

3 Advantages of E-MLC-based Solid State Drives 3

3.1 Write Amplification 3

3.2 JEDEC 64.8 Endurance Calculation 5

3.3 SSD Endurance for Sample Applications 6

3.4 Ensuring Uncorrectable Bit Error Rate (UBER) 7

4 Summary 7

1 Overview

An Enterprise grade SSD is defined as much by its reliability as by its performance and endurance. Mission-critical data must be saved, stored over time, and returned intact. Likewise, an Enterprise grade SSD must be able to achieve this goal throughout its expected life time. This dual mandate requires an Enterprise grade SSD to use highly reliable flash components that have the right endurance characteristics.

This application note discusses the multiple reliability benefits of Enterprise MLC (E-MLC) flash and the reasons why it is used in SMART Storage Systems' XceedIOPS SSDs.

2 E-MLC vs. MLC NAND Flash

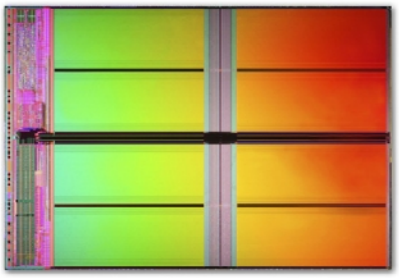
Multi-Level Cell (MLC) NAND flash refers to NAND flash capable of storing two bits per memory cell. Normally, the acronym "MLC" is used to denote commodity-grade memory, and it is sometimes also called out as "cMLC" (for "commodity MLC" or "consumer MLC"). The advantages of MLC over SLC (Single Level Cell) technology are high data density per flash memory device, and lowest cost per gigabyte. The comparative limitations of MLC flash are its lower endurance and relaxed reliability specifications. For mission-critical data and enterprise applications, there is a clear mismatch between the reliability and endurance requirements and what MLC NAND flash can offer.

By contrast, Enterprise MLC (or "E-MLC") NAND flash is designed to provide the cost and density advantages of MLC, while additionally delivering the reliability and endurance required in enterprise environments. E-MLC NAND flash offers densities identical to those of MLC, and up to ten times the Program/Erase cycles (P/E cycles) specifications of MLC. This endurance boost translates directly into better life expectancy for SSDs used in write-intensive applications. E-MLC cost is slightly higher than MLC, but does not approach the large cost delta of Single-Level Cell (SLC) flash.

The two key advantages of E-MLC (when compared to MLC) are:

- Higher P/E cycle specification, resulting in higher endurance and longer lifetime
- Higher reliability through improved manufacturing test processes

Figure 1: Micron NAND flash



2.1 Flash Endurance: Program/Erase Cycles

NAND flash devices specify their lifetime in terms of flash cell endurance, which is measured in Program/Erase cycles. Whenever a flash cell is erased and re-programmed, it sustains a small amount of damage, the accumulation of which eventually renders the cell unreliable over time. Current 3xnm MLC NAND flash devices are specified for between 1,500 and 5,000 Program/Erase cycles, depending on the manufacturer. Note that these specifications can only be supported if the SSD controller employs a high level of ECC correction (i.e., 12 bits correctable per 512 Bytes).

By contrast, E-MLC NAND flash that is available today can sustain up to 30,000 Program/Erase cycles. This primary difference results in a major boost in drive life, and a much higher reliability of the SSD in write-intensive enterprise applications.

Table 1 below shows an overview of the P/E cycles for 3xnm NAND flash technologies, offered by the different flash manufacturers.

Table 1: P/E Cycles Specification for 3xnm NAND Flash Technologies

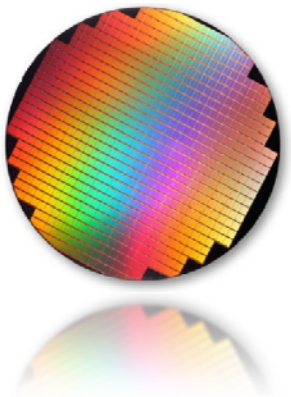
	MLC	E-MLC	SLC
Micron	5K	30K	100K
Samsung	1.5K	N/A	100K
Toshiba	3K	N/A	100K
Intel	5K	N/A	100K

SMART's XceedIOPS SSD product line takes full advantage of E-MLC technology to ensure superior storage drive endurance in high reliability, high availability applications.

2.2 Manufacturing Process

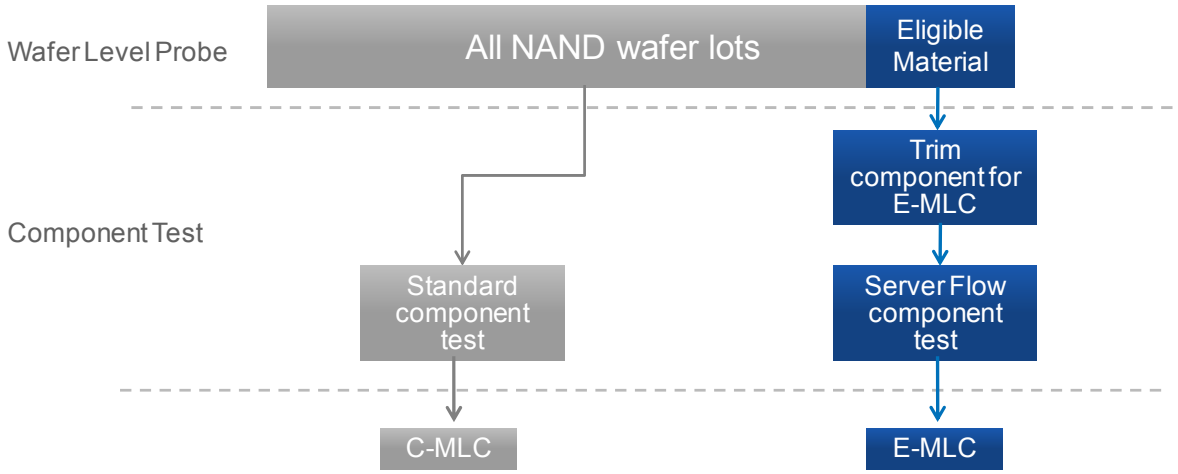
An often misunderstood fact is that E-MLC and MLC NAND flash originate on the same wafer, albeit from a different part of the wafer. In modern high density MLC NAND flash manufacturing, there is a significant data reliability disparity between flash die, even across the same wafer.

Figure 2: Micron NAND Wafer



During the manufacturing process, the flash manufacturer characterizes the flash die on each wafer and identifies those premium die that will provide the highest measure of data reliability and device life—specifications worthy of Enterprise-class systems. These die are trimmed and subjected to a server flow component test cycle, more rigorous than the standard MLC test process. The die that passes this test is specified as E-MLC NAND flash; the rest is sold as MLC NAND flash and specified for greatly reduced reliability and endurance specifications.

Figure 3: NAND Manufacturing of E-MLC vs MLC



Enterprise MLC (E-MLC) flash die differs from MLC die in the following ways:

- Program/Erase Cycle Specification (30K P/E cycles vs 5K P/E cycles for Micron 34nm)
- Slower program/erase timing (tPROG and tBERS)
- Shorter data retention at 100% P/E cycles (3 months vs. 1 year for MLC)

3 Advantages of E-MLC-based Solid State Drives

The primary Enterprise-class reliability advantages an SSD gains from E-MLC NAND flash are:

- Greatly extended SSD drive life under enterprise workloads
- Greater data integrity under those workloads as a result of the superior endurance, which translates to fewer RAISE™ data rebuild events.

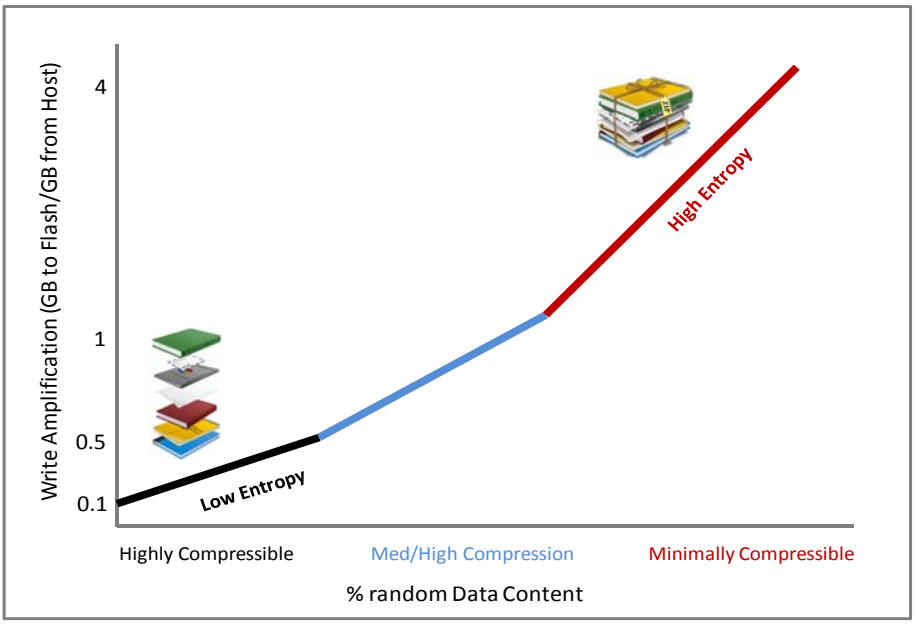
3.1 Write Amplification

Because flash memory must be erased before it can be rewritten, the process to perform these operations results in moving (or rewriting) user data and metadata more than once. This multiplying effect is called Write Amplification, or the ratio of data written to the flash over the amount of data written by the host. Keeping the Write Amplification as low as possible will proportionally increase the lifetime of the SSD.

XceedIOPS SATA SSDs employ a data "awareness" hardware engine to both increase performance and decrease write amplification, resulting in better overall drive endurance.

Figure 4 below shows how the data content, or entropy (a measure for data randomness), effects Write Amplification. Typical database log files and text fields with no or low efficiency encoding are considered to have low entropy. Highly efficient encoded file formats such as MPEG-4 video files that are highly compressed are considered to have high entropy. The Write Amplification of high entropy data content is higher since not much further compression can be achieved and all data needs to be written to the flash.

Figure 4: Write Amplification of Various Data Types



SMART Storage Systems has validated the above approximate Write Amplification graph across multiple drives over hundreds of hours and many hundreds of thousands of I/Os. Below a list of example applications and what write amplification can be expected in these environments:

1. **Mixed random write workload**—Medium degree of compressibility, write operations aligned on 4K boundaries, random starting LBAs. With a 28% over-provisioned XceedIOPS SSD, the average Write Amplification is approximately 1.0.
2. **Database write workload**—Highly compressible data, write operations aligned on 4K boundaries, random starting LBAs. With a 28% over-provisioned XceedIOPS SSD, the average Write Amplification is approximately 0.75.
3. **Video Server workload**—Minimal compressible data, write operations aligned on 4K boundaries, random starting LBAs. Represents a generic worst-case write workload. With a 28% over-provisioned XceedIOPS SSD, the average Write Amplification is approximately 4.0.

The example applications listed above only consider write traffic from the host system. Read traffic intermixed with write traffic does not influence the wear of an SSD, and impacts overall endurance positively because it occupies SSD bandwidth that could otherwise be used for wear-causing write operations.

3.2 JEDEC 64.8 Endurance Calculation

Because E-MLC NAND flash exhibits 6x to 20x better endurance than MLC flash, it extends the SSD lifetime or endurance by an equivalent amount, all other things being equal. According to the JEDEC JESD218 specification[1], endurance ratings of SSDs are represented in terabytes written (TBW) over the SSD's lifetime, provided the following conditions are met:

1) The SSD maintains its capacity
2) The SSD maintains the required Uncorrectable Bit Error Rate (UBER) for its application class (Enterprise or Client SSD)
3) The SSD meets the required functional failure requirement (FFR) for its application class
4) The SSD retains data with power off for the required time for its application class

Table 2 below lists the requirements for both application classes.

Table 2: JEDEC JESD218 application class requirements

Application Class	Active Use (power on)	Data Retention (power off)	Functional Failure Requirement (FFR)	UBER
Client	40˚C 8hrs/day	30˚C 1 year	≤ 3%	≤ 10^{-15}
Enterprise	55˚C 24hrs/day	40˚C 3 months	≤ 3%	≤ 10^{-16}

The endurance calculation used by JEDEC is shown below:

$$\textbf{Teraybtes Written (TBW)} = \frac{\textbf{User Capacity in Bytes} * \textbf{Endurance in PE Cycles}}{\textbf{2} * \textbf{Write Amplification} * \textbf{1,000}}$$

With user capacity in Gigabytes = (User-addressable LBA count – 21,168) / 1,953,504.

An example calculation is given below for an XceedIOPS 200GB SSD (User capacity = 390,721,968 – 21,168/1,953,504 = 200) that uses 28% over provisioning, and is subjected to a database workload (Write Amplification = 0.75):

$$\textbf{Terabytes Written} = \frac{\textbf{200} * \textbf{30,000}}{(\textbf{2} * \textbf{0.75} * \textbf{1,000})} = \frac{\textbf{600,000}}{(\textbf{2} * \textbf{0.75} * \textbf{1,000})} = \textbf{4,000 TBW}$$

[1] JEDEC JESD218 specification is available for download http://www.jedec.org/

3.3 SSD Endurance for Sample Applications

For the three example applications described in section 3.1, calculations for amounts of Terabytes written for both Client and Enterprise SSD are provided, illustrating the positive effect of E-MLC's superior endurance on overall TBW.

The calculations compare a well-known Client SSD (160GB, MLC) with an XceedIOPS SSD (200GB, E-MLC). Table 3 below shows the parameters of the SSDs that are being used in the calculations.

Table 3: Client SSD vs Enterprise SSD parameters

	Client SSD	XceedIOPS
Flash Technology	MLC	E-MLC
Capacity	160GB	200GB
User Capacity (per JEDEC JESD218)	1.6×10^{11} Bytes	2.0×10^{11} Bytes
P/E cycles	5,000	30,000
Write Amplification	1.1 (data pattern independent)	0.75 to 4.0 (data pattern dependent)
Warranty	3 years	5 years

Table 4 below shows the endurance calculations for the various workloads.

Table 4: Lifetime calculation for various workloads

	160GB Client SSD MLC NAND flash 3-year warranted life	XceedIOPS 200GB SSD E-MLC NAND flash 5-year warranted life
Database Write Workload	**WA = 1.1**	**WA = 0.75**
TBW over lifetime	363	4,000
Allowed TBW per day to warranted lifetime	0.33	2.19
SSD Lifetime at 2:1 read/write workload, 100% duty cycle	0.60 years	9.1 years
Normal Random Write Workload	**WA = 1.1**	**WA = 1.0**
TBW over lifetime	363	3,000
Allowed TBW per day to warranted lifetime	0.33	1.64
SSD Lifetime at 2:1 read/write workload, 100% duty cycle	0.60 years	6.8 years
Video Server Workload	**WA = 1.1**	**WA = 4.0**
TBW over lifetime	363	750
Allowed TBW per day to warranted lifetime	0.33	0.41
SSD Lifetime at 2:1 read/write workload, 100% duty cycle	0.60 years	1.6 years

Assumptions: Random 4K-aligned IOs.

Table 4 reveals the following:

- XceedIOPS SSDs outperform the 160GB MLC Client SSD in SSD lifetime across all workloads;
- XceedIOPS SSDs leverages E-MLC flash to exceed 5-year warranted life in 24/7 duty across most workloads (For worst-case high-WA duty, XceedIOPS SSD monitors workload and work done to ensure warranted life).
- 160GB MLC Client SSD cannot sustain 24/7 duty under any workload, even for its lesser 3-year warranty period.

3.4 Ensuring Uncorrectable Bit Error Rate (UBER)

E-MLC flash exhibits 6x to 20x the endurance of MLC because it is selected and trimmed to incur much lower damage each P-E cycle. Since damage to a flash cell means risk or damage to the data that cell holds, it follows that E-MLC flash provides higher data integrity. This in turn means less data correction, fewer uncorrectable errors, and fewer scenarios requiring high-latency second-tier data rebuilds (such as the RAISE™ all-page rebuild capability provided by the XceedIOPS SSD). The higher overall flash memory reliability of E-MLC yields three benefits related to bit errors: superior data integrity in general, fewer extreme error recovery scenarios, and fewer unrecoverable data scenarios.

SMART's XceedIOPS SSD is specified for an Uncorrectable Bit Error Rate of 10E-18. This is two orders of magnitude better than the JEDEC specification of 10E-16 for Enterprise SSD.

4 Summary

Client SSD that use MLC flash memory are targeted to consumer and light-workload client applications. SMART's XceedIOPS SSD achieves true enterprise-level reliability and true sustainable enterprise-class performance and endurance through the use of E-MLC NAND flash. It derives the following benefits from its use of E-MLC flash, as compared to Client SSDs:

- Superior data integrity at all stages of SSD life, due to higher overall flash memory reliability, verified via the rigorous E-MLC manufacturing test process flow;
- Minimized high-latency error correction events and unrecoverable data scenarios, due to higher overall flash memory reliability;
- Longer SSD Endurance: 6x to 20x longer SSD life for any given application or work performed.

Disclaimer:

No part of this document may be copied or reproduced in any form or by any means, or transferred to any third party, without the prior written consent of an authorized representative of SMART Storage Systems. ("SMART"). The information in this document is subject to change without notice. SMART assumes no responsibility for any errors or omissions that may appear in this document, and disclaims responsibility for any consequences resulting from the use of the information set forth herein. SMART makes no commitments to update or to keep current information contained in this document. The products listed in this document are not suitable for use in applications such as, but not limited to, aircraft control systems, aerospace equipment, submarine cables, nuclear reactor control systems and life support systems. Moreover, SMART does not recommend or approve the use of any of its products in life support devices or systems or in any application where failure could result in injury or death. If a customer wishes to use SMART products in applications not intended by SMART, said customer must contact an authorized SMART representative to determine SMART's willingness to support a given application. The information set forth in this document does not convey any license under the copyrights, patent rights, trademarks or other intellectual property rights claimed and owned by SMART. The information set forth in this document is considered to be "Proprietary" and "Confidential" property owned by SMART.

ALL PRODUCTS SOLD BY SMART ARE COVERED BY THE PROVISIONS APPEARING IN SMART'S TERMS AND CONDITIONS OF SALE ONLY, INCLUDING THE LIMITATIONS OF LIABILITY, WARRANTY AND INFRINGEMENT PROVISIONS. SMART MAKES NO WARRANTIES OF ANY KIND, EXPRESS, STATUTORY, IMPLIED OR OTHERWISE, REGARDING INFORMATION SET FORTH HEREIN OR REGARDING THE FREEDOM OF THE DESCRIBED PRODUCTS FROM INTELLECTUAL PROPERTY INFRINGEMENT, AND EXPRESSLY DISCLAIMS ANY SUCH WARRANTIES INCLUDING WITHOUT LIMITATION ANY EXPRESS, STATUTORY OR IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

©2012 SMART Storage Systems. All rights reserved.

Corporate Headquarters: 39870 Eureka Dr., Newark, CA 94560, USA ♦ Tel:(510) 623-1231 ♦ Fax:(510) 623-1434 ♦ E-mail: info@smartstoragesys.com
Flash Design Center: 2 Robbins Road, Westford, MA 01886, USA ♦ Tel:(978) 303-8500 ♦ Fax:(978) 303-8757
Flash Design Center: 2600 W. Geronimo, Chandler, AZ 85244, USA ♦ Tel:(480) 792-8900 ♦ Fax:(480) 792-8901
Asia: Plot 18, Lrg Jelawat 4, Kawasan Perinudstrian Seberang Jaya 13700, Prai, Penang, Malaysia ♦ Tel:+604-3992909 ♦ Fax:+604-3992903