

# Enabling MLC NAND Flash for Cost-Effective, High-Capacity Data Storage

*Douglas Wong, Toshiba America Electronic Components, Inc., 9775 Toledo Way, Irvine, CA 92618, +001-949-455-2000, [Doug.Wong@taec.toshiba.com](mailto:Doug.Wong@taec.toshiba.com)*

*Raz Dan, M-Systems, 8371 Central Ave., Suite A, Newark CA 94560, +81 90-2768-9603, [raz.dan@m-sys.com](mailto:raz.dan@m-sys.com)  
Rochelle Singer, M-Systems, 7 AtirYeda, Kfar Saba 44425 Israel, +972-9-764-5131, [rochelle.singer@m-sys.com](mailto:rochelle.singer@m-sys.com)*

## **ABSTRACT**

The more functionality that OEMs offer users to support applications such as moving maps, tactical computers and airborne systems, the more memory capacity that is required to store additional applications and user data – reliably, cost-effectively, and space-effectively. To meet these demands, OEMs are encouraging flash memory manufacturers to implement advanced technologies that can provide the performance and reliability of a big system in a low-cost, small-as-possible package.

Among the technologies that flash memory manufacturers are investigating is Multi-Level Cell (MLC) technology. MLC technology offers double the bit storage capacity per cell, reducing the total die size. This has a major impact on the silicon cost.

On the surface, MLC looks like the perfect solution. But without highly reliable and efficient algorithms to guarantee data reliability and some mechanism to boost MLC performance, much lower than standard NAND-based performance, MLC technology can meet neither reliability nor performance requirements.

Working closely with Toshiba, M-Systems developed and patented *x2 technology* to enable MLC as a local memory storage solution, without compromising reliability and achieving performance rates that rival and even exceed standard NAND-based flash technology rates. This paper explains how M-Systems' *x2 technology* achieves this.

## **INTRODUCTION**

Multi Level Cell (MLC) technology greatly reduces flash die size to achieve a breakthrough cost structure. It does this by storing 2 bits of data per physical cell instead of the traditional 1 bit per cell, using Binary Level Cell (BLC) technology. But the increased density of the MLC flash media has grave consequences in terms of accuracy and performance. A number of flash vendors, with varying degrees of success, have made attempts to implement MLC technology on selected flash platforms while overcoming its limitations.

Recently, Toshiba introduced MLC NAND technology. Although NAND constitutes a particularly good building block for MLC technology due to its high erase and write performance and high density (small size), MLC NAND is practically unusable for local data and code storage due to the degradation in accuracy and performance. *x2 technology*, developed by M-Systems at Toshiba's request, combines advanced hardware and customized software algorithms to boost reliability and performance levels to rival those of 1 bit per cell devices.

This paper discusses the MLC solutions available on the market today, their benefits and limitations, and the hardware and software innovations of *x2 technology* that overcome MLC NAND limitations, enabling it to function as the most highly cost-effective, fast and reliable local storage solution.

## COMPARING BLC AND MLC FLASH TECHNOLOGIES

### Basic Flash Technology

A basic flash memory cell consists of a transistor with a floating gate that is used to store electrons (see Figure 1). The number of electrons stored in the floating gate of every cell determines the identity of the data stored in it (either '1' or '0'). Writing '1', equivalent to erasing a cell, is done by charging the floating gate with electrons. Writing '0', or programming a cell, is done by discharging electrons from the floating gate.

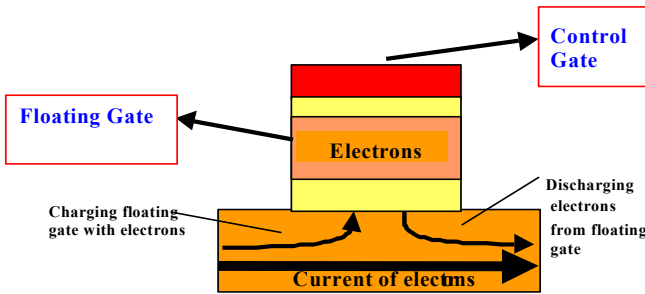


Figure 1: A Basic Flash Cell (Transistor with a Floating Gate)

Sensing the voltage level of the cell determines if it stores a '1' or a '0'. Above a pre-defined voltage ( $V_t$ ), the cell is recognized as '0'. Below this pre-defined value, it is recognized as '1'.

### BLC Technology

Figure 2 shows the voltage distribution and the  $V_t$  that distinguishes between programmed and erased states in BLC flash technology. Note that there are two voltage states per cell.

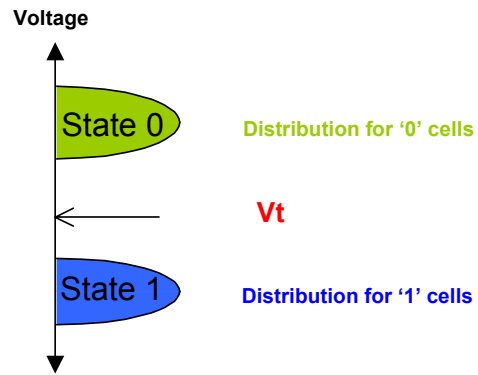


Figure 2: Bit Distribution in BLC Flash Technology

### MLC Technology

The basic method used in MLC technology is identical to BLC technology, except that it enables double the number of bits to be stored in a single physical cell by charging the floating gate of a transistor to four different voltage levels, instead of two. Instead of using one  $V_t$  to distinguish between two voltage states (high and low charge), MLC technology uses three  $V_t$  levels to distinguish between four states. Figure 3 below illustrates  $V_t$  level placement for two bits per cell MLC technology.

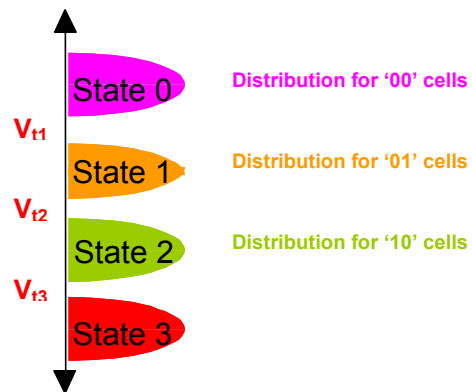


Figure 3: Bit Distribution in MLC Technology

## **MLC BENEFITS AND LIMITATIONS**

MLC high-density design innovations reduce the silicon die size, which is the major element contributing to overall device cost. For MLC NAND, this reduction in size and cost is greatest in capacities of 32MByte and higher, where the die can be as small as 50% of the size required to provide the same capacity BLC flash device. The savings must be measured both in dollars and space, particularly for the cell phone market where every millimeter of board real estate can have an impact on the size of the end-user product and, ultimately, on market success.

But these very same high-density design innovations introduce three, major areas of design limitations as compared with BLC flash:

- Data Reliability
- Performance
- Flash management

This section discusses these areas in order to lay the groundwork for understanding how x2 technology overcomes the associated problems.

### **Data Reliability**

As shown in Figure 2 and Figure 3, a BLC flash cell must distinguish between 2 voltage levels, whereas an MLC flash cell must distinguish between 4. Since both BLC and MLC-based devices use the same size voltage window, the distance between adjacent voltage levels in MLC is much smaller than in BLC. This reduced distance has an impact on data reliability. Detecting the voltage levels in an MLC flash cell is a more precise and complex task than in a BLC cell, subject to a higher probability of error that can affect data reliability in both the short and long term.

The probability of all types of errors in BLC flash is on the order of  $10^{-10}$ . The overall

probability of MLC flash errors, mainly caused by the program disturb effect that is described below, is two orders of magnitude worse than BLC technology.

### **Long-Term Data Errors**

Flash memory cells must provide long-term data content capabilities to function reliably as a non-volatile memory device. In order to do this, the long-term stability of their voltage levels is critical. Drifting voltage effects, which tend to slowly change the cell's voltage level from its initial level to a different level after cell programming or erasing, may incorrectly be interpreted as a different logical value. Due to the smaller distance between MLC levels than BLC levels, MLC flash cells are more likely to be affected by drifting voltage effects and, consequently, more potentially prone to errors.

Initial measurements indicate that the impact of MLC on data retention is not as severe as had been feared. However, there is still cause for concern in this area due to the architectural density associated with MLC technology.

### **Program Disturb Errors**

The program disturb effect, also called over program effect, causes a programming operation on one page to induce a change in bit value on another, unrelated page. In BLC flash technology based on a 0.16 $\mu$  manufacturing process, the typical program disturb error rate is on the order of 1 bit error per  $10^{10}$  bits programmed. This compares with an error rate on the order of 1 bit error per  $10^8$  bits programmed with MLC flash technology.

### **Read Disturb Errors**

The read disturb effect causes a page read operation to induce a permanent, bit value change in one of the read bits. In BLC flash technology based on a 0.16 $\mu$  manufacturing

process, the typical read disturb error rate is on the order of 1 bit error per  $10^6$  repetitive reads of the page containing the bit.

Although MLC cells are more prone to such errors, the effect in actual measurements is less severe than in program disturb errors. The measured rate is on the order of 1 bit error per approximately  $10^5$  repetitive reads of the page.

### **Performance**

MLC technology requires more time than BLC technology for completing the basic flash operations of reading a page into the flash buffer, writing a flash buffer into a page, and erasing a flash unit. Especially for write operations, raw flash comparisons indicate that MLC performance is only 25% that of BLC. But many factors other than raw flash speed influence performance, including: host CPU bus timing issues, error detection and correction, software algorithms employed by the device driver, file system overhead, patterns of file access by the user, bus cycles and more.

In fact, from the user point of view, raw read or write times are totally irrelevant. What the user “feels” is how long it takes from when, for example, a long sequence of write commands is issued to the file system, until the requests are completed. To get a “true” measure of these times, the measurements should be performed under scenarios that duplicate the real world as closely as possible. This implies first filling the disk to almost full capacity, and then performing the measurements, taking into account the hidden mechanisms of the software interfacing the flash to the user (file system, device driver, etc.).

### **Sustained Read**

When comparing sustained read performance values in real-world scenarios for BLC with MLC, the gap lessens considerably: MLC

performance is 98% of BLC performance. Operations that both BLC and MLC require to support a sustained read operation – such as running the driver code and the file system code, and accumulating bus cycles to support address, command, error correction code and control information – account for closing the gap.

### **Sustained Write**

A comparison of sustained write performance for both technologies in real-world scenarios must take into account an additional factor: making room for new data when no free space is available. This means adding to the calculation the time it takes to erase a flash unit and, depending on the time it takes to manage the flash (using M-Systems’ TrueFFS, for instance, adds 5% of the time required to write a unit), this time as well. For BLC, these calculations result in a sustained write performance rate of 250KBytes per second on a low MIPS platform, or 4  $\mu$ sec per byte for a typical mix of files, as compared with 172KBytes per second for MLC. (Note that the number of sectors per unit for MLC is twice the corresponding number for BLC.) When these figures are translated into percentages, MLC sustained write performance is approximately 69% of BLC write performance.

Write performance greatly varies according to the user’s access patterns, mainly the average file size. For large files the rate is much higher (up to approximately 600 KBytes per second); for very small files it is much lower. Here, unlike in read operations, the time that is required for file system handling is more significant than device driver time, especially when dealing with small files. Bus cycle time for writing is practically the same as for reading. All the remaining time is spent on software overhead.

## Flash Management

Because of MLC’s unique architecture, pages can only be written sequentially, whereas in BLC they can be written randomly within the erase block. MLC also makes partial page programming impossible, as opposed to BLC technology that enables it. This means that the existing translation layers used by TrueFFS to support BLC devices, NFTL and INFTL, are unusable, since they rely on random page access. Sequential write only and the lack of partial page programming impose limitations on MLC that affect reliability as well as performance.

## OVERCOMING MLC LIMITATIONS

Because MLC technology can potentially bring the industry breakthrough cost and size benefits for local data and code storage, M-Systems chose to take on the challenge of perfecting it by providing solutions to overcome MLC reliability, performance and flash management limitations.

x2 technology, customized by M-Systems specifically to meet this challenge, is a combination of algorithms, performance-enhancing innovations and flash management capabilities. Developed in cooperation with Toshiba, x2 technology is integrated seamlessly into the different modules of M-Systems’ standard DiskOnChip® architecture and fully compatible with its TrueFFS® technology for flash management. It includes reliability and performance improvements integrated into TrueFFS, the thin controller and the flash media itself, as shown in Figure 4. x2 technology cleverly balances software and hardware to keep reliability and performance at their peak while maintaining MLC cost and size benefits.

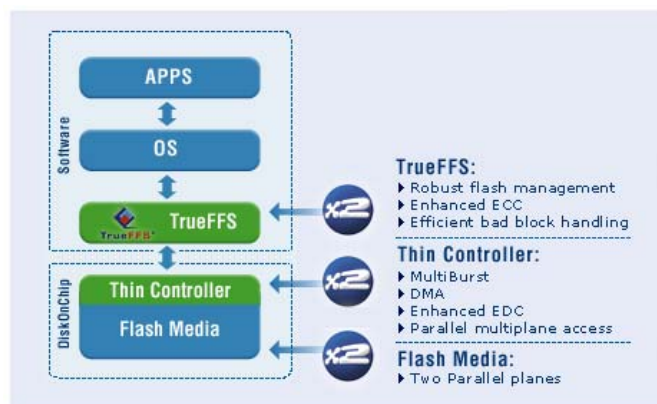


Figure 4: x2 Technology Seamless Integration into M-Systems’ MLC DiskOnChip

Table 1 maps the various features of x2 technology against the three major areas of MLC limitations that they overcome. The remainder of this section explains how each feature achieves these enhancements.

Table 1: Overcoming MLC Limitations with x2 Technology

x2 Technology Feature	Areas of MLC Enhancement		
	Reliability	Performance	Flash Management
<b>TrueFFS</b>			
Robust flash management	✓ □	✓	✓
Enhanced EDC	✓	✓	✓
Enhanced ECC	✓	✓	✓
Efficient bad block handling	✓		✓
<b>Thin Controller</b>			
MultiBurst		✓	
DMA		✓	
Parallel multiplane access		✓	
<b>Flash Media</b>			
Two parallel planes		✓	✓

## **Robust Flash Management**

To overcome MLC flash access and partial programming limitations that affect all three areas of MLC limitations, x2 technology uses a specially customized translation layer called Sequential Access Flash Translation Layer (SAFTL). SAFTL is incorporated seamlessly into M-Systems' TrueFFS. It maps each virtual unit into a chain of physical units, much in the same way that translation layers for BLC operate. However, unlike traditional translation layers, SAFTL does not implement one-to-one simple mapping between the virtual sector offset in the virtual unit and its physical location in the physical units. Instead, the data of a virtual sector can be in any location within the physical unit chain of its virtual unit. Each physical sector containing data also contains the offset of its corresponding virtual sector in its virtual unit.

SAFTL enables each physical unit to be filled sequentially, as required by MLC flash, starting from the first sector to the last. Each write request to the corresponding virtual unit is written to the next free physical sector, regardless of the virtual sector number requested to be written. When a physical unit is full and a new write request arrives, a new free physical unit is allocated and added to the chain. New unit allocation always occurs concurrently with writing a sector, so that sector data and unit control data can be written in one operation to improve performance.

## **Enhanced EDC and ECC**

The Error Detection Code (EDC) and Error Correction Code (ECC) developed for x2 technology is based on M-Systems' highly effective combination used in DiskOnChip. This system contains hardware-embedded EDC mechanism to detect errors on-the-fly and software-embedded ECC mechanism to reduce silicon size and cost. The combination of hardware and software results in the most

cost-effective data reliability in the industry by detecting and correcting 2 errors per page without imposing performance penalties.

The EDC and ECC enhancements for MLC are capable of correcting up to 4 errors per page, using two industry-standard error codes: an extended Hamming code and a BCH (Bose, Chaudhuri and Hocquenghem) code.

The Hamming code can detect 2 errors per page and correct 1 error per page. The BCH code can detect 4 errors per page and correct an equal number. This combination of codes provides an even higher rate of coverage than 2 bits per page provide for BLC technology, as shown in Table 2 and Table 3.

It also ensures that the minimal amount of code required is used for detection and correction to deliver the required reliability without degrading performance.

*Table 2: BLC Bit Error Probability*

<b>EDC/ECC Capability</b>	<b>Error Probability</b>
None	$10^{-7}$
1 bit error/page	$10^{-14}$
2 bit errors/page	$10^{-20}$

*Table 3: MLC Bit Error Probability*

<b>EDC/ECC Capability</b>	<b>Error Probability</b>
None	$10^{-5}$
1 bit error/page	$10^{-10}$
2 bit errors/page	$10^{-14}$
3 bit errors/page	$10^{-19}$
4 bit errors/page	$10^{-24}$

## **Efficient Bad Block Handling**

x2 technology handles bad blocks, which can be randomly present in flash media, by enabling unaligned block access to two planes. Bad blocks are mapped individually on each plane, as shown in Figure 5. Good units can

therefore be aligned or unaligned, thereby minimizing the effects of bad blocks on the media. Without this capability, a bad block in one plane would cause a good block in the second plane to be tagged as a bad block, making it unusable. This customized method of bad block handling for two planes enhances data reliability without adversely affecting performance.

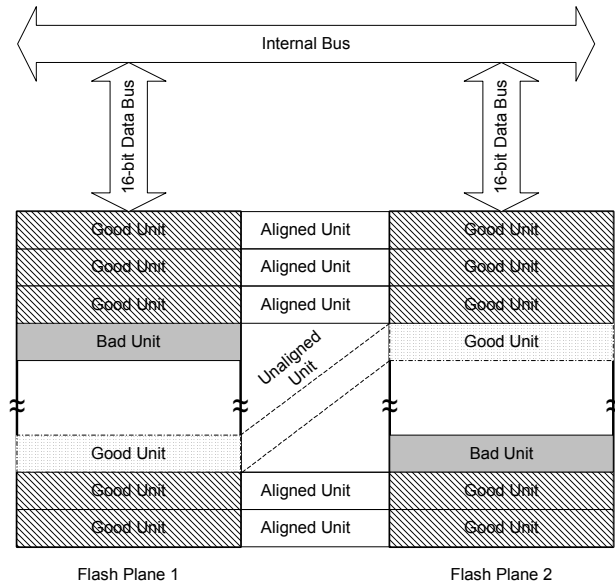


Figure 5: Unaligned Multiplane Bad Block Access

**MultiBurst**

To improve MLC read performance rates, x2 technology incorporates a feature called MultiBurst. MultiBurst enables parallel read access from two 16-bit planes to the flash controller, thereby achieving the desired output data rate for the host. The host accesses the first word of a page with a relatively slow access time, but each subsequent word with a very fast access time. Two cycles of 16 bits each are sent to the host at a clock rate set by the host rather than limited by flash operation, as shown in Figure 6.

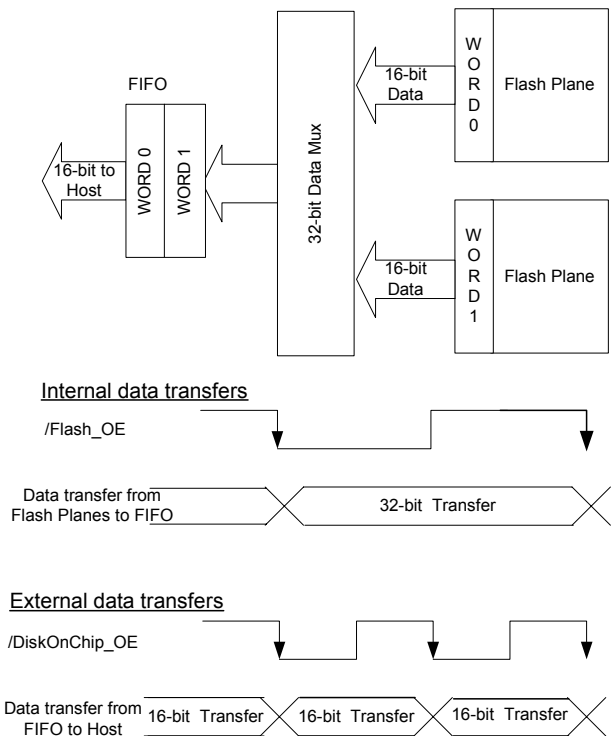


Figure 6: MultiBurst Operation

**DMA**

By enabling DMA operation, x2 technology reduces the CPU overhead. This is a particularly useful feature for transferring large files in support of Real-Time Operating Systems (RTOS). In addition, it can be used to enhance overall system performance by reducing boot time. In this case, the DMA mechanism is used to quickly transfer large blocks of code from the DiskOnChip into shadow RAM.

When comparing future generations of DiskOnChip products that will use MLC and x2 technology to raw flash products, such as Intel StrataFlash or AMD MirrorBit, this capability has at least a threefold benefit: increased performance, easier integration, and reduced external part count by allowing direct connection to a DMA controller without additional hardware.

### Parallel Multiplane Access

As discussed earlier, the MLC flash media is built of two planes that can operate in parallel. This architecture is one of the most powerful, x2 technology innovations, doubling read, write and erase performance. Two pages on different planes can be concurrently read or written if they have the same offset within their respective blocks, even if the blocks are unaligned.

### POWER CONSUMPTION

M-Systems' DiskOnChip was designed for mobile systems that require very low power consumption. Therefore, the design incorporates power management features, such as Deep Power-Down mode, which consumes only 10  $\mu$ A.

Since the design is completely static (requiring no free-running clocks), it automatically goes into standby mode when not accessed. In addition, TrueFFS places DiskOnChip in Deep

Power-Down mode at the end of every sector transfer. This design provides for a quick transition from Deep Power-Down mode to operational mode with minimal latency to minimize performance penalties.

### SUMMARY

The major improvements in flash NAND devices brought about by MLC technology are: much smaller size per bit, and consequently, much lower price per bit. These advantages come with added complexity in both device hardware architecture and device driver software. However, this document shows that x2 technology, by cleverly customizing the thin controller, TrueFFS and the flash media, can provide a flash disk storage device based on MLC technology that is as reliable and as fast as BLC devices in common use today.

*Table 4: Comparing NAND Flash Alternatives*

	<b>BLC NAND</b>	<b>MLC NAND</b>	<b>MLC NAND with x2 Technology</b>
Size	100%	~50% of BLC	~53% of BLC
Sustained write performance	~800KB/sec	~350KB/sec	~700KB/sec
Reliability	Occasional random errors	Frequent random errors	Perfect device

## **BIOGRAPHIES**

Douglas Wong is a member of the technical staff for memory products at Toshiba America Electronic Components, Inc. (TAEC). His responsibilities include system engineering and new product definition. Prior to joining Toshiba in 1993, Wong served as an electronics engineer for electronic warfare at VSE Corporation and was responsible for testing electronic counter measures (ECM) systems in military aircraft. Wong is a mentor for the Future Scientists and Engineers of America. He has presented papers on NAND flash and small form factor memory cards at various conferences including Portable by Design, WESCON IC Expo, Portable Computer Components and Convergence Tech/IC Expo. He has also authored articles for *ChipCenter*, *Communication Systems Design* and *Portable Design*. Wong earned a B.Sc. in Electrical Engineering from California Polytechnic State University, San Luis Obispo, and an M.Sc. in Electrical Engineering (semiconductor physics concentration) from the University of California at Los Angeles.

Raz Dan is the VP of Customer Support for M-Systems Inc., Newark CA, a company that has pioneered storage solutions based on flash memory and TrueFFS technology for over thirteen years. He is currently responsible for promoting the company's hardware and software technology and helping customers integrate these products into new designs on a daily basis. Raz has been involved in the design and implementation of flash disks for ten years. Previously, he worked as both a software and hardware engineer on various types of embedded systems. Raz received his B.Sc. in Electrical Engineering in 1990 from Tel Aviv University, Tel Aviv, Israel.

Rochelle Singer is the Technical & Marketing Documentation Manager of M-Systems. She has been documenting high-tech hardware and software products for the past 20 years for various industries, including: semiconductor, military and aerospace, medical instrumentation and research, and printing. She also occasionally freelances, and has contributed articles to a number of journals, including *Link* and the *Journal of the Society for Technical Communication*. Rochelle received her B.A. in English from Barnard College (Columbia University) and her M.F.A. from Brooklyn College in New York.