Flash Memory Technology Direction

Jim Cooke - Director, Applications Engineering
Micron Technology, Inc.

April 30, 2007

Abstract

A number of key new computing system initiatives—including Windows Vista™ ReadyBoost™ and Windows ReadyDrive™ technologies—exploit the performance, economy, power efficiency, and high reliability of NAND Flash memory. System designers implementing NAND Flash-based traditional hard disk drives (HDDs) or newer hybrid hard drives (HHDs) and solid state disks (SSDs) face a number of decisions. This paper explains the trade-offs associated with available disk caching methods, the differences between various types of Flash memory, and the advantages that NAND offers when superior performance is critically important.

This information applies for the following operating systems:
Windows Vista™

Contents

Author's Disclaimer and Copyright: Products and specifications discussed herein are subject to change by Micron without notice. Products are warranted only to meet Micron's production data sheet specifications. All information discussed herein is provided "AS IS" and without warranties of any kind. Micron, the M logo, and the Micron logo are trademarks of Micron Technology, Inc. All other trademarks are the property of their respective owners. © 2007 Micron Technology, Inc. All rights reserved. Rev. 04/07

Windows Hardware Engineering Conference - WinHEC Sponsors' Disclaimer: The contents of this document have not been authored or confirmed by Microsoft or the WinHEC conference co-sponsors (hereinafter "WinHEC Sponsors"). Accordingly, the information contained in this document does not necessarily represent the views of the WinHEC Sponsors and the WinHEC Sponsors cannot make any representation concerning its accuracy. THE WinHEC SPONSORS MAKE NO WARRANTIES, EXPRESS OR IMPLIED, WITH RESPECT TO THIS INFORMATION.

Microsoft, Windows, Windows NT, Windows Server, and Windows Vista are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. The names of actual companies and products mentioned herein may be the trademarks of their respective owners.

Flash Memory Technology Direction

A number of key new computing system initiatives—including Windows Vista™ ReadyBoost™ and Windows ReadyDrive™ technologies—exploit the performance, economy, power efficiency, and high reliability of NAND Flash memory. NAND Flash memory is enabling new types of storage products that are driving an improved memory-caching hierarchy. This hierarchy essentially re-orders how and where storage tasks are performed in the system, fills a critical gap between main memory and hard disk drives, and optimizes performance and reliability by partitioning storage tasks across traditional hard disk drives (HDDs) and newer hybrid hard drives (HHDs) and solid state disks (SSDs).

System designers implementing these new NAND Flash-based storage technologies face a number of decisions. They must understand the trade-offs associated with available disk caching methods, the differences between various types of Flash memory, and the advantages that NAND offers when superior performance is critically important.

More Storage Choices Than Ever

Today's three most widely used primary storage technologies include HDDs, HHDs, and SSDs.

HDDs use ultra-sophisticated magnetic recording and playback technologies and are the primary data storage component in notebooks, desktops, servers, and dedicated storage systems. They offer significant advantages in terms of capacity and cost per bit, but they also require considerable time to boot the operating system and read applications.

HDD technology has not kept pace with processor performance because of inherent challenges related to mechanical head and rotational latencies. For instance, the rotational latency—the amount of time it takes for platters to rotate 360 degrees—of a high-end, 7,200-rotations-per-minute (RPM) drive is approximately 8 milliseconds. Add to that the seek latency—the amount of time it takes for the heads to position themselves properly over the cylinder—and the result is considerable initial latency that inhibits performance. This initial latency might not be a concern when amortized over large files, but it can become significant when accessing small files, such as directory structures or file allocation tables (FAT).

These HDD shortcomings can, in many cases, be solved with the addition of NAND Flash technology. NAND Flash memory offers advantages for PC storage, including faster boot times, decreased heat and noise, and improved power consumption, performance, and reliability. There are three major areas where NAND Flash memory is playing an increasingly important role in overcoming the limitations of traditional HDD technology.

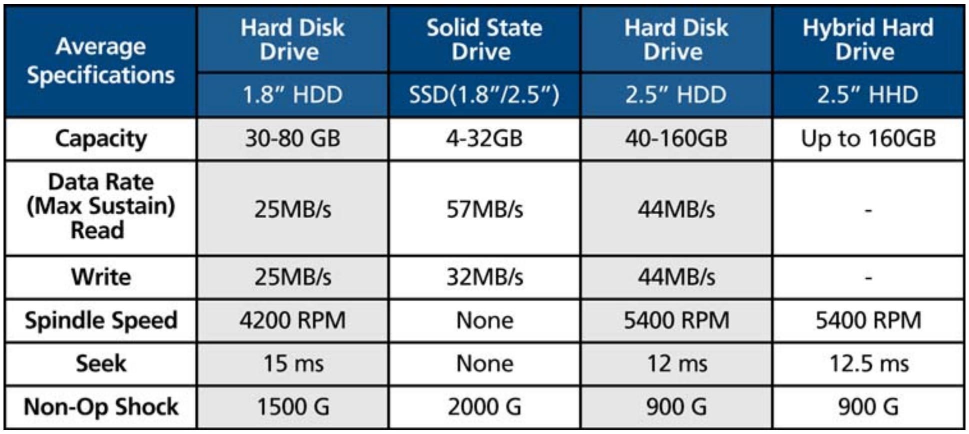
The first area is HHDs, a new type of hard drive that behaves for the most part like a traditional hard disk drive with the addition of a nonvolatile NAND cache, which addresses some of the traditional latencies of the standard mechanical HDD. HHDs differ from standard hard drives in that they employ a large buffer of nonvolatile Flash memory with which to cache data during normal use. This enables the HDD's platters to remain idle much of the time, which decreases power consumption and improves reliability. In addition, this cache can speed the boot process since the system does not have to wait for the platters to spin up. HHDs represent an incremental upgrade to HDDs, building on the basic HDD structure but adding a nonvolatile cache that enables faster read/write capability even when the spindle has stopped. Even though the mechanical rotational and seek latencies associated with traditional HDDs still exist, the ability to cache data provides attractive performance benefits.

HHD-like performance can also be achieved through a second implementation of NAND Flash technology interfacing to the PC chipset itself. Intel refers to this implementation as Robson. Robson is the code name for a new Intel® platform technology that uses nonvolatile Flash memory to increase system responsiveness, speed multi-tasking, and extend battery life. Robson will roll out in the first half of 2007 with the Santa Rosa notebook chipset platform and will be implemented with Microsoft's Vista operating system.

Initially, it is expected that the standard Santa Rosa Robson chipset configuration will include 512MB of NAND, but will offer the option of larger capacities. The NAND Flash will interface with the Southbridge portion of Intel's core-logic chipset, also known as the integrated controller hub (ICH). The NAND Flash can either be soldered on to the motherboard or offered as part of a mezzanine card. Robson will enable the operating system to use NAND Flash memory to cache various data and make it more accessible, thereby offering many of the advantages of the hybrid hard drive with conventional SATA HDD hardware rather than the more specialized HHD.

The third area where NAND plays an increasingly important role in PC storage is in SSDs. Here, the entire "drive" is comprised of solid state NAND Flash memory. Although generally referred to as a "disk," SSDs have no moving parts, they consume less power than traditional disks (translating into less heat dissipation, better reliability, and improved battery life), and they are significantly faster and more rugged. SSDs are implemented using a printed circuit board with NAND Flash memory and a controller, such as the newer serial ATA (SATA) controller or its parallel ATA (PATA) predecessor. Based on recent advances in NAND lithography, SSD densities have reached the necessary capacities for mass-market appeal, and they offer many features that have led to improved user experience.

Figure 1 shows the comparative advantages of SSDs, including power efficiency, reliability, and performance.



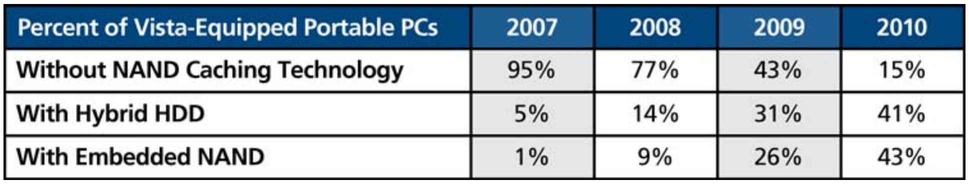
Average Specifications	Hard Disk Drive	Solid State Drive	Hard Disk Drive	Hybrid Hard Drive
	1.8" HDD	SSD(1.8"/2.5")	2.5" HDD	2.5" HHD
Capacity	30-80 GB	4-32GB	40-160GB	Up to 160GB
Data Rate (Max Sustain) Read	25MB/s	57MB/s	44MB/s	-
Write	25MB/s	32MB/s	44MB/s	-
Spindle Speed	4200 RPM	None	5400 RPM	5400 RPM
Seek	15 ms	None	12 ms	12.5 ms
Non-Op Shock	1500 G	2000 G	900 G	900 G

Figure 1

SSDs continue to rise in popularity. Web-feet research predicts shipments to increase at a 146% CAGR from 2005 to 2010. Data processing applications are expected to drive a majority of SSD unit shipments. And while the majority of today's portable computers feature an HDD, it is expected that by 2011 one-third will feature an SSD.

Finally, NAND Flash can also be used to boost performance in other storage implementations, including add-on USB flash disks (UFDs), add-on ExpressCards, and add-on SD/MMC cards—or any other media through the use of Microsoft's Windows ReadyBoost technology. ReadyBoost technology enables users to increase the speed of frequently used applications by using system memory devices (Robson, USB drives, SD cards, CompactFlash, etc.) as a write-through virtual memory OS page cache. Users can determine how much of the Flash memory is used as a performance cache.

The adoption of NAND cache is growing, with IDC estimating that more than 85 percent of Vista-equipped PCs will have some sort of NAND-based cache by 2010. (See Figure 2.) HDDs will remain attractive thanks to their low cost per byte, but NAND-based solutions will narrow the gap. The cost delta between NAND-based technologies and sub-2.5" HDDs will shrink considerably, from a $40 cost/GB delta in 2005 to just over $1 cost/GB in 2011 (IDC, 2007). Hybrid NAND-based drives will provide an incremental upgrade to HDDs, while SSDs will become increasingly popular given their significant advantages.



Percent of Vista-Equipped Portable PCs	2007	2008	2009	2010
Without NAND Caching Technology	95%	77%	43%	15%
With Hybrid HDD	5%	14%	31%	41%
With Embedded NAND	1%	9%	26%	43%

Figure 2

A Closer Look at NAND Flash Technology

There are two primary types of NAND Flash technology. Historically, the majority of the market was served by single-level cell (SLC) NAND. Beginning in 2006, the market migrated to multi-level cell (MLC) NAND, and by the end of the year there was a nearly even distribution in shipments between the two. During 2007 and beyond, it is expected that the majority of shipments will be MLC NAND. However, SLC will retain much of the market for high-performance, high-reliability applications.

There are several key differences between the two technologies. MLC NAND tends to lead in low-cost consumer applications, including media players, MP3 devices, media cards, and USB flash drives. Meanwhile, professional products and SSDs will continue to demand the higher performance and reliability of SLC NAND Flash memory.

Figures 3A and 3B illustrate the comparative differences between MLC and SLC NAND Flash cells, respectively. SLC NAND stores two binary states (either a binary 1 or a binary 0) in a single cell, whereas MLC NAND can store four states: 00, 01, 10, and 11. What might not be obvious to the casual observer is that there is considerably more design margin with SLC, which leads to greater robustness, reliability, and endurance compared to MLC.

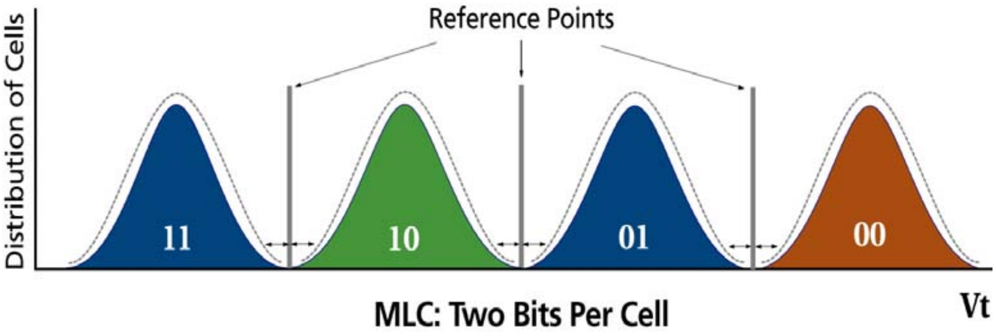


Figure 3A

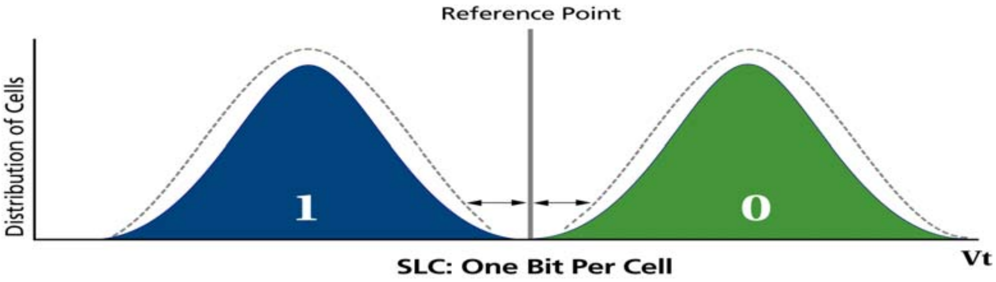
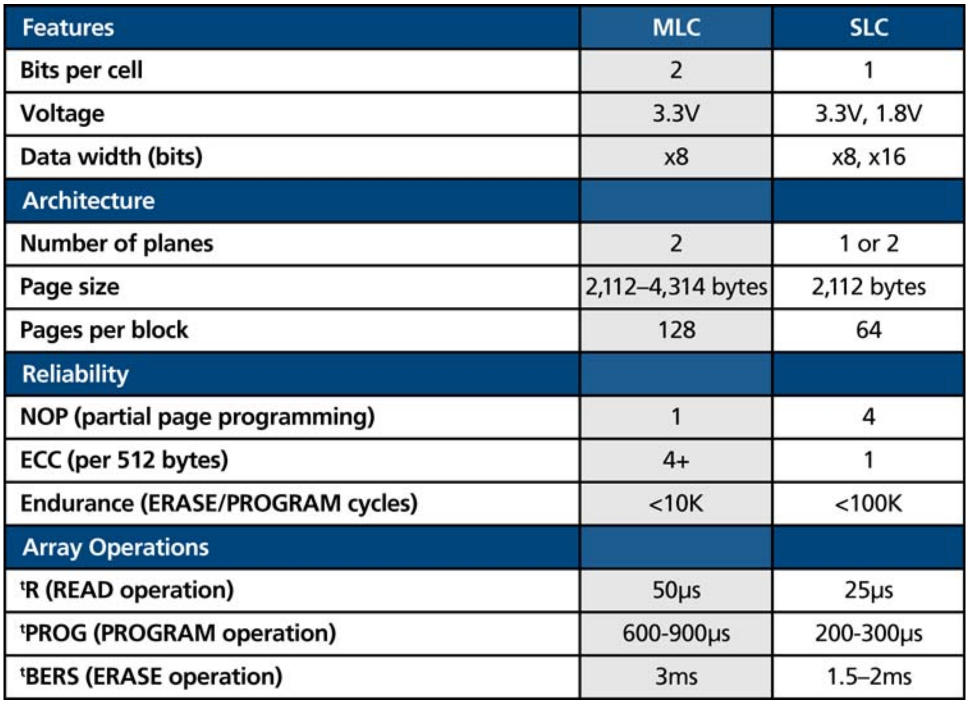


Figure 3B

Figure 4 shows the major differences between MLC and SLC devices, which are density, reliability, and performance.



Features	MLC	SLC
Bits per cell	2	1
Voltage	3.3V	3.3V, 1.8V
Data width (bits)	x8	x8, x16
Architecture		
Number of planes	2	1 or 2
Page size	2,112–4,314 bytes	2,112 bytes
Pages per block	128	64
Reliability		
NOP (partial page programming)	1	4
ECC (per 512 bytes)	4+	1
Endurance (ERASE/PROGRAM cycles)	<10K	<100K
Array Operations		
^{t}R (READ operation)	50µs	25µs
^{t}PROG (PROGRAM operation)	600-900µs	200-300µs
^{t}BERS (ERASE operation)	3ms	1.5–2ms

Figure 4

Figure 5 shows a typical 2Gb, 2K-page SLC NAND architecture. The device is comprised of 2,048 independent blocks. A block is the smallest erasable entity. Each block contains 64 pages that consist of 2,112 bytes—2,048 bytes of data and 64 bytes of spare area for ECC and other software overhead. A page is the smallest programmable unit. The device includes an input/output shift register, known as the cache register, which is used for double buffering. Data is shifted into and out of the cache register byte-by-byte. When a READ operation is requested, the array is accessed and the data is loaded into the cache register then shifted out. For a programming operation, data is shifted into the registers before ultimately programming the data into the array.

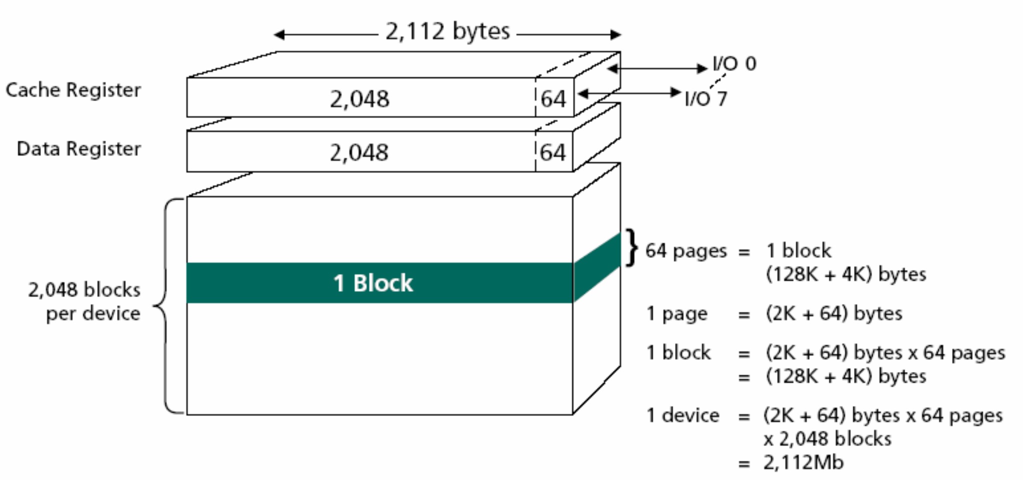
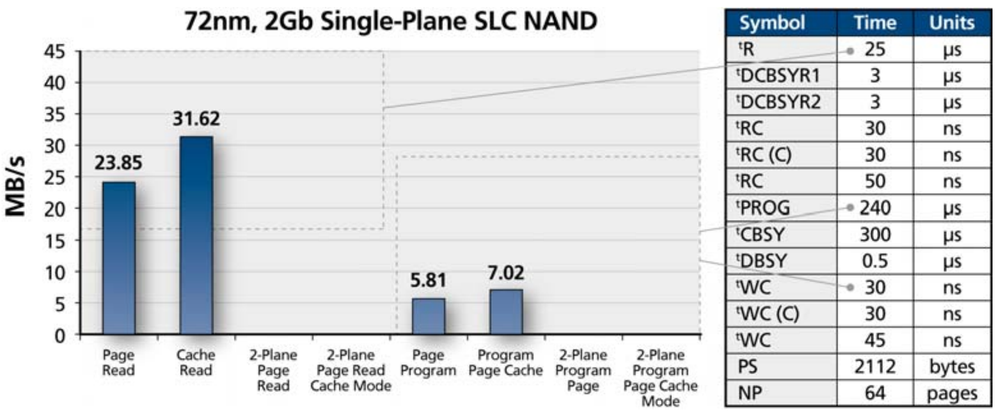


Figure 5

Figure 6 shows the major timing parameters and their effect on overall performance. The read performance of NAND Flash memories is affected by three timing parameters: ^{t}R – the time to transfer data from the NAND array to the data register; ^{t}RC – the read clock cycle time; and ^{t}RC(C) – the read cycle time for cache operations. A normal 2Gb, 2K-page SLC NAND page read using 30ns clocks and 25µs ^{t}R equates to 23.85 MB/sec read performance, but in cache read mode (operating much like a pipelined read), it is possible to achieve almost 32 MB/sec performance.

For program performance, the key timing parameters are clocking data into the registers (^{t}WC) and programming the array (^{t}PROG). Normal programming performance can reach 5.8 MB/sec, while cache programming performance can reach more than 7 MB/sec.



Symbol	Time	Units
^{t}R	25	µs
^{t}DCBSYR1	3	µs
^{t}DCBSYR2	3	µs
^{t}RC	30	ns
^{t}RC (C)	30	ns
^{t}RC	50	ns
^{t}PROG	240	µs
^{t}CBSY	300	µs
^{t}DBSY	0.5	µs
^{t}WC	30	ns
^{t}WC (C)	30	ns
^{t}WC	45	ns
PS	2112	bytes
NP	64	pages

Figure 6

It is possible to almost double performance by moving to a two-plane SLC NAND architecture. These devices are divided into two physical planes with odd and even blocks. Shift registers are twice as big to accommodate the dual-plane data. It is possible to access two pages for read or program or to erase two blocks concurrently.

Figures 7A and 7B show the architecture and performance graphs for a 4Gb SLC NAND Flash. In a 4Gb SLC dual-plane device, program performance increases from 13 MB/sec for 2-plane operation to more than 19 MB/sec in cache mode—clearly a significant performance gain.

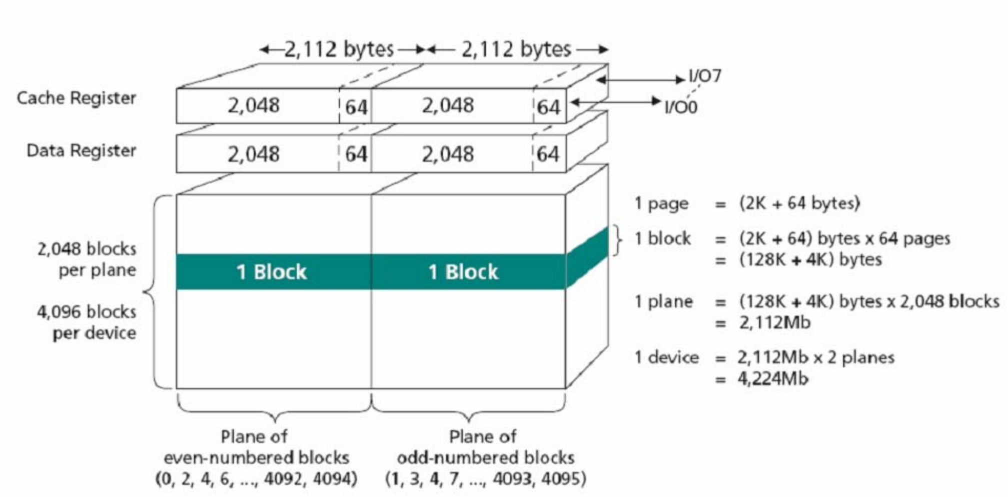
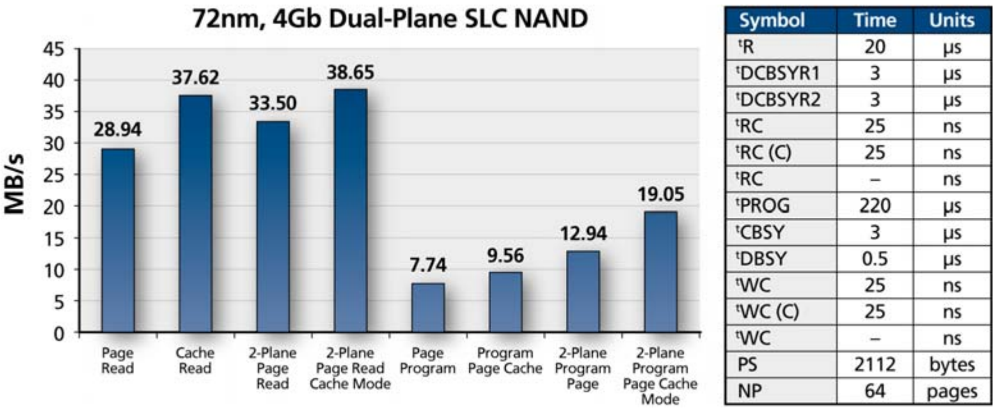


Figure 7A



Symbol	Time	Units
^{t}R	20	µs
^{t}DCBSYR1	3	µs
^{t}DCBSYR2	3	µs
^{t}RC	25	ns
^{t}RC (C)	25	ns
^{t}RC	–	ns
^{t}PROG	220	µs
^{t}CBSY	3	µs
^{t}DBSY	0.5	µs
^{t}WC	25	ns
^{t}WC (C)	25	ns
^{t}WC	–	ns
PS	2112	bytes
NP	64	pages

Figure 7B

Figures 8A and 8B show the architecture and performance graphs for an 8Gb MLC NAND Flash. Read performance ranges from 20 MB/sec to 37 MB/sec, and program performance ranges from 3 MB/sec to 6.5 MB/sec. Recall that MLC programming time is more than twice as slow as with an SLC NAND device.

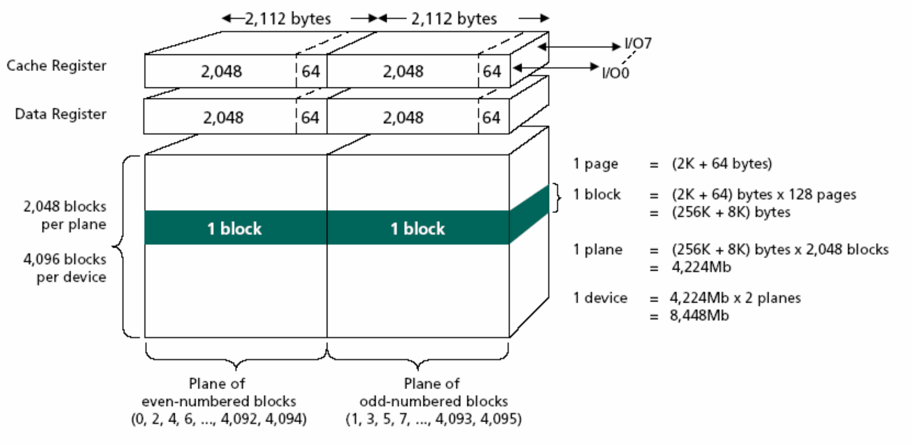


Figure 8A

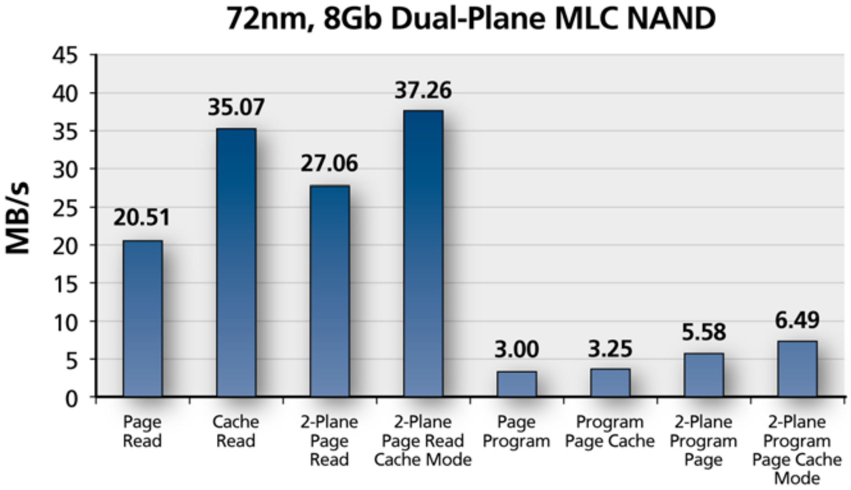


Figure 8B

Figures 9A and 9B show the architecture and performance graphs for a 16Gb MLC device. This device's 2-plane, 4K-page MLC architecture yields excellent read performance. Program performance is significantly better—up to 4.28 MB/sec for normal program commands and up to 9.5 MB/sec using the program cache mode. The 4K page has 4,096 bytes of data and 218 bytes of spare area, compared to the 8Gb device's 2,048-byte page with 64 bytes of spare area. The extra spare area is available for enhanced ECC coding that the controller can take advantage of.

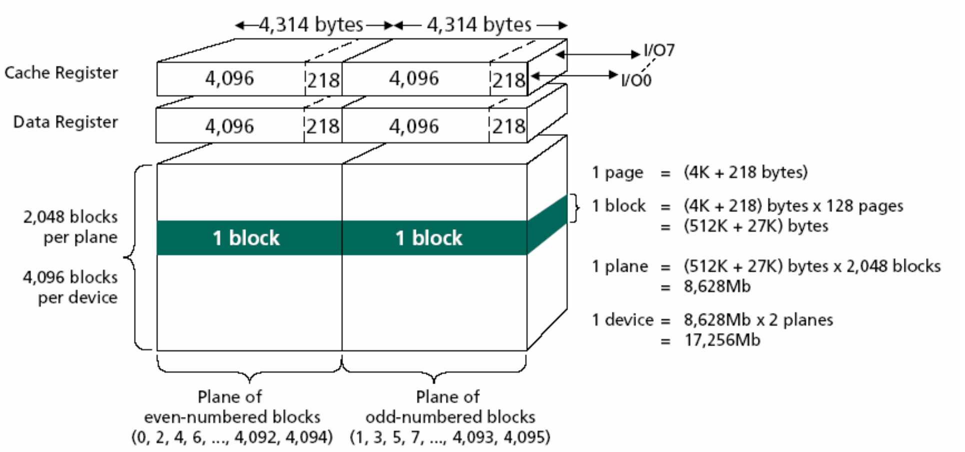


Figure 9A

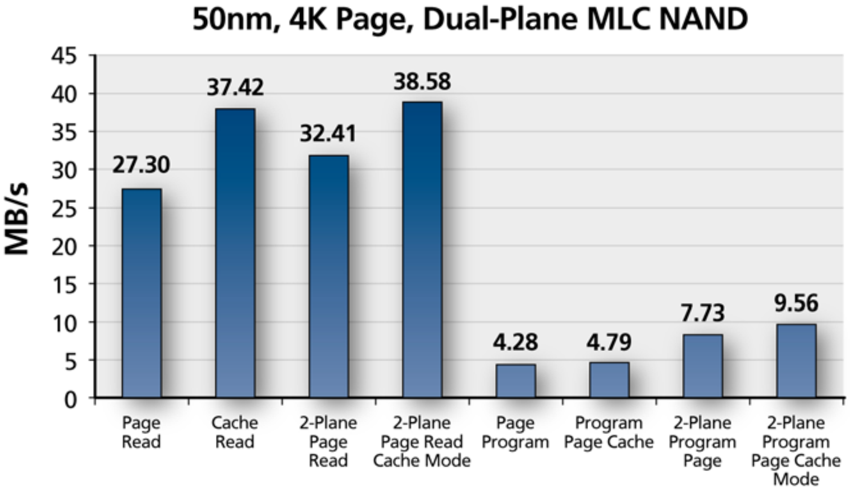


Figure 9B

NAND Error Modes

Error modes for NAND include program disturb, read disturb and endurance. Each error-mode issue is well understood and easily addressed. It is mandatory to use the minimum ECC specified for reliable systems. Using more robust ECC schemes will provide additional system reliability.

In some cases, Program and Read operations may cause electrons to move to or from other cells within the block.

To reduce program disturb, it is recommended to sequentially program pages in a block. It is also important to minimize partial-page programming operations in SLC devices, and it is mandatory to restrict page programming to a single operation in MLC.

Read disturb can be reduced by minimizing excessive reads. The rule of thumb is no more than 1 million READ cycles (per block) for SLC, and a maximum of 100,000 READ cycles for MLC. If possible, the data should be read equally from pages within the block. If it is necessary to exceed the “rule of thumb” cycle count, then the data should be moved to another block and the original block should be erased. Each erase resets the read disturb cycle count.

Endurance for MLC devices are typically specified at 10,000 PROGRAM/ERASE cycles, and SLC devices are specified at 100,000 PROGRAM/ERASE cycles. Also, it is possible to meet the extended data retention by limiting PROGRAM/ERASE cycles in blocks that require long retention. In this way, infrequently cycled blocks will have longer retention and frequently cycled blocks will have shorter retention.

As a general rule, to improve endurance, the pass/fail status (SR0) for PROGRAM and ERASE operations should always be checked. If there is a fail status after program, the data should be moved to an available block and the failed block should be marked as bad.

It is important to employ wear leveling, which ensures that data is written equally to all good blocks rather than cycling the same block. Wear leveling provides additional benefits on SLC devices where blocks can support up to 100,000 PROGRAM/ERASE cycles, but is imperative on MLC devices where blocks can typically support less than 10,000 cycles. Consider this: if a block were to be erased and programmed each minute, the 10,000 cycling limit would be exceeded in just 7 days (60 x 24 x 7 = 10,080 cycles). Now consider an 8Gb MLC device that contains 4,096 independent blocks. Using the previous example and distributing the cycles over all 4,096 blocks, each block would be programmed fewer than three times (versus the 10,800 cycles involved with cycling the same block). If one were to provide perfect wear leveling on a 4,096 block device every minute of every day, it would take 77 years to reach the specified PROGRAM/ERASE cycle limit for the device.

Some believe that ECC can fix anything and it is true—to a point. It is still critically important to understand the target data-error rate for a particular system and the intended use model for the system (including wear leveling and the number of partial page programs), because these all have a dramatic effect on the overall reliability of the system. The ECC circuit must be designed to improve the raw bit error rate (BER) of the NAND Flash device under intended use conditions so as to meet the system's target BER. As the raw NAND Flash BER increases, matching the ECC to the application's target BER becomes more important.

Another Option: Embedded MultiMedia Card (eMMC)

Interfacing to NAND directly will always provide the lowest-cost solution. However, future MLC devices will have complexities that will require increased attention. For example, the ECC necessary to support MLC is moving from 4+ bits to 8+ bits and higher in the future. When interfacing to NAND, the processor needs to understand the architectural details of the NAND such as flash block sizes, page size, number of pages in a block, number of planes, and advanced features such as cache operations, timing, etc.

eMMC offers the next logical step in the NAND Flash evolution for embedded applications. eMMC addresses all of the NAND complexities and offers a simple write/read memory with a standard MMC interface. The availability of a standard interface saves valuable time for the system designer. The high-speed interface supports clock rates of up to 52 MHz and is offered in host-selectable bus widths of 1, 4, or 8 bits. eMMC is offered in a standard BGA package and handles all ECC, wear leveling, and block management, off-loading these tasks from the processor.

Figures 10A and 10B show the differences between a standard interface and the new eMMC managed interface. With the traditional solution on the left, the processor interfaces to the NAND directly. A typical processor would include the buffers, ECC, and logic necessary to interface with the NAND directly. The system software must perform the block management, wear leveling, and block replacement algorithms. In contrast, the managed eMMC interface, shown in Figure 11B, only requires that the processor include a low-level driver with which to implement a simple read/write memory interface. All of the complexities of the NAND interface are handled by the on-chip controller.

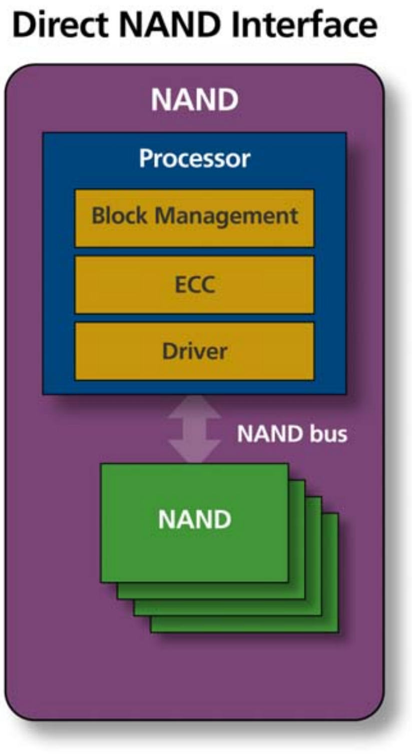


Figure 10A

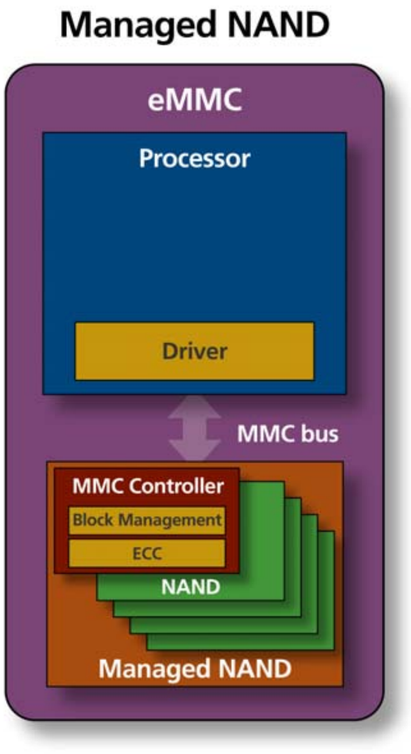


Figure 10B

Conclusion

The storage industry is embarking on a new chapter in system design and hierarchical memory allocation and partitioning. Thanks to many advances in NAND performance, reliability, and system integration, this storage technology is rapidly emerging as the preferred nonvolatile memory for SSDs, HHDs, and mobile devices. As NAND Flash technology evolves, the market is splitting between low-cost consumer applications that demand MLC solutions, and professional and prosumer products, including UFDs and SSDs, which are best served by higher-performance SLC devices. As MLC devices continue to grow in popularity, however, designers are coming face to face with their increased complexity and the additional hardware and software design considerations they involve. For mobile or embedded applications, eMMC provides an attractive alternative for managing all of these complexities and provides the next logical step in the NAND Flash evolution.