

TrueFFS[®]: Enabling the Latest NAND Technology for Growing Personal Storage Demands

October 2006

DC-000234, Rev 2.0



personal storage. made smarter™

msystems

Table of Contents

Scope	3
Background.....	3
TrueFFS [®] flash management technology.....	4
What's wrong with NAND?.....	4
Limited P/E cycle	4
Bit flipping.....	4
Bit pairing	4
Data retention errors	5
Randomly scattered bad blocks	5
Limited flash lifespan.....	5
MLC and 4-bit/cell flash technology constraints.....	6
MLC data reliability	6
MLC performance.....	7
Overcoming flash constraints with TrueFFS.....	7
Virtual dynamic mapping.....	7
Dynamic bad block management	8
Dynamic and static wear-leveling	8
Optimized erase algorithms	9
Error detection/correction	9
Power failure immunity	9
Detects and tracks dynamic bad blocks	9
Conclusion.....	10
How to Contact Us	11

Scope

Storage-hungry applications are proliferating in the mobile arena and in embedded systems. As the market demands better performance, capacity features and design, NAND flash media is becoming more compact and cost-effective. But by increasing densities, reducing size and cost, the quality of NAND flash is gradually dropping, as well. Although NAND flash is recognized as the best technology available today for high-capacity storage, NAND flash is an inherently difficult material to work with. NAND suffers from random errors as well as reliability, endurance and data retention problems that require complex management solutions. These problems are multiplied as NAND's bit per cell density increases from SLC, to MLC and to the latest processes.

This white paper discusses the problems specific to NAND flash, and explains how msystems patented TrueFFS[®] flash management technology overcomes these problems, making all the benefits of the latest NAND flash technology available for easy integration at the capacities, levels of performance, reliability and at the cost demanded by OEMs and their customers.

Background

It was not too long ago when NOR flash was the only non-volatile memory (NVM) solution available to designers. However, as the storage requirements evolved, with new needs for high capacity and high performance, NOR's poor cost structure, inability to scale to large capacities and very slow write performance have limited its usage to applications requiring small storage capacities.

This opened the door for NAND to become today's preferred non-volatile media solution for new applications such as feature-rich mobile handsets, consumer electronic devices and now solid state disks (SSDs) for laptops which take advantage of NAND's ability to provide both high-performance, high-capacity data storage and bootable OS and application code. But flash solution vendors and their OEM customers have discovered that the benefits of NAND do not come without a price; Reliability, endurance and data retention problems, many of which are described below, can render raw multi-level cell (MLC) NAND flash very complex to manage.

With every new generation of NAND flash, as bit per cell density increases and process nodes shrink, NAND foundries produce raw material which is more and more prone to errors, yields to faster media degradation and requires complex management technologies to enable acceptable levels of performance, reliability and data retention. Nearly 80% of the applications using NAND today either employ a strong controller embedded flash management software or do not require high reliability (low-end MP3 players, for example). With this in mind, fabs concentrate on delivering the highest density NAND as quickly as possible to keep up with the latest market prices without concern for compatibility from one generation/process to the next. Due to the nature of NAND flash, and this manufacturing strategy, the problems are different from one silicon vendor to the next and from one process to another, each delivering its own flavor of NAND, with unique properties which also change from generation to generation.

Ironically, as the demands increase on the OEMs to rapidly develop and ship new, reliable products to market with high-capacity, cost-effective storage, the NAND flash media that they have to work with is getting worse and worse. Advanced flash management technology such as msystems TrueFFS can solve all of the performance, reliability, data retention and compatibility problems associated with the latest NAND flash, making them transparent to OEM application designers and their customers, even for the most recent MLC-based products. Using TrueFFS flash management technology, the latest NAND flash becomes the most accessible, easy-to-implement and cost-effective NVM solution for designs demanding high-capacity data storage and bootable OS and application code.

TrueFFS[®] flash management technology

TrueFFS (True Flash File System) is msystems' patented flash management technology and has become the de facto market leader for NAND flash management. Essentially, TrueFFS is a software driver and controller combination that resides between the operating system and the flash media. TrueFFS provides the operating system with full block-device functionality so that the flash device appears to the operating system as a standard hard drive. At the same time, TrueFFS transparently manages the flash media, masking all of the intrinsic problems that plague raw NAND flash. TrueFFS is a critical element in overcoming the OEM's challenge to deliver high-performance applications that demand high-capacity data storage and bootable OS and application code. TrueFFS makes the latest NAND flash accessible while reducing time to market and development costs.

What's wrong with NAND?

NAND flash technology uses a unique method to organize stored data, which does not enable it to be accessed by means of standard file system calls. In addition, the various flash technologies are plagued, to lesser and greater extents, by a number of limitations that reduce the flash media lifespan and overall reliability.

With the introduction of new NAND flash technologies such as the 4 bit/cell NAND, these constraints and limitations are expected to become exponentially more complex, even more so than they have with the transition from SLC to MLC (1-bit/cell to 2-bit/cell).

Limited P/E cycle

Before writing, or programming new data to flash media, previously written data must be erased to free space. This is known as the program/erase (P/E) cycle. As silicon manufacturing geometries are reduced to squeeze more cells onto each wafer, and bit per cell density increases, there is a greater likelihood of errors, requiring an even more complex P/E cycle. This process causes the flash to reach its P/E cycle limit faster (described in *Limited flash lifespan*, below) and increases power consumption

Bit flipping

In flash architectures sometimes a bit can either be reversed or be reported as reversed. One such reversal may seem insignificant; For example when viewing a photo or listening to a song, if a bit is corrupted, the user will not notice the problem. However, if the bit flip affects the system OS, the configuration files or other sensitive information, it may cause the system to be corrupted and to hang completely. When faulty reporting is the problem, repeating the read operation may solve it. But if the bit was actually reversed, an elaborate error detection/correction (EDC/ECC) algorithm must be applied.

Bit pairing

In MLC NAND, each of the two bits stored in a single physical cell is, by design, written to a different logical memory page than its pair. As in any NVM media, a single bit can be corrupted for a number of reasons including a disturbance during writing caused by a power or software failure. When one bit in the cell is corrupted, its MLC pair is guaranteed to be corrupted, as well. For example, an OS might write a bit of OS code to an MLC cell. Then a day later a relatively insignificant bit from a multimedia file is corrupted during the write operation in the other bit in the cell. Although the multimedia file can sustain corruption of a single bit without notice, the previously written paired bit will also be corrupted. Without flash management such as TrueFFS, the effects of bit pairing could lead to several problems such as unexpected file corruptions, the mobile address book disappearing or even an OS freeze or crash. To make things worse, bit-

pairing schemes differ between NAND flash vendors and even between different NAND generations from the same vendor.

Data retention errors

Flash memory cells must retain the stability of their voltage level to ensure data retention for an acceptable period according to the needs of the application. In order to do this, the long-term stability of voltage levels is critical. Leakage to/from the floating gate, called charge drift, tends to slowly change the cell's voltage level from its initial level to a different level after cell programming or erasing. This new level may incorrectly be interpreted as a different logical value.

Randomly scattered bad blocks

Due to yield considerations, MLC NAND media are knowingly shipped from the fab with up to 5% randomly scattered bad blocks. In addition to the 'initial' bad blocks that are present when the flash is produced, normal wear from writing and erasing causes additional blocks to be corrupted, and to therefore become unusable.

Working with NAND devices requires initially scanning the media for bad blocks, and then mapping them all out to ensure they are not used. Failing to do so in a reliable manner may result in a high failure rate of the final device, and even a total recall.

Limited flash lifespan

The individual cells in all flash devices, regardless of the technologies on which they are based, are limited in the number of P/E cycles they can sustain before the probability of error rises to unacceptable levels. This number depends on the specific flash technology employed. A region of flash that is close to its cycle limit typically displays sporadic write failures that become more frequent and more critical. Eventually, the sector becomes unusable. Without flash management technology such as TrueFFS, more and more sectors become unusable. Capacity is then gradually reduced, existing data becomes prone to random errors and application and OS code can quickly become corrupted, increasing likelihood of total application failure.

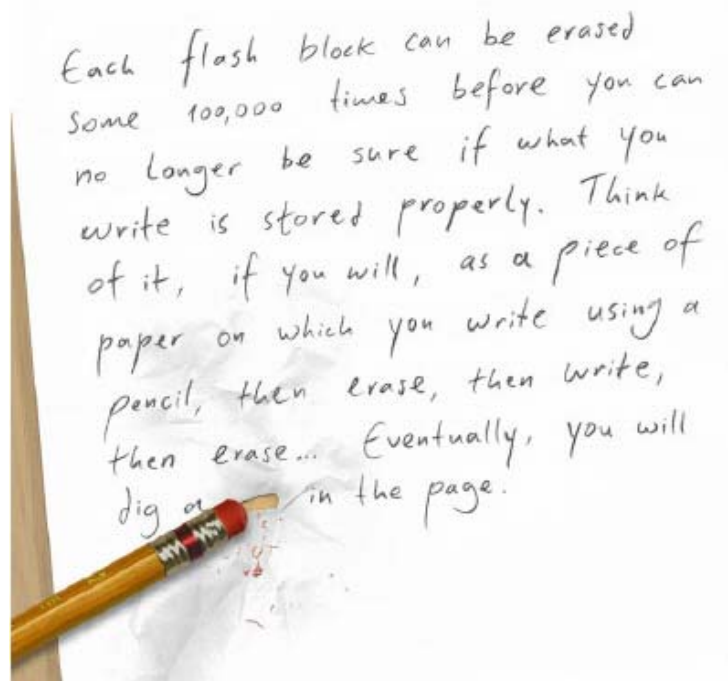


Figure 1: Like paper, there is a limit to the number of times flash media can be erased

A system usually has several relatively small but critical files, which are constantly being updated. If these files are programmed and erased at the same location, that location may reach its upper limit of P/E cycles fairly quickly, corrupting the data in these files. This phenomenon is more critical with MLC and 4-bit/cell flash which is far more prone to errors than SLC flash devices.

MLC and 4-bit/cell flash technology constraints

Increasing NAND flash bits per cell density reduces cost and size, but also exacerbates standard flash limitations, potentially degrading reliability, performance and flash lifespan, and further complicating ease-of-integration.

Single Level Cell (SLC) technology stores 1 bit of data per cell using two voltage levels, Multi-Level Cell (MLC) stores 2 bits of data per cell, using four voltage levels, and 4-bit/cell technology will store 4 bits per cell, using 16 voltage levels (see Figure 2). Although 4-bit/cell represents the most advanced, cost-effective NAND technology to date, it also presents unsurpassed complexity.

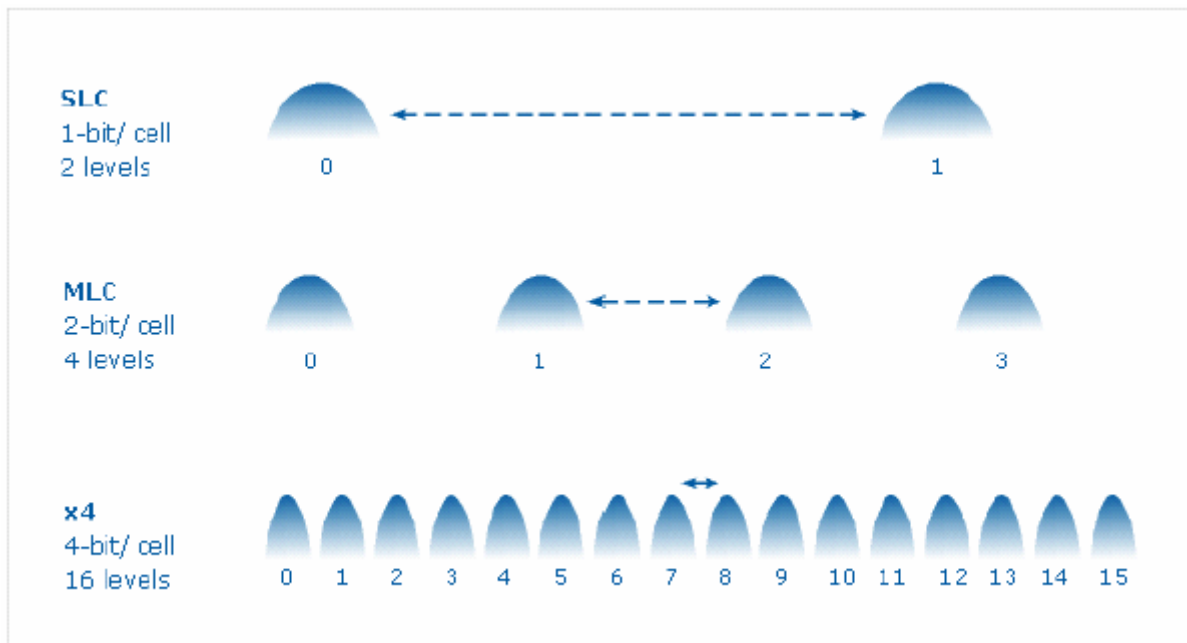


Figure 2: Voltage Level Comparison between SLC, MLC and 4-bit/cell Flash Technologies

Increased density results in increased complexity. The programming and erase processes become more complicated since the circuits must maintain tighter tolerances. As a result, multiple-bit-per-cell technology affects all constraints for standard flash technology.

Without advanced flash management technology such as TrueFFS, implementing cost-reducing innovations such as 4-bit/cell NAND would be impossible.

MLC data reliability

As shown in Figure 2, an SLC flash cell must distinguish between 2 voltage levels, whereas an MLC flash cell must distinguish between 4 or even 16 levels. Since both SLC and MLC-based devices use a voltage window with a similar size, the distance between adjacent voltage levels in MLC is much smaller than in SLC flash. 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 an SLC flash cell, subject to a higher probability of read and write errors. MLC and 4-bit/cell flash

media are more likely to be affected by current drift/leakage and increased parasitic capacitance effects which cause unpredictable changes in voltage levels that can impact data reliability, performance and power consumption.

This degradation is caused not only by new advanced technologies but also by continually decreasing lithography (also known as manufacturing process, or geometry). The size reduction in the manufacturing process, from 400nm in 2000 to down to 60nm today, exacerbates reliability problems, making it even more difficult to model and simulate and therefore hard to mitigate without adding considerable circuitry.

MLC performance

Due to the increased complexity of managing MLC NAND technology, basic flash operations of reading a page into the flash buffer, writing a flash buffer into a page, and erasing a flash unit require more time than SLC technology. It is difficult to create a solution without compromising either performance or reliability. Advanced flash management technology such as TrueFFS can balance the complexity of maintaining the reliability of data access without reducing system performance.

Overcoming flash constraints with TrueFFS

TrueFFS employs mechanisms that overcome all of the limitations that are, in essence, manufactured into SLC, and to a greater extent MLC NAND flash media. Brief descriptions of these mechanisms follow.

Virtual dynamic mapping

The flash media is organized in physical blocks (also called erase units) that are further divided into physical sectors, the smallest storage area in a read/write block. Standard file system calls specifying hard disk sector numbers and cylinders cannot be used to access data on the flash media.

TrueFFS dynamic virtual mapping is the process of mapping the OS storage model to the flash physical model (see Figure 3). This ensures consistently fast performance when accessing data. TrueFFS achieves this by clustering related data, such as sectors of a file, into one unit and distributing unrelated data evenly among other units. This approach reduces both fragmentation and the number of calls required to map different physical blocks into the memory window.

The TrueFFS Flash Translation Layer (FTL) is the industry standard flash data recording format. It enables flash to emulate the behavior of a standard disk drive by mapping the OS model to the physical flash model (see Figure 3); specifically, to the physical location at which the data is to be written. This enables it to plug into almost any type of file system, without learning the internal structure of the flash media and its unique command interface.

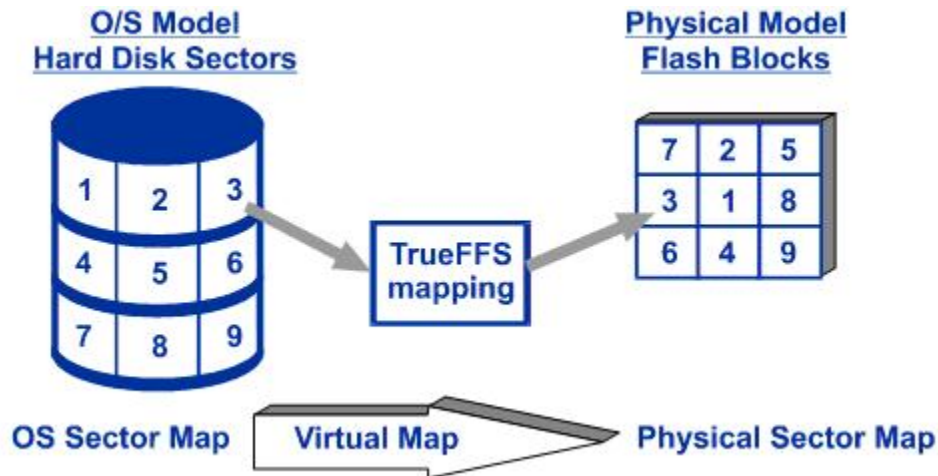


Figure 3: TrueFFS Dynamic Virtual Mapping to Emulate Hard Disk

To emulate a hard disk, an erase unit is evenly divided into one or more equally sized read/write blocks, which are the same size as a hard disk data sector (see Figure 4).

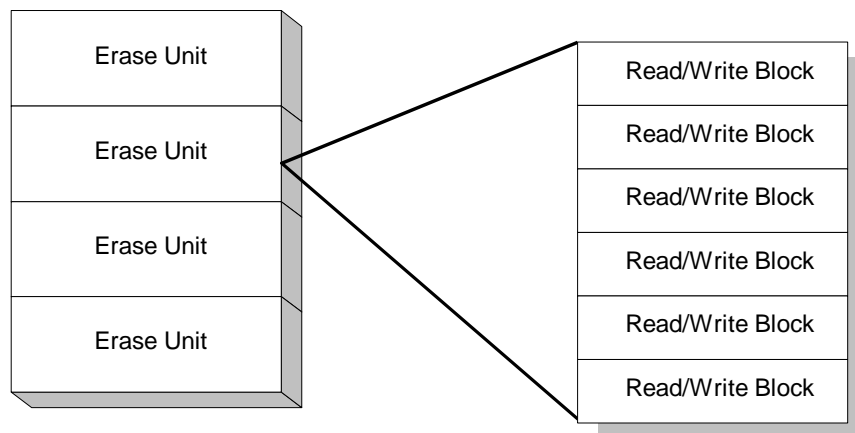


Figure 4: TrueFFS Hard Disk Emulation Mechanism

Dynamic bad block management

TrueFFS maps out all NAND flash bad blocks dynamically, making sure that they are not accessed for reliable data storage, and remembering the bad block locations so that they are not accessed in subsequent operations. This not only ensures data integrity, but also enhances performance by eliminating the need for repeated write operations as a result data being repeatedly mapped to the same bad block.

Dynamic and static wear-leveling

The number of P/E cycles for each physical erase block in NAND flash media is limited by the material itself. This limit – once estimated at 100K for SLC NAND – drops as bit/cell density increases; Current MLC media is often estimated at ~10K P/E cycles. An advanced mechanism is required to ensure that the P/E cycles are distributed evenly across the entire flash media in order to extend the lifetime of the device.

TrueFFS uses two types of wear-leveling, dynamic and static:

- **Dynamic wear-leveling.** TrueFFS uses statistical allocation to perform dynamic wear-leveling on newly written data. Logical sectors are dynamically mapped to any available physical sector. This results in blocks that are used evenly while not affecting high-performance levels.
- **Static wear-leveling.** Flash regions may contain static files, characterized by sectors of data that remain unchanged for very long periods of time. If wear-leveling were applied only to newly written data, static areas would never be cycled, at the expense of the remaining areas that would wear out rapidly. Applying only dynamic wear-leveling would lower data retention levels even more significantly in cases where flash memory contains large static areas. To overcome this limitation, TrueFFS forces data transfer into static areas as well as into dynamic areas, thereby applying wear-leveling over the entire media.

Optimized erase algorithms

TrueFFS includes algorithms for minimizing the number of erase operations and optimizing folding operations, which enable "garbage collection" for future usage of blocks that no longer contain valid data. During garbage collection operations, TrueFFS also performs space reclamation, which is similar to defragmentation in a hard disk.

Error detection/correction

TrueFFS implements sophisticated mathematical algorithms to ensure high data reliability and solve MLC NAND bit-flipping issues without degrading performance. Error detection is implemented in hardware by the controller to achieve faster results.

Power failure immunity

TrueFFS uses algorithms including "erase after write" instead of "erase before write" to ensure data integrity during normal operation and in the event of a power failure. Used areas are reclaimed only after an operation is complete by erasing and updating the virtual maps. The "erase after write" algorithm is also used to update and store mapping information on the flash memory. This keeps the mapping information coherent, even during power failures. The only mapping information held in RAM is a table pointing to the location of the actual mapping information. This table is reconstructed from the information stored in the flash memory during power-up or after reset.

Additionally, if a write operation is interrupted by a power failure causing a sector to be corrupted, TrueFFS then checks the sector and marks it as though it was never written to. This enables the sector to be overwritten without adding programming time at the time of the next write.

Power failure immunity is more complicated in MLC NAND due to the bit pairing effect. A previous, successfully written bit could be damaged by a new write operation due to bit pairing. To make matters worse, each MLC NAND flash media has its own unique characteristics. TrueFFS overcomes these new challenges.

Detects and tracks dynamic bad blocks

The first time that a write operation is performed to a block, TrueFFS reads the data and verifies that the block is operative. On subsequent write operations to the same block, self-adapting algorithms "learn" the block status from the previous operation, enabling the verify operation to be omitted and thereby enhancing performance.

Conclusion

Raw NAND flash is an unpredictable material, becoming more and more problematic as bit/cell density increases and silicon geometries shrink from one generation to the next. This presents design engineers with increased integration and material costs, delays and risks. These obstacles make it difficult to develop solutions that can effectively take advantage of the latest, most advanced materials while rushing new and innovative applications to market.

msystems TrueFFS flash management software makes the most advanced NAND flash available to OEMs by eliminating all major NAND flash challenges. By incorporating advanced flash management technologies, including error detection code (EDC) and error correction code (ECC), automatic bad block detection and advanced wear leveling algorithms TrueFFS enhances performance, reliability and endurance, accelerates design and integration time and reduces overall costs to enable OEMs to meet the growing demands of their customers.

msystems' strong IP portfolio, including numerous innovations employed in TrueFFS, mitigates the risks that can face OEMs who attempt to integrate NAND flash with 3rd party management software and drivers. When NAND flash, an advanced controller and the advanced flash management technologies of TrueFFS are all developed and manufactured to work together as an organic unit, OEMs can count on unsurpassed reliability and easy integration of the latest flash into applications that require high-capacity data and bootable code storage.

How to Contact Us

USA

msystems, Inc.
555 North Mathilda Avenue
Suite 220
Sunnyvale, CA 94085
Tel: +1-408-470-4440
Fax: +1-408-470-4470

China

msystems (Shenzhen) Trading Ltd.
Room 121-122
Bldg. 2, International Trade & Commerce Bldg.
1001 HongHus Rd.
Futian Free Trade Zone
Shenzhen, China
Tel: +86-755-8348-5218
Fax: +86-755-8348-5418

Japan

msystems Japan, Inc.
Asahi Seimei Gotanda Bldg., 3F
5-25-16 Higashi-Gotanda
Shinagawa-ku Tokyo, 141-0022
Tel: +81-3-5423-8101
Fax: +81-3-5423-8102

Korea

msystems Asia Ltd. Korea Representative Office
#1002 BYC Building
648-1 Yeoksam-dong
Kangnam-ku, Seoul, Korea
Tel: +82-2-3452-9079
Fax: +82-2-3452-9145

Taiwan

msystems Asia, Ltd.
Room 14F
No. 6, Sec. 3
Minquan East Rd
Zhongshan District
Taipei, Taiwan, 104
Tel: +886-2-2515-2522
Fax: +886-2-2515-2295

Europe

msystems Ltd.
7 Atir Yeda St.
Kfar Saba 44425, Israel
Tel: +972-9-764-5000
Fax: +972-3-548-8666

Internet

www.m-systems.com

Information

info@m-systems.com

This document is for information use only and is subject to change without prior notice. msystems Ltd. assumes no responsibility for any errors that may appear in this document. No part of this document may be reproduced, transmitted, transcribed, stored in a retrievable manner or translated into any language or computer language, in any form or by any means, electronic, mechanical, magnetic, optical, chemical, manual or otherwise, without prior written consent of msystems.

msystems' products are not warranted to operate without failure. Accordingly, in any use of the Product in life support systems or other applications where failure could cause injury or loss of life, the Product should only be incorporated in systems designed with appropriate and sufficient redundancy or backup features.

Contact your local msystems' sales office or distributor, or visit our website at www.m-systems.com to obtain the latest specifications before placing your order.

© 2006 msystems Ltd. All rights reserved.

msystems, DiskOnChip, DiskOnKey, DiskOnKey MyKey, FFD, Fly-By, iDiskOnChip, iDOC, mDiskOnChip, mDOC, MegaSIM, SuperMAP, TrueFFS, uDiskOnChip, uDOC, and Xkey are trademarks or registered trademarks of msystems Ltd. Other product names or service marks mentioned herein may be trademarks or registered trademarks of their respective owners and are hereby acknowledged. All specifications are subject to change without prior notice.