Flash Memory Technology Direction

Jim Cooke (jcooke@micron.com)
Director of Applications Engineering
Memory Products Group
Micron Technology, Inc.
Agenda

- Disk drive definitions: HDD, HHD, SSD
- NAND Flash benefits for drives
  - Hybrid
  - ReadyBoost™
  - Intel® Robson technology
  - Chipset adoption
  - SSD
Agenda (continued)

- NAND Flash market
- MLC versus SLC
  - Architecture
  - Performance
- Error modes
- Embedded MMC (eMMC)
Disk Drive Definitions

- Hard disk drives (HDD)
  HDDs utilize ultra-sophisticated magnetic recording and playback technologies. They are used as the primary data storage component in notebooks, desktops, servers, and dedicated storage systems.

- Hybrid hard drives (HHD)
  HHDs are a new type of large-buffer computer hard drive. They are different from standard hard drives in that they employ a large buffer (up to 1GB) of nonvolatile flash memory used to cache data during normal use. By using this large buffer, the platters of the hard drive are at rest almost all times, instead of constantly spinning as is the case in HDDs. This feature offers numerous benefits, such as decreased power consumption, improved reliability, and a faster boot process.

- Solid state drives (SSD)
  SSDs are data storage devices that use nonvolatile memory (Flash) and volatile memory (SDRAM) to store data. While technically not "disks," these devices are referred to this way because they are typically used as replacements for HDDs.

Source: Gartner, wikipedia.org
SSD versus HDD

<table>
<thead>
<tr>
<th></th>
<th>SSD</th>
<th>HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Performance</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Endurance</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Power</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Weight</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Cost per bit</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

- Based on recent advances in NAND lithography, SSD densities have reached capacities for mass market appeal.
- SSD offers many features improve user experiences.
- Early concerns about reliability and endurance have been overcome.

NAND solid state drives are ready for deployment in many applications
Why Add NAND to a PC?

Issues in PC architecture today

- Long boot times for OS and applications
- Unacceptable boot-up times for applications
- Hard disk drive (HDD) latency falling behind processor performance—HDD maximizes GB, not performance
Why Add NAND to a PC? (continued)

- Industry wants extended notebook battery lifetime
- HDD access (and motors) degrade battery life
- NAND accesses save power
PC Opportunities for NAND

1. PC chipset/add-in card
   - Intel® Robson technology in future platforms
   - Add-in card or soldered onto motherboard

2. Hybrid HDD with cache
   - Add NAND to the HDD chipset
   - Microsoft approach

3. Solid state drive (SSD)
   - Flash replacement for HDD
Hard Disk Drive (HDD)

- Rotational latency
- Seek latency

Source: Web-Feet Research 2006
Hybrid hard drives have the same basic structure as standard HDDs, but also have a nonvolatile cache. This feature enables near instantaneous read/write capability even when the spindle has stopped.
Additional Ways NAND Flash Boosts Performance

- ReadyBoost can be implemented as:
  - Add-on USB Flash disk
  - Add-on ExpressCard™
  - Add-on SD/MMC card or any other media

- Users determine how much of the Flash is used as a performance cache
ReadyBoost Setup

Speed up your system by utilizing the available space on this device.

- Do not use this device.
- Use this device.

Space to reserve for system speed:

While the device is being used for system speed the reserved space will not be available for file storage.

Intel Robson Technology

- First ships with the Santa Rosa notebook chipset platform
- Rolls out toward the end of 1Q07 and will be implemented with Microsoft's® Vista™
- Standard Santa Rosa Robson chipset configuration expected to include 512MB of NAND, but to offer 1GB as an option

Source: Gartner, 15 December 2006
# Vista-Equipped Portable PCs with NAND Caching Solutions

(Attach Rate by Year)

<table>
<thead>
<tr>
<th>Technology</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither NAND caching technology</td>
<td>95%</td>
<td>77%</td>
<td>43%</td>
<td>15%</td>
</tr>
<tr>
<td>Hybrid HDD</td>
<td>5%</td>
<td>14%</td>
<td>31%</td>
<td>41%</td>
</tr>
<tr>
<td>Embedded NAND</td>
<td>1%</td>
<td>9%</td>
<td>26%</td>
<td>43%</td>
</tr>
</tbody>
</table>

Source: IDC 2006
NAND Chipset Adoption in Notebook PCs (2007–2010)

Source: Gartner Dataquest (October 2006)
NAND Chipset Adoption in Desktop PCs (2007–2010)

Source: Gartner Dataquest (October 2006)
Hybrid HDDs in PCs and Associated NAND Flash Consumption

Source: Gartner Dataquest (November 2006)
Solid State Drives are Different

- No moving parts
- Lower power (less heat, longer battery life)
- More rugged
- Faster
### Solid State Drives are Rugged

<table>
<thead>
<tr>
<th>Average Specifications</th>
<th>Hard Disk Drive</th>
<th>Solid State Drive</th>
<th>Hard Disk Drive</th>
<th>Hybrid Hard Drive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.8in HDD</td>
<td>SSD (1.8/2.5in)</td>
<td>2.5in HDD</td>
<td>2.5in HHD</td>
</tr>
<tr>
<td>Capacity</td>
<td>30–80GB</td>
<td>4–32GB</td>
<td>40–160GB</td>
<td>Up to 160GB</td>
</tr>
<tr>
<td>Data rate (max sustain)</td>
<td>25 MB/s</td>
<td>57 MB/s</td>
<td>44 MB/s</td>
<td></td>
</tr>
<tr>
<td>Read</td>
<td>25 MB/s</td>
<td></td>
<td>44 MB/s</td>
<td></td>
</tr>
<tr>
<td>Write</td>
<td>25 MB/s</td>
<td>32 MB/s</td>
<td>44 MB/s</td>
<td></td>
</tr>
<tr>
<td>Spindle speed</td>
<td>4,200 RPM</td>
<td>None</td>
<td>5,400 RPM</td>
<td>5,400 RPM</td>
</tr>
<tr>
<td>Seek</td>
<td>15ms</td>
<td>None</td>
<td>12ms</td>
<td>12.5ms</td>
</tr>
<tr>
<td>Non-operating shock</td>
<td>1,500G</td>
<td>2,000G</td>
<td>900G</td>
<td>900G</td>
</tr>
</tbody>
</table>

- Both SSD and HHD provide power savings in various applications, but the exact power savings fluctuates from application to application.
- In a test of a 32GB SSD drive, the power savings of the SSD was 1 watt better than the closest tested HDD.

Source: Web-Feet Research, Seagate, Tom’s Hardware
SSDs Forecasted to Grow at a 146 Percent CAGR from 2005–2010

Worldwide Disk Drives by Type
(Millions of units shipped)

- Solid state drives
- Hybrid hard drives
- Hard disk drives

Source: Web-Feet Research, Gartner
Data Processing Applications Expected to Drive Majority of SSD Unit Shipments

Worldwide SSD Shipments by Market
(Millions of units shipped)

- Military/aerospace
- Industrial/medical
- Consumer
- Data processing

Source: Web-Feet Research
2.5in is the “Sweet Spot” Form Factor for SSDs

Worldwide SSD Shipments by Form Factor
(Millions of units shipped)

Source: Web-Feet Research
Demand for SSDs and HDDs in Portable Computers

Source: Web-Feet Research 2006
HDD versus SSD

2006

2011

Source: 2006 Web-feet Research
HDD, NAND Flash Pricing (Log Chart)

- HDD 0.85in, 1.0in, 1.8in Combined
- NAND Flash
- Mobile HDD 2.5in (portable PCs)

Source: IDC 2007
Solid State Drive Meets or Exceeds HDD Reliability

**Application**
- Hard Disk Drive: Application Data Error Rate $10^{-15}$ to $10^{-14}$
- Solid State Drive: Application Data Error Rate $\sim10^{-15}$

**Data management**
- Bad block management
- Channel and block coding

**Raw media**
- Typical raw error rate $>10^{-4}$
- Typical raw error rate $<10^{-5}$
Key Takeaways

- Hybrid hard drives represent an incremental upgrade to HDDs
- Solid state drives are significantly different and offer several advantages
- As NAND becomes increasingly competitive in the densities offered and the price, the adoption rate will increase
- 2.5in is the “sweet spot” form factor for SSDs
All signs point to ... NAND
MLC versus SLC

- Multi-level cell NAND (MLC) stores four states per memory cell and enables two bits programmed/read per memory cell.

- Single-level (SLC) NAND stores two states per memory cell and enables one bit programmed/read per memory cell.
MLC versus SLC (continued)

- MLC NAND Flash will lead in the lowest cost for consumer applications
  - Media players
  - MP3/camera phones
  - Media cards
- Professional products, ReadyBoost UFDs, and solid state drives (SSDs) will still demand the higher performance and higher reliability of single-level cell (SLC) NAND Flash
## MLC versus SLC (continued)

<table>
<thead>
<tr>
<th>Features</th>
<th>MLC</th>
<th>SLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bits per cell</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Voltage</td>
<td>3.3V</td>
<td>3.3V, 1.8V</td>
</tr>
<tr>
<td>Data width (bits)</td>
<td>x8</td>
<td>x8, x16</td>
</tr>
</tbody>
</table>

### 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

<table>
<thead>
<tr>
<th>tR (MAX)</th>
<th>PROG (TYP)</th>
<th>BERS (TYP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50µs</td>
<td>600–900µs</td>
<td>3ms</td>
</tr>
<tr>
<td>25µs</td>
<td>200–300µs</td>
<td>1.5–2ms</td>
</tr>
</tbody>
</table>
2Gb, 2K Page SLC NAND Architecture

- Cache Register: 2,048 bytes
- Data Register: 2,048 bytes
- 64 blocks per page
- 64 pages per block
- 64 blocks per device

64 pages = 1 block (128K + 4K) bytes
1 page = (2K + 64) bytes
1 block = (2K + 64) bytes x 64 pages = (128K + 4K) bytes
1 device = (2K + 64) bytes x 64 pages x 2,048 blocks = 2,112Mb
NAND Architecture

- NAND architecture is based on independent blocks.
- Blocks are the smallest erasable units.
- Pages are the smallest programmable units.
- Partial pages can be programmed in some devices.

* Typical for 4Gb SLC
2Gb, SLC 72nm, 2K Page Performance

72nm, 2Gb Single-Plane SLC NAND

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Time</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>tR</td>
<td>25</td>
<td>us</td>
</tr>
<tr>
<td>tDCBSYR1</td>
<td>3</td>
<td>us</td>
</tr>
<tr>
<td>tDCBSYR2</td>
<td>3</td>
<td>us</td>
</tr>
<tr>
<td>tRC</td>
<td>30</td>
<td>ns</td>
</tr>
<tr>
<td>tRC (C)</td>
<td>30</td>
<td>ns</td>
</tr>
<tr>
<td>tRC</td>
<td>50</td>
<td>ns</td>
</tr>
<tr>
<td>tPROG</td>
<td>240</td>
<td>us</td>
</tr>
<tr>
<td>tCBSY</td>
<td>300</td>
<td>us</td>
</tr>
<tr>
<td>tDBSY</td>
<td>0.5</td>
<td>us</td>
</tr>
<tr>
<td>tWC</td>
<td>30</td>
<td>ns</td>
</tr>
<tr>
<td>tWC (C)</td>
<td>30</td>
<td>ns</td>
</tr>
<tr>
<td>tWC</td>
<td>45</td>
<td>ns</td>
</tr>
<tr>
<td>PS</td>
<td>2112</td>
<td>Byte</td>
</tr>
<tr>
<td>NP</td>
<td>64</td>
<td>Pages</td>
</tr>
</tbody>
</table>
Two-Plane Features

- Device is divided into two physical planes, odd/even blocks
- Provides ability to:
  - Concurrently access two pages for read
  - Erase two blocks concurrently
  - Program two pages concurrently
- The page addresses of blocks from both planes must be the same during two-plane read/program/erase operations
4Gb, Two-Plane 2K Page SLC NAND Architecture

- 2,112 bytes per page
- 2,048 blocks per plane
- 4,096 blocks per device

1 page = (2K + 64 bytes)
1 block = (2K + 64) bytes x 64 pages = (128K + 4K) bytes
1 plane = (128K + 4K) bytes x 2,048 blocks = 2,112Mb
1 device = 2,112Mb x 2 planes = 4,224Mb
4Gb, 2K Page SLC NAND Performance

72nm, 4Gb Dual-Plane SLC NAND

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Time</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>tR</td>
<td>20</td>
<td>us</td>
</tr>
<tr>
<td>tDCBSYR1</td>
<td>3</td>
<td>us</td>
</tr>
<tr>
<td>tDCBSYR2</td>
<td>3</td>
<td>us</td>
</tr>
<tr>
<td>tRC</td>
<td>25</td>
<td>ns</td>
</tr>
<tr>
<td>tRC (C)</td>
<td>25</td>
<td>ns</td>
</tr>
<tr>
<td>tRC</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>tPROG</td>
<td>220</td>
<td>us</td>
</tr>
<tr>
<td>tCBSY</td>
<td>3</td>
<td>us</td>
</tr>
<tr>
<td>tDBSY</td>
<td>0.5</td>
<td>us</td>
</tr>
<tr>
<td>tWC</td>
<td>25</td>
<td>ns</td>
</tr>
<tr>
<td>tWC (C)</td>
<td>25</td>
<td>ns</td>
</tr>
<tr>
<td>PS</td>
<td>2112</td>
<td>Byte</td>
</tr>
<tr>
<td>NP</td>
<td>64</td>
<td>Pages</td>
</tr>
</tbody>
</table>
8Gb, Two-Plane 2K Page
MLC NAND Architecture

- 2,112 bytes
- 2,112 bytes
- I/O7
- I/O0
- Cache Register
- Data Register
- 2,048 blocks per plane
- 4,096 blocks per device
- 1 block = (2K + 64 bytes) x 128 pages = (256K + 8K) bytes
- 1 plane = (256K + 8K) bytes x 2,048 blocks = 4,224Mb
- 1 device = 4,224Mb x 2 planes = 8,448Mb
8Gb, 2K Page MLC Performance

72nm, 8Gb Dual-Plane MLC NAND

- Page Read: 20.51 MB/s
- Cache Read: 35.07 MB/s
- 2-Plane Page Read: 27.06 MB/s
- 2-Plane Page Read Mode: 37.26 MB/s

Symbol | Time | Units
--- | --- | ---
tR | 50 | us
tDCBSYR1 | 7 | us
tDCBSYR2 | 7 | us
tRC | 25 | ns
tRC (C) | 25 | ns
tRC | 25 | ns
tPROG | 650 | us
tCBSY | 30 | us
tDBSY | 0.5 | us
tWC | 25 | ns
tWC (C) | 25 | ns
tWC | 25 | ns
PS | 2112 | Byte
NP | 128 | Pages

Microsoft WinHEC 2007
16Gb, Two-Plane 4K Page MLC NAND Architecture

- **Cache Register:** 4,096 bytes
- **Data Register:** 4,096 bytes

- **2,048 blocks per plane**
- **4,096 blocks per device**

**1 block**
- **1 page** = (4K + 218) bytes
- **1 block** = (4K + 218) bytes x 128 pages = (512K + 27K) bytes
- **1 plane** = (512K + 27K) bytes x 2,048 blocks = 8,628Mb
- **1 device** = 8,628Mb x 2 planes = 17,256Mb
Two-Plane, 4K Page MLC NAND Architecture

50nm, 4K Page, Dual-Plane MLC NAND

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Time</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>tR</td>
<td>50</td>
<td>us</td>
</tr>
<tr>
<td>tDCBSYR1</td>
<td>7</td>
<td>us</td>
</tr>
<tr>
<td>tDCBSYR2</td>
<td>7</td>
<td>us</td>
</tr>
<tr>
<td>tRC</td>
<td>25</td>
<td>ns</td>
</tr>
<tr>
<td>tRC (C)</td>
<td>25</td>
<td>ns</td>
</tr>
<tr>
<td>tRC</td>
<td>50</td>
<td>ns</td>
</tr>
<tr>
<td>tPROG</td>
<td>900</td>
<td>us</td>
</tr>
<tr>
<td>tCBSY</td>
<td>3</td>
<td>us</td>
</tr>
<tr>
<td>tDBSY</td>
<td>0.5</td>
<td>us</td>
</tr>
<tr>
<td>tWC</td>
<td>25</td>
<td>ns</td>
</tr>
<tr>
<td>tWC (C)</td>
<td>35</td>
<td>ns</td>
</tr>
<tr>
<td>PS</td>
<td>4314</td>
<td>Byte</td>
</tr>
<tr>
<td>NP</td>
<td>128</td>
<td>Pages</td>
</tr>
</tbody>
</table>
Future Micron NAND Flash devices support the open NAND Flash interface (ONFI) specification.

Micron is a founding member of ONFI.

The ONFI 1.0 specification is available at [http://www.onfi.org/](http://www.onfi.org/).
NAND Error Modes

- Program disturb
- Read disturb
- Data retention
- Endurance

These issues are understood well and can be addressed.
ECC Can Fix Everything (Well Almost)

- You must understand your target data error rate for your particular system.
- Understand the use model that you intend for your system.
- Design the ECC circuit to improve the raw bit error rate (BER) of the NAND Flash, under your use conditions, to meet your system’s target BER.
ECC Code Selection Becoming More Important

As the raw NAND Flash BER increases, matching the ECC to the application’s target BER becomes more important.

For SLC:
A code with a correction threshold of one is sufficient

$t = 4$ required (as a minimum) for MLC
Reducing Program Disturb

- Program pages in a block sequentially, from page 0 to page 63 (SLC) or to page 127 (MLC)
- Minimize partial-page programming operations (SLC)
- It is mandatory to restrict page programming to one single operation (MLC)
- Use ECC to recover from program disturb errors
Reducing Read Disturb

- "Rule of thumb" for excessive reads per block between ERASE operations
  - SLC – 1,000,000 READ cycles
  - MLC – 100,000 READ cycles
- If possible, read equally from pages within the block
- If exceeding the "rule of thumb" cycle count, then move the block to another location and erase the original block
- Erase resets the READ DISTURB cycle count
- Use ECC to recover from read disturb errors
Improving Data Retention

- Limit PROGRAM/ERASE cycles in blocks that require long retention
- Limit reads to reduce read disturb

Example:

Retention required (arbitrary time)

- 5 year
- 2 year
- 0.5 year

Block cycles (arbitrary cycles)

- 10 cyc
- 1,000 cyc
- 10,000 cyc

Infrequently cycled blocks have longer retention
Frequently cycled blocks have shorter retention
Endurance Recommendations

- Always check pass/fail status (SR0) for PROGRAM and ERASE operations
  - Note: READ operations do not set SR0 to fail status
- If fail status after program, move all block data to an available block and mark the failed block bad
Endurance
Recommendations (continued)

- Use ECC to recover from errors
- Write data equally to all good blocks (wear leveling)
- Protect block management/metadata in spare area with ECC
Wear Leveling

- Wear leveling is a plus on SLC devices where blocks can support up to 100,000 PROGRAM/ERASE cycles.
- Wear leveling is imperative on MLC devices where blocks can typically support fewer than 10,000 cycles.
- If you erased and reprogrammed a block every minute, you would exceed the 10,000 cycling limit in just 7 days!
  \[60 \times 24 \times 7 = 10,080\]
- Rather than cycling the same block, wear leveling involves distributing the number of blocks that are cycled.
Wear Leveling (continued)

An 8Gb MLC device contains 4,096 independent blocks.

If we took the previous example and distributed the cycles over all 4,096 blocks, each block would have been programmed fewer than three times (versus the 10,800 cycles when you cycle the same block).

If you provided perfect wear leveling on a 4,096 block device, you could erase and program a block every minute, every day, for 77 years!

\[
\frac{10,000 \times 4,096}{60 \times 24} = \frac{40,960,000}{1,440} = 28,444 \text{ days} = 77.9 \text{ years}
\]
Embedded MMC (eMMC)

- Direct NAND interface will always provide the lowest cost solution
- The complexities of future MLC require increased attention
- ECC algorithm is becoming more complex, moving from 4+ bits to 8+ bits in the future
- A managed interface addresses the complexities of current and future NAND Flash devices
Embedded MMC (eMMC) (continued)

- The host does not need to know the details of NAND Flash block, such as sizes, page sizes, planes, new features, process generation, MLC vs. SLC, wear leveling, and ECC requirements.

- Embedded MMC (eMMC) is the next logical step in the NAND Flash evolution for embedded applications because it turns a program/erase/read device with bad blocks and bad bits (NAND Flash) into a simple write/read memory.
The Need for a Managed Solution

- NAND page size, number of planes, and block size are technology dependent.
- ECC and number of partial page program operations are technology and vendor dependent.
- Commands and interface inconsistencies between vendors.
What is eMMC?

- MLC NAND + MMC 4.2 version controller device
- High-speed solution
  - Host selectable x1, x4, and x8 I/Os
  - 52 MHz clock speed (MAX)
- Backward-compatible with previous MMC systems
- Handles ECC, wear leveling, and block management

12 x 16 x 1.3mm BGA package
What is eMMC?
Conclusions

- NAND Flash is the lowest cost, nonvolatile memory available today
- Major applications are SSD and mobile devices
- Complexities of MLC NAND require increased hardware and software design
- For embedded applications, all these complexities are addressed using the controller included with eMMC