NAND Flash Memory

Jin-Soo Kim (jinsookim@skku.edu)
Computer Systems Laboratory
Sungkyunkwan University
http://csl.skku.edu
Flash Memory
Modern Storage Media

Timeline: Data Backup Storage

- **Blu-ray disk & HD-DVD**: The timeline illustrates the evolution of storage strategies for backups. It shows when this or that storage device or technology began/stopped to be used for backup purposes.
- **Flash Drive**:
- **DVD**: Optical
- **CD - R/RW**: Optical
- **Floppy Disk**: Optical
- **Hard Disk**: Optical
- **Magnetic Tape**: Magnetic
- **Punch Card**: Mechanical

http://www.backuphistory.com
Storing Data

▪ Change the state of something
▪ Read the current state
▪ Maintain the state without any power (non-volatility)
▪ Better if we can change the state multiple times (overwrite)

▪ Having just two states (0 or 1) is simplest and most reliable
Flash Memory Basics

- Two states based on the presence of electrons

0 = Electrons present
1 = No electrons

- Challenges
  - How to attract or expel electrons?
  - How to find whether there are electrons or not?
  - How to keep electrons without any power?
Flash Memory Cell

- Transistor with floating gate
  - The floating gate is insulated all around with an oxide layer
  - Electrons trapped in the floating gate can remain for up to years
Flash Memory Operations

- **Write (or program)**
  - Apply a high voltage at the CG
  - Trap electrons inside the FG
  - Once programmed, the cell can not be reprogrammed until it is erased

- **Erase**
  - Apply a large voltage in the opposite direction
  - Pull the electrons away from the FG
Flash Memory Operations

- **Read**
  - Electrons in the FG partially cancel the electric field from the CG, increasing the threshold voltage of the cell.
  - A higher voltage must be applied to the CG to make the channel conductive.

![Flash Memory Diagram](image-url)
Flash Memory Characteristics

- **Erase-before-write**
  - Read
  - Write or Program: 1 → 0
  - Erase: 0 → 1

- **Bulk erase**
  - Program unit:
    - NOR: byte or word
    - NAND: page
  - Erase unit: block
NOR Flash

- Random, direct access interface
- Fast random reads
- Slow erase and write
- Mainly for code storage
- Spansion, Micron, Macronix, …
NAND Flash

- I/O mapped access
- Smaller cell size, lower cost
- Better performance for erase and write
- Mainly for data storage
- Samsung, Toshiba, SanDisk, Micron, SK Hynix, …
Semiconductor Memory Hierarchy

- **MOS Memory**
  - **Volatile**
    - Random Access Memory (RAM)
      - Static RAM (SRAM)
        - 1970 by Intel
      - Dynamic RAM (DRAM)
        - 1970 by Intel
    - Dynamic RAM (DRAM)
  - **Non-Volatile**
    - Read Only Memory (ROM)
      - Programmable ROM (PROM)
        - EPROM
          - 1971 by Intel
      - Mask ROM
        - 1970 by Intel
    - Non-Volatile Memory
      - NOR Flash
      - NAND Flash
        - 1984 by Toshiba
Memory Types

**FLASH**
- High-density
- Low-cost
- High-speed
- Low-power
- High reliability

**EPROM**
- Non-volatile
- High-density
- Ultraviolet light for erasure

**EEPROM**
- Non-volatile
- Lower reliability
- Higher cost
- Lowest density
- Electrically byte-erasable

**DRAM**
- High-density
- Low-cost
- High-speed
- High-power

**ROM**
- High-density
- Reliable
- Low-cost
- Suitable for high production with stable code

Source: Intel Corporation.
NAND Flash Memory
Making it Smaller

- Hwang’s law
  - The density of the top-of-the-line flash memory chips will double every 12 months
Density Growth

Source: Samsung Electronics
Cost Trends

Source: IEEE Computer, 2011
Recent Cost Trends

J. Handy, Flash Technology: Annual Update, FMS 2015.
NAND Global Market Share

Source: DRAMeXchange & Statista, 2015.
## NAND Technology by Company

<table>
<thead>
<tr>
<th>Makers</th>
<th>20 nm Class (2D)</th>
<th>10 nm Class (2D)</th>
<th>3D NAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMSUNG</td>
<td><img src="image1.png" alt="27nm, 21nm" /></td>
<td><img src="image2.png" alt="19nm, 16nm" /></td>
<td><img src="image3.png" alt="24L" /></td>
</tr>
<tr>
<td>TOSHIBA SanDisk</td>
<td><img src="image4.png" alt="24nm" /></td>
<td><img src="image5.png" alt="19nm, 15nm" /></td>
<td><img src="image6.png" alt="32L MLC, 32L TLC" /></td>
</tr>
<tr>
<td>Micron</td>
<td><img src="image7.png" alt="25nm, 20nm" /></td>
<td><img src="image8.png" alt="16nm" /></td>
<td></td>
</tr>
<tr>
<td>SK Hynix</td>
<td><img src="image9.png" alt="26nm" /></td>
<td><img src="image10.png" alt="16nm" /></td>
<td></td>
</tr>
</tbody>
</table>

J. Choe, Comparison of 20nm & 10nm-class 2D Planar NAND and 3D V-NAND Architecture, FMS, 2015.
NAND by Applications

Source: Samsung Electronics, 2014.
NAND Flash Architecture
Logical View of NAND Flash

- A collection of blocks
- Each block has a number of pages
- The size of a block or a page depends on the technology (but, it’s getting larger)
NAND Flash Example

- 2Gb NAND flash device organization

Source: Micron Technology, Inc.
Plane

- Each plane has its own page register and cache register
- Pages can be programmed or read at once
- Optional feature: 1, 2, 4, 8, … planes

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
Die / Chip

- Each chip has multiple dies (can be stacked)
- Extra circuits, chip enable signal, ready/busy signal

NAND Flash Types

- **SLC NAND**
  - Single Level Cell
  - 1 bit/cell
- **MLC NAND**
  - Multi Level Cell (misnomer)
  - 2 bits/cell
- **TLC NAND**
  - Triple Level Cell
  - 3 bits/cell
- **3D NAND**

Source: Micron Technology, Inc.
Samsung V-NAND
Characteristics of NAND Flash
Erase Before Write

- In-place update (overwrite) is not allowed
- Pages must be erased before new data is programmed
- The erase unit is much larger than the read/write unit
  - Read/write unit: page (4KB, 8KB, 16KB, …)
  - Erase unit: block (64-512 pages)

- What if there are live pages in the block we wish to erase?
Limited Lifetime

- The number of times NAND flash blocks can reliably programmed and erased (P/E cycle) is limited
  - SLCs: 50,000 ~ 100,000
  - MLCs: 1,500 ~ 5,000
  - eMLCs (Enterprise MLCs): 10,000 ~ 30,000
  - TLCs: < 1,000
- High voltage applied to cell degrades oxide
  - Electrons are trapped in oxide
  - Break down of the oxide structure
- Requires wear leveling
Writing Letters and Erasing Paper

Old technology
Thick Paper

"Because thin paper wears down more easily than thick paper."

Thin Paper
Advanced Semiconductor technology

2000

Flash Endurance

E. Grochowski et al., Future Technology Challenges for NAND Flash and HDD Products, FMS, 2012.
Asymmetric Read/Write Latency

- Reading a page is faster than programming it
- Usually more than 10x
  - e.g. 1ynm MLC\(^1\): Read 45\(\mu\)s, Program 1350\(\mu\)s, Erase 4ms
- Programming a page should go through multiple steps of Program & Verify phases

- As the technology shrinks, read/write latency tends to increase
- MLC and TLC make it even worse

\(^1\) D. Sharma, System Design for Mainstream TLC SSD, FMS, 2014.
MLC Programming

- LSB programmed first
  - Cell cannot move to the lower voltage before erase

Program: “1” (Erase) → “0” (Program)

LSB Program: 1) Erase → Erase, 2) Erase → LSB
MSB Program: 1) Erase → Erase, 2) Erase → PV1, 3) LSB → PV2, 4) LSB → PV3
Paired Pages in MLC/TLC

- One cell represents two or three bits in paired pages
  - LSB: low voltage, fast program, less error
  - MSB: high voltage, slow program, more error

- Performance difference
  - LSB page can be corrupted when MSB page programming is interrupted

L. M. Grupp et al., The Harey Tortoise: Managing Heterogeneous Write Performance in SSDs, USENIX ATC, 2013.
# MLC vs. TLC

<table>
<thead>
<tr>
<th></th>
<th>MLC NAND 1ynm 128Gb</th>
<th>TLC NAND 1ynm 128Gb</th>
<th>TLC NAND (SLC mode)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_R$ (read)</td>
<td>45 μs</td>
<td>80 μs (1st)</td>
<td>50 μs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>105 μs (2nd)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>80 μs (3rd)</td>
<td></td>
</tr>
<tr>
<td>$t_{PROG}$ (program)</td>
<td>1350 μs</td>
<td>550 μs (1st)</td>
<td>350 μs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1700 μs (2nd)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4650 μs (3rd)</td>
<td></td>
</tr>
<tr>
<td>$t_{BERS}$ (erase)</td>
<td>4 ms</td>
<td>10 ms</td>
<td>10 ms</td>
</tr>
</tbody>
</table>

*D. Sharma, System Design for Mainstream TLC SSD, FMS, 2014.*
Bit Errors

- Bits are flipping frequently
- Error Correction Code (ECC) in spare area

### Bits Required in the NAND Flash Spare Area

<table>
<thead>
<tr>
<th>Error Correction Level</th>
<th>Bits Required in the NAND Flash Spare Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hamming</td>
</tr>
<tr>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>N/A</td>
</tr>
<tr>
<td>8</td>
<td>N/A</td>
</tr>
<tr>
<td>9</td>
<td>N/A</td>
</tr>
<tr>
<td>10</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Source: Micron Technology, Inc.
ECC Requirements

- Endurance continues to deteriorate
- Stronger ECCs are required: RS, BCH, LDPC

Sources of Error

- **Write disturbance**
  - When a page is programmed, adjacent calls receive elevated voltage stress

- **Read disturbance**
  - Repeated reading from one page can alter the values stored in other unread pages

- **Retention error**
  - Threshold voltage shifts down due to charge leakage from the floating gate
Bad Blocks

- **Initial bad blocks**
  - Due to production yield constraints and the pressure to keep costs low
  - SLCs: up to 2%
  - MLCs: up to 5%

- **Run-time bad blocks**
  - Read, write, or erase failure
  - Permanent shift in the voltage levels of the cells due to trapped electrons

- Requires run-time bad block management
Page Programming Constraints

▪ **NOP**
  - The number of partial-page programming is limited
  - 1 / sector for most SLCs (4 for 2KB page)
  - 1 / page for most MLCs and TLCs

▪ **Sequential page programming**
  - Pages should be programmed sequentially inside a block
  - For large block SLCs, MLCs, and TLCs

▪ **SLC mode**
  - Possible to use only LSB pages in MLCs and TLCs
  - Faster and more reliable, higher P/E cycles
### Comparisons

<table>
<thead>
<tr>
<th></th>
<th>SLC</th>
<th>MLC</th>
<th>TLC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bits per cell</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><strong>Performance</strong></td>
<td>★★★</td>
<td>★★</td>
<td>★</td>
</tr>
<tr>
<td><strong>Endurance</strong></td>
<td>★★★</td>
<td>★★</td>
<td>★</td>
</tr>
<tr>
<td><strong>Capacity</strong></td>
<td>★</td>
<td>★★</td>
<td>★★★</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td>★★★</td>
<td>★★</td>
<td>★</td>
</tr>
<tr>
<td><strong>Cost / GB</strong></td>
<td>$$$</td>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td><strong>Applications</strong></td>
<td>Enterprise</td>
<td>Enterprise / Consumer</td>
<td>Consumer</td>
</tr>
</tbody>
</table>
Beauty and the Beast

- NAND Flash memory is a beauty
  - Small, light-weight, robust, low-cost, low-power non-volatile device

- NAND Flash memory is a beast
  - Much slower program/erase operations
  - No in-place-update
  - Erase unit > write unit
  - Limited lifetime
  - Bit errors, bad blocks, …

- Software support is essential for performance and reliability!