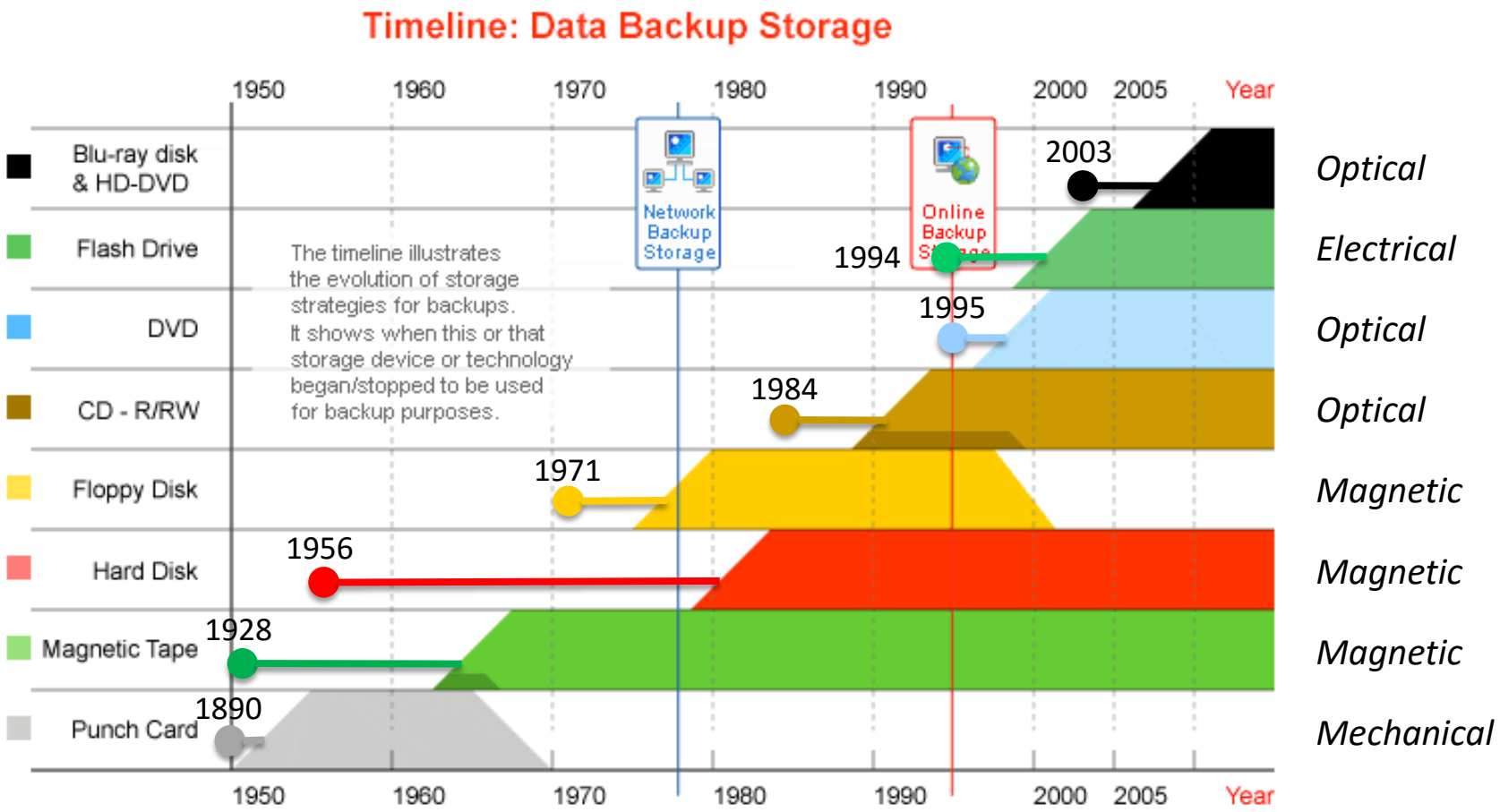
NAND Flash Memory

Jin-Soo Kim (*jinsookim@skku.edu*)
Computer Systems Laboratory
Sungkyunkwan University
http://csl.skku.edu

Flash Memory

Modern Storage Media



http://www.backuphistory.com

Storing Data

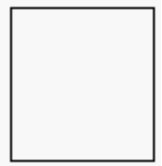
- Change the state of something
- Read the current state
- Maintain the state without any power (non-volatility)
- Better if we can change the state multiple times (overwrite)

- 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

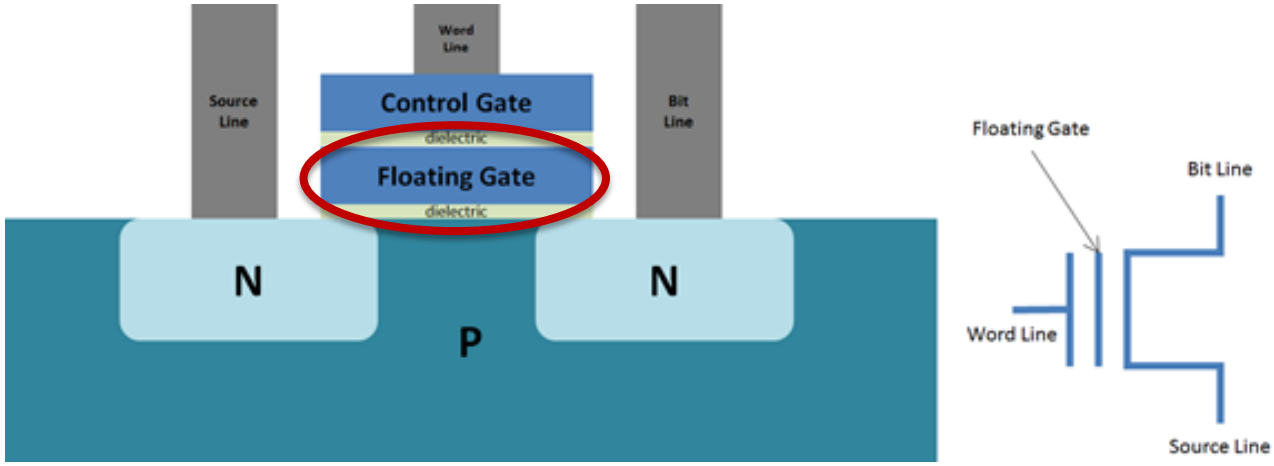


1 = No electrons

- 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

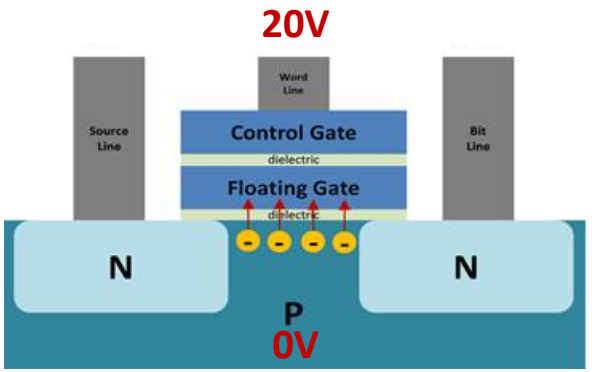
- Transistor with floating gate
  - The floating gate is insulated all around with an oxide layer
  - Electrons trapped in the floating gate can remain for up to years



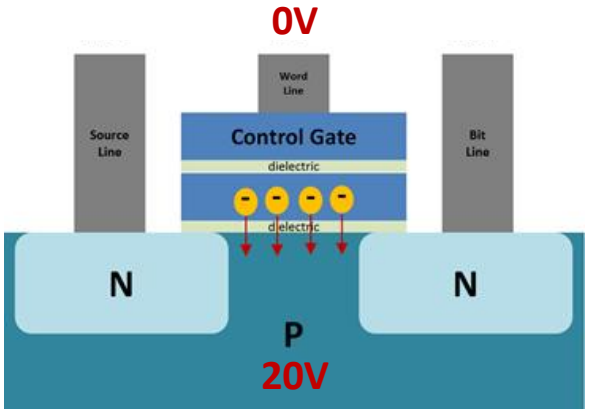
http://www.thenandflash.com

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

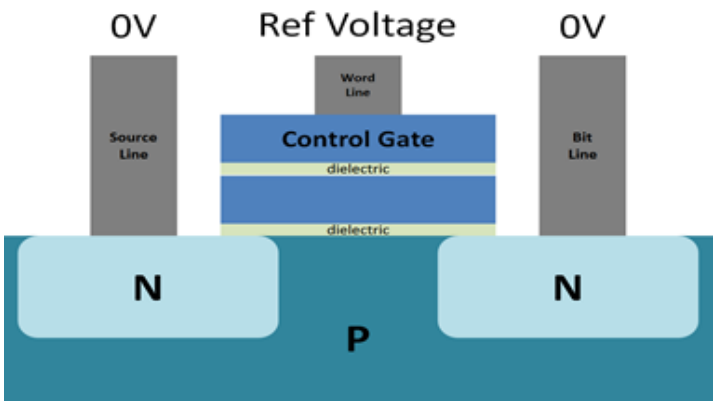


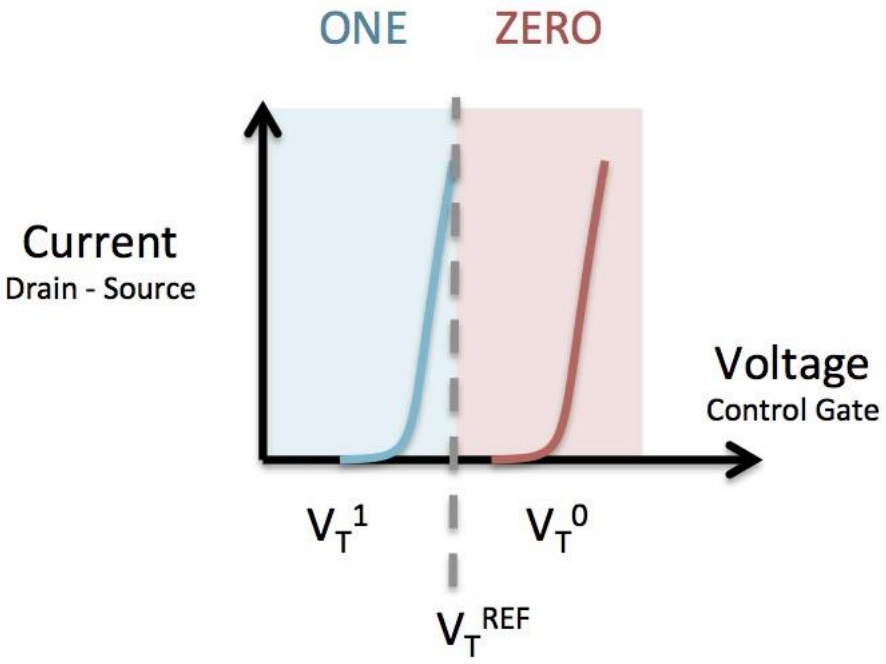
- Erase
  - Apply a large voltage in the opposite direction
  - Pull the electrons away from the FG



Flash Memory Operations

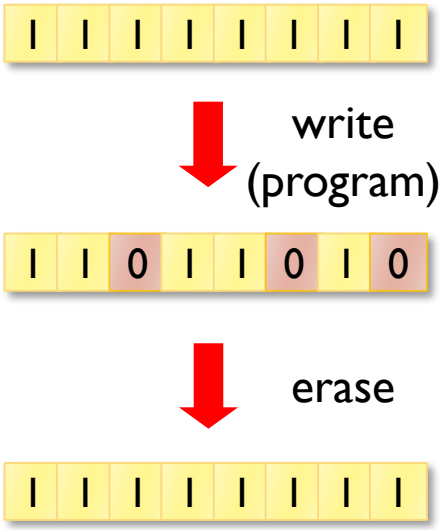
- Read
  - Electrons in the FG partially cancel the electric field from the CG, increasing the threshold voltage of the cell
  - A higher voltage must be applied to the CG to make the channel conductive





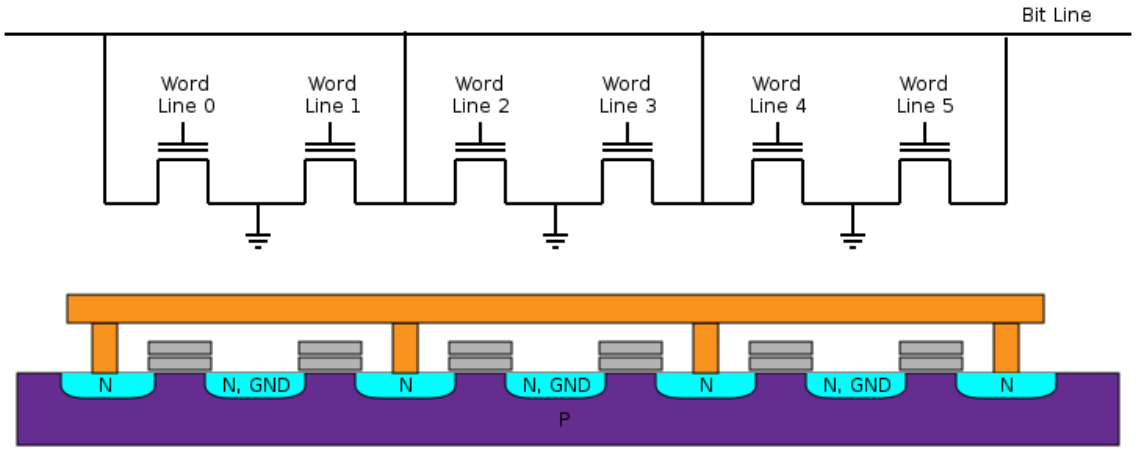
Flash Memory Characteristics

- 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



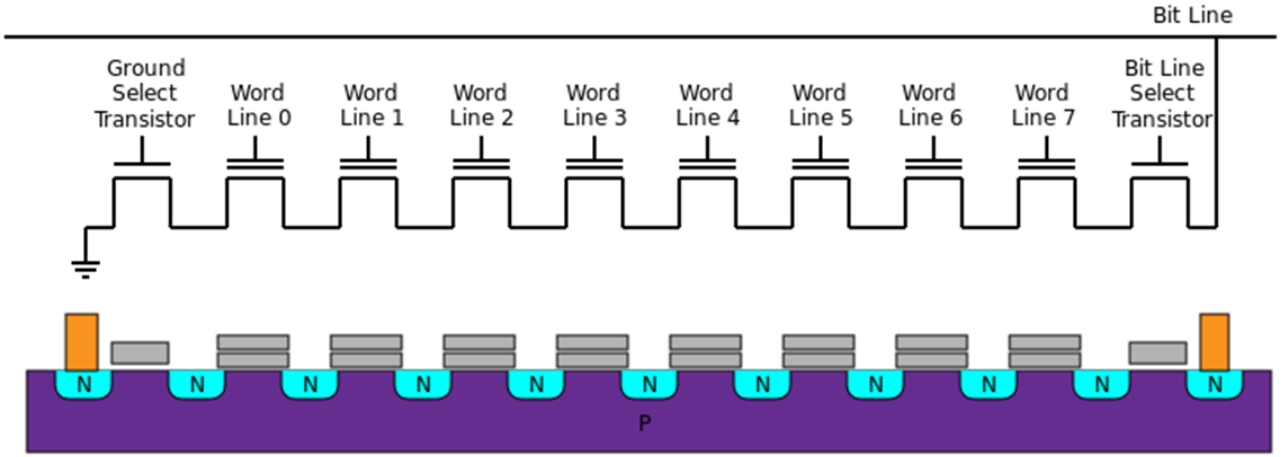
NOR Flash

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

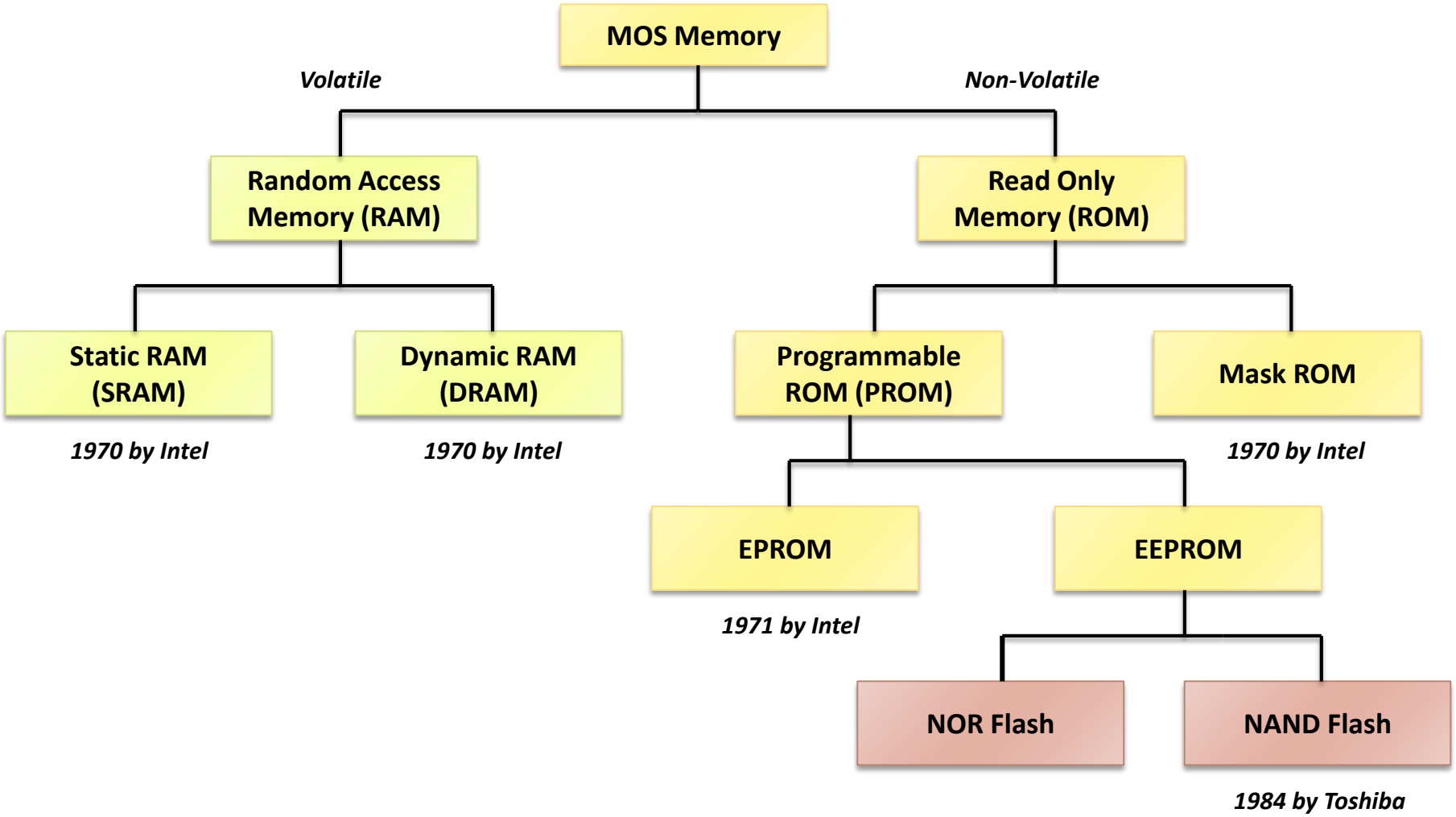


NAND Flash

- 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



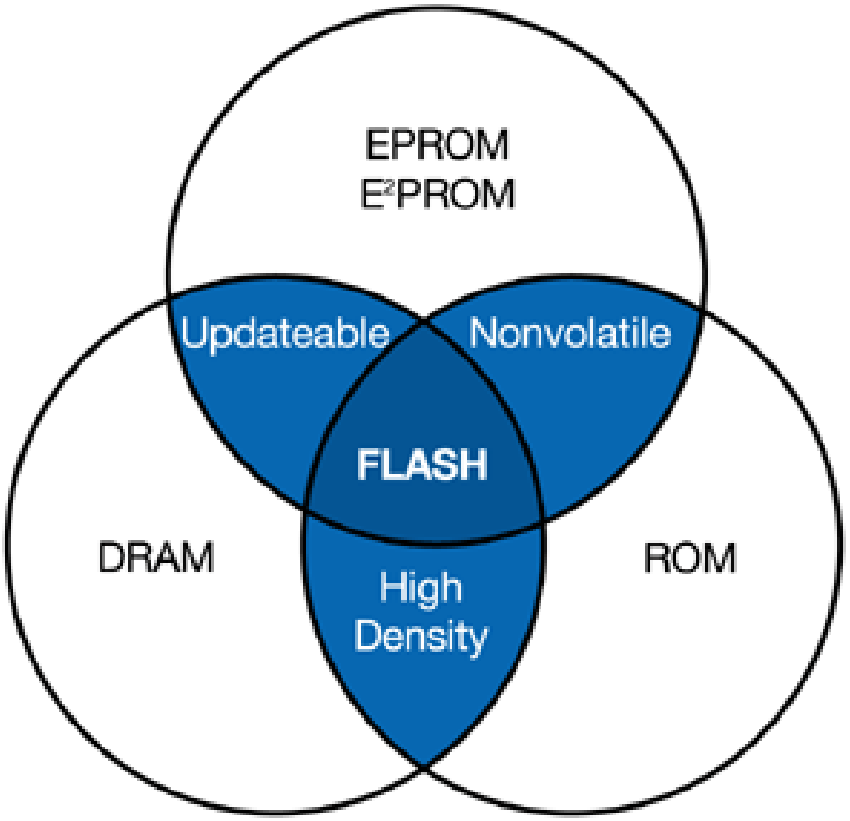
Memory Types

FLASH

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

DRAM

- High-density
- Low-cost
- High-speed
- High-power



EPROM

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

EEPROM

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

ROM

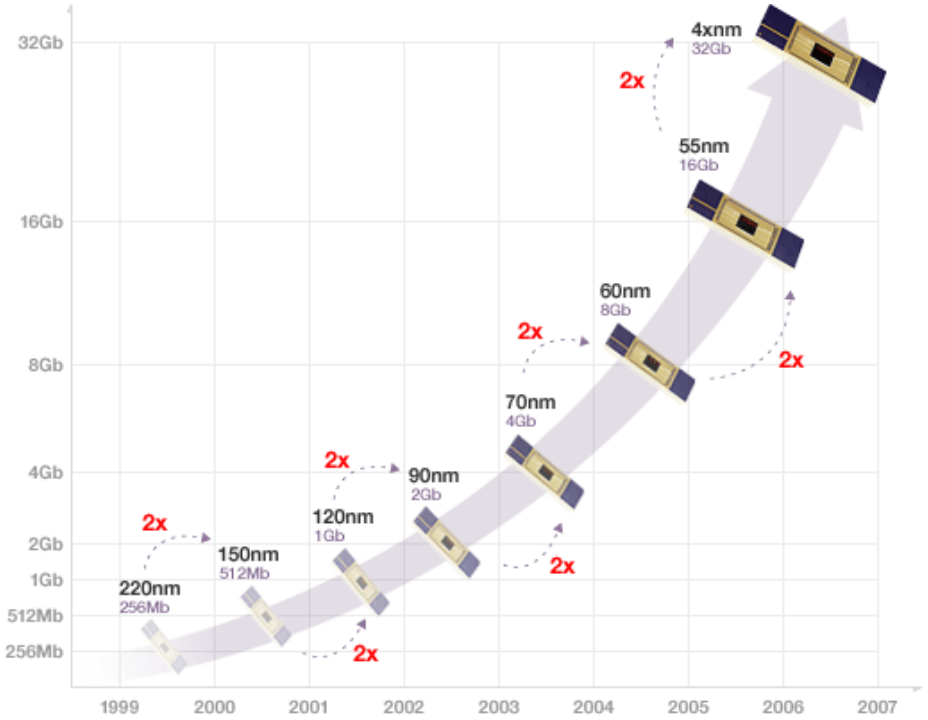
- High-density
- Reliable
- Low-cost
- Suitable for high production with stable code

Source: Intel Corporation.

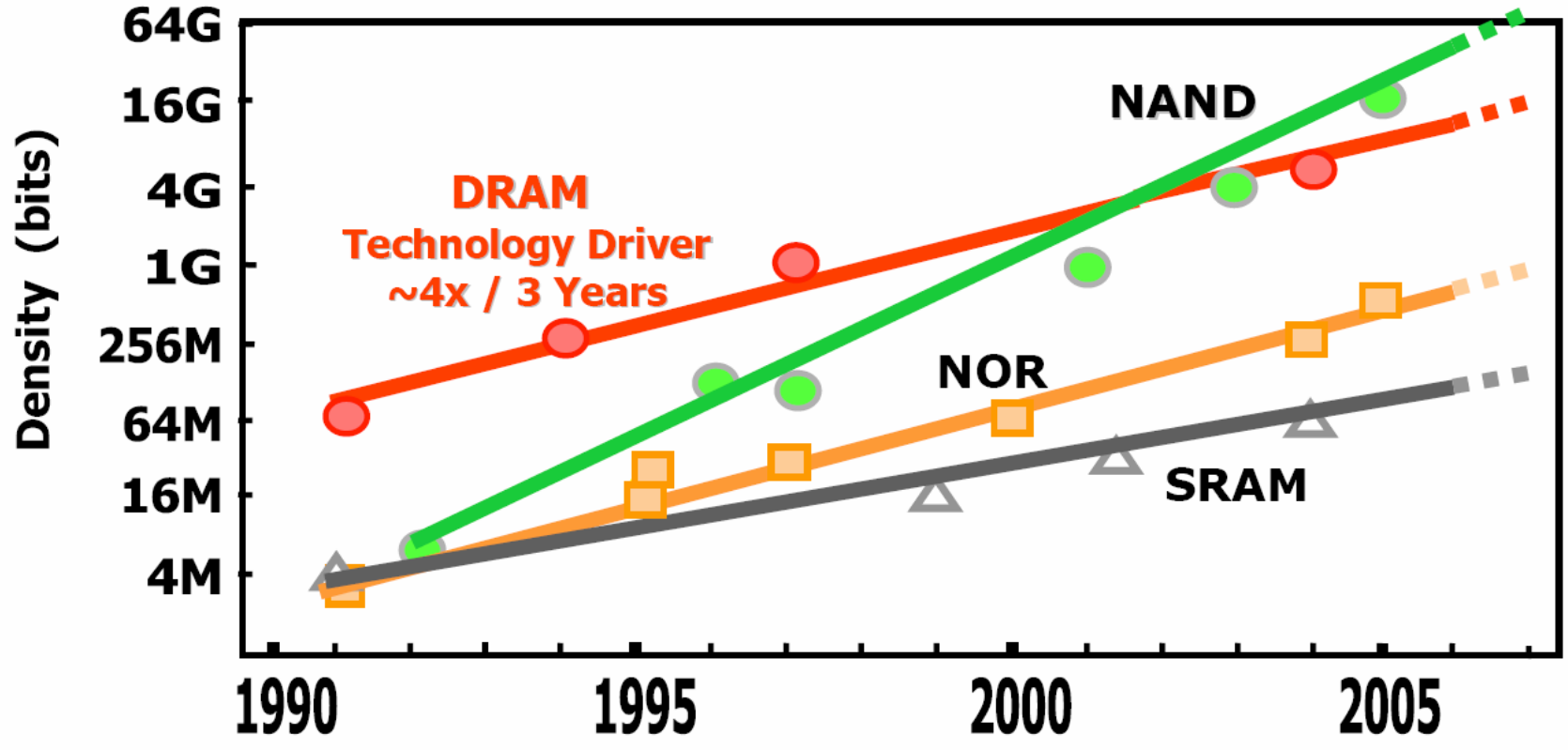
NAND Flash Memory

Making it Smaller

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

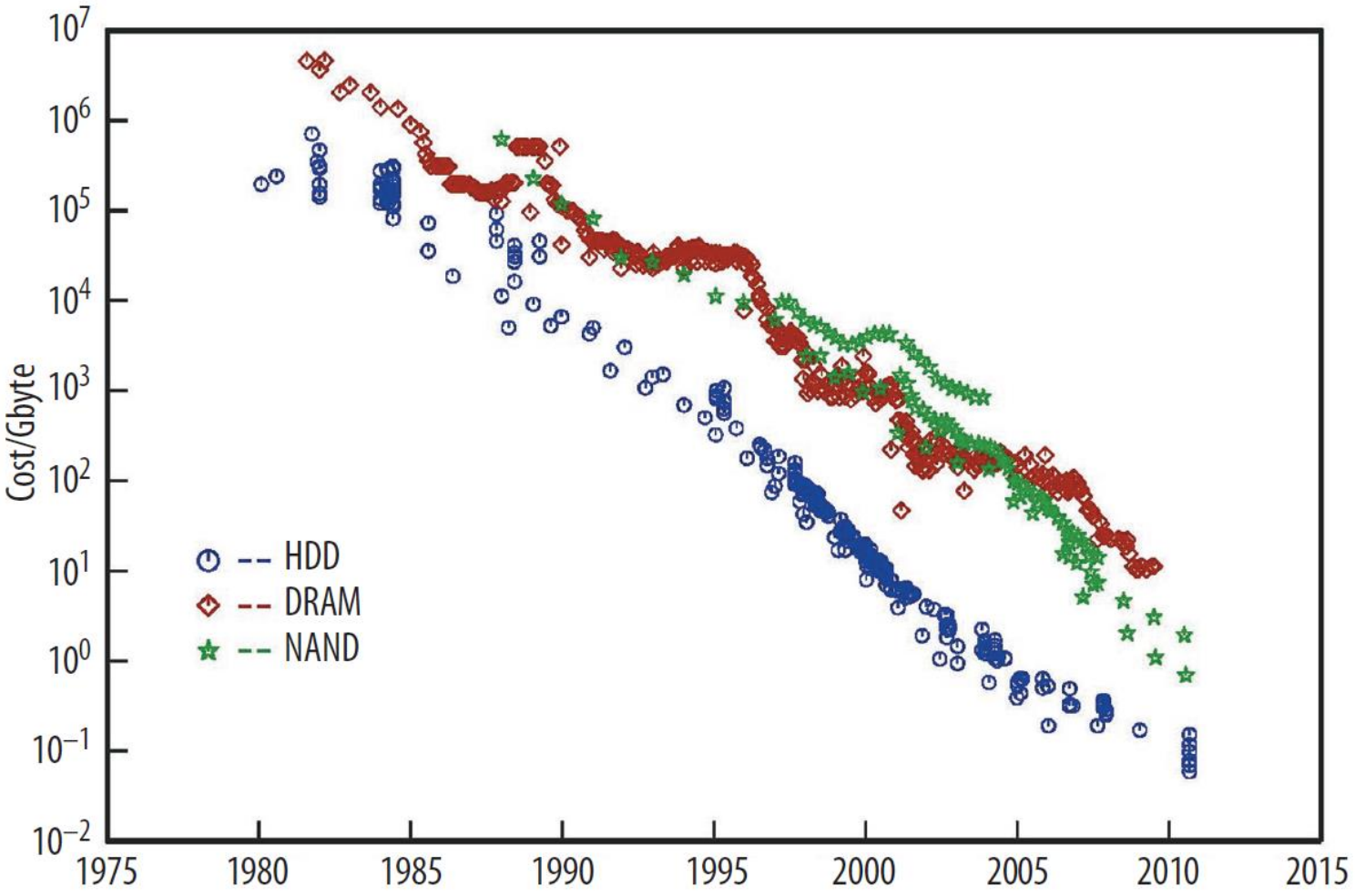


Density Growth



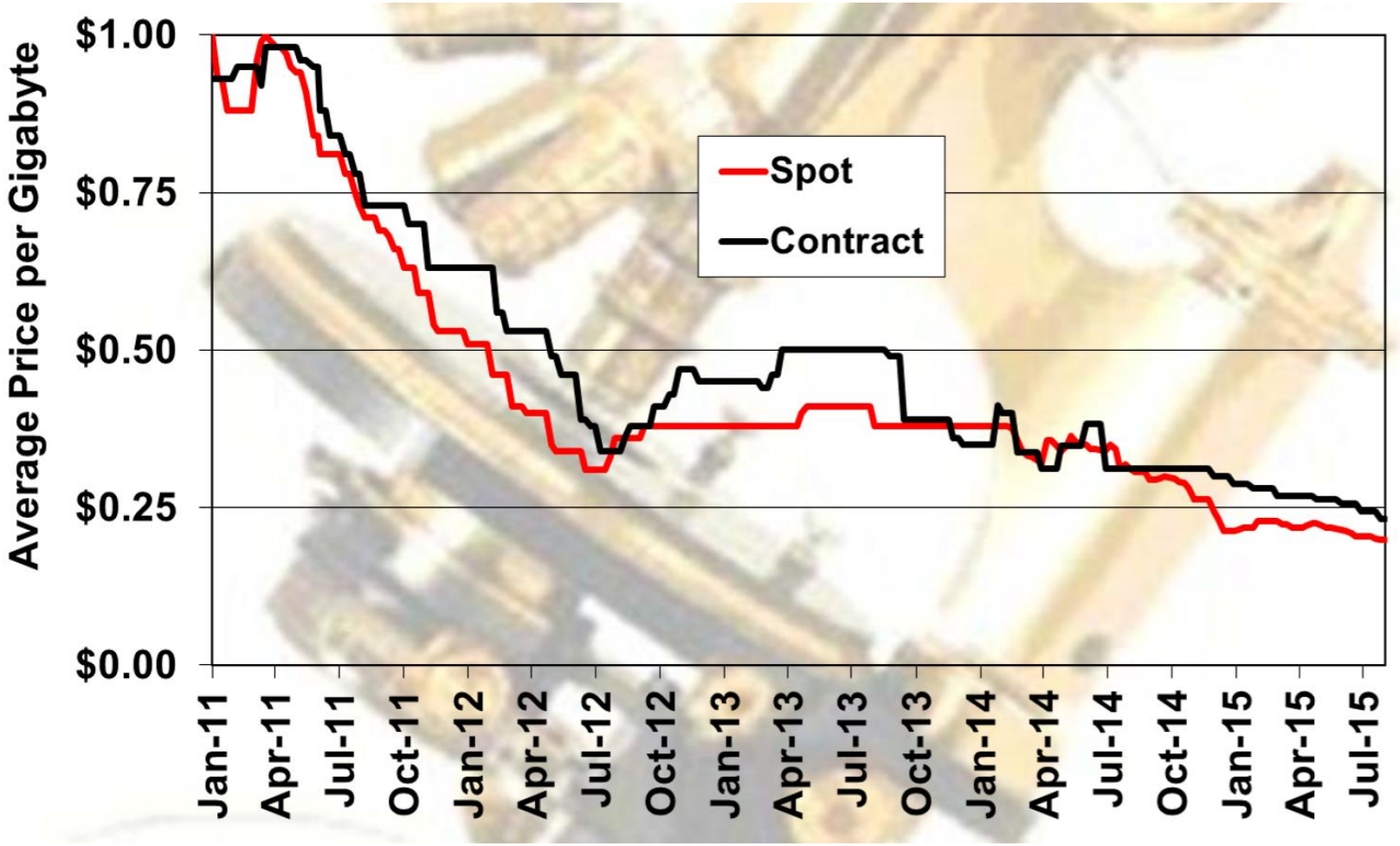
Source: Samsung Electronics

Cost Trends



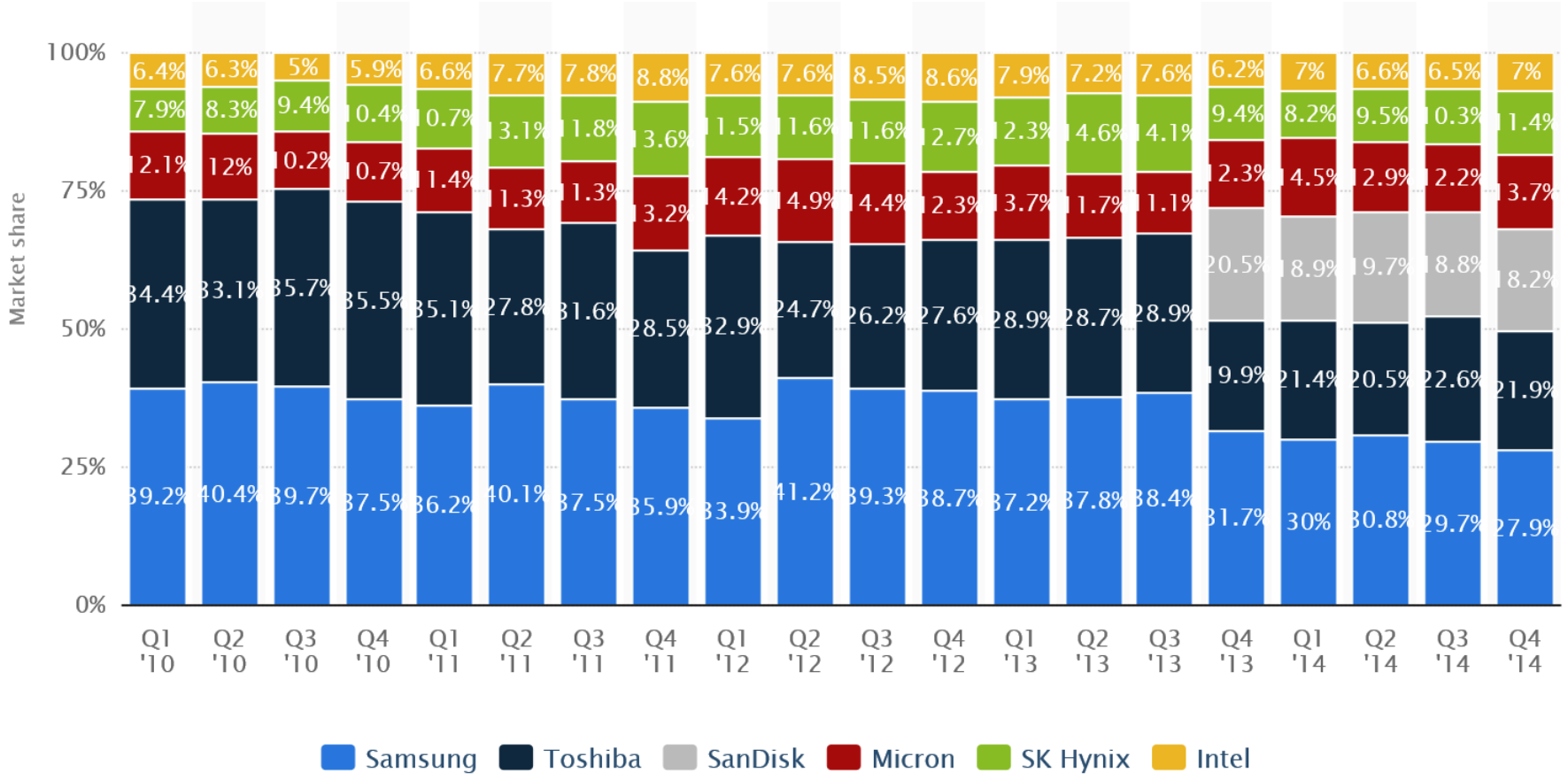
Source: IEEE Computer, 2011

Recent Cost Trends



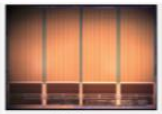
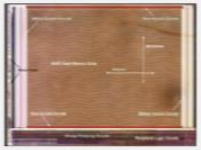
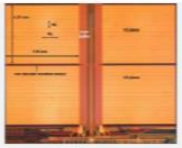
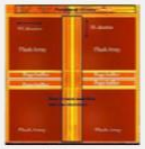
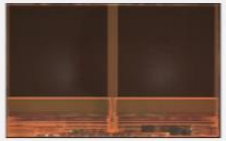
J. Handy, Flash Technology: Annual Update, FMS 2015.

NAND Global Market Share



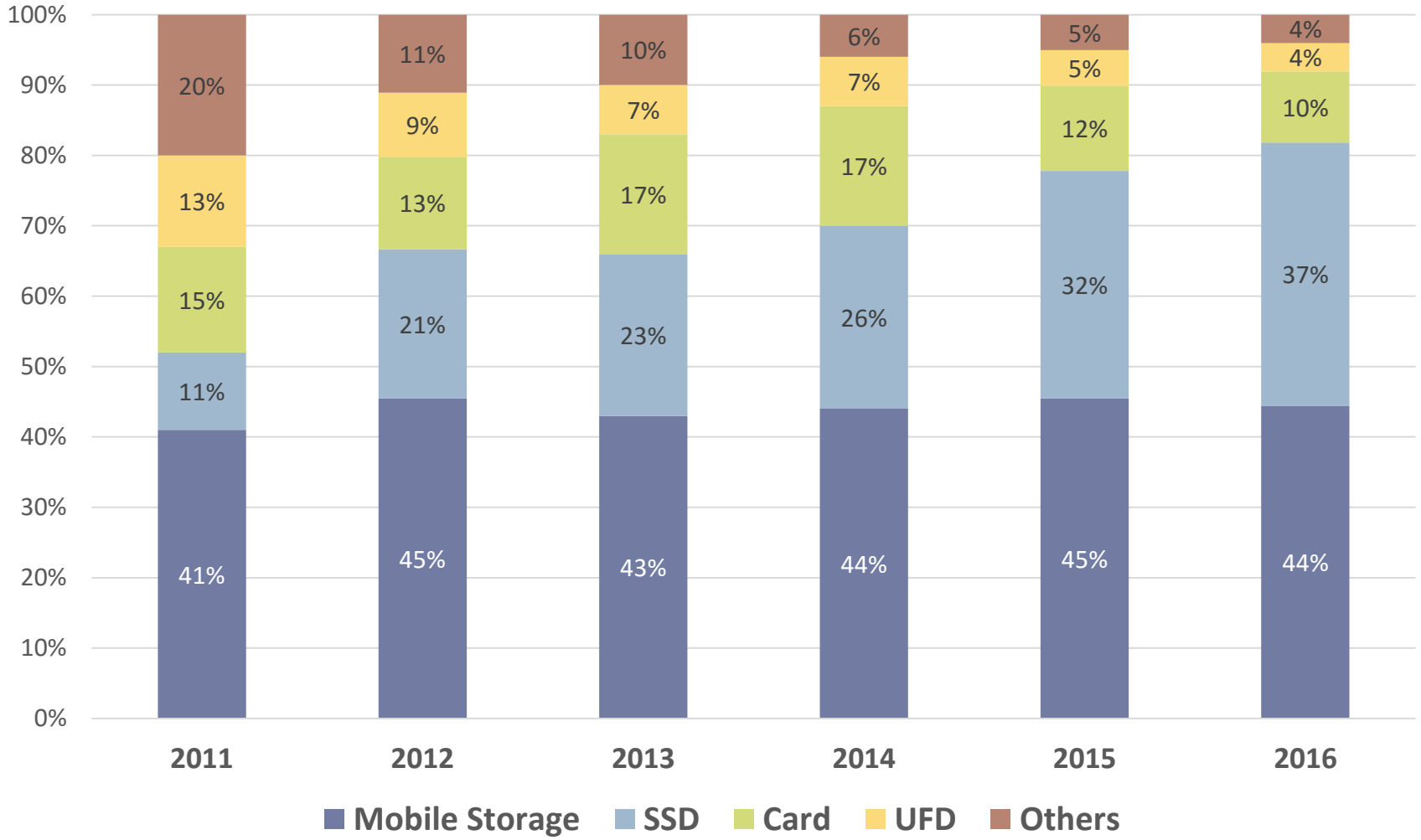
Source: DRAMeXchange & Statista, 2015.

NAND Technology by Company

Makers	20 nm Class (2D)	10 nm Class (2D)	3D NAND
SAMSUNG	 27nm  21nm	 19nm  16nm	 24L  32L MLC  32L TLC
TOSHIBA SanDisk	 24nm	 19nm  A-19nm  15nm	
Micron IM FLASH TECHNOLOGIES	 25nm  20nm  16nm		
SK hynix	 26nm	 16nm	

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

NAND by Applications

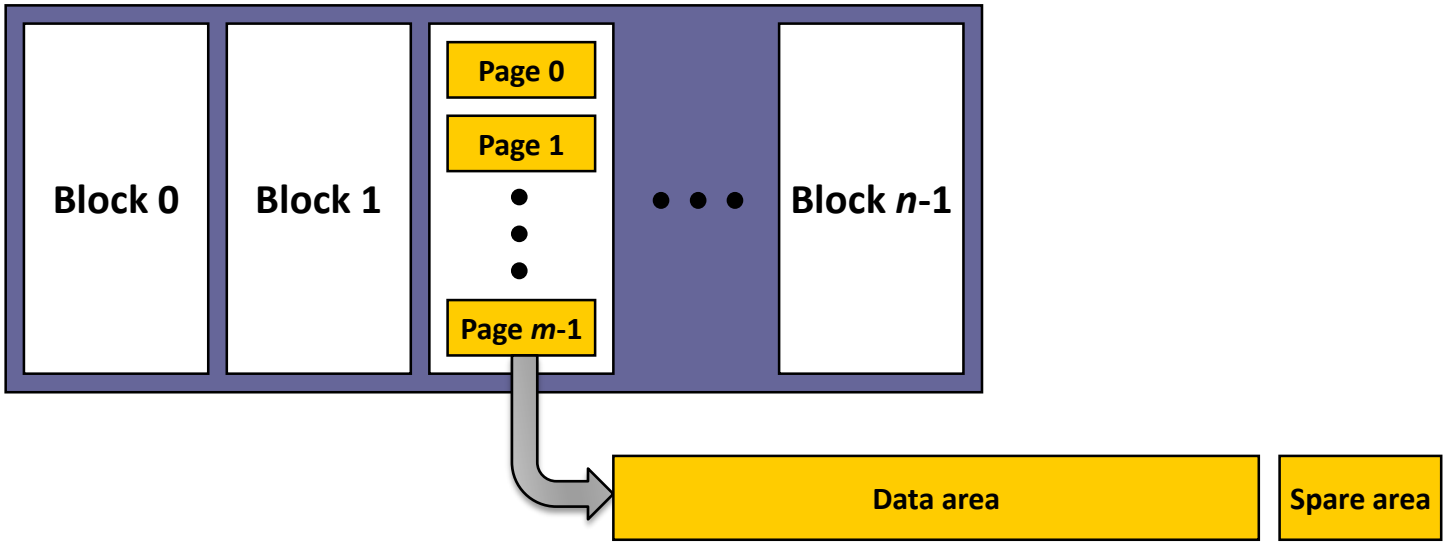


Source: Samsung Electronics, 2014.

NAND Flash Architecture

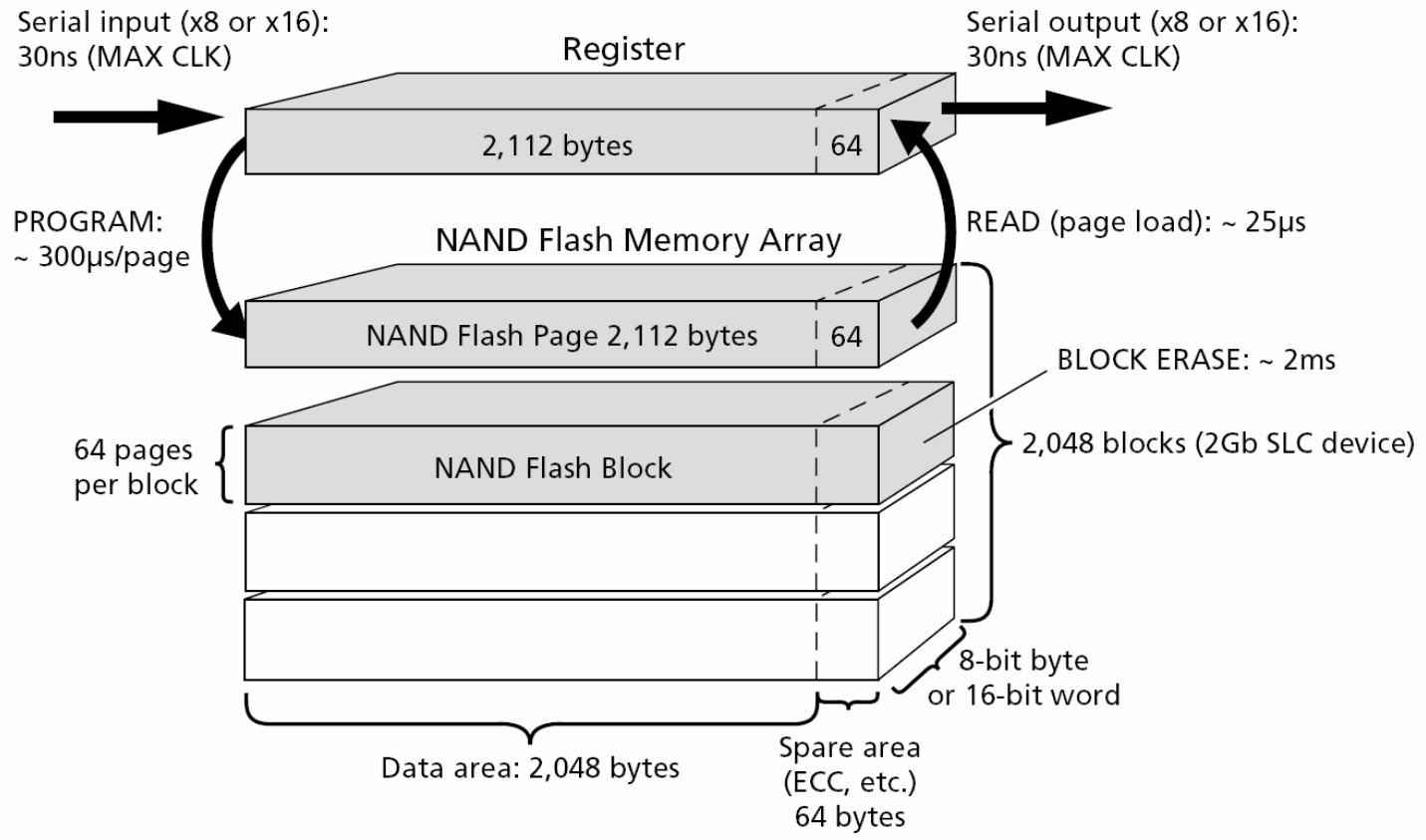
Logical View of NAND Flash

- 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

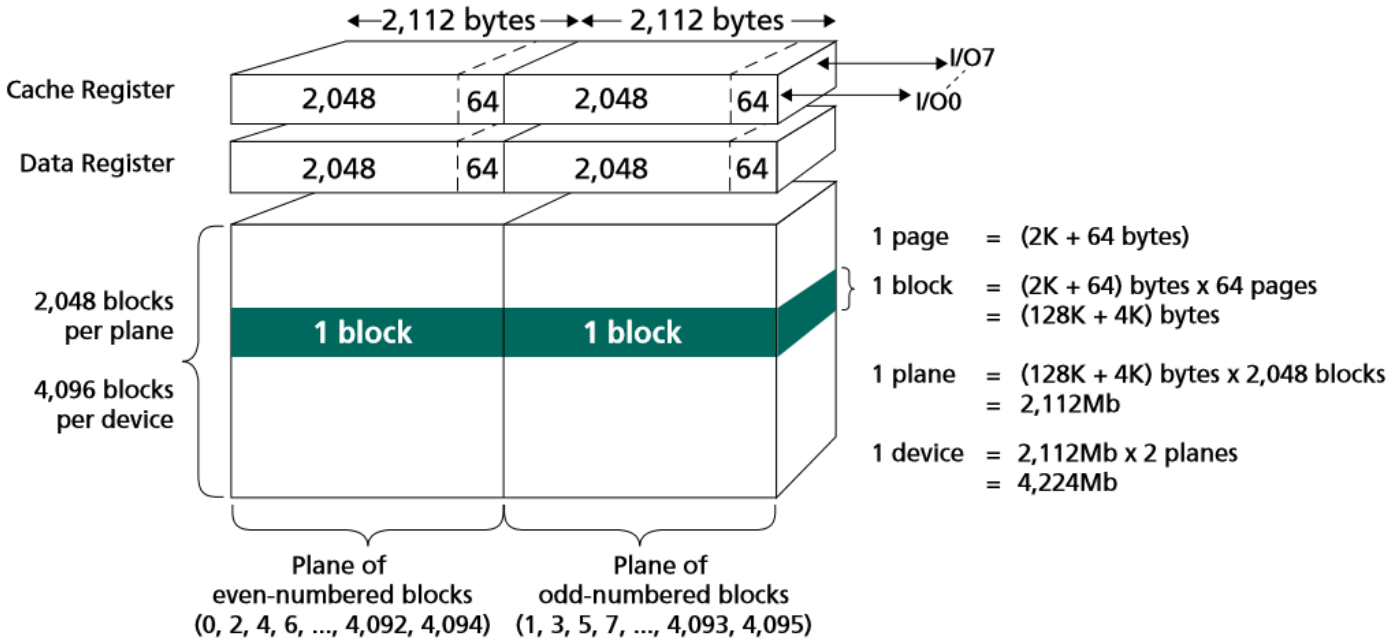
- 2Gb NAND flash device organization



Source: Micron Technology, Inc.

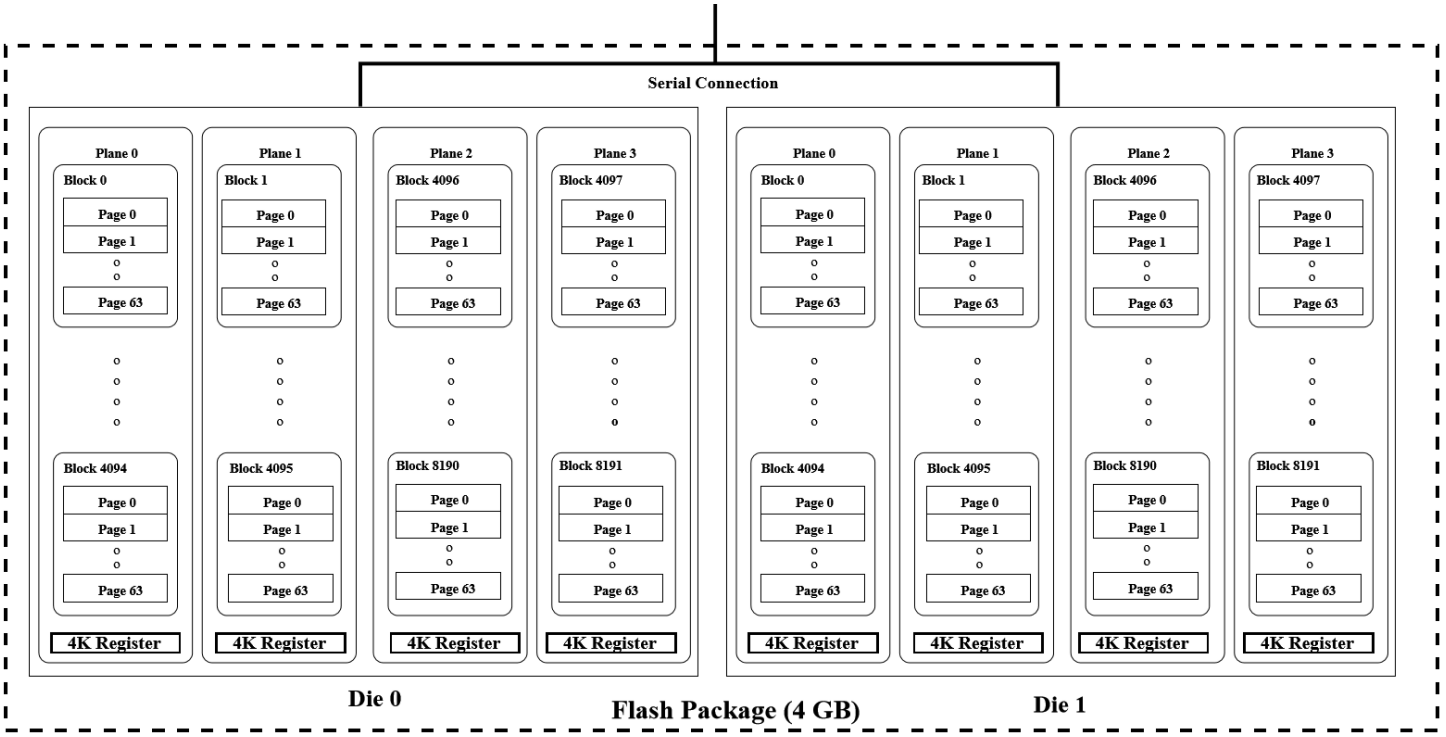
Plane

- 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



Die / Chip

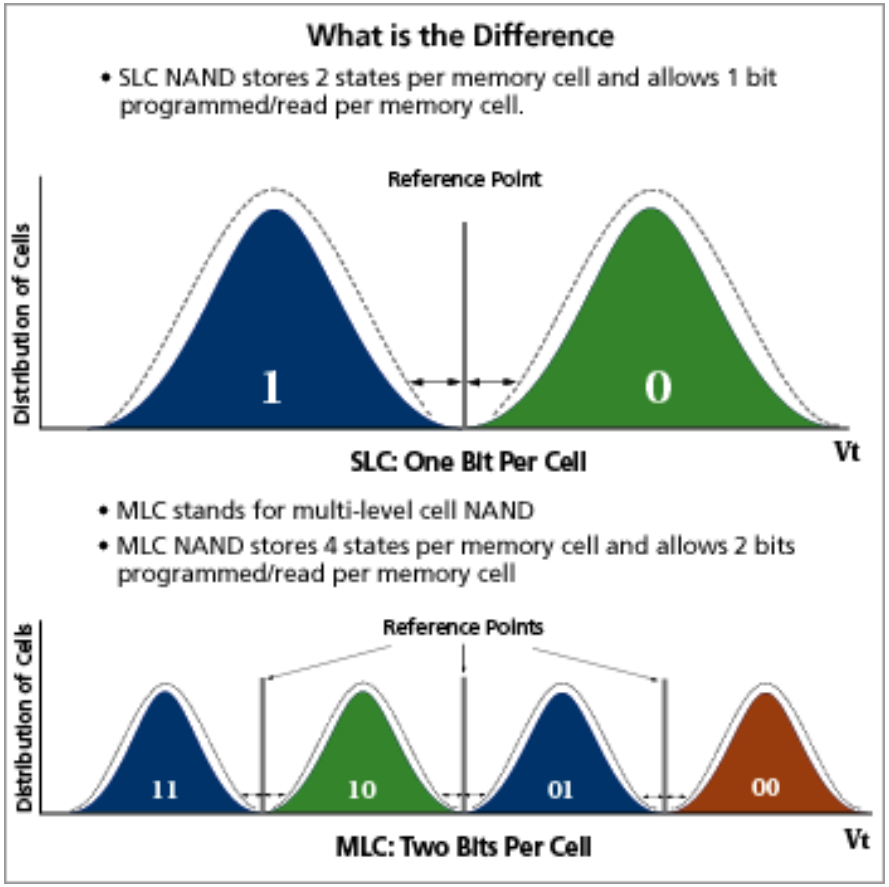
- Each chip has multiple dies (can be stacked)
- + extra circuits, chip enable signal, ready/busy signal



N. Agrawal, Design Tradeoffs for SSD Performance, USENIX ATC, 2008.

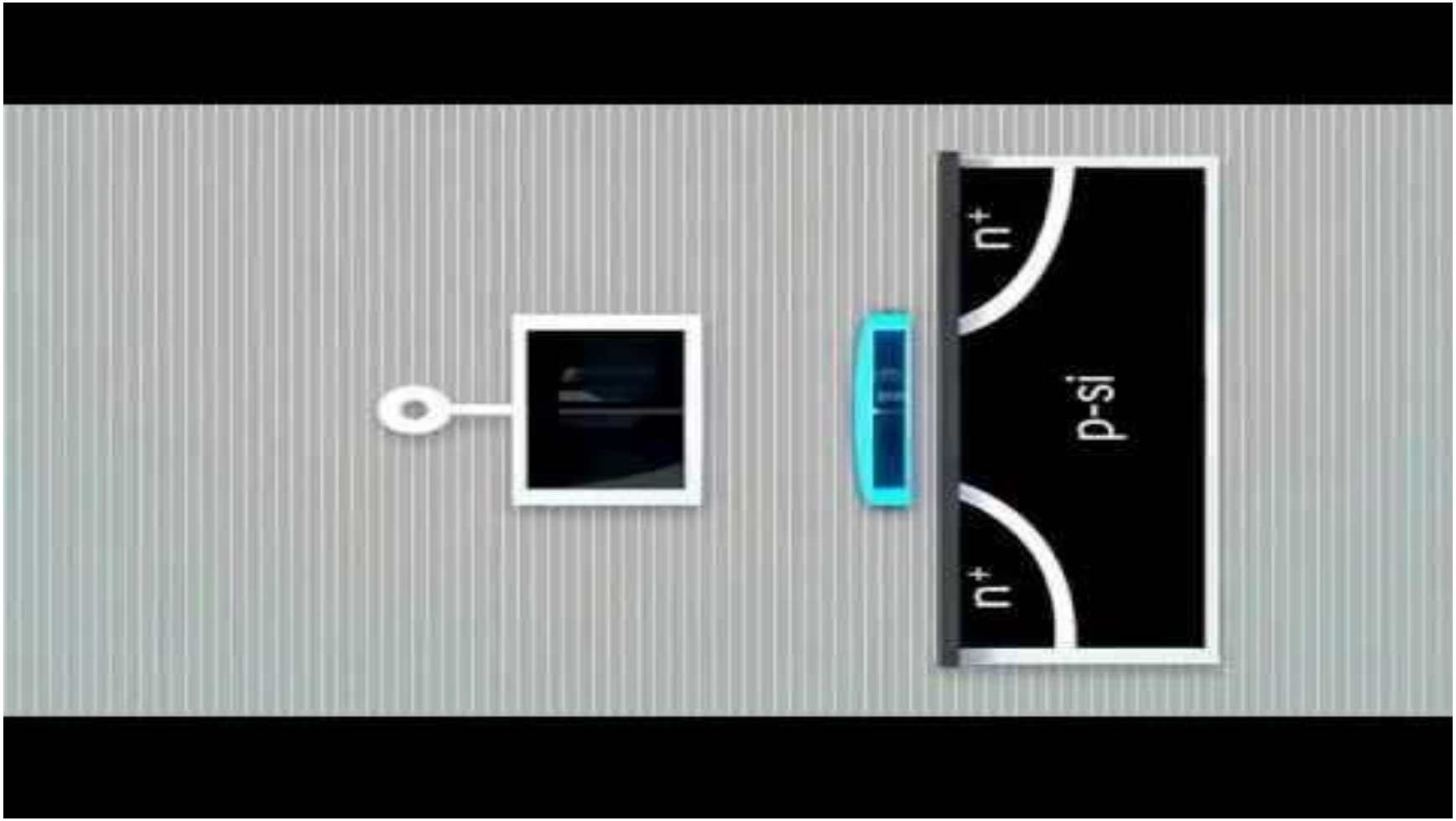
NAND Flash Types

- 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

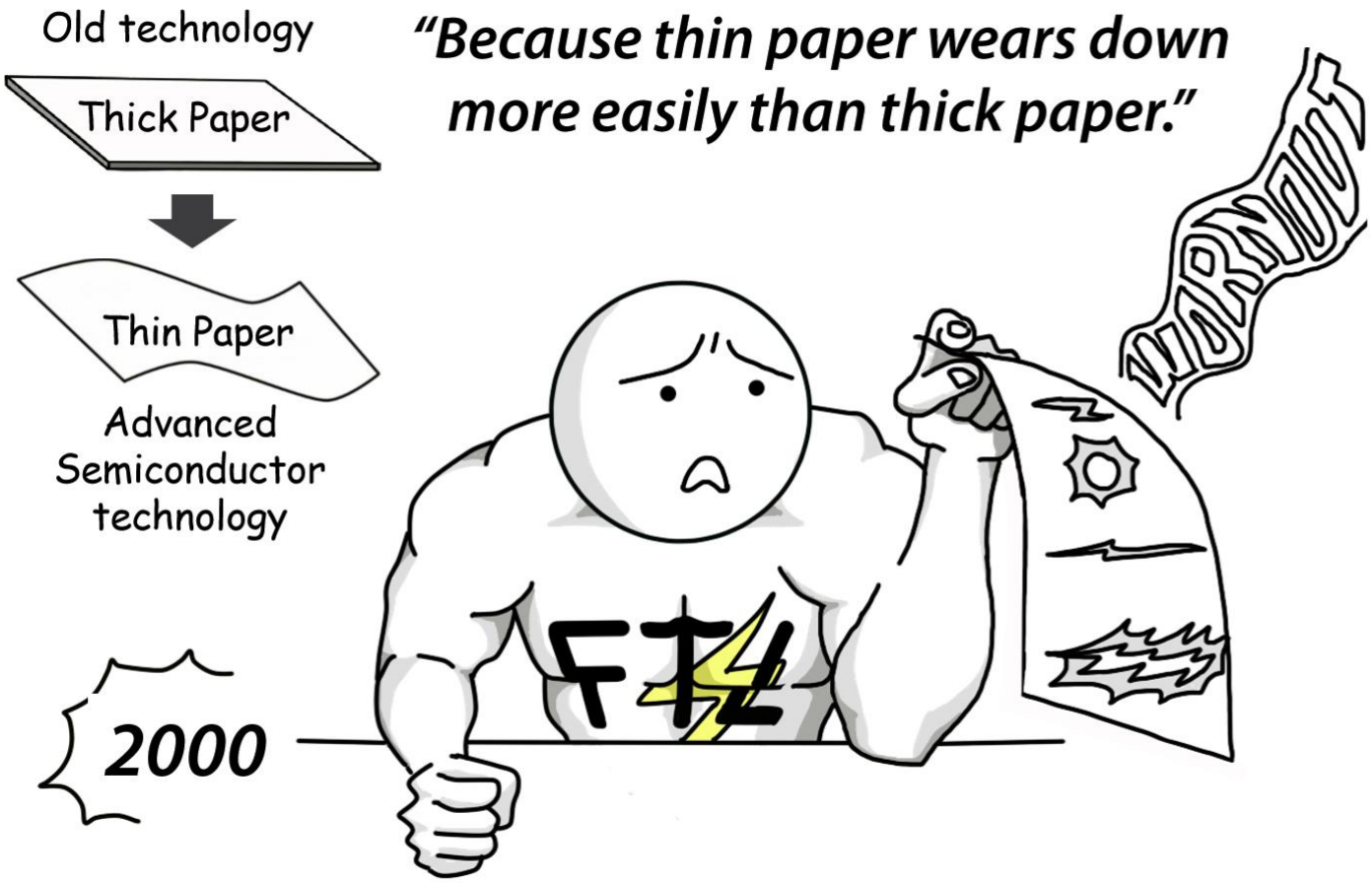
- In-place update (overwrite) is not allowed
- Pages must be erased before new data is programmed
- The erase unit is much larger than the read/write unit
  - Read/write unit: page (4KB, 8KB, 16KB, …)
  - Erase unit: block (64-512 pages)

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

Limited Lifetime

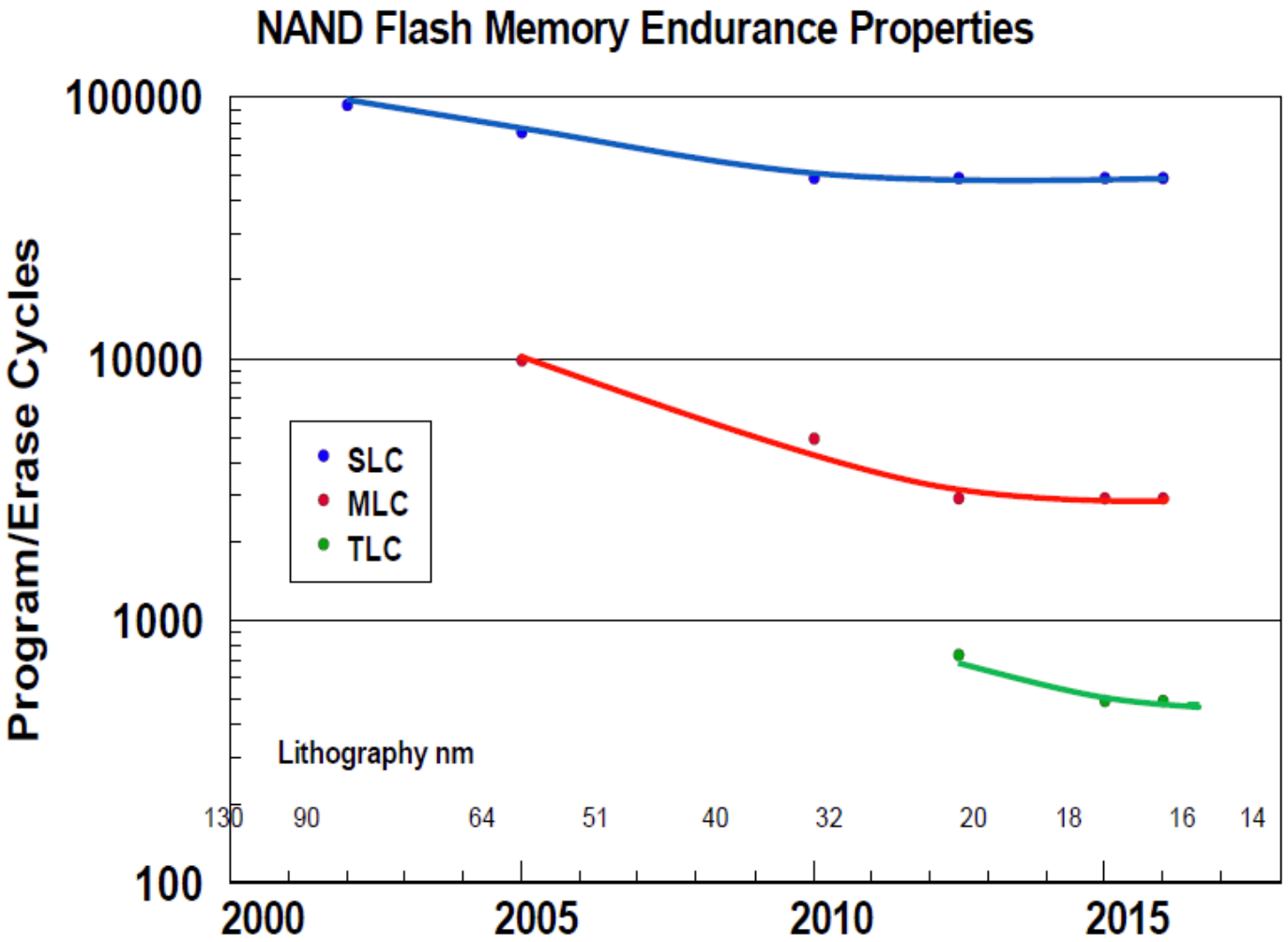
- The number of times NAND flash blocks can reliably programmed and erased (P/E cycle) is limited
  - SLCs: 50,000 ~ 100,000
  - MLCs: 1,500 ~ 5,000
  - eMLCs (Enterprise MLCs): 10,000 ~ 30,000
  - TLCs: < 1,000
- High voltage applied to cell degrades oxide
  - Electrons are trapped in oxide
  - Break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

Writing Letters and Erasing Paper



J. Jeong et al., Lifetime Improvement of NAND Flash-based Storage Systems Using Dynamic Program and Erase Scaling, FAST, 2014.

Flash Endurance



E. Grochowski et al., Future Technology Challenges for NAND Flash and HDD Products, FMS, 2012.

Asymmetric Read/Write Latency

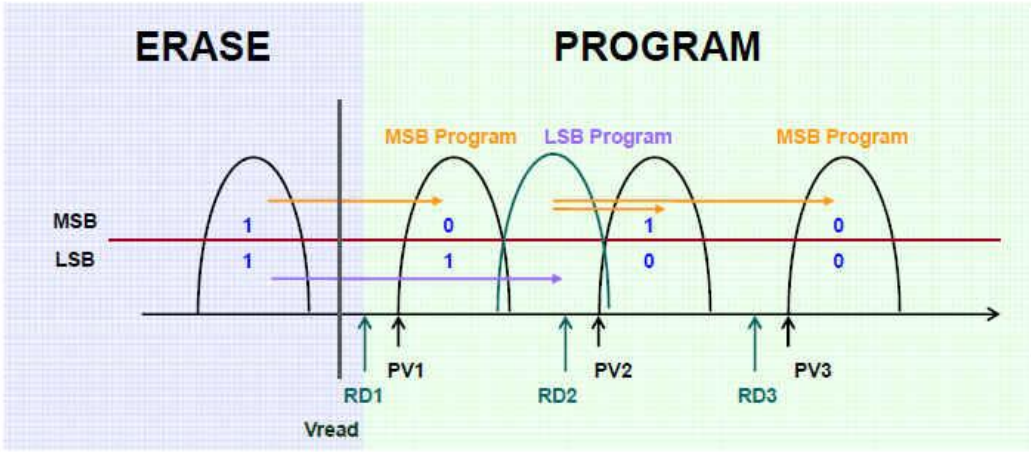
- Reading a page is faster than programming it
- Usually more than 10x
  - e.g. 1ynm MLC[1]: Read 45μs, Program 1350μs, Erase 4ms
- Programming a page should go through multiple steps of Program & Verify phases

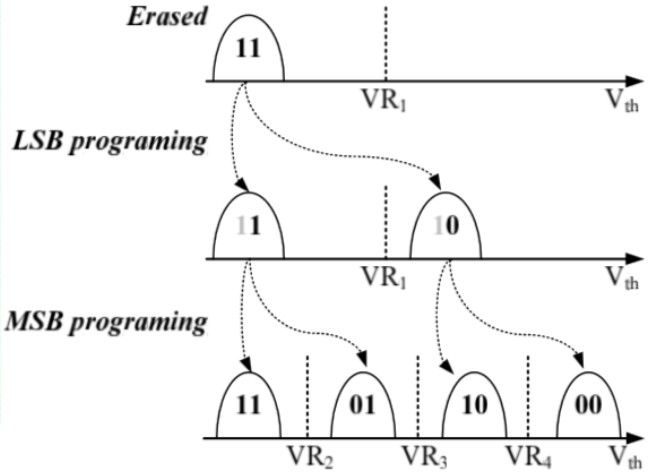
- As the technology shrinks, read/write latency tends to increase
- MLC and TLC make it even worse

[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





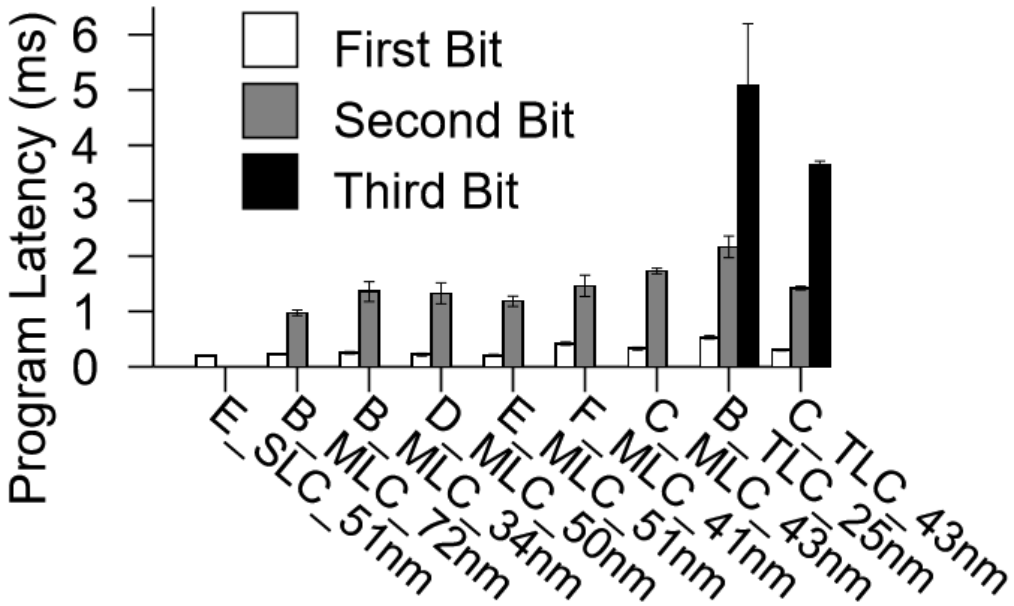
Program : "1"(Erase) → "0"(Program)

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.

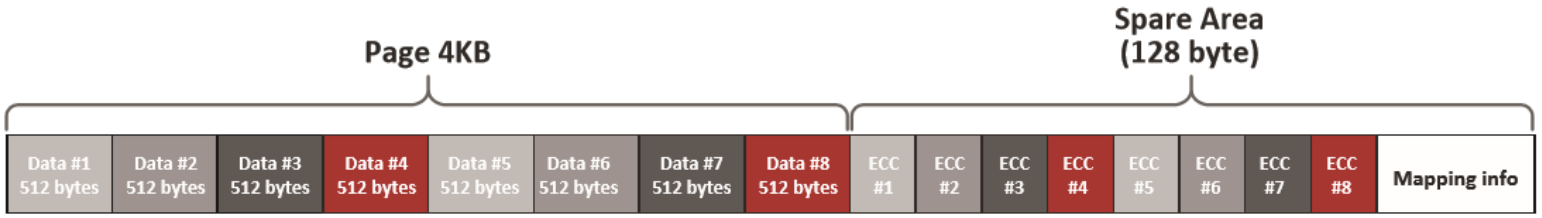
MLC vs. TLC

	MLC NAND 1ynm 128Gb	TLC NAND 1ynm 128Gb	TLC NAND (SLC mode)
t_R (read)	45 µs	80 µs (1st)	50 µs
		105 µs (2nd)	
		80 µs (3rd)	
t_{PROG} (program)	1350 µs	550 µs (1st)	350 µs
		1700 µs (2nd)	
		4650 µs (3rd)	
t_{BERS} (erase)	4 ms	10 ms	10 ms

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

Bit Errors

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

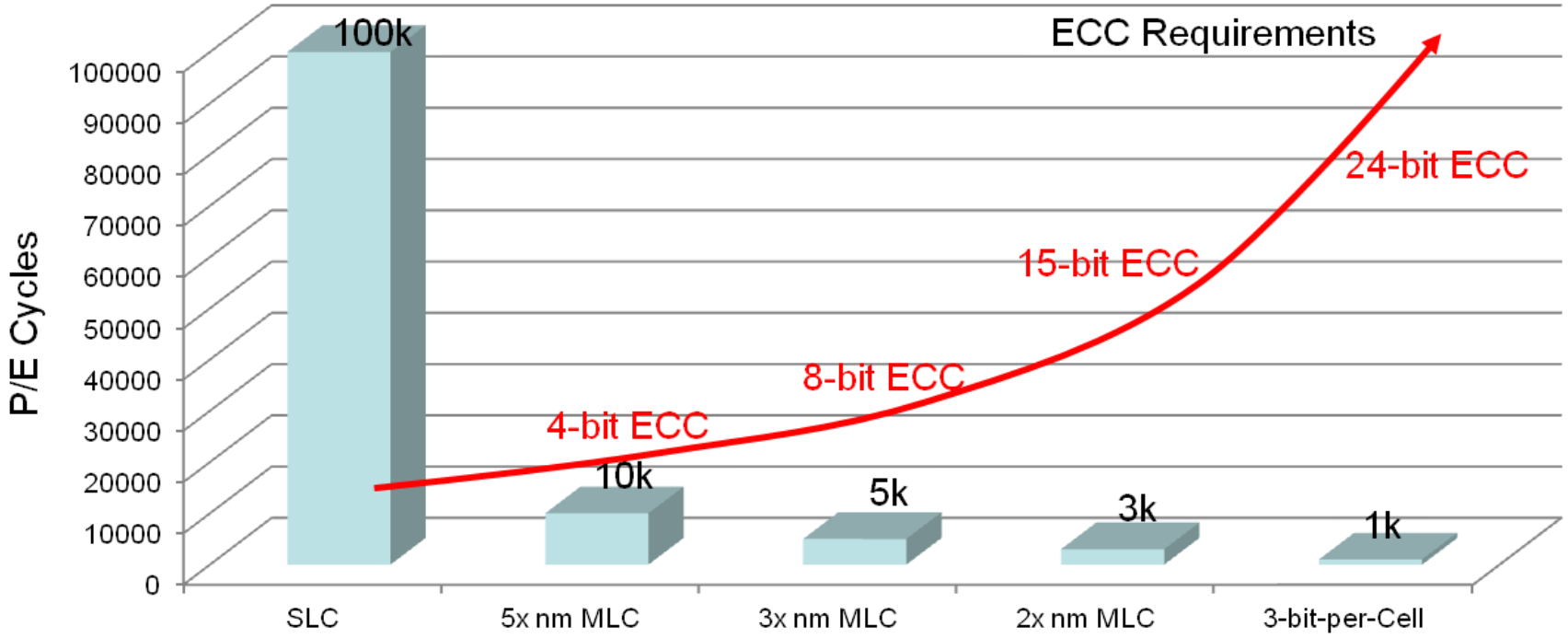


Error Correction Level	Bits Required in the NAND Flash Spare Area		
	Hamming	Reed-Solomon	BCH
1	13	18	13
2	N/A	36	26
3	N/A	54	39
4	N/A	72	52
5	N/A	90	65
6	N/A	108	78
7	N/A	126	91
8	N/A	144	104
9	N/A	162	117
10	N/A	180	130

Source: Micron Technology, Inc.

ECC Requirements

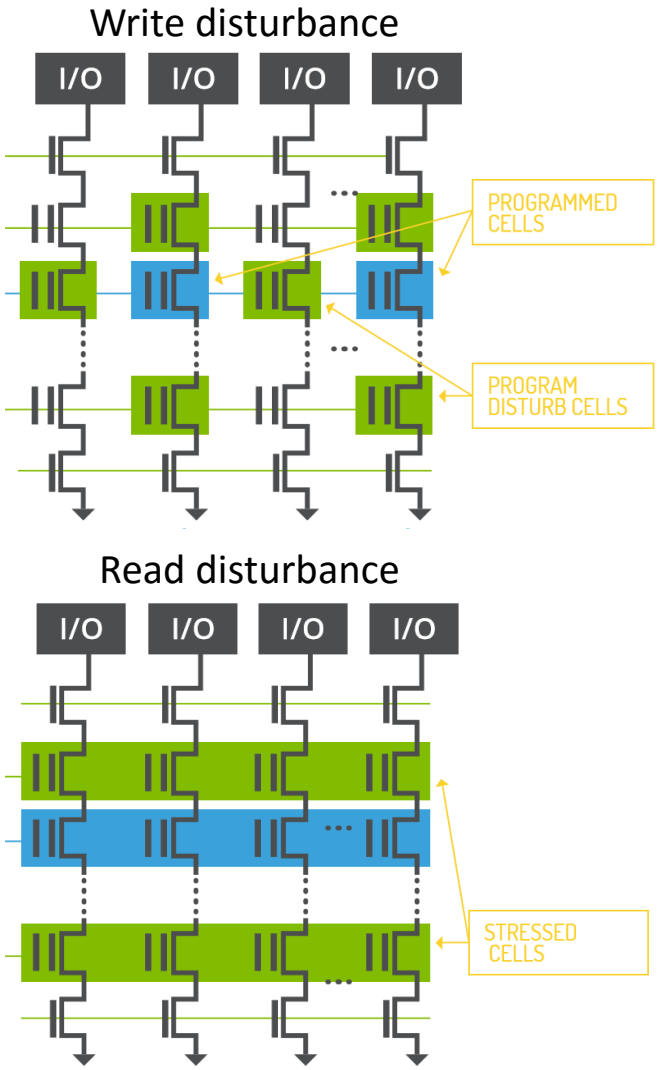
- Endurance continues to deteriorate
- Stronger ECCs are required: RS, BCH, LDPC



Y. Cai et al., Error Patterns in MLC NAND Flash Memory: Measurement, Characterization, and Analysis, DATE, 2012.

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

	SLC	MLC	TLC
Bits per cell	1	2	3
Performance	★★★	★★	★
Endurance	★★★	★★	★
Capacity	★	★★	★★★
Reliability	★★★	★★	★
Cost / GB	$$$	$$	$
Applications	Enterprise	Enterprise / Consumer	Consumer

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!