NAND Flash-based Storage

Jo, Heeseung

Today's Topics

NAND flash memory

Flash Translation Layer (FTL)

OS implications

Flash Memory Characteristics

Flash memory

- Non-volatile, updateable, high-density
- Low cost, low power consumption, high reliability

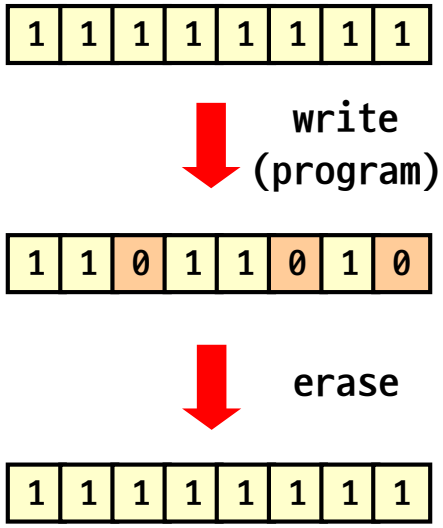
Erase-before-write

- Read
- Write(Program): 1 -> 0
- Erase: 0 -> 1

Read faster than write/erase

Bulk erase

- Erase unit: block
- Program unit: byte or word (NOR), page (NAND)



NOR Flash

NOR flash

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

- Intel, Spansion, STMicro, ...

NAND Flash

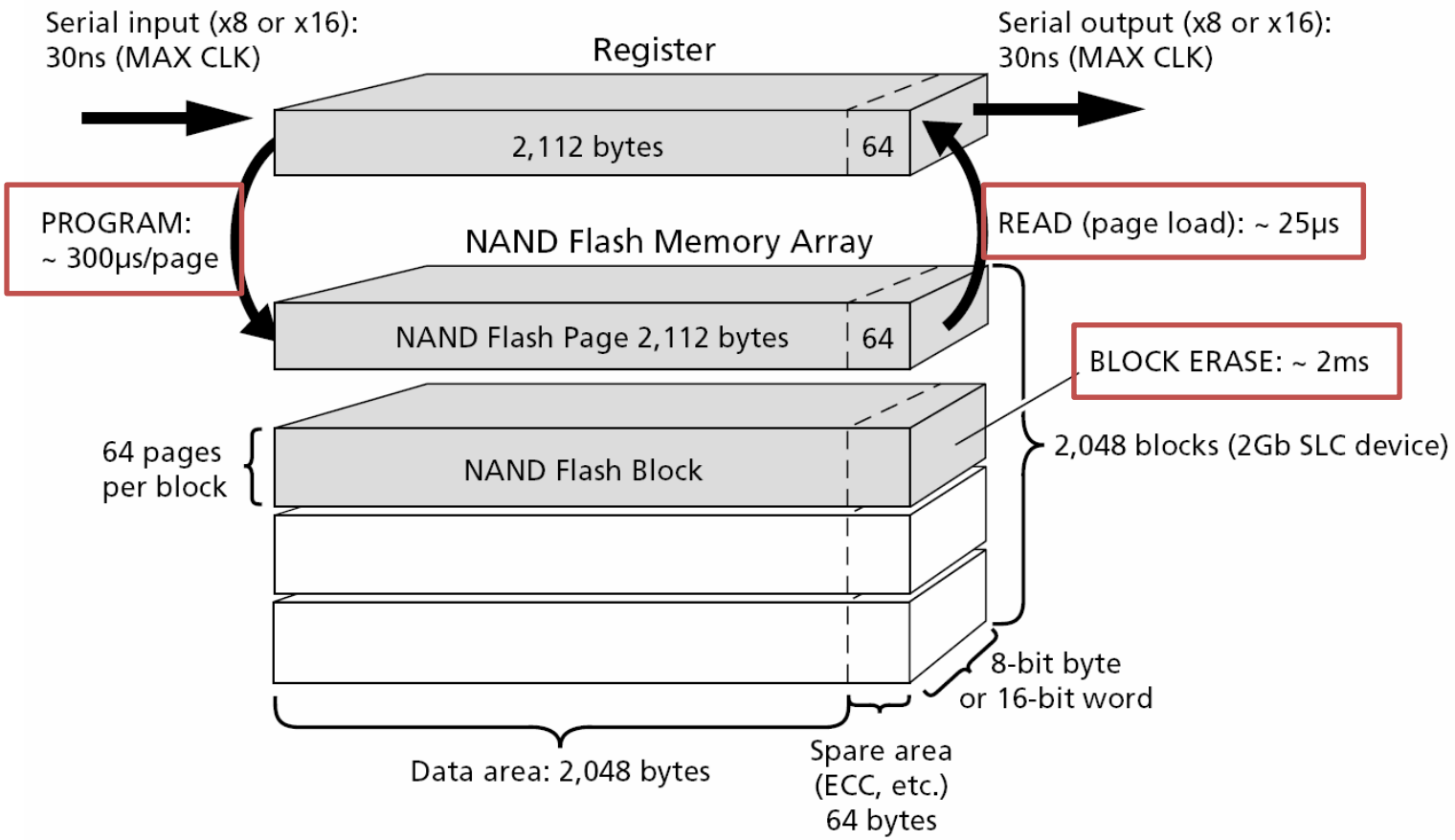
NAND flash

- I/O mapped access
- Smaller cell size
- Lower cost
- Smaller size erase blocks
- Better performance for erase and write(program)
- Mainly for data storage

- Samsung, Toshiba, Hynix, ...

NAND Flash Architecture

2Gb NAND flash device organization



Source: Micron Technology, Inc.

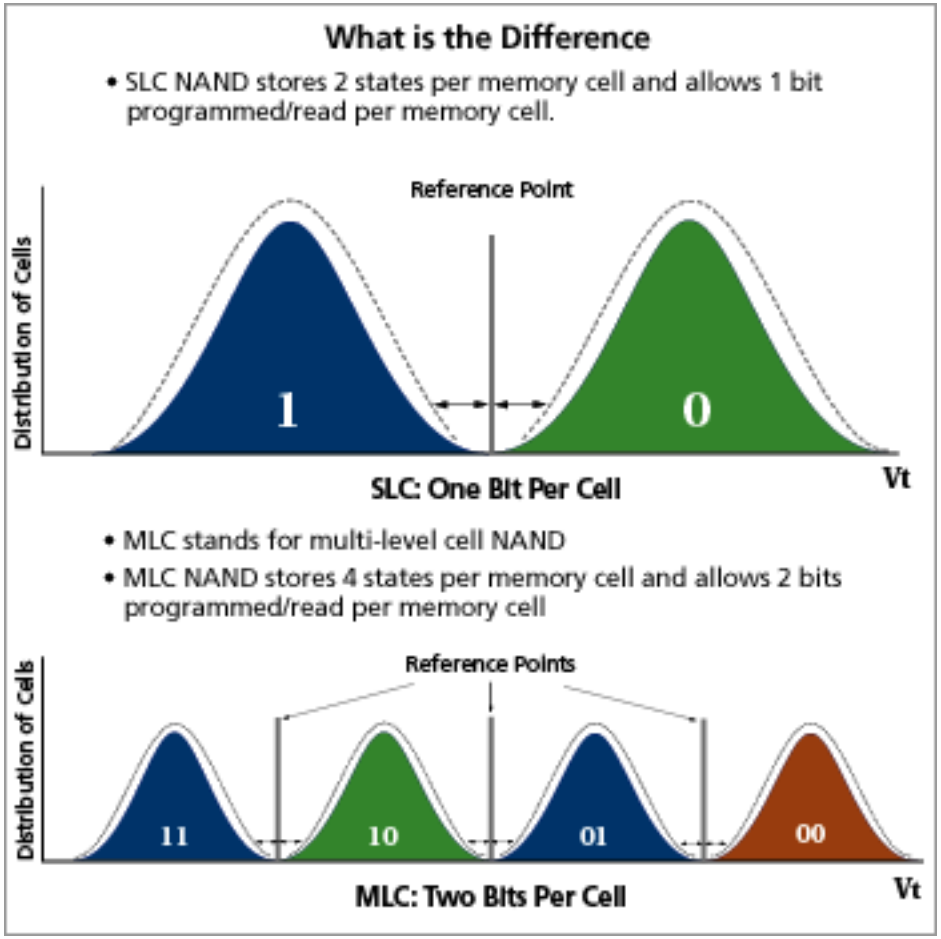
NAND Flash Types (1)

SLC NAND Flash

- Small block (≤ 1Gb)
- Large block (≥ 1Gb)

MLC NAND Flash

TLC NAND Flash



Source: Micron Technology, Inc

NAND Flash Types (2)

	SLC NAND[1] (small block)	SLC NAND[2] (large block)	MLC NAND[3]
Page size (Bytes)	512+16	2,048+64	4,096+128
Pages / Block	32	64	128
Block size	16KB	128KB	512KB
t_R (read)	15 µs (max)	20 µs (max)	50 µs (max)
t_{PROG} (program)	200 µs (typ) 500 µs (max)	200 µs (typ) 700 µs (max)	600 µs (typ) 1,200 µs (max)
t_{BERS} (erase)	2 ms (typ) 3 ms (max)	1.5 ms (typ) 2 ms (max)	3 ms (typ)
NOP	1 (main), 2 (spare)	4	1
Endurance Cycles	100K	100K	10K
ECC (per 512Bytes)	1 bit ECC 2 bits EDC	1 bit ECC 2 bits EDC	4 bits ECC 5 bits EDC

[1] Samsung K9F1208X0C (512Mb) *[2] Samsung K9K8G08U0A (8Gb)* *[3] Micron Technology Inc.*

NAND Applications

Universal Flash Drives (UFDs)

Flash cards

- CompactFlash, MMC, SD, Memory stick, …

Embedded devices

- Cell phones, MP3 players, PMPs, PDAs,
  Digital TVs, Set-top boxes, Car navigators, …

Hybrid HDDs

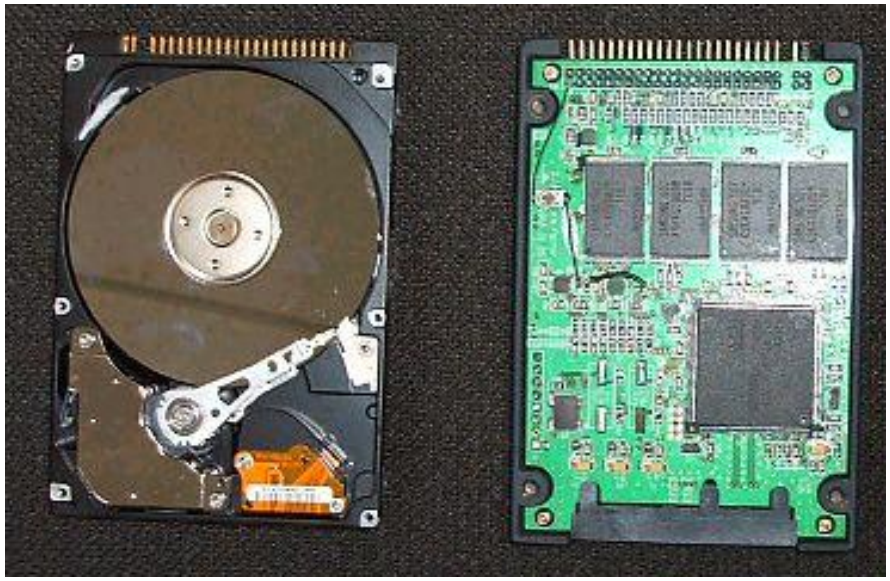
Intel Turbo Memory

SSDs (Solid-State Disks)

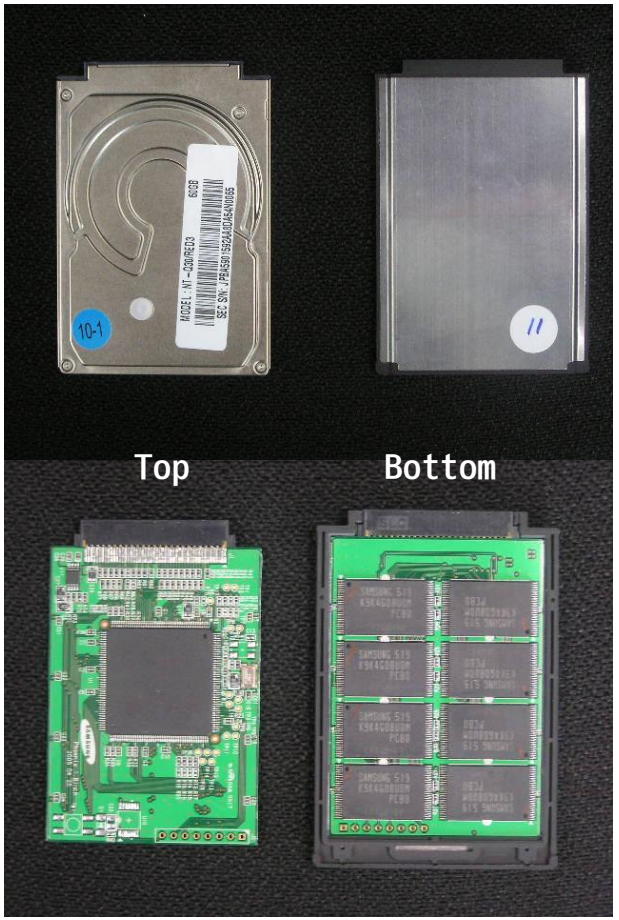
SSDs (1)

HDDs vs. SSDs

2.5" HDD Flash SSD
(101x70x9.3mm)

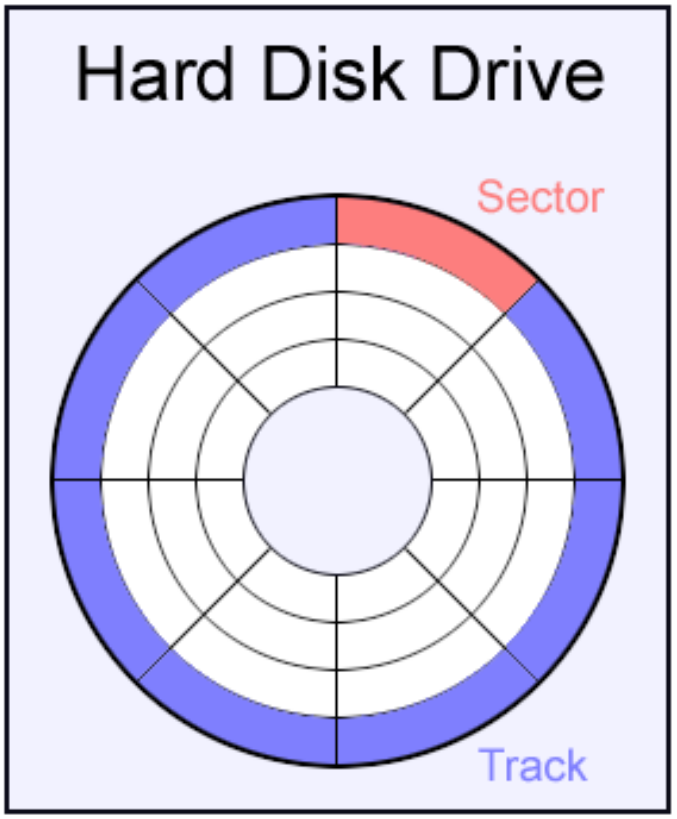


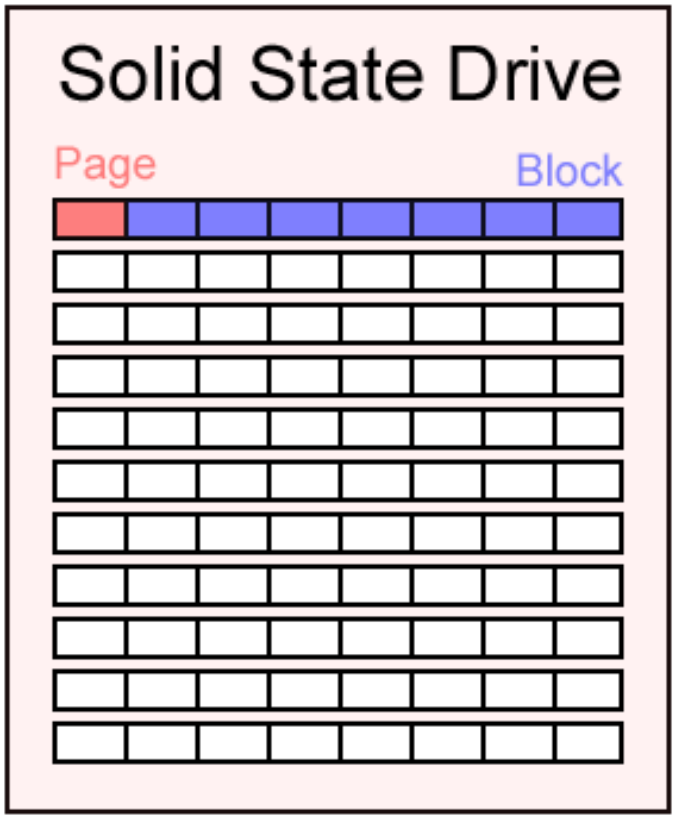
1.8" HDD Flash SSD
(78.5x54x4.15mm)



SSDs (2)

HDD vs. SSD





SSDs (3)

Feature	SSD (Samsung)	HDD (Seagate)
Model	MMDOE56G5MXP (PM800)	ST9500420AS (Momentus 7200.4)
Capacity	256GB (16Gb MLC x 128, 8 channels)	500GB (2 Discs, 4 Heads, 7200RPM)
Form factor	2.5” Weight: 84g	2.5” Weight: 110g
Host interface	Serial ATA-2 (3.0 Gbps) Host transfer rate: 300MB	Serial ATA-2 (3.0 Gbps) Host transfer rate: 300MB
Power consumption	Active: 0.26W Idle/Standby/Sleep: 0.15W	Active: 2.1W (Read), 2.2W (Write) Idle: 0.69W, Standby/Sleep: 0.2W
Performance	Sequential read: Up to 220 MB/s Sequential write: Up to 185 MB/s	Power-on to ready: 4.5 sec Average latency: 4.17 msec
Measured performance[1] (On MacBook Pro, 256KB for sequential, 4KB for random)	Sequential read: 176.73 MB/s Sequential write: 159.98 MB/s Random read: 10.56 MB/s Random write: 2.93 MB/s	Sequential read: 86.07 MB/s Sequential write: 84.64 MB/s Random read: 0.61 MB/s Random write: 1.28 MB/s
Price[2]	583,770 won	88,800 won

[1] *Source: http://forums.macrumors.com/showthread.php?t=658571*

[2] *Source: http://www.danawa.com (As of Nov. 21, 2010)*

NAND Constraints

Different read/write performance

- Sequential read: Up to 220 MB/s
- Sequential write: Up to 185 MB/s

No in-place update

- Require sector remapping (or address translation)

Bit errors

- Require the use of error correction codes (ECC)

Bad blocks

- Factory-marked & run-time bad blocks
- Require bad block remapping

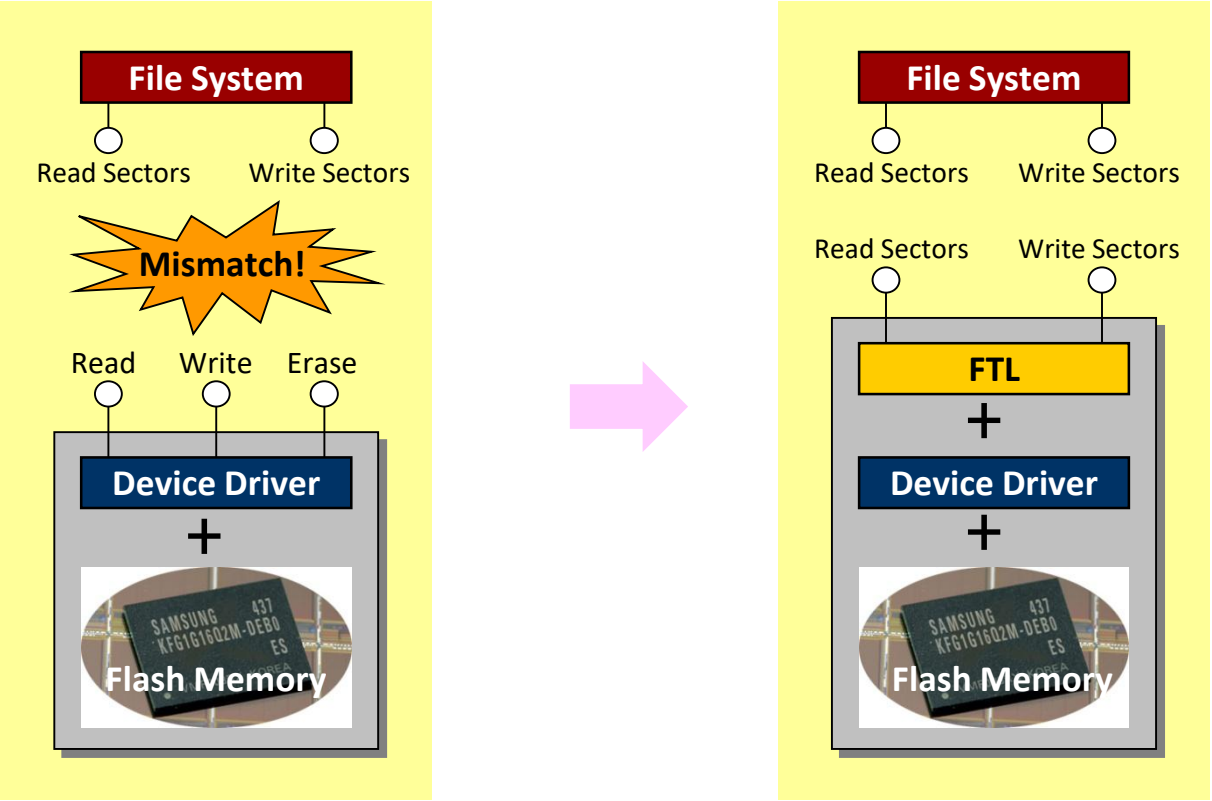
Limited program/erase cycles

- < 100K for SLCs
- < 10K for MLCs
- Require wear-leveling

FTL (1)

What is FTL?

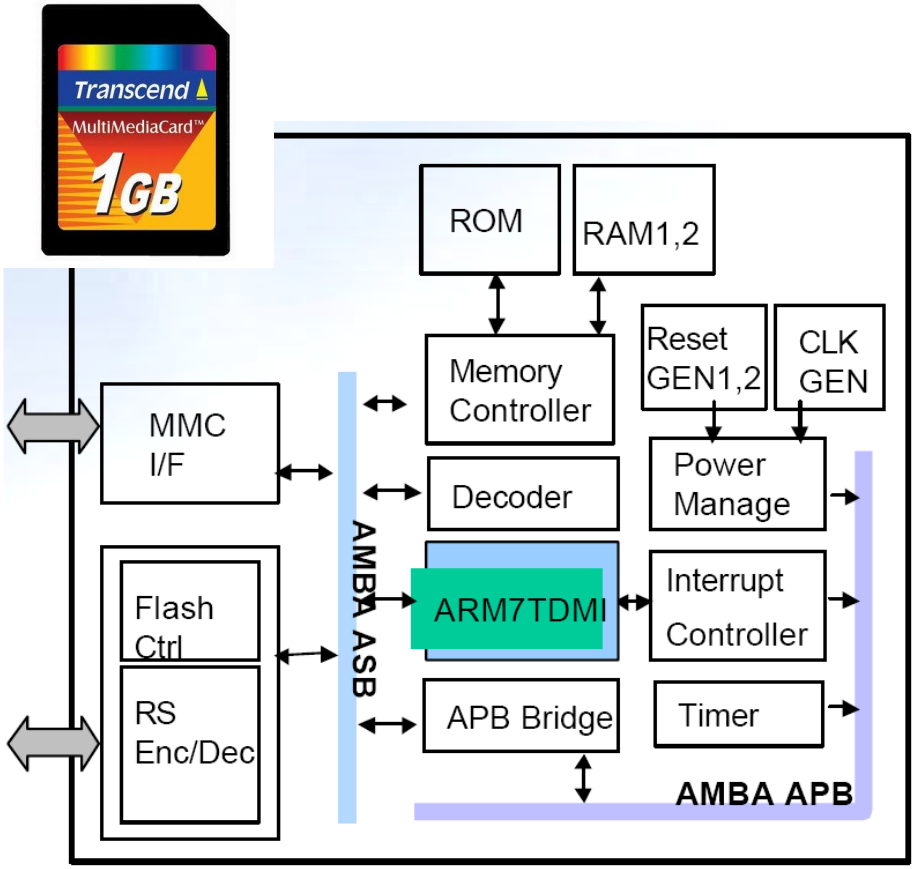
- Flash translation layer
- A software layer to make NAND flash fully emulate traditional block devices (e.g., disks).



Source: Zeen Info. Tech.

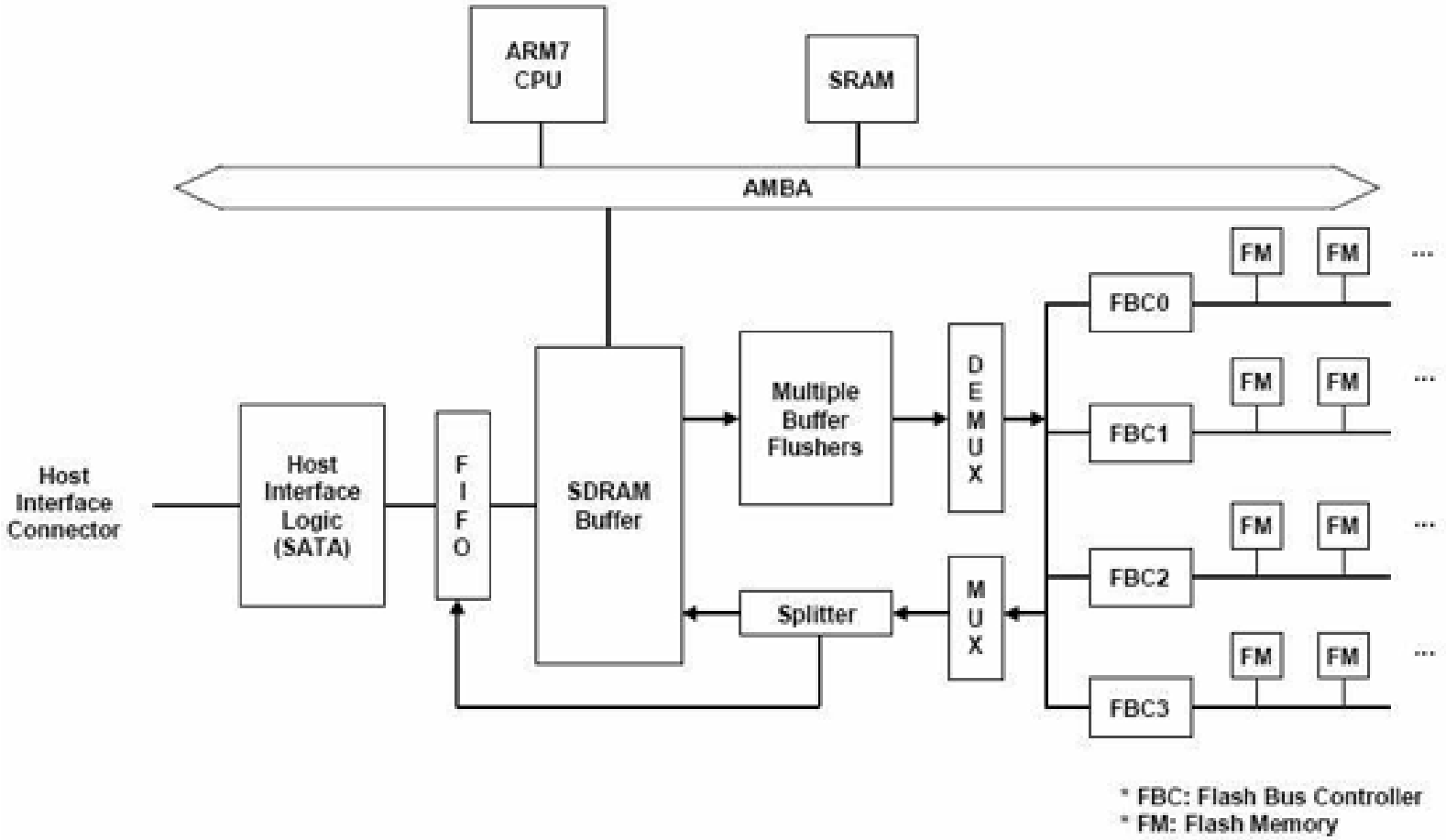
FTL (2)

Flash cards internals



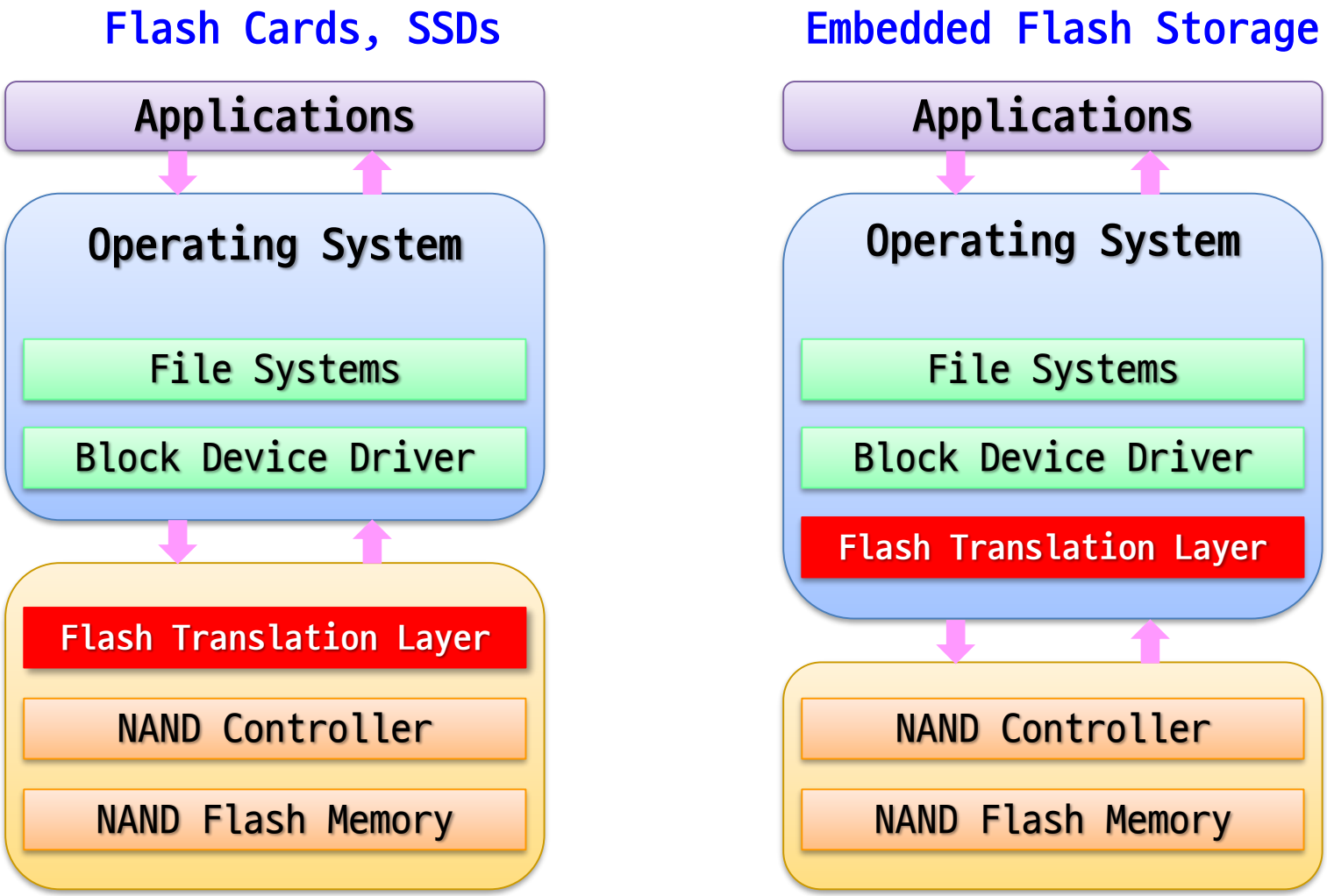
FTL (3)

SSDs internals



Source: Mtron Technology

FTL (4)


Flash Cards, SSDs
Applications
Operating System
File Systems
Block Device Driver
Flash Translation Layer
NAND Controller
NAND Flash Memory
Embedded Flash Storage
Applications
Operating System
File Systems
Block Device Driver
Flash Translation Layer
NAND Controller
NAND Flash Memory

FTL (5)

For performance

- Address translation
- Garbage collection
- Hot/cold data identification/separation
- Interleaving over multiple channels & flash chips
- Request scheduling
- Buffer management

FTL (6)

For reliability

- Bad block management
- Wear-leveling
- Power-off recovery
- Error correction code (ECC)

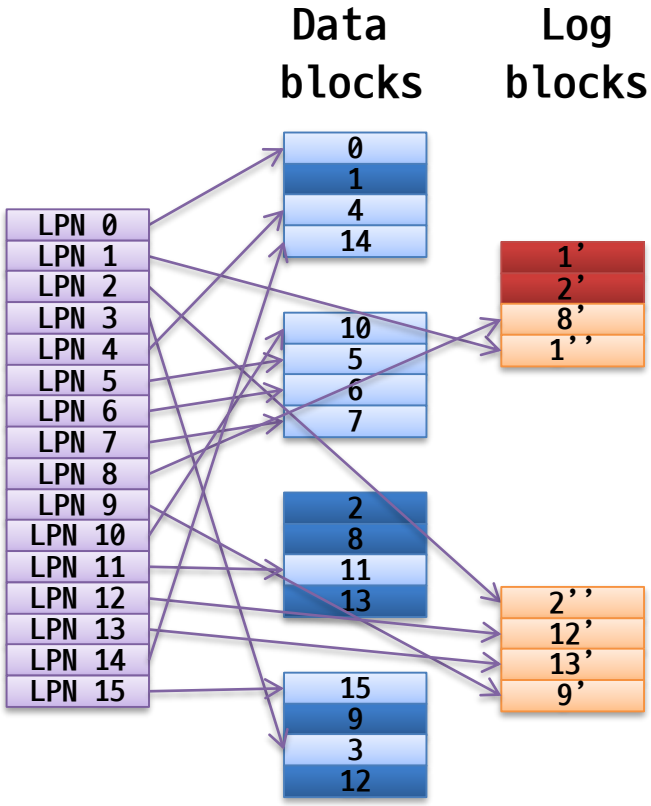
Other features

- Encryption
- Compression
- Deduplication

Sector Mapping (1)

General page mapping

- Most flexible
- Efficient handling of small writes
- Large memory footprint
  - One mapping entry per page:
    32MB for 32GB MLC (4KB page)
  - Bitmap for page validity
  - Per-block invalid page counter
- Sensitive to the amount of reserved blocks
- Performance affected as the system ages



invalid in data block

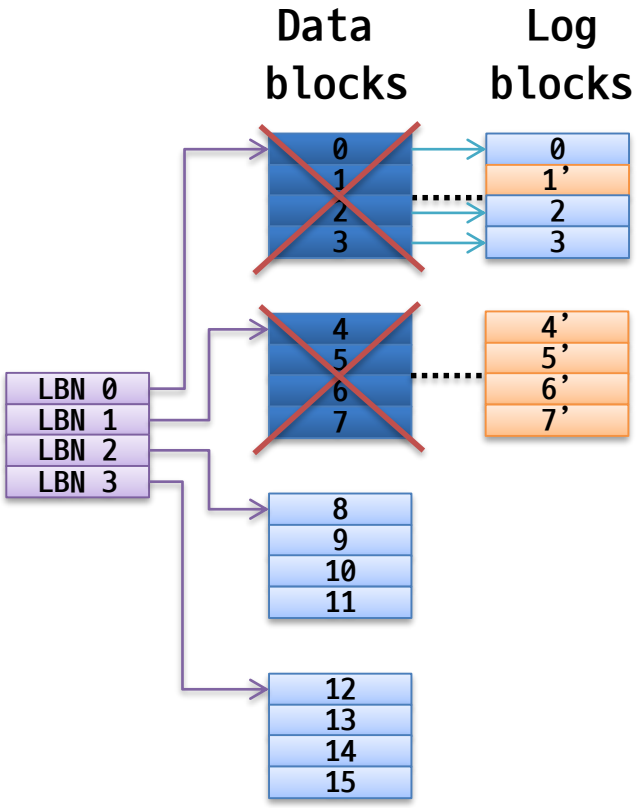
invalid in log block

W = <1, 2, 8, 1, 2, 12, 13, 9>

Sector Mapping (2)

Naïve block mapping

- Each table entry maps one block
- Small RAM usage
- Inefficient handling of small writes

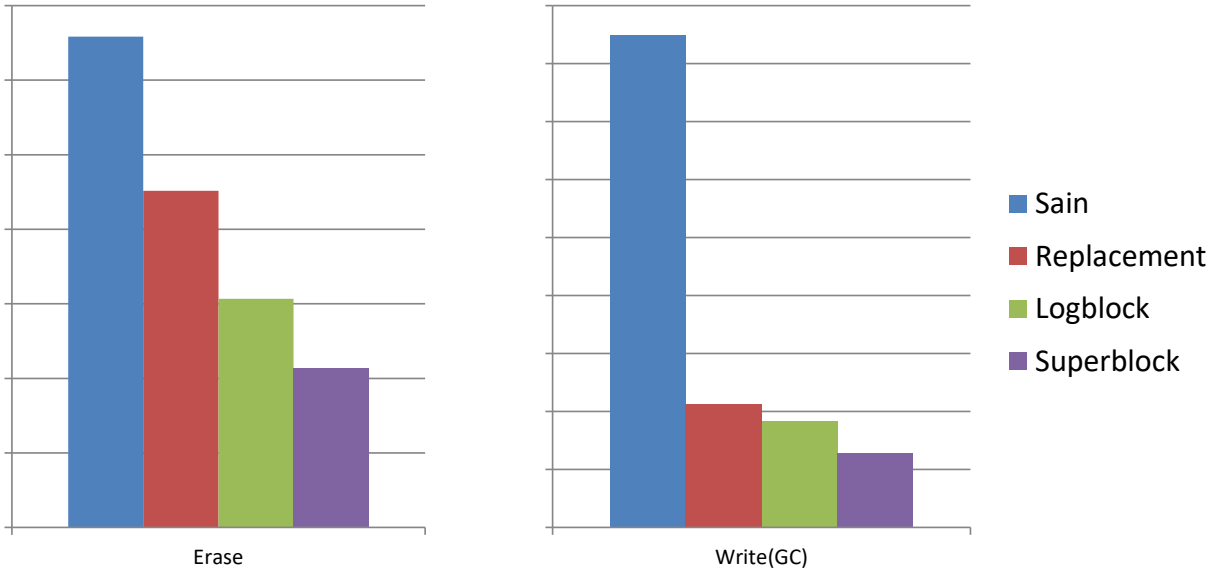


W = <4, 5, 6, 7, 1>

Performance (1)

Simulation environment

- 4GB flash memory
  - Large block SLC NAND (2KB page, 128KB block)
- FTL schemes
  - Naïve block mapping
  - Replacement block
  - Log block
  - Superblock
- Workload
  - Trace from PC using NTFS



OS Implications (1)

NAND flash has different characteristics compared to disks

- No seek time
- Asymmetric read/write access times
- No in-place-update
- Good sequential read/sequential write/random read performance
- But bad random write performance
- Wear-leveling

- Traditional operating systems have been optimized for disks

OS Implications (2)

SSD support in Microsoft Windows 7

- Turn off "defragmentation" for SSDs
- New "TRIM" command
  - Remove-on-delete
- Align file system partition with SSD layout
- Larger block size proposal (4KB)

Beauty and the Beast

NAND Flash memory is 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 (10K~100K program/erase cycles)
- Bad blocks

Software support for NAND flash memory is very important for performance & reliability

=> FTL is crucial for performance and reliability