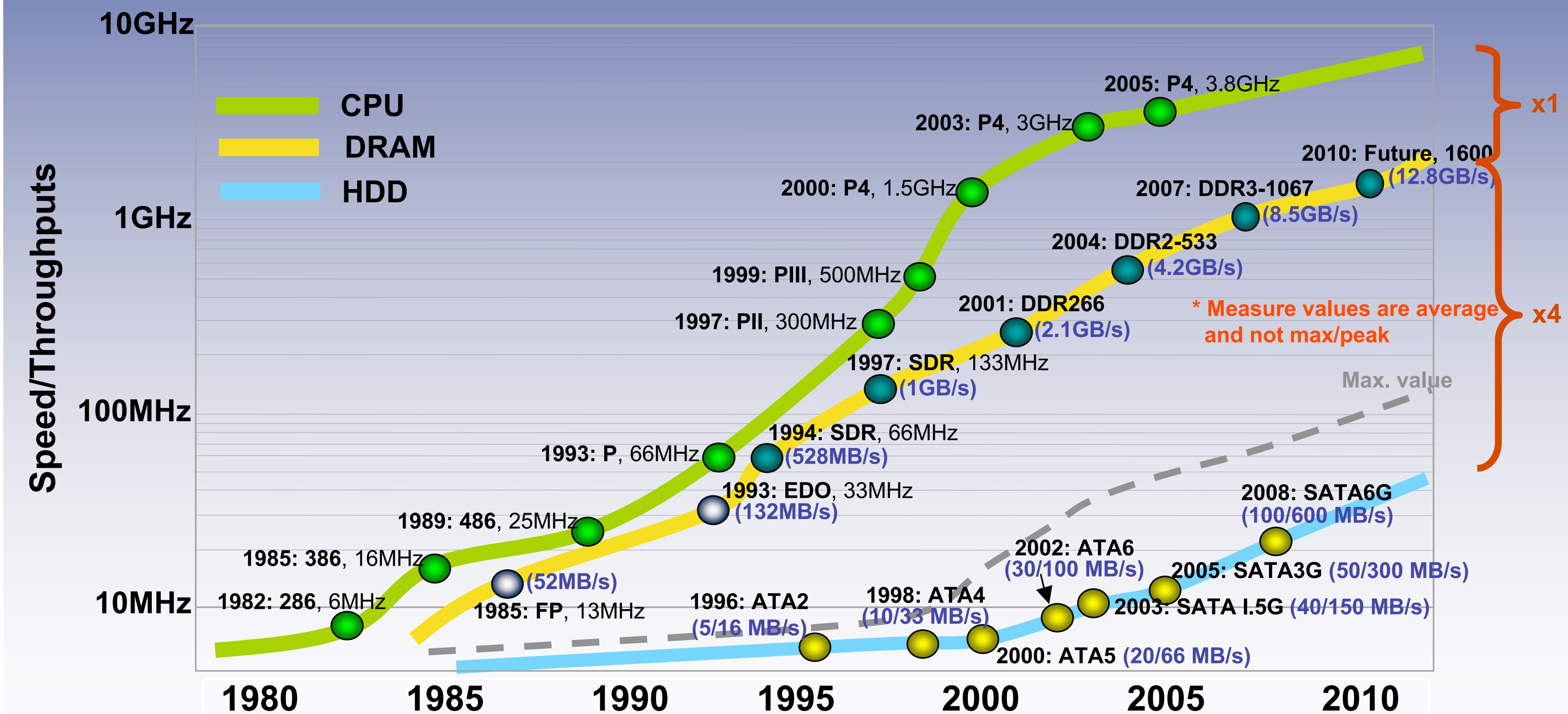
Solid State Drives
Moving into Design

Y.R. Kim
Director, Technical Marketing
SAMSUNG Semiconductor Inc.

System Performance Bottleneck

Performance gap between DRAM & Storage is 4X greater than between CPU & DRAM

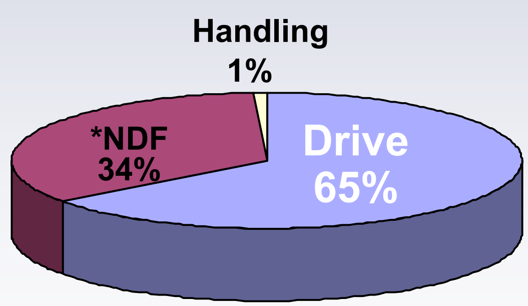


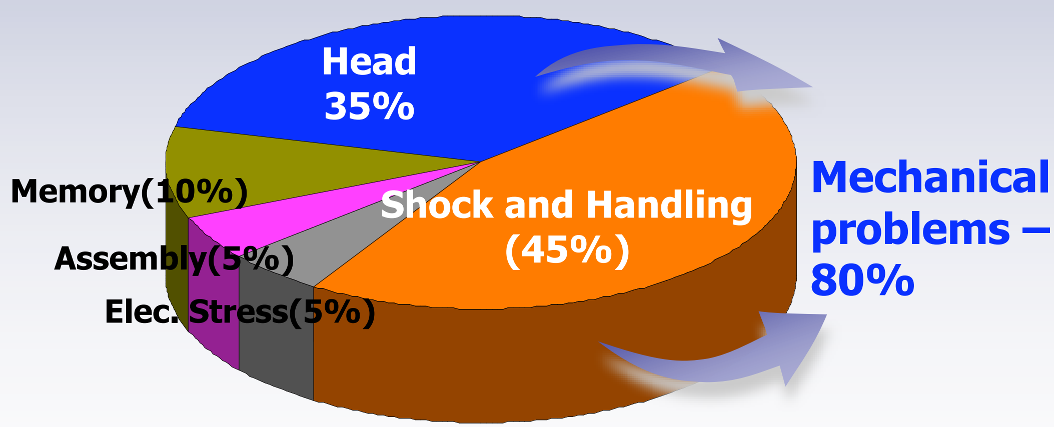
Source: Samsung

Other Existing System Hurdles?

- Booting time
- System power budget/Mobility/Battery life
- Mechanical failures of HDD
  - No spindle motor
  - No moving parts

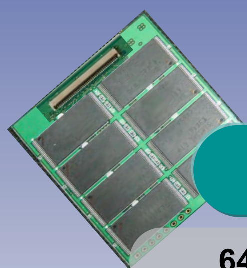
NotePC Failure Report from Vendor (465 EA)
Source: anonymous PC OEM

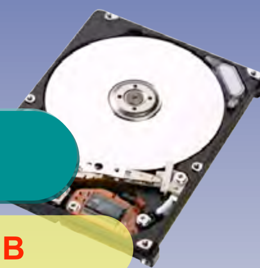




*NDF : Not Define Failure

Why NAND Flash Better than HDD?

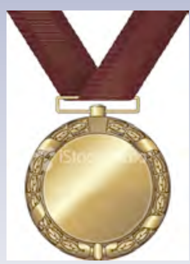




	1.8" SSD		1.8" HDD
	64GB/32GB/16GB/8GB	Density	40GB/60GB/80GB
	~ $500 (32GB)	Cost	< $150 (80GB)
	54 x 71 x 3 (mm)	Dimension	54 x 78.5 x 8.2 (mm)
1/5x	13 g	Weight	61 g
2.6x	R: 53MB/s, W: 32MB/s	Sequential Access	R/W: 22 ~ 48MB/s
11x	R: 53MB/s, W: 13MB/s	Random Access	R/W: 1 ~ 5MB/s
1/3	Active: 0.5W	Power (Active)	1.5 W
20x	20G (10~2000Hz)	Operating Vibration	1.0G (22~500Hz)
3.3x	MTBF: > 2M hours (TBD)	Endurance	MTBF: < 300K hours

SSD Wins

- Reliability



- Power Consumption

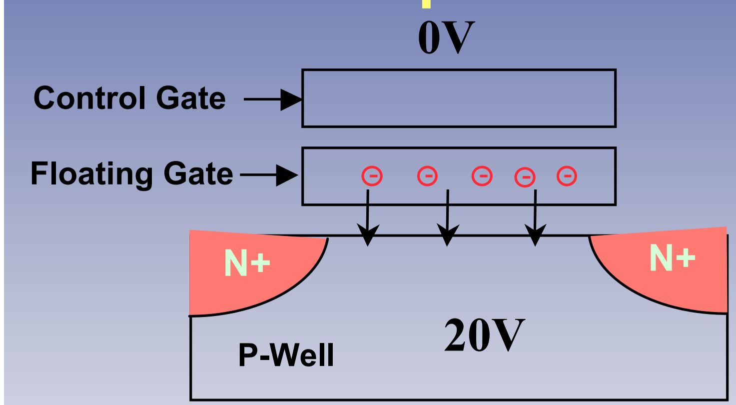
- Performance

Reliability Factors

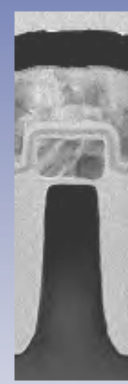
- Understanding of NAND Flash endurance
- ECC/EDC
- Wear-leveling
- Lifetime estimation – worst case scenario

Cell Operation of NAND Flash

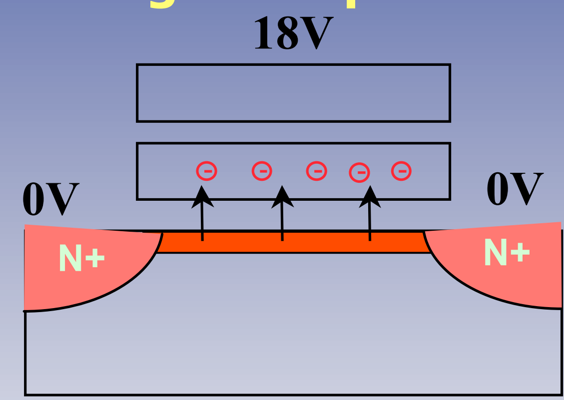
❑ Erase Operation



60nm Cell

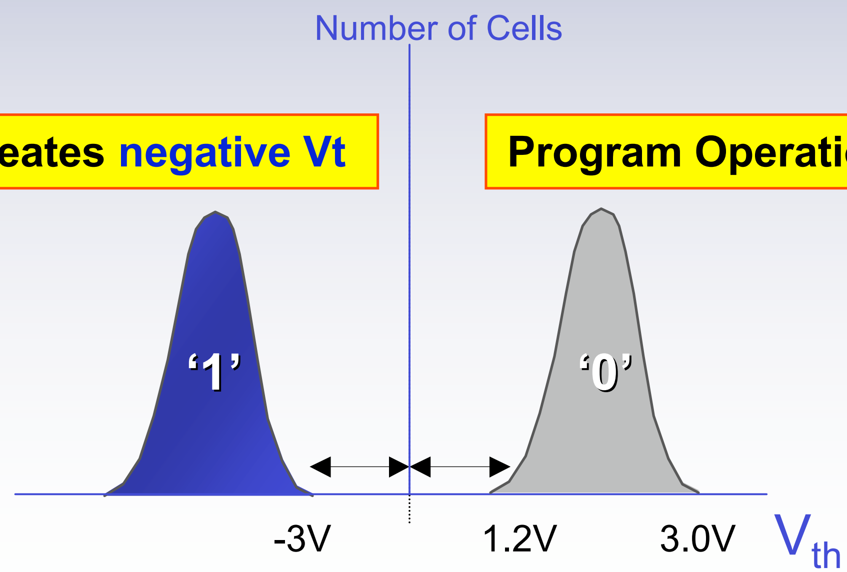


❑ Program Operation

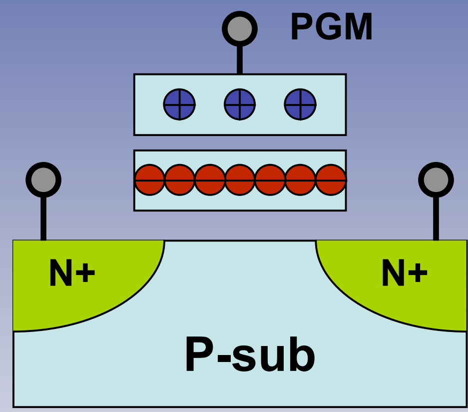


Erase Operation creates negative Vt

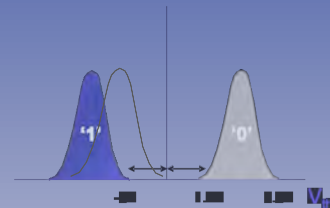
Program Operation creates positive Vt

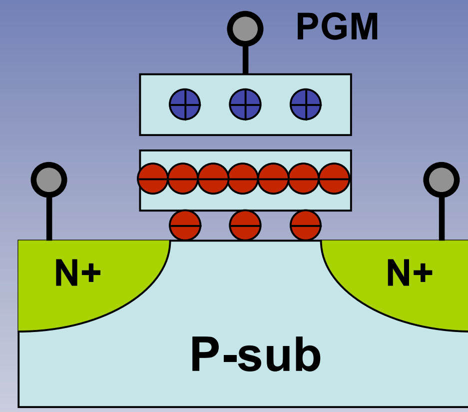


Program & Erase (P/E) Cycling

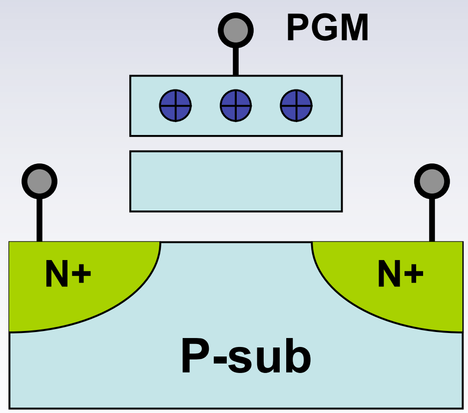


Before Cycle, Programmed Cell

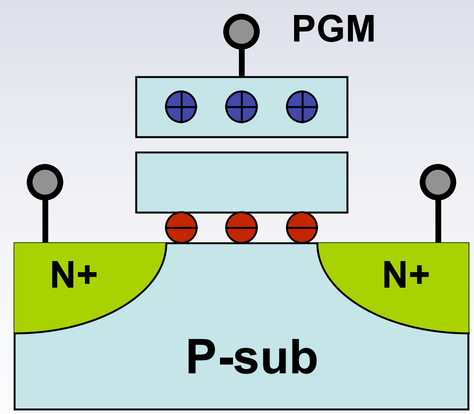




After Cycle, Programmed Cell



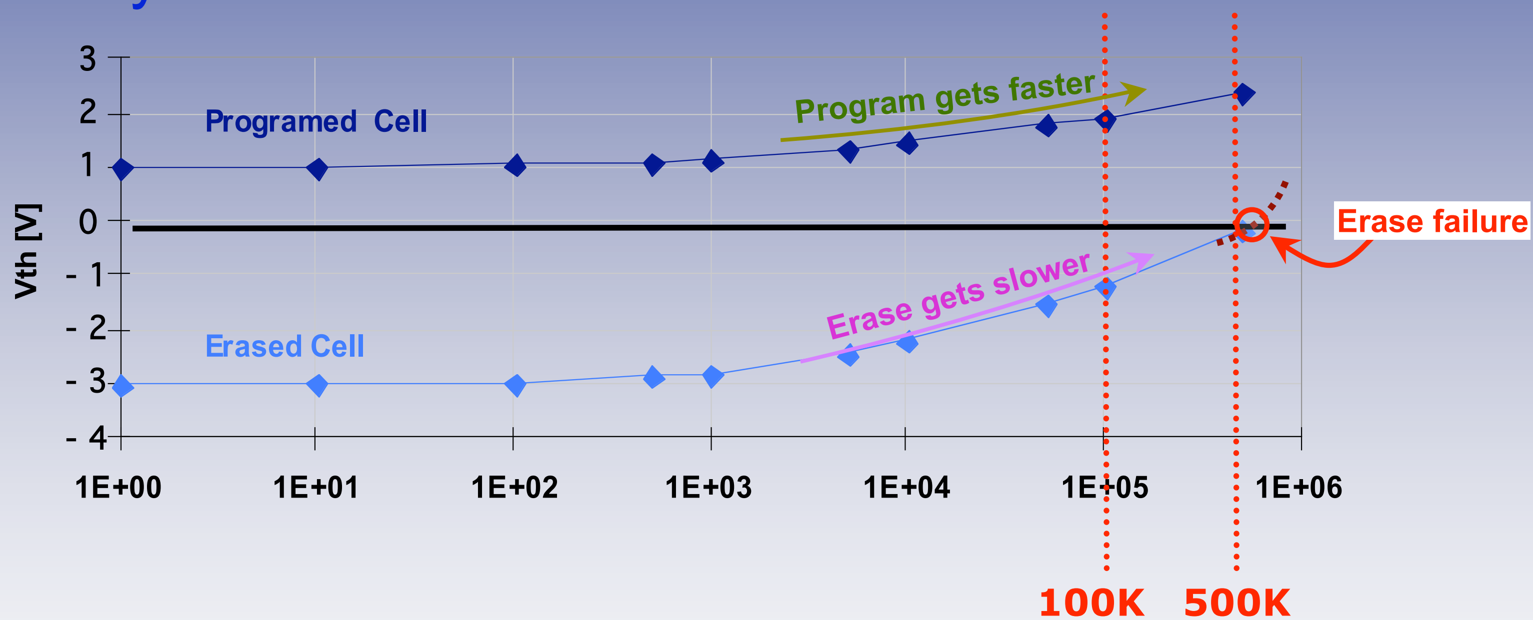
Erased Cell



Erased Cell

Inherent Endurance Characteristics

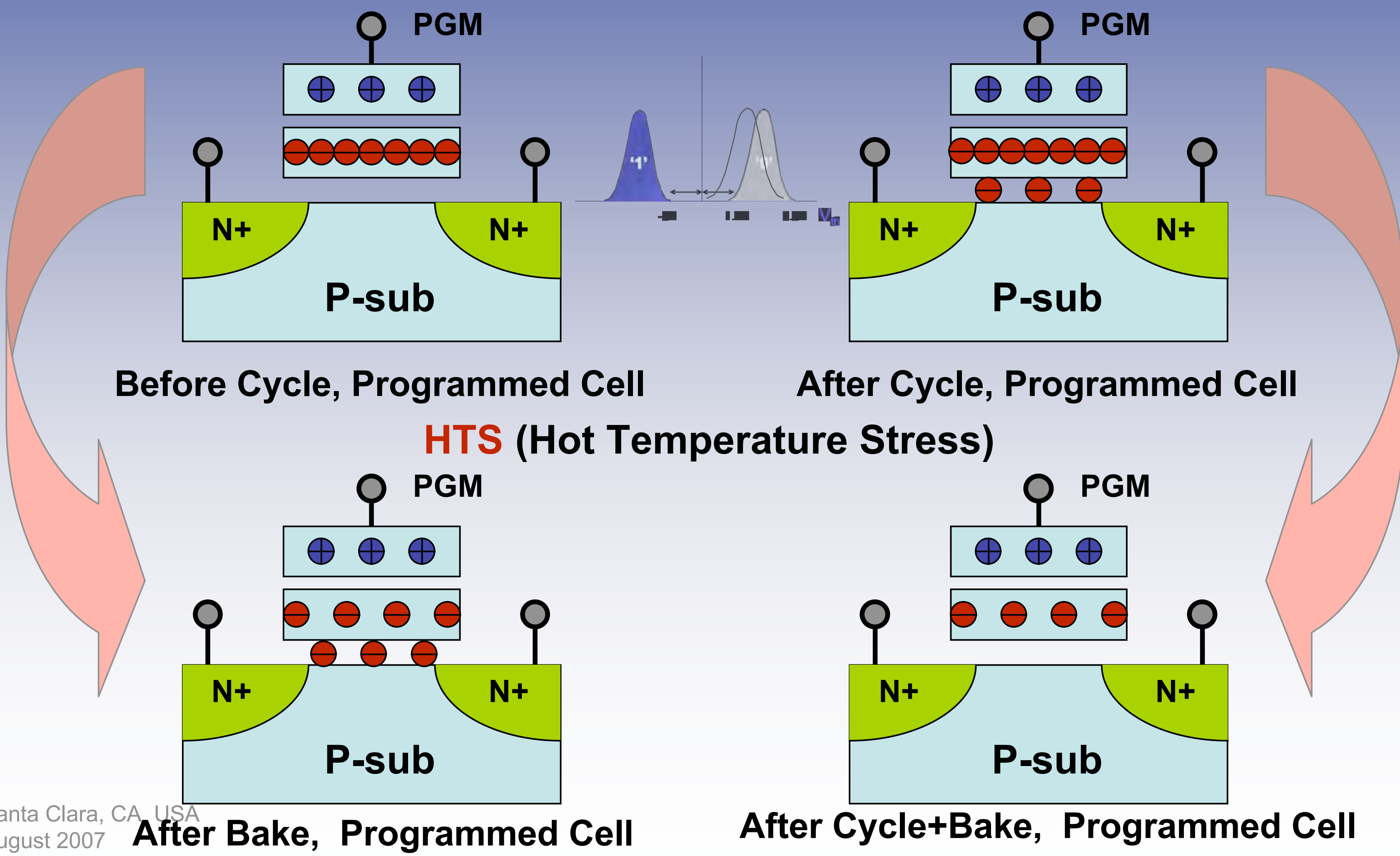
- Memory Cell Vth vs. Endurance



- Usually erase failure first occurs over 100K P/E cycles, resulting in an invalid block without any data loss
- The number of invalid blocks only gradually increases after 100K

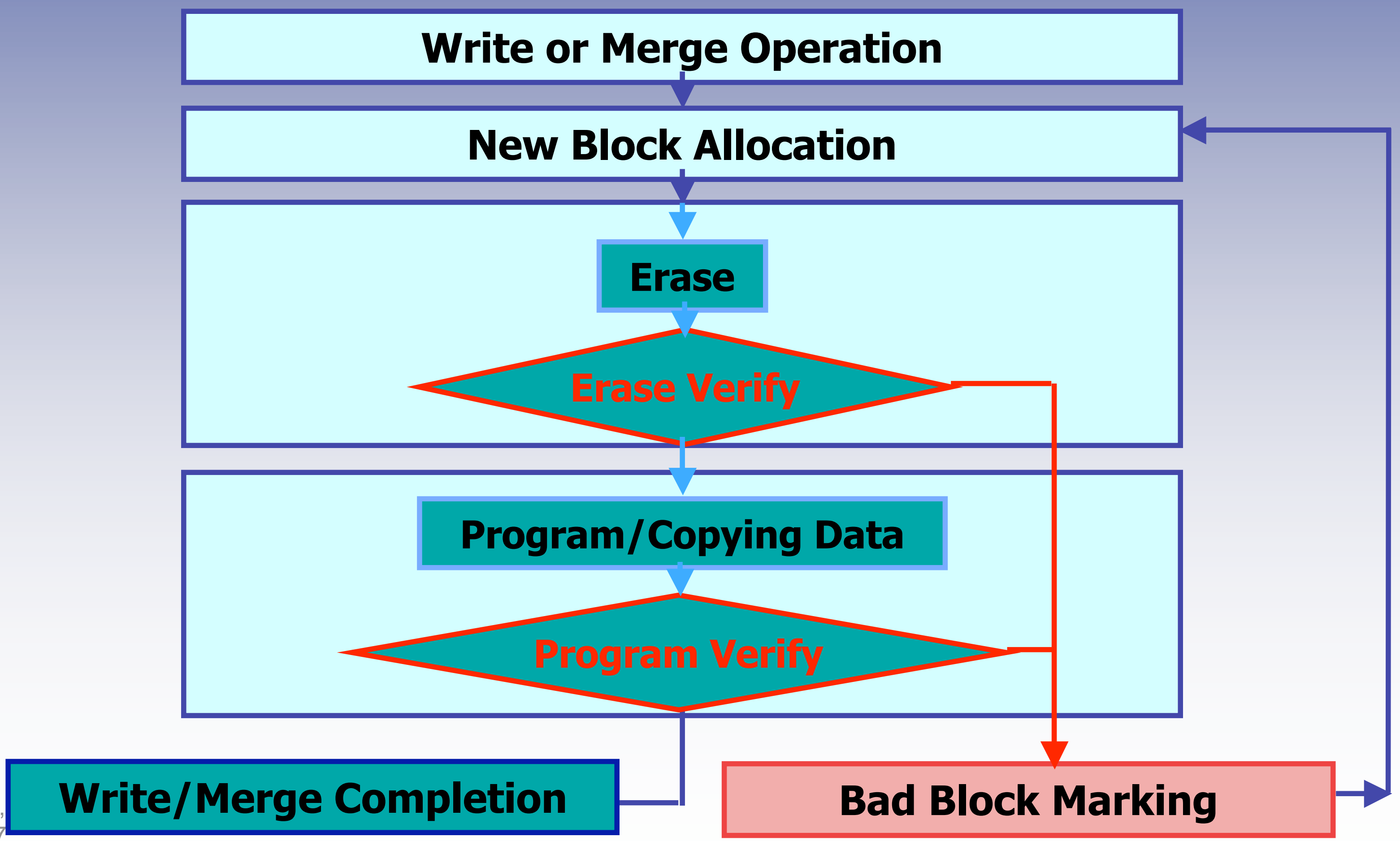
Read Retention Issue

- HTS Effect – Why phenomenon occurs



Erase/Program Verification on SSD

- Bad Block Detection during SSD Operation



Failure Modes of NAND Flash

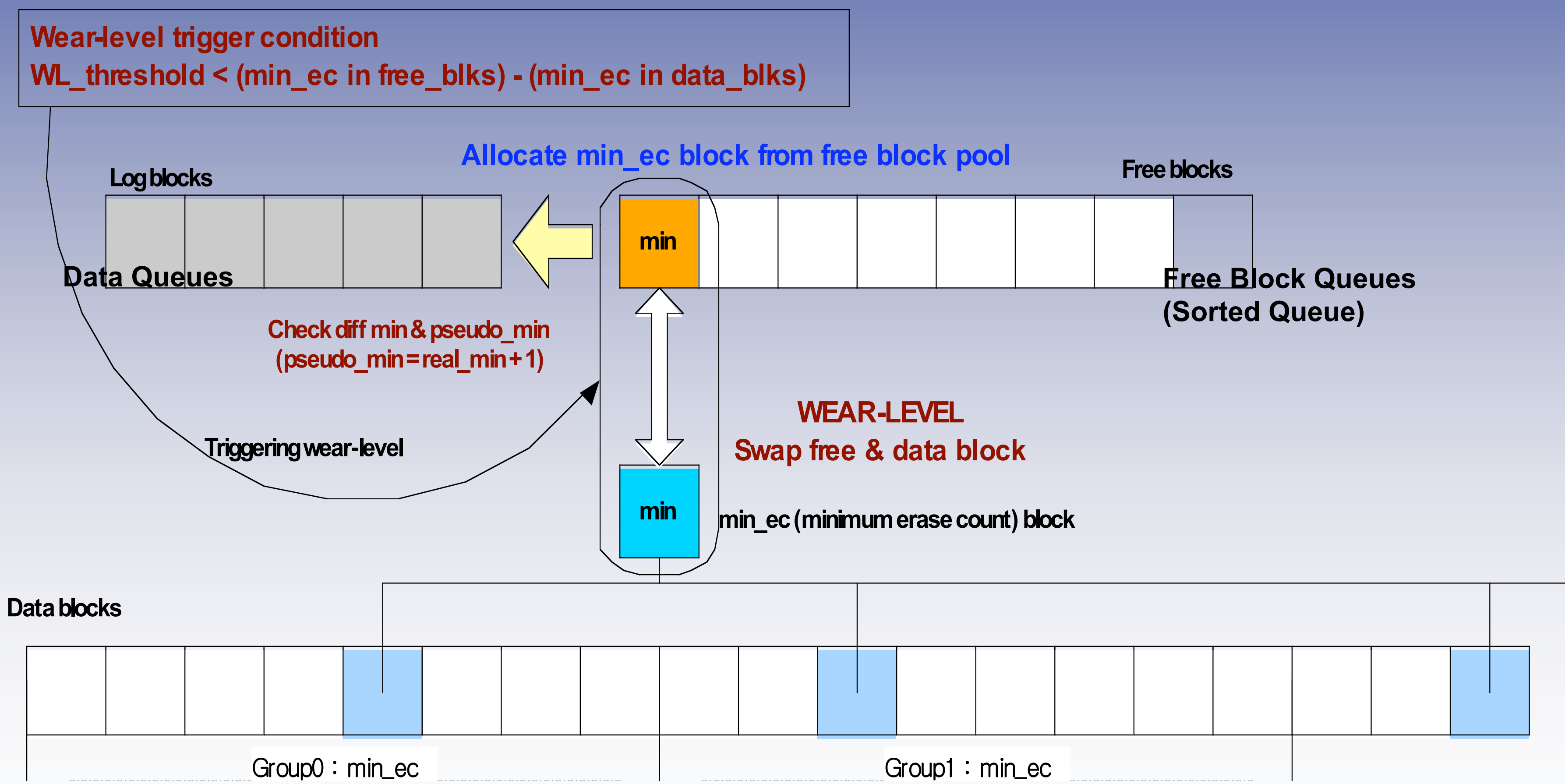
Failure Mode	Condition	Solution	Implication
Run-time write failure	Erase/program fail status	Block replacement	No data error; Bad blocks increase
Read failure	1 bit error	ECC-correction	No data error; No bad block increase

- Max. 2% of invalid blocks guaranteed up to 100K endurance
- Invalid blocks might be more than 2% after 100K endurance
- 2-bit failure rate irrelevant to write performance, only related to read
  - Failures too low to be observed; can only be calculated
  - Single bit error rate/512B(SFR): < 0.01 PPM @100K Cycle (60nm NAND)
  - Failure rate = $(SFR)^2$ x (# of sectors / chip [4Gb])
    = 1 x 10^{-16} x 1048576 * 64 chips = 0.07 PPM
  - SATA-2 has high error resilience due to 6-bit ECC per sector (512B)

Wear-Leveling

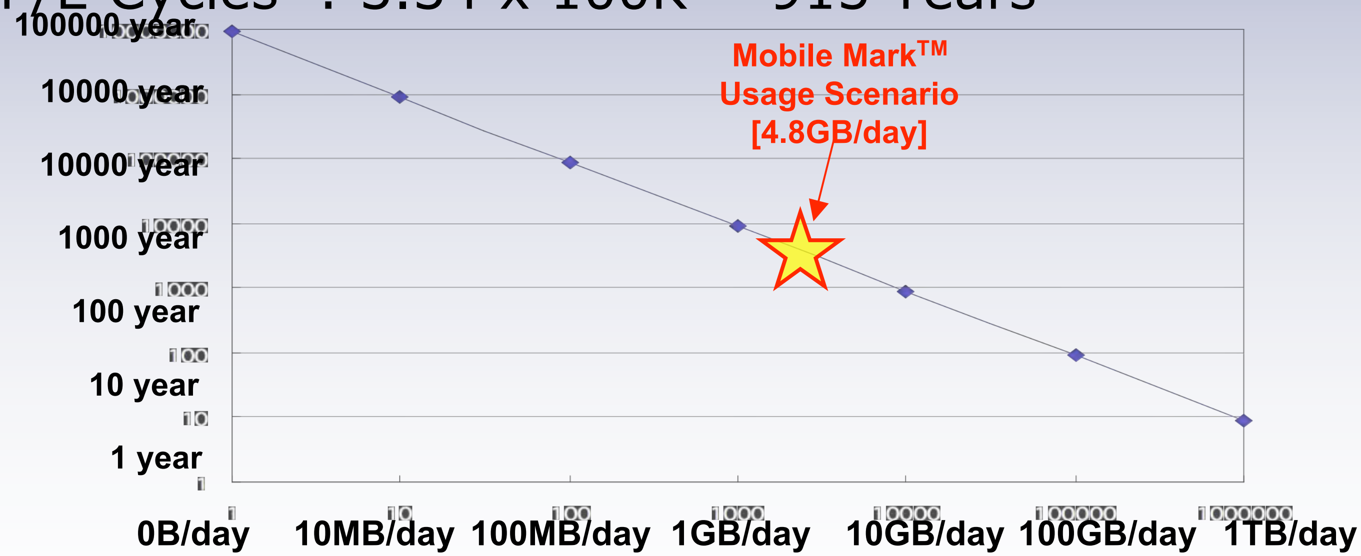
- Data wear-leveling
  - Wear-leveling on data blocks
- Meta wear-leveling
  - Wear-leveling on ATL context blocks (map blocks)
- Erase count
  - Each physical block has its own erase count. (Static wear-leveling)
  - FTL format makes block erase count zero
- Wear-level threshold value
  - Triggering value for wear-level operation
  - Configurable at compilation time
  - If threshold value becomes smaller, wear-leveling happens more frequently
  - Too frequent wear-leveling may cause performance to drop
- Group-based, vertical ordered block mapping
  - Full scanning too expensive to use to find blocks with minimum erase count
  - First round: find the minimum block within each group
  - Second round: find the minimum within each group's minimum
  - Vertical ordered block mapping required to avoid excessive dependence upon a specific group

Wear Leveling (Group-based)



SSD's Lifetime Estimation

- Assuming xGB/day usage scenario, can estimate life expectancy
- MobileMark™ usage scenario: 4.8GB/day
  - Workload: 4.8GB write during one day
  - Assuming 16GB, SSD lifetime estimated
    - – 1 P/E Cycle : 16GB/4.8GB = 3.34 Days
- 100K P/E Cycles : 3.34 x 100K = 913 Years



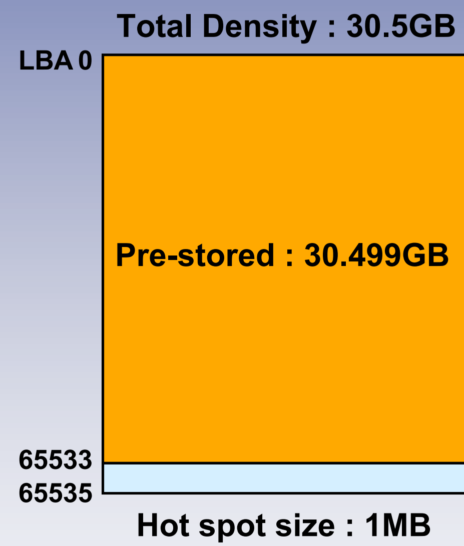
Lifetime Estimation (Worst Case)

❑ Usage scenario

- 1MB-sized hot spot area (LBAs) & remains are pre-stored

- **Continuous hot spot updating**
- Full 64KB sequential writing speed

  Assumption: 100% running 24h X 7d
  Workload : 25 MB/sec
  32GB SSD lifetime
  - 1 P/E Cycle : 32GB/25MB/s = 22 Min
  - 100K P/E Cycles : 22Min x 100K = 4.2 Years



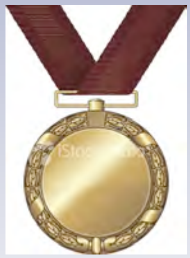
❑ Real endurance measurement

- Sample size : 32GB 10 EA

	1M P/E Cycle	2M P/E Cycle	3M P/E Cycle
Test Result	Pass	Pass	Pass

SSD Wins

- Reliability



- Power Consumption

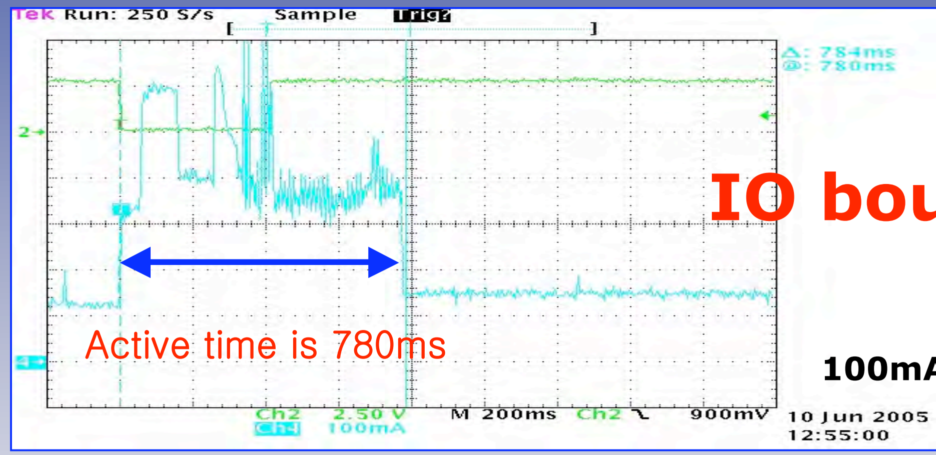
- Performance

Power Comparison -Read

HDD
1 Sector Read
(2.5” HDD)

/DASP

Curren
t



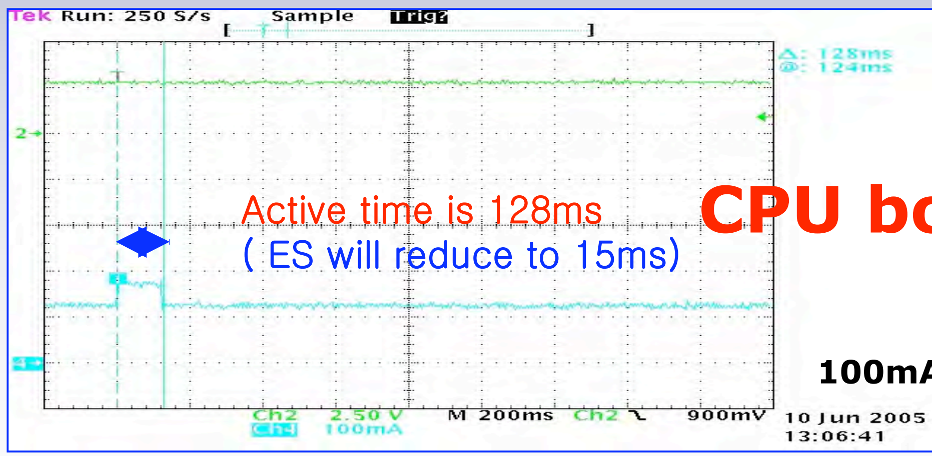
IO bound

100mA/Div.

SSD
1 Sector Read
(SAMSUNG 4GB SSD)

/DASP

Current

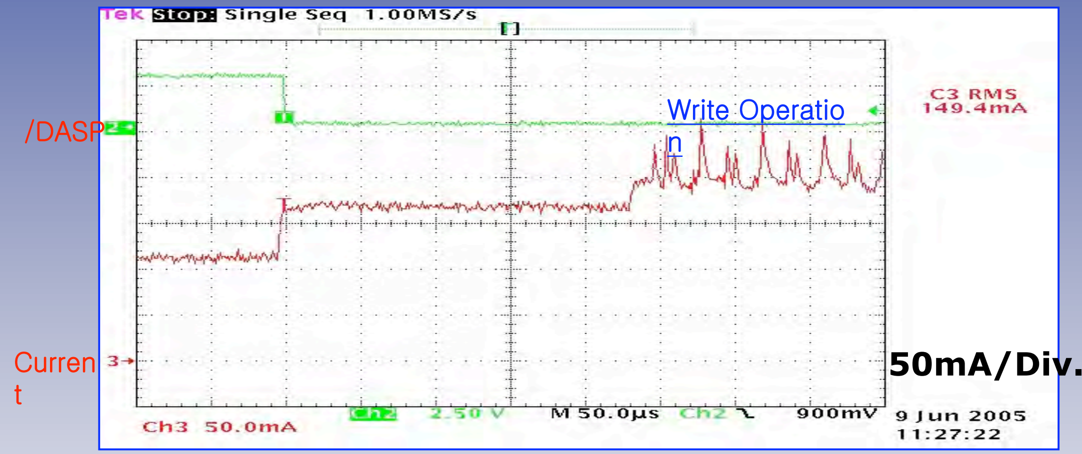


CPU bound

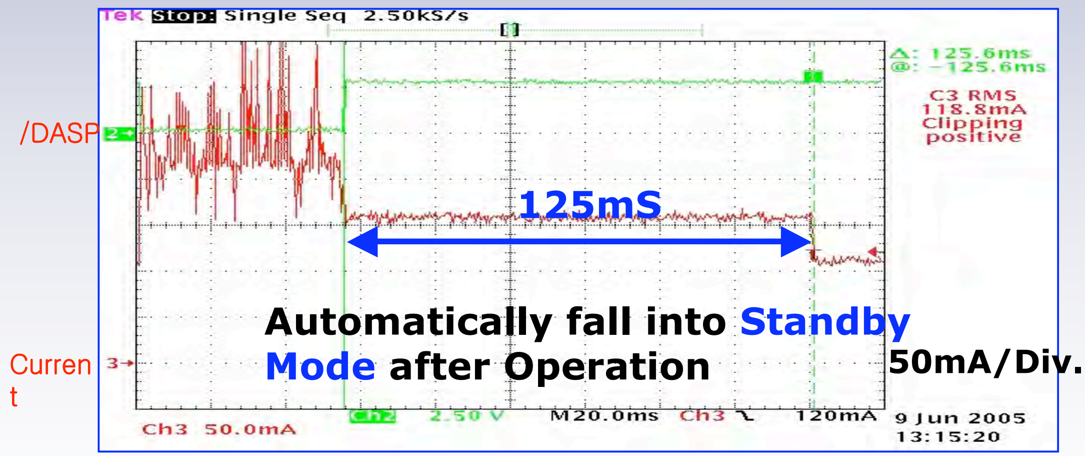
100mA/Div.

Power Consumption - Write

SSD Write Start

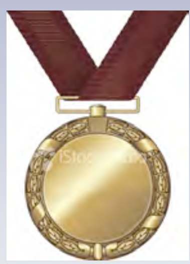


SSD Write End



SSD Wins

- Reliability

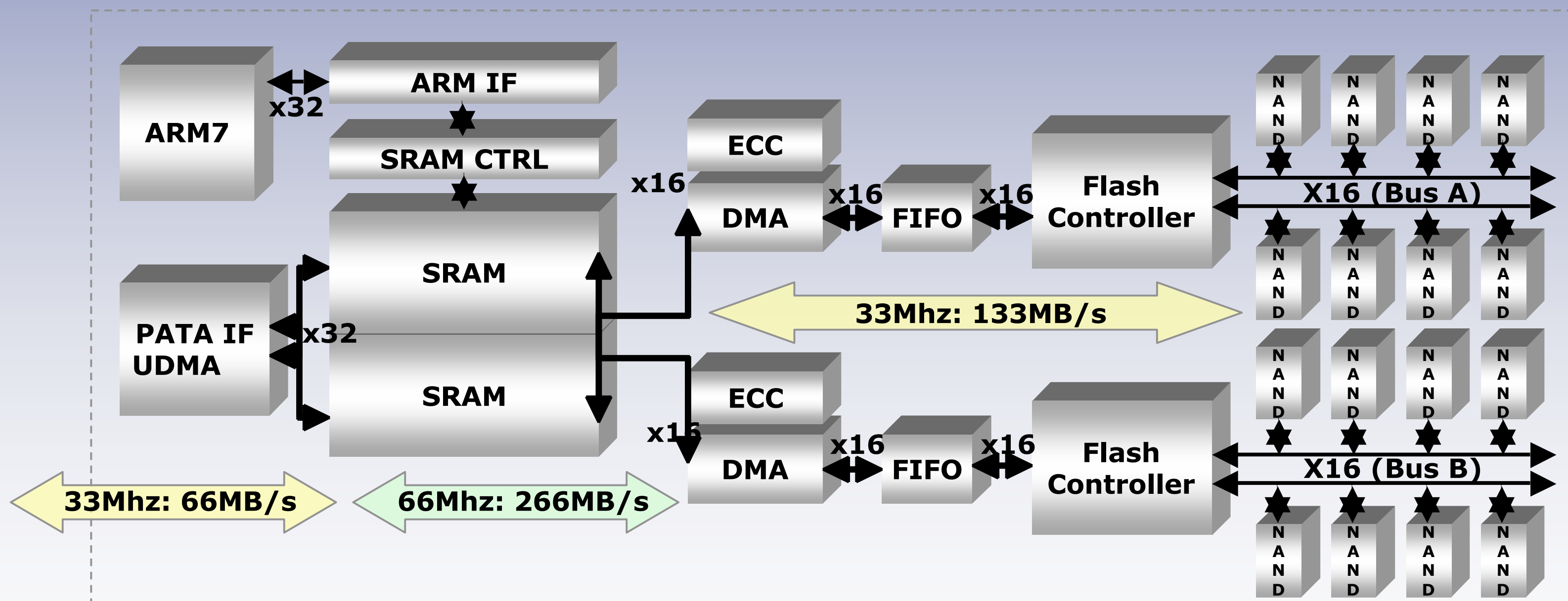


- Power Consumption

- Performance

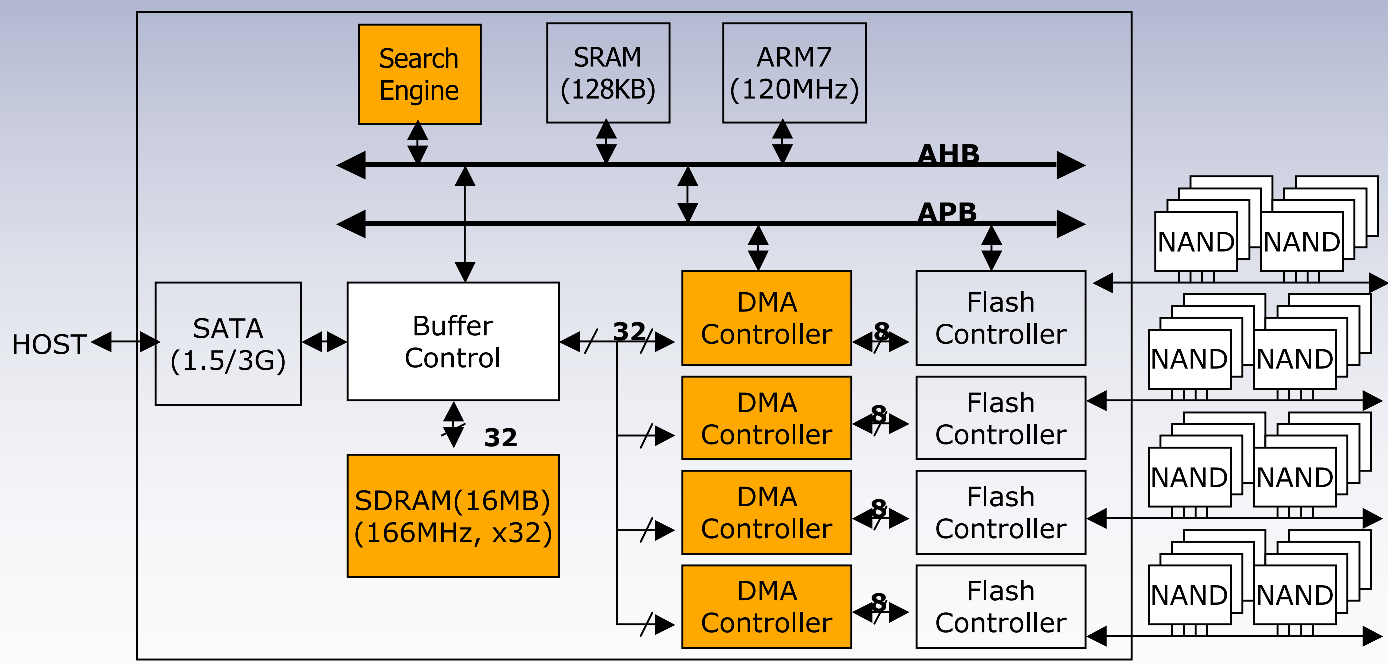
SAMSUNG PATA/SATA-I SSD Architecture

- Implements optimized buffering architecture using SRAM (64KB)
- Concurrent operation of 16 NAND chips (4-channel/4-way interleaving)
- 4 Parallel HW ECC (RS-2 bit)



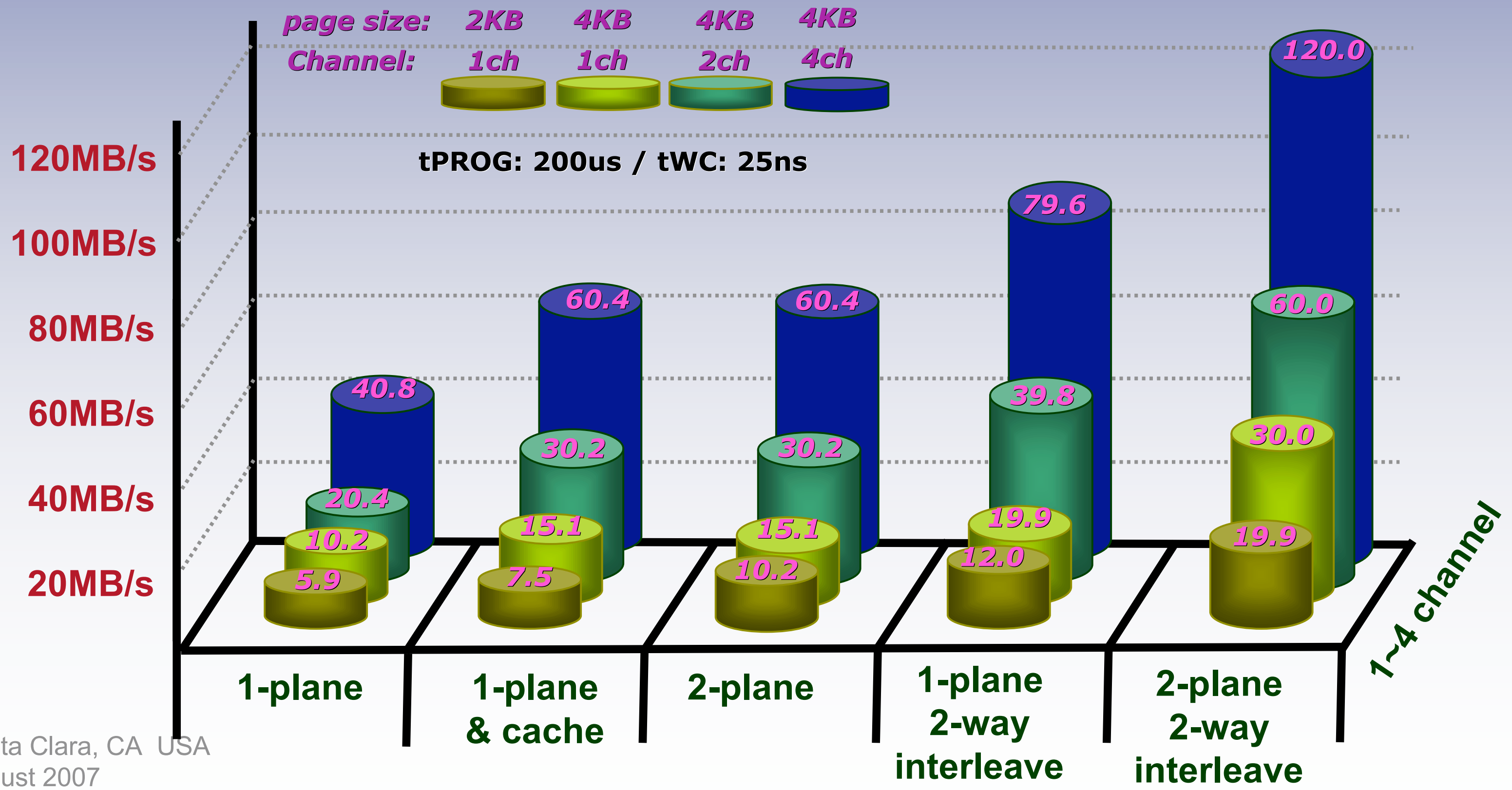
SAMSUNG SATA-II SSD Architecture

- SLC/MLC NAND support
- Independent 4-channel, 4-way interleaving (CRC supports data integrity)
- 6-bit ECC per sector (for SLC& MLC; can support MLC)
- 16MB DRAM buffer (w/ CRC) & dedicated HW indexing engine



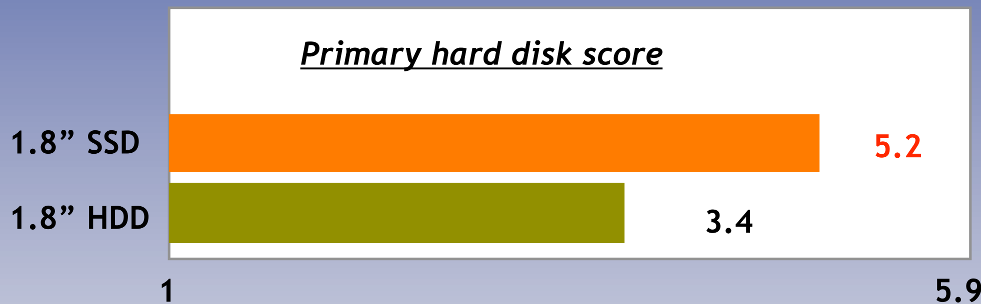
How to improve Performance?

- 15MB/s with a single 50nm 8Gb SLC in two plane mode
- Using multi-plane, multi-way/channel interleaving, read/write throughput increase possible!
- Up to now, 2-lane/4-way/4-channel interleaving supported



Performance on Windows Vista

- Window Express Index*

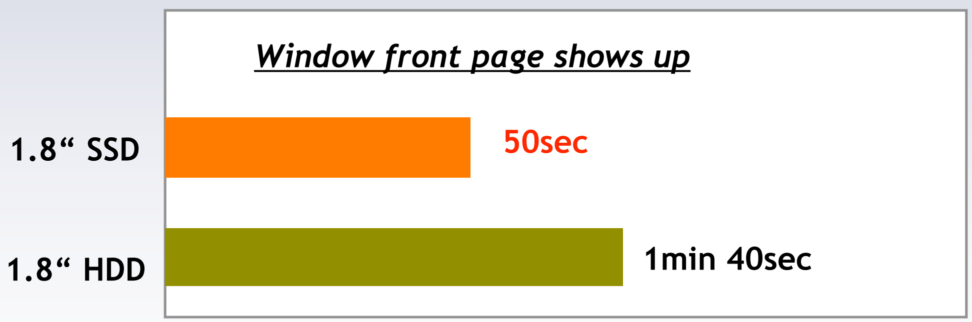


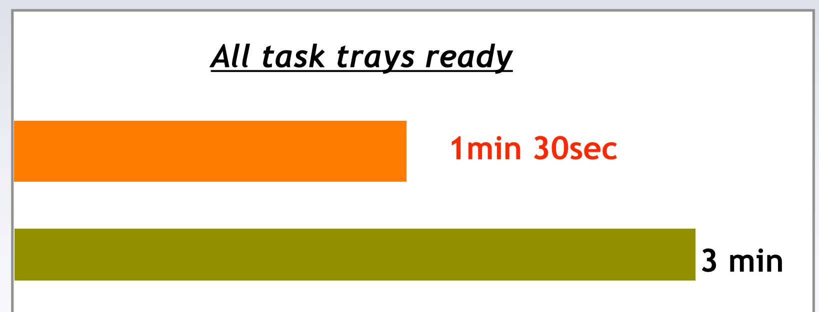
Test Environment

System: Sony VAIO type G
CPU: Core Solo U1400(1.2 GHz),
OS: Windows Vista Business
SSD: 32GB, HDD: 60GB

* Window Express Index: Standard tool for performance diagnosis in Window Vista OS

- Booting Time





(Source : Itmedia.co.jp)

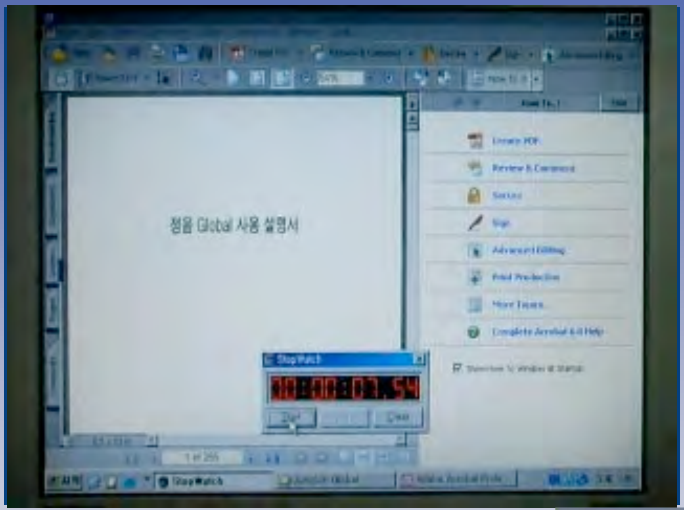
Application Launching Time

- Acrobat Launching Time Comparison (PATA/SATA-I SSD)

Device	Location of files			Acrobat launching time
	Windows XP	Acrobat	PDF file	
SSD	Flash	Flash	Flash	3.2s (4.4x)
H-HDD or Robson	HDD	Flash	Flash	5.0s (2.8x)
	HDD	Flash	HDD	7.5s (1.8x)
HDD	HDD	HDD	HDD	14.0s (1.0x)

- **Application: Acrobat Professional 6.0**
- **Data: JungUmGlobal Manual_Ko.pdf (21.9MB)**
- **System: Pentium 4 2.8 GHz, 512MB**
- **PATA/SATA-I SSD 32GB, PATA HDD 80GB**

Fast Application Launch

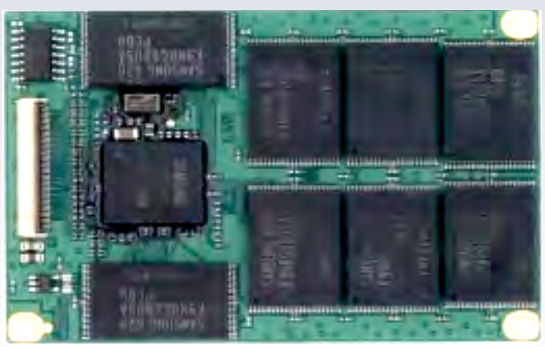


HHDD (7초)

SSD (3초)

HDD(14초)

SSD Future Direction



Fast BIOS Required for SSD

- General Software developed Fast BIOS for H-HDD

GENERAL
SOFTWARE

http://www.marketwire.com/mw/release_html_b1?release_id=234679

General Software, Inc. Announces Faster Boot Time for Windows Vista With Its Embedded BIOS Firmware and Microsoft's Windows ReadyDrive Technology

"Power On" to Desktop Display Measured at 24 Seconds, Compared to 72 Seconds With Traditional BIOS

SAN JOSE, CA -- (MARKET WIRE) -- April 03, 2007 -- EMBEDDED SYSTEMS CONFERENCE SILICON VALLEY -- General Software, Inc. demonstrated what it believes is one of the world's fastest boot times for Windows Vista, using a commercially available VIA mainboard and a Samsung Hybrid Hard Disk Drive accelerated by General Software's Embedded BIOS® firmware with StrongFrame™ Technology, at the Embedded Systems Conference in San Jose today.

The boot time for Windows Vista on the motherboard, from power on to the desktop display was measured at 24 seconds, contrasted to 72 seconds (over a minute) with the traditional one-size-fits-all BIOS installed on the board. The time achieved for resume from hibernation is 11 seconds.

The H-HDD operates similar to a traditional hard disk drive, but avoids the time it takes for the drive to spin-up after being powered on, before its data can be accessed. In traditional drives, this "power on time" can be expensive -- 10 to 20 seconds, depending on the mechanics and diagnostics performed by the drive each time it starts.

General Software's special Embedded BIOS firmware was able to achieve a POST time of 0.838ms, well under one second.

Windows Vista™ Optimization

- More sequential read command
  - Command: 60%, Size: 80%
  - Over last 15 years, OS optimized to be HDD-friendly
- Increased CPU operation time not related to Disk I/O

- Boot time decreases 10 seconds, even though SSD actually reduces disk service time by 20 seconds
- I/O bound with HDD, while CPU bound with SSD
- More OS optimization required to attain SSDs' maximum performance benefits

10 sec

		HDD	SSD
Boot time		34 s	24 s
Read	Size	210 MB	295 MB
	Time	26.59 s	7.06 s
Write	Size	7.19 MB	5.98 MB
	Time	1.18 s	0.93 s
Total disk time		27.77 s	7.99 s

20 sec

Conclusions

- **SSD to bridge the gap between system memory and HDD**
  - Performance boosting by using multiple chips
- **SSD as a primary storage system**
  - Overall System performance improvements
  - Less power consumption (battery life and energy savings)
  - Resistant to shock and vibration
  - High MTBF
  - Smaller form factor
  - TCO benefits
- **Robust wear leveling for long life expectancy**
  - Intelligent algorithm of Samsung proprietary wear-leveling
- **OS optimization for SSD**
  - Cooperation with OS vendors to maximize benefits of Samsung SSD