

Managing the evolution of Flash : beyond memory to storage

Tony Kim

**Director, Memory Marketing
Samsung Semiconductor Inc.**

**Nonvolatile Memory Seminar
Hot Chips Conference
August 22, 2010
Memorial Auditorium
Stanford University**

Contents

- **NAND Flash technology**
- **Flash storage management**
- **Flash storage architecture by apps**
- **Future trend**
- **Conclusions**

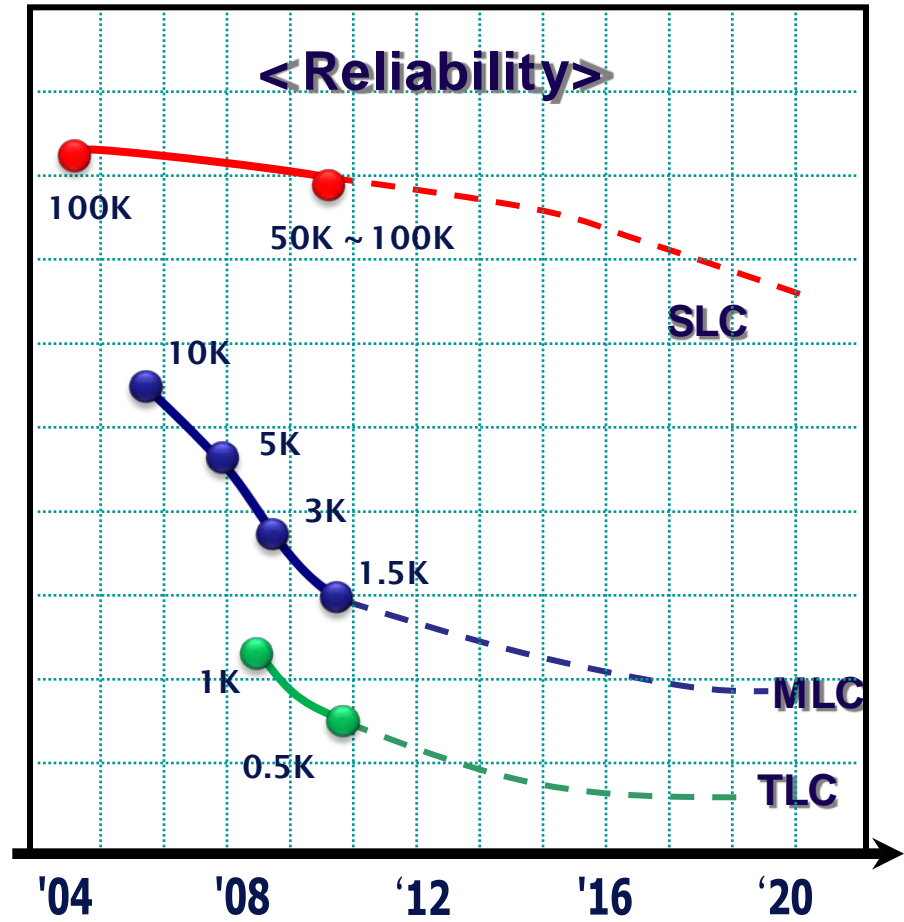
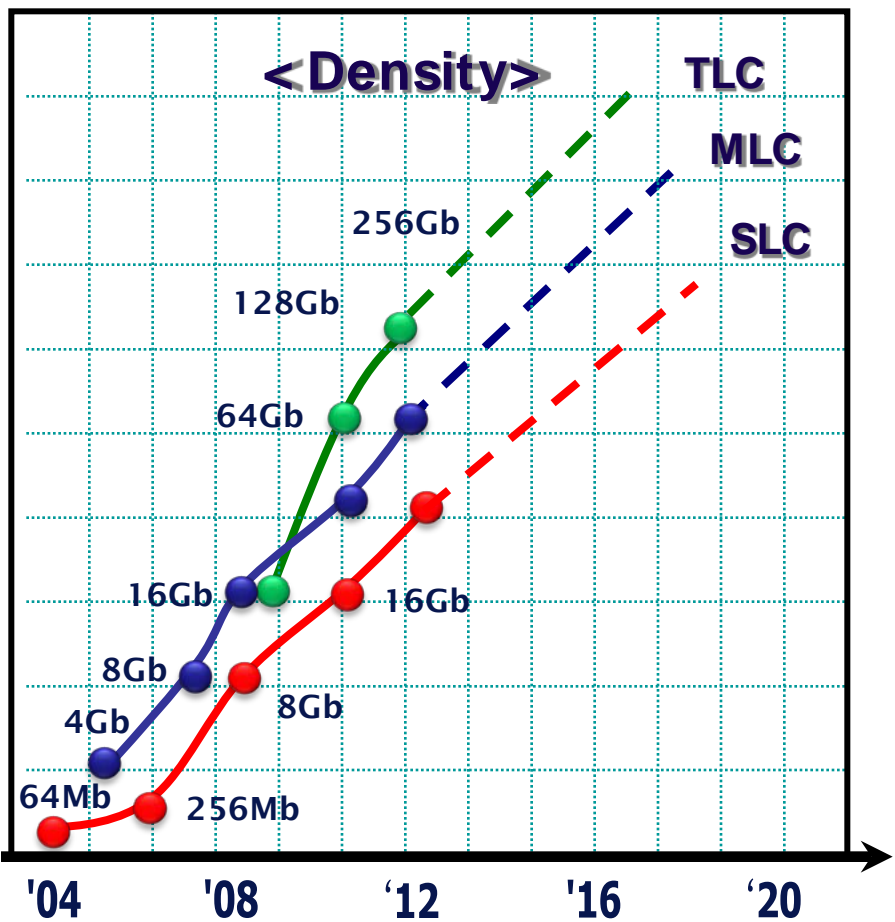
Contents

- **NAND Flash technology**
- Flash storage management
- Flash storage architecture by apps
- Future trend
- Conclusions



NAND Technology Shifts Smoothly

- Density has doubled every year since 2004
- There are breaking points on key technology beyond 2013
- Lithography shrink slows, NAND Reliability degrades

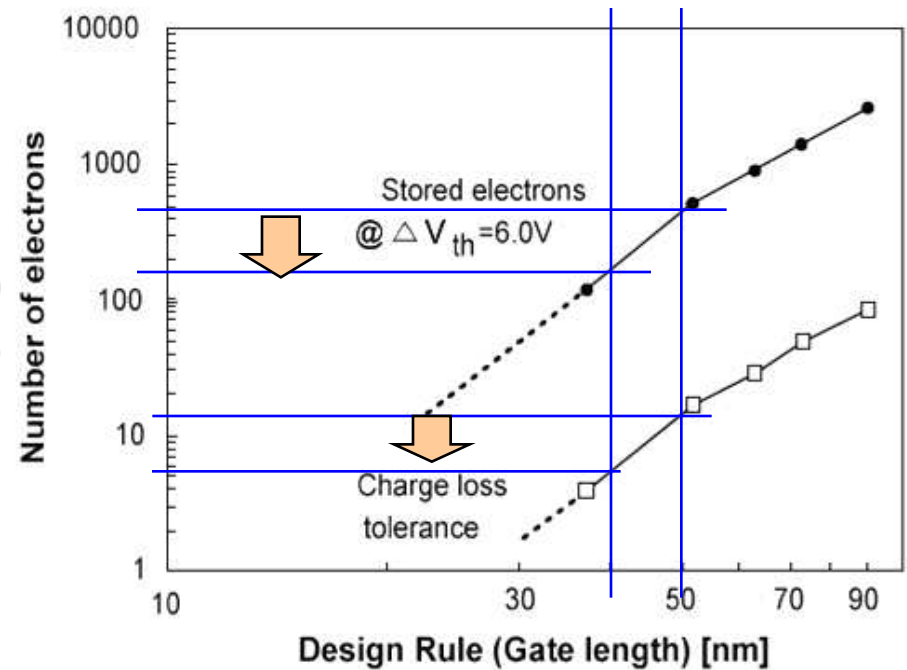
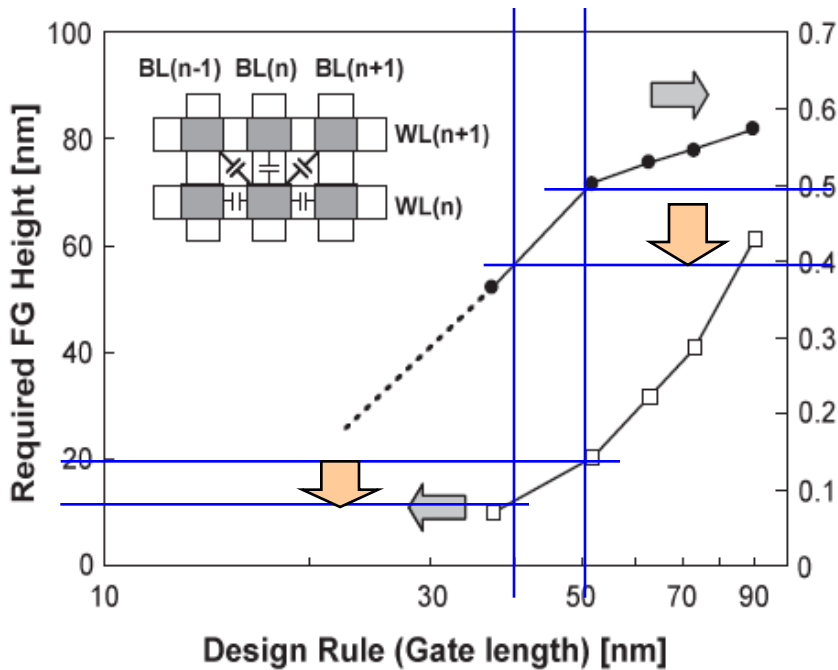


* SLC : Single level cell, MLC : Multi level cell, TLC : Triple level cell



Reliability Tolerance by Technology

■ Influence on Scaling-down of Floating Gate



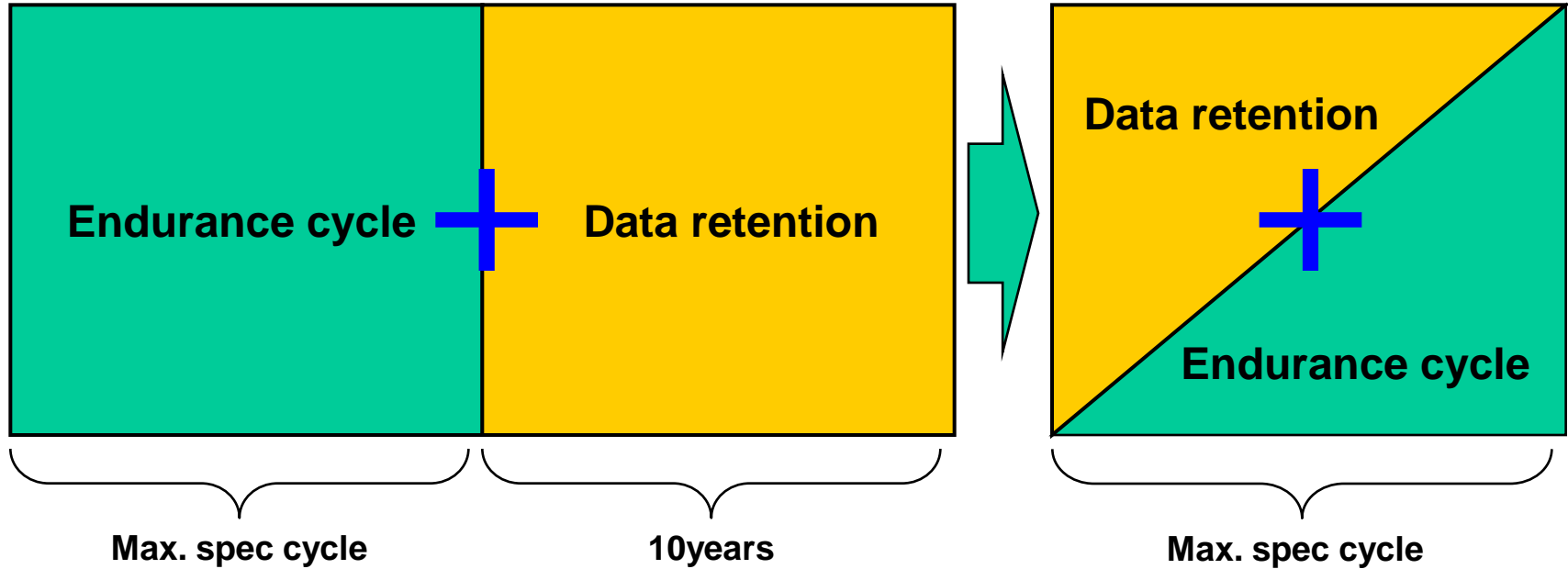
- Increase of F-Poly interference
→ Cell interference↑
- Decrease of coupling ratio
→ $V_{PGM/ERS}$ ↑

- Reduction of charge loss damage tolerance → Charge loss↑

Ref: YunSeung Shin, Symposium on VLSI Circuits, pp.156 – 159, 2005

JEDEC Standard : Cycle & Data Retention

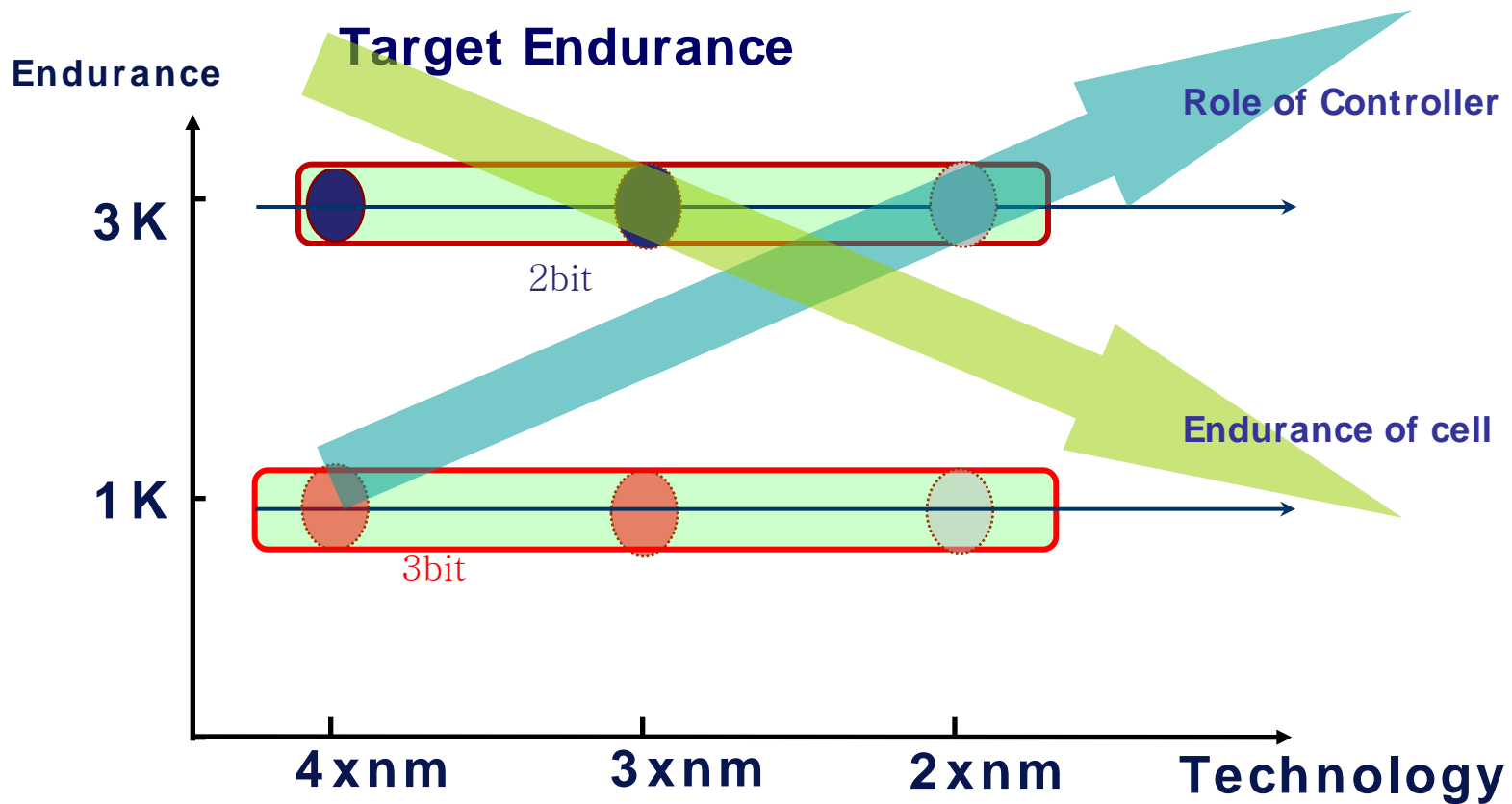
- Density Write operations should occur across device life time
- 10 year retention after lifetime write cycle is unpractical



Until 2006	As of 2007
100% END Cycle + 10y DTN	10% END Cycle + 10y DTN 100% END Cycle + 1y DTN

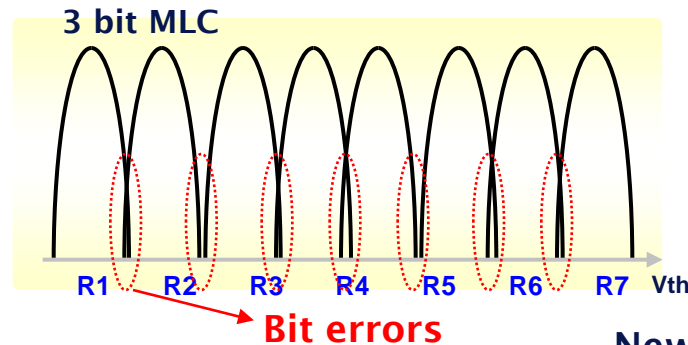
Controller : Critical for Maintaining Reliability

- More intelligence of controller can offset some of the generic degradation of NAND reliability from scaling

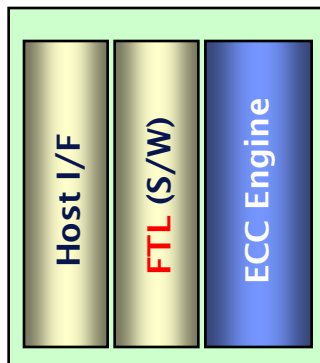


Technology Engine with ECC/ Bit Error/ FTL

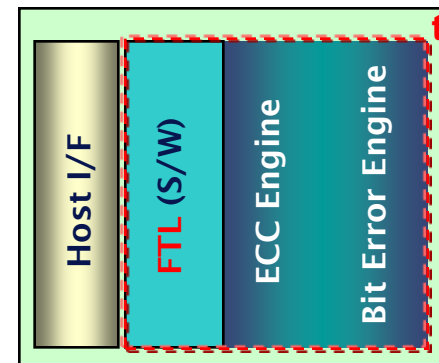
- Legacy controller only with ECC cannot reliably handle 2 bit in 3xnm and beyond, let alone 3 bit
- Optimization with Flash cell characteristics in mind is crucial
- New metric for reliability measurement is needed such as lifetime data amount with standard pattern



Legacy Controller



New Controller



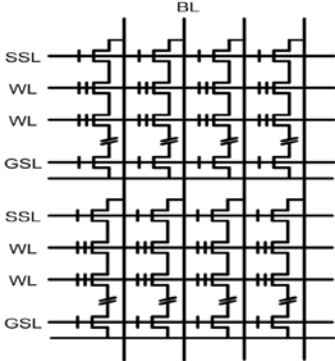
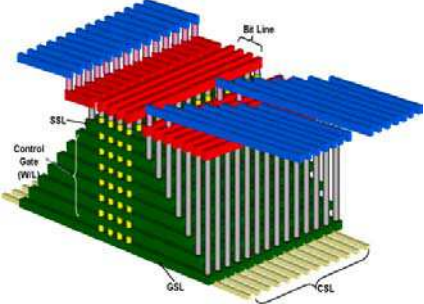
Melded together





Samsung's Innovative Flash Technology

- Samsung is exploring new technology to break status-quo
- Samsung believes 3D-NAND is the most likely successor for Planar NAND in the coming future

	Planar NAND	3D-NAND
Cell architecture		
Advantages	Easy to produce with simple process	High reliability Process reuse
Challenges	Hard to shrink under 1xnm	Hard to stack

3D-NAND details

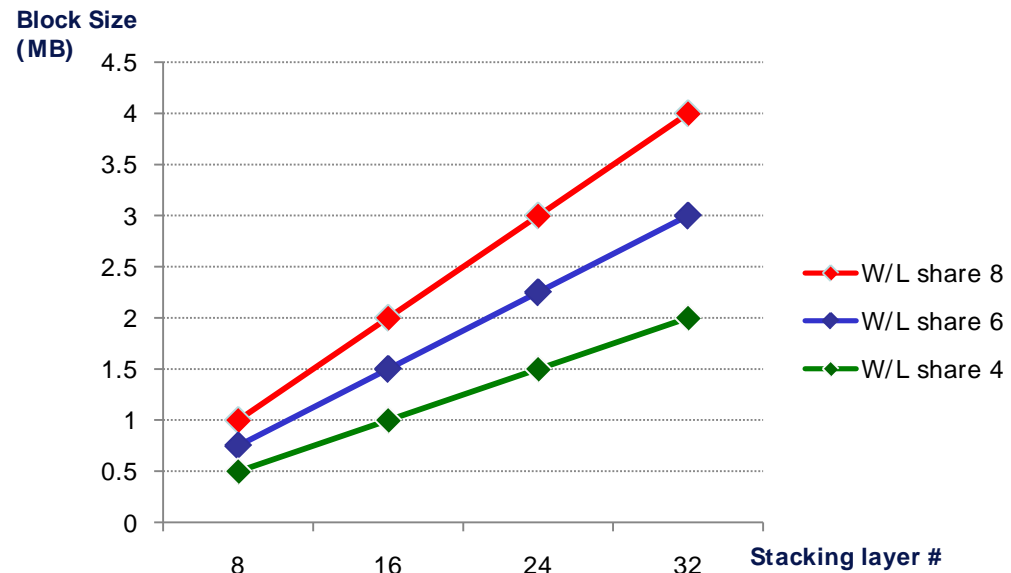
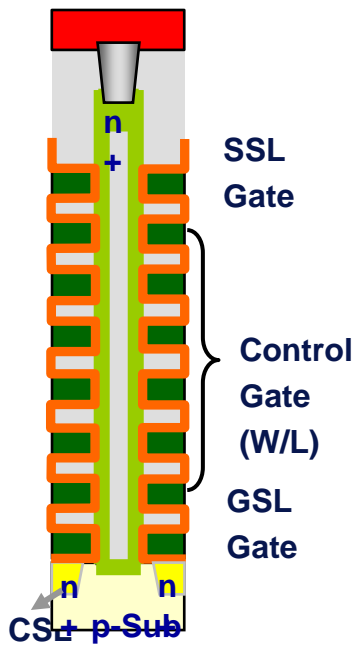


■ Potential benefit

- Better reliability/ endurance than planar since cell design rule is much more relaxed.

■ Issues in future scaling

- Bit cost reduction is done by increasing the stacking layer, thereby increasing by 2X per each generation becomes more difficult and unlikely
- Block size will be larger than the planar-equivalent



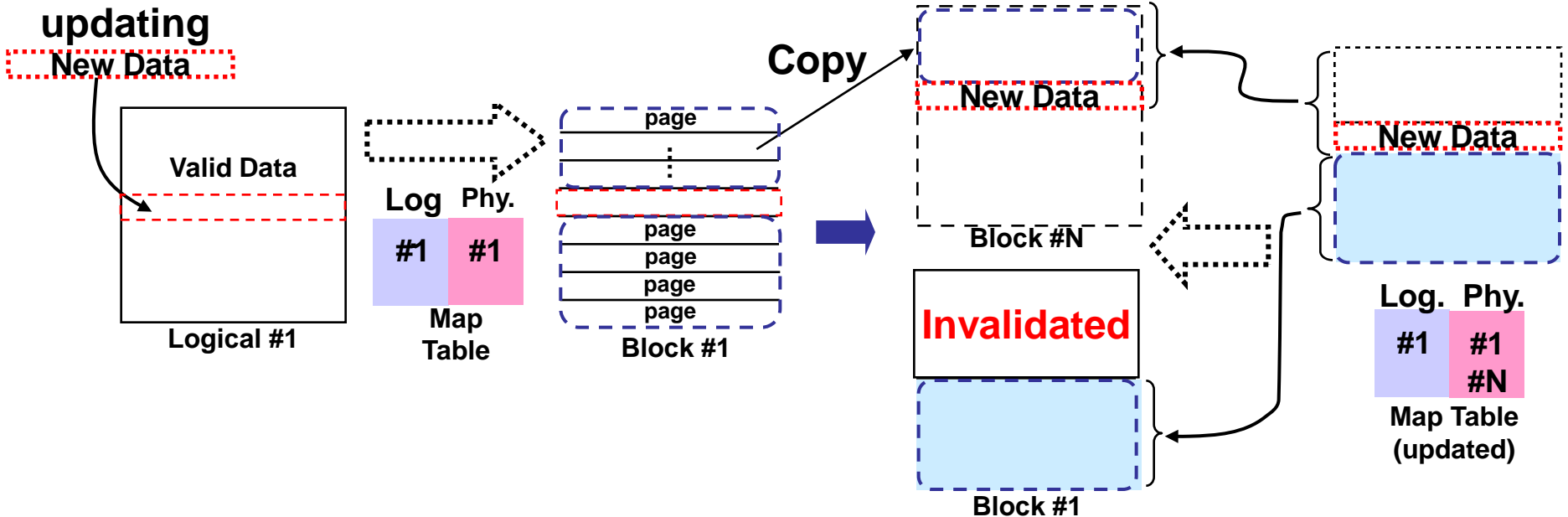
Contents

- NAND Flash technology
- **Flash storage management**
- Flash storage architecture by apps
- Future trend
- Conclusions

Why Software Is Needed for Flash?



- Small data unit in Program and the large unit in Erase requires another block newly allocated through mapping

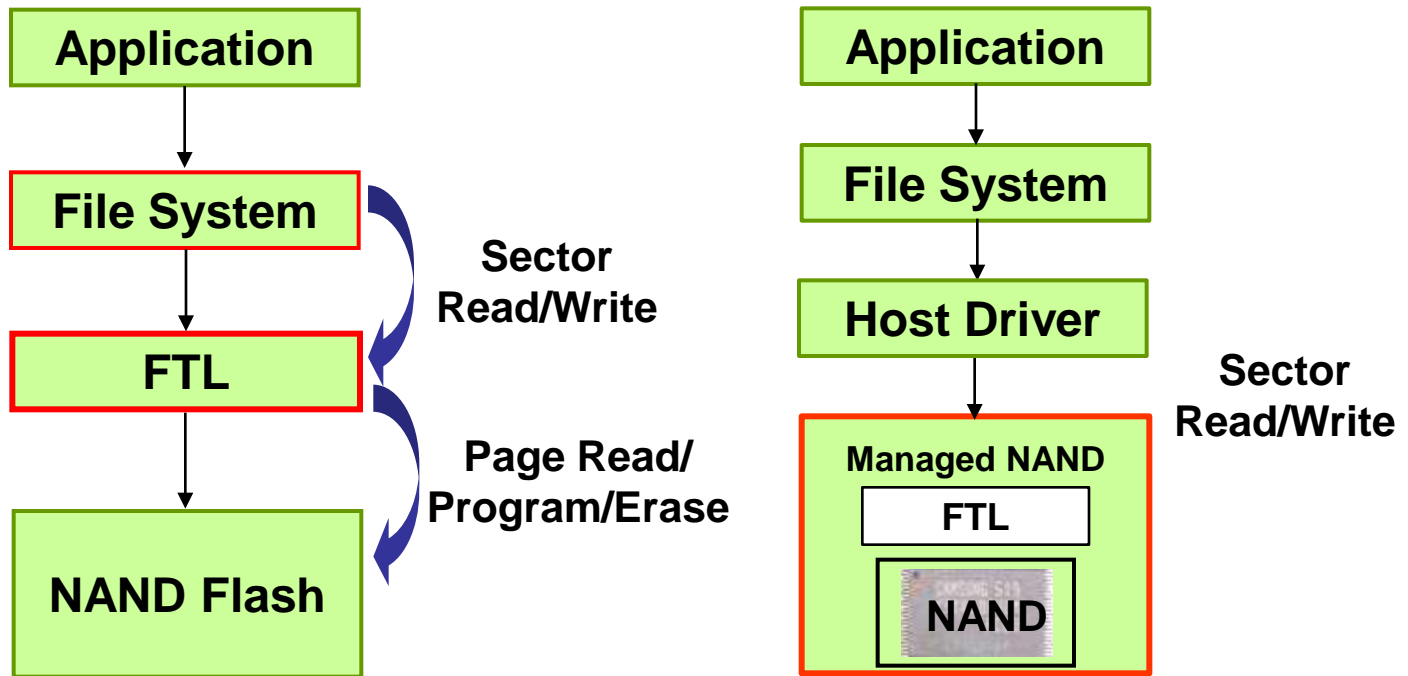


Logical block #1 is mapped to physical block #1 and #N after updating

FTL (Flash Translation Layer)

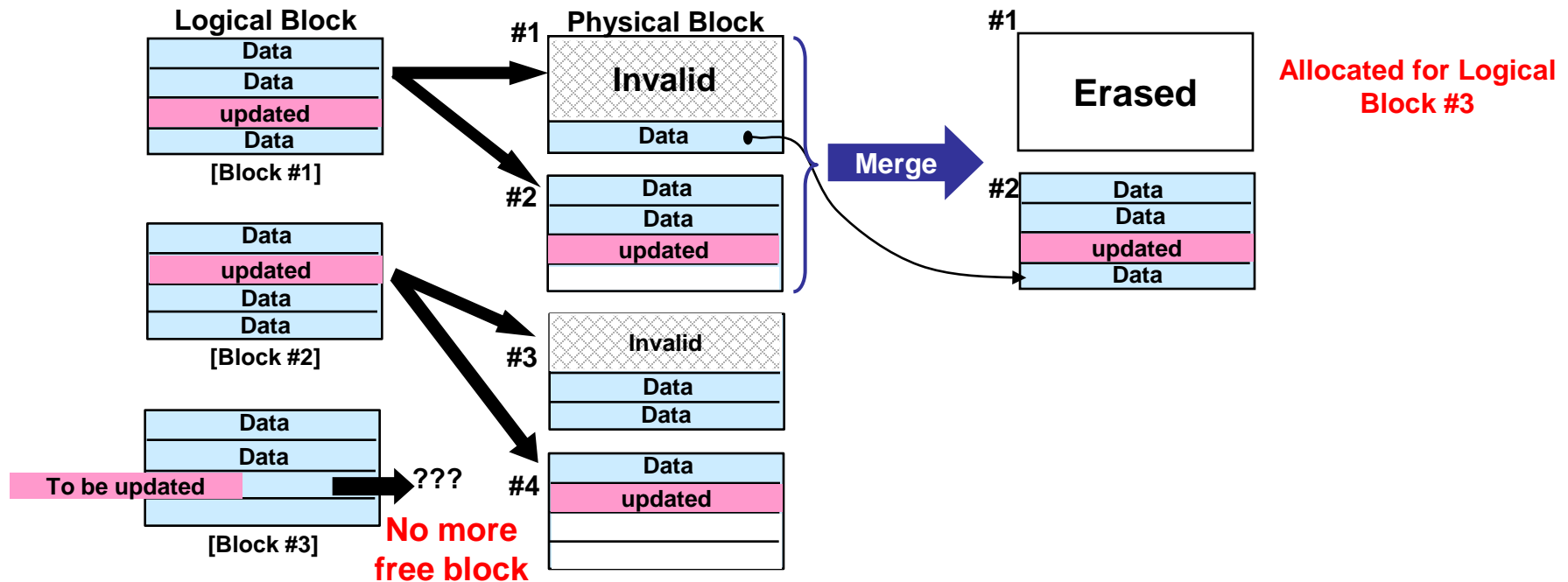


- Manages mapping from logical to physical address
- Detects and maps out Bad Blocks
- Does Wear-leveling for life extension



Reclaiming Valid Data : Garbage Collection

- Over time Flash is mixed with valid and invalid data
- Free space needs to be reclaimed to write new data
- Garbage collection merges the valid data from the scattered blocks



What is WAI (Wear Acceleration Index)?



- **WAI** : The index that represent how much FTL accelerate the wear-out of NAND

$$WAI = \frac{EraseCount_{Total} \times BlockSize_{Bytes}}{WriteSize_{Bytes}}$$

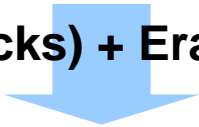
Example:

Write Request (2 blocks)



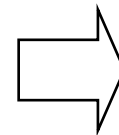
FTL

Write (4 blocks) + Erase (4 blocks)



NAND

WAI		NAND Life time		Meaning
High		Short		Bad
Low		Long		Good



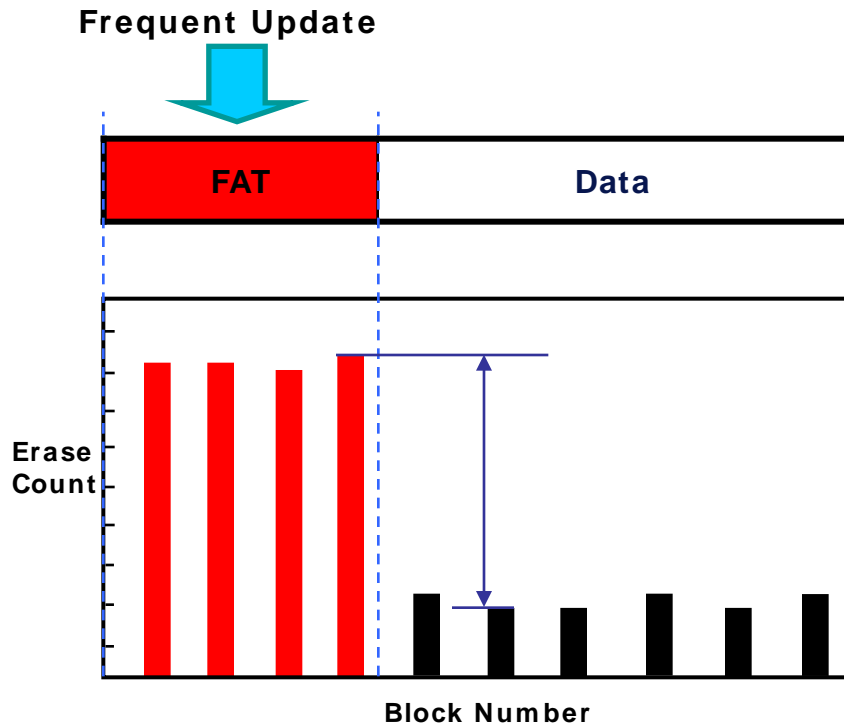
$$WAI = \frac{4 \times 128KB}{256KB} = 2$$

* Block Size: 128KB

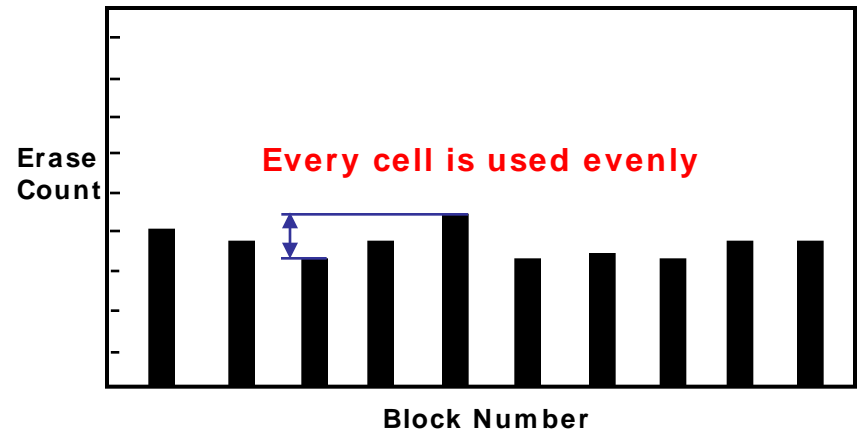
What is Wear-leveling ?



- Hot data like FAT can wear out certain portion of cell array
- Wear-leveling maximizes the life span of NAND flash as each cell is used evenly



Before Wear-leveling



After Good Wear-leveling

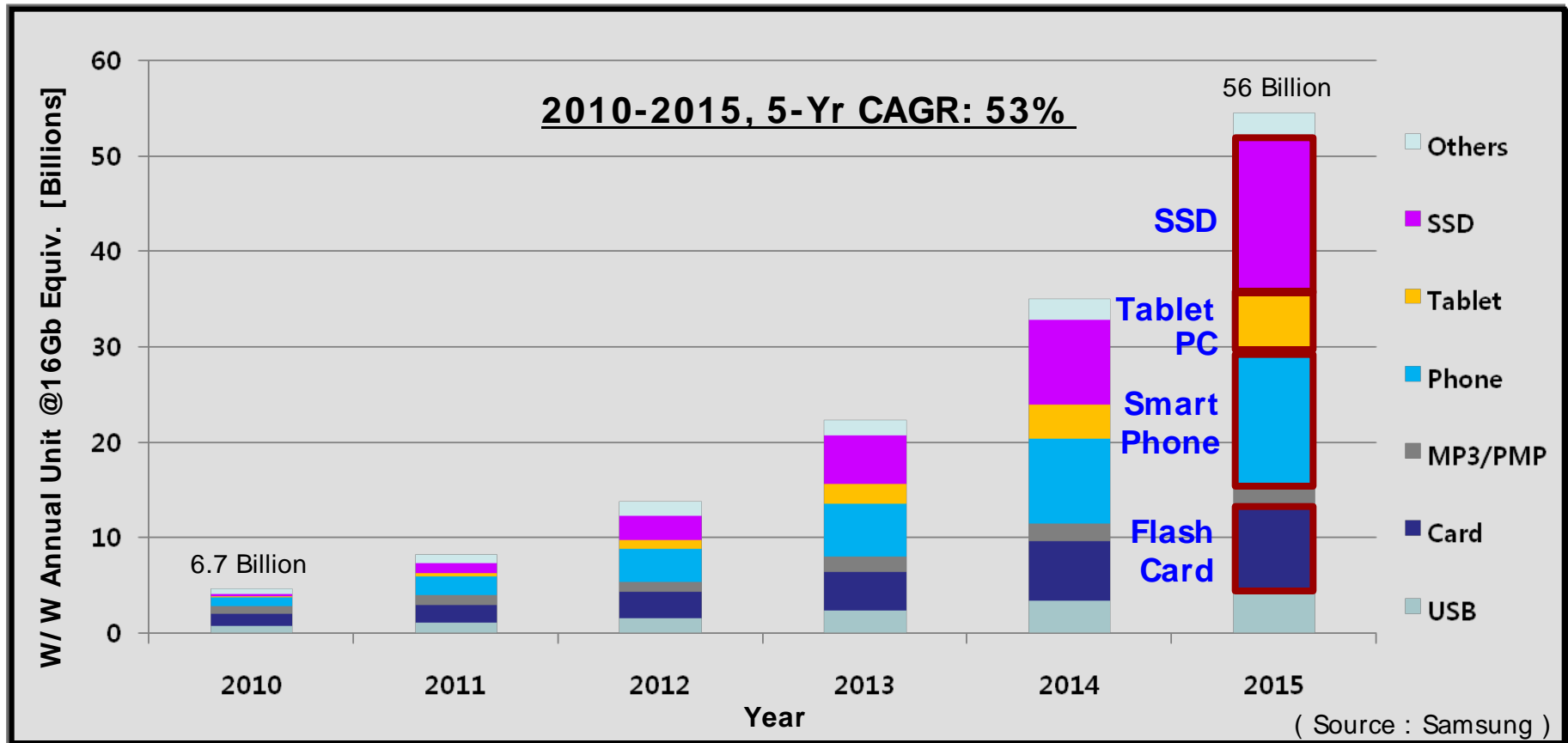
Contents

- NAND Flash technology
- Flash storage management
- **Flash storage architecture by apps**
- Future trend
- Conclusions



NAND Flash Market Outlook (`10 ~ `15)

- Samsung expected NAND Market CAGR of 53% between 2010 and 2015
- Key applications for NAND market growth for next decade are : Flash Card, Smart Phone, Tablet PC and SSD

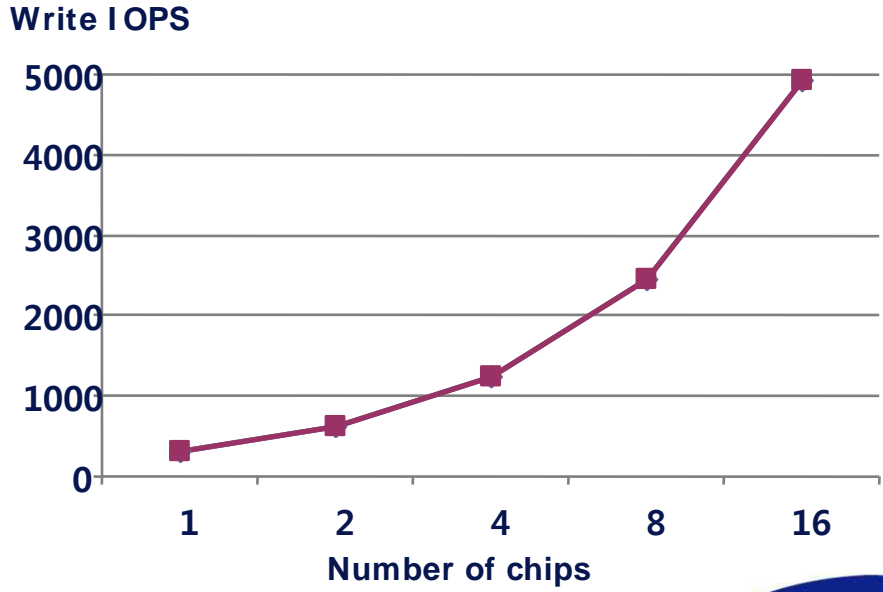
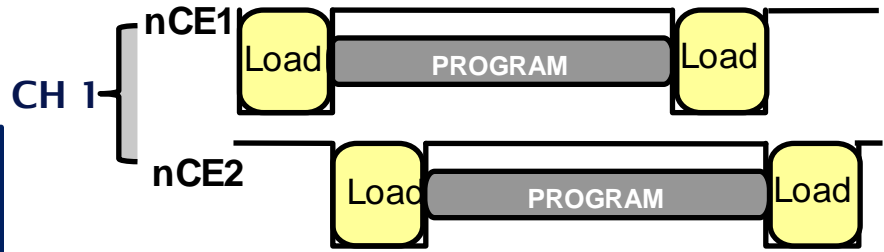
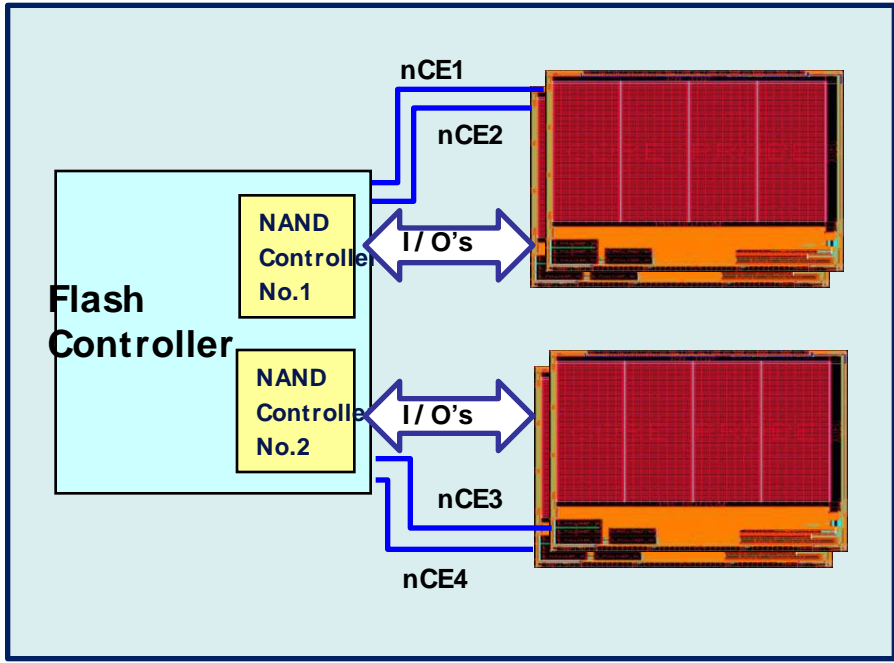




Performance Multiplied with Multi-Chips

- Flash storage performance can be easily expanded with multi-way write interleaving along with multi-channels
- During program Busy, data can be loaded and programmed into other NAND devices on the same bus

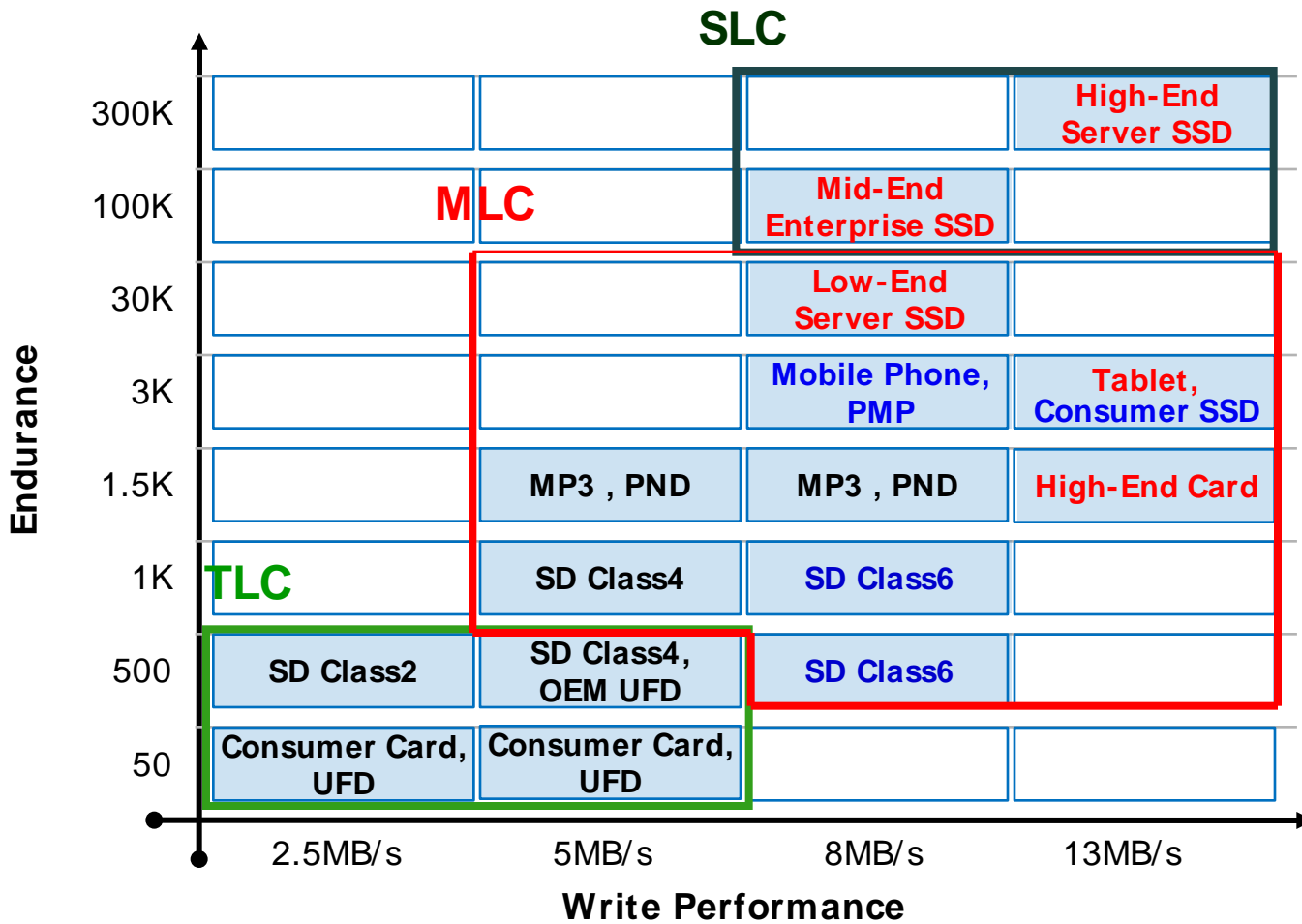
Flash Storage architecture





Application Segmented by Technology

- Applications will be fragmented by performance and reliability
- Closer communications needed to understand user requirements and tailor appropriate solutions



Contents

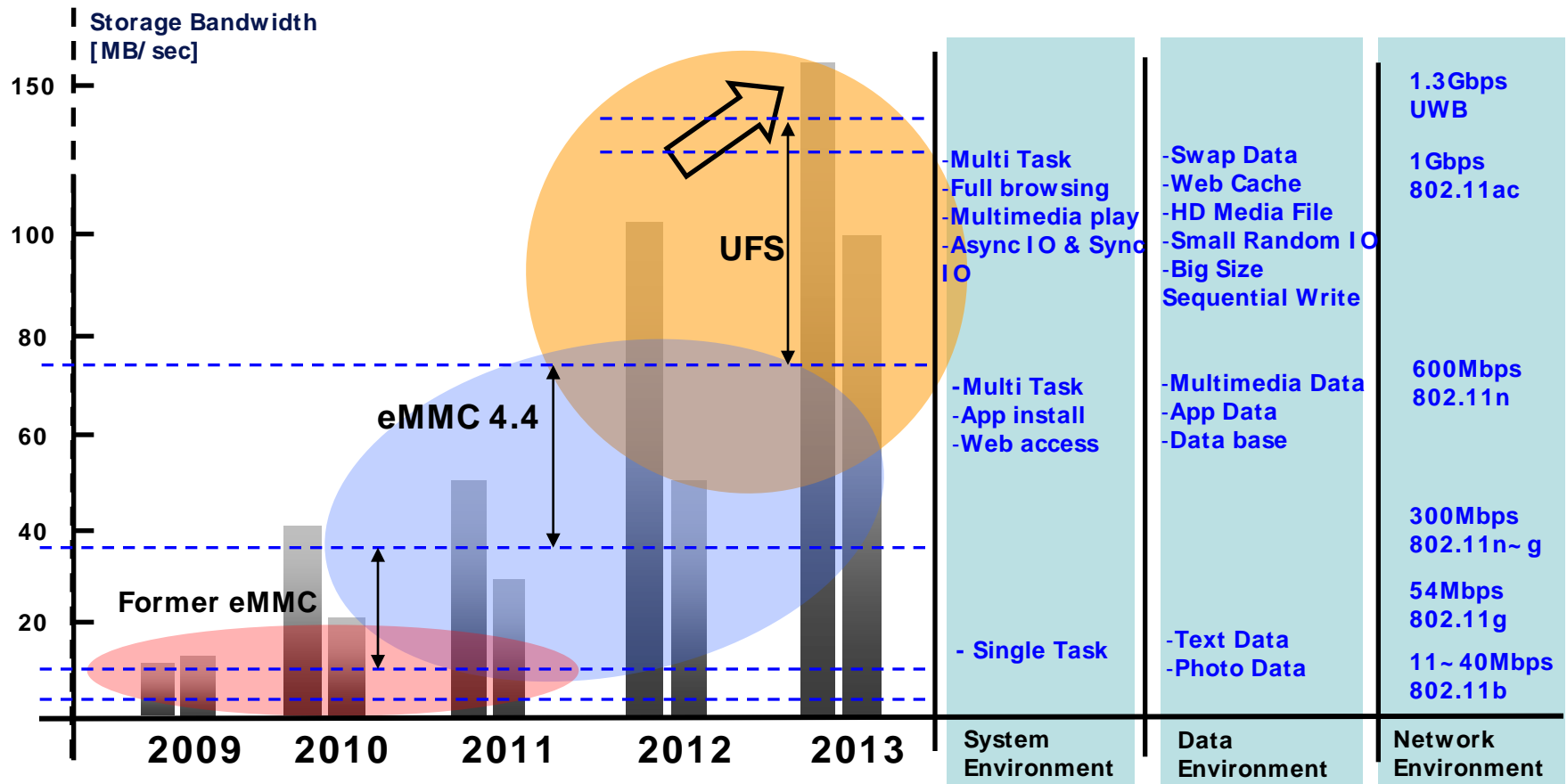
- NAND Flash technology
- Flash storage management
- Flash storage architecture by apps
- **Future trend**
- **Conclusions**



Storage Architecture Evolution : Technical Trend

■ New environment on the storage drives higher performance

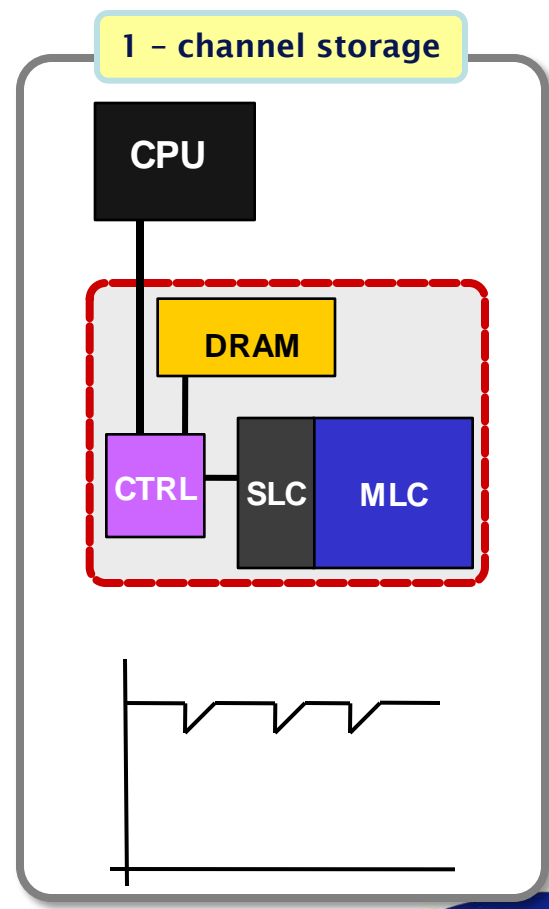
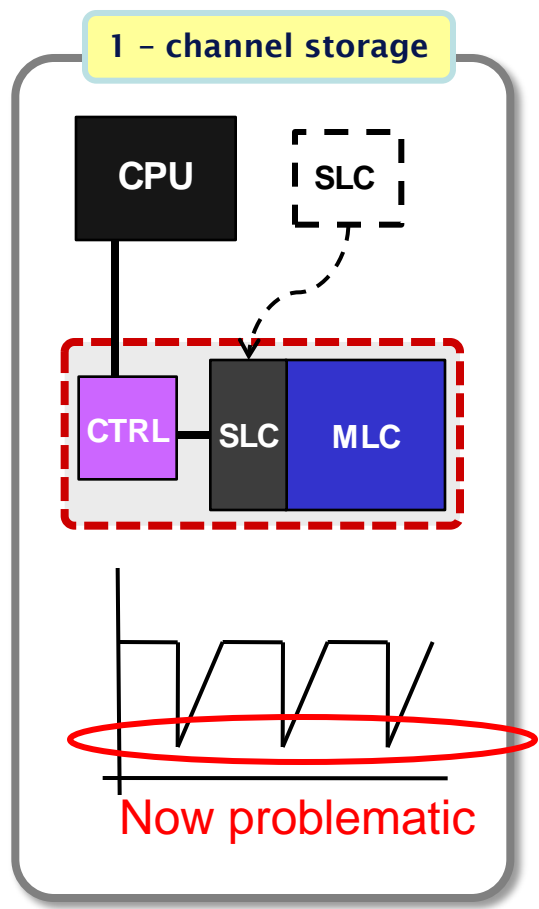
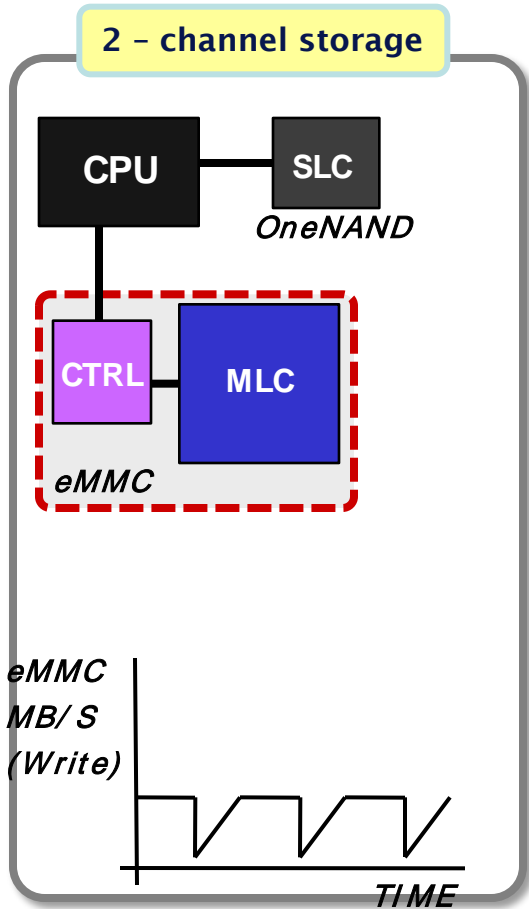
- Multi tasking with Swap IO
- High performance App with fast data IO
- High bandwidth Network with User Data IO





Architecture for high Performance and Low Latency

- Architecture of “OneNAND + moviNAND” can be unified to “moviNAND” only with fast random write and low latency
- Implementing HPI can resolve the problem



Preparing Host for New Features of eMMC & UFS

- The enhancements in read/ write performance, data integrity at sudden system-power failure and lower power consumption at idle are only realized with the relevant support of file-system, and OS in some cases
- Linux open source community is far behind in aligning to them
- Chipset and handset vendors should work out the responsibility details

Features	Purpose	eMMC	UFS
Trim	Write speed up	√	√
Reliable write	Data integrity at power-loss	√	√
HPI	Write-suspend for fast read access	√	√
Command queuing	Read/Write speed up		√
Sleep Power	Lower power at idle	√	√

Data Attributes in Mobile System



■ Data in mobile device has different data attribute

• System data/ Code data / Swap data

- System working, Multi-tasking, System application working
- High speed random I/O performance & data reliability are mainly required

• User data

- High density Multimedia data read & write
- Sequential I/O is mainly required

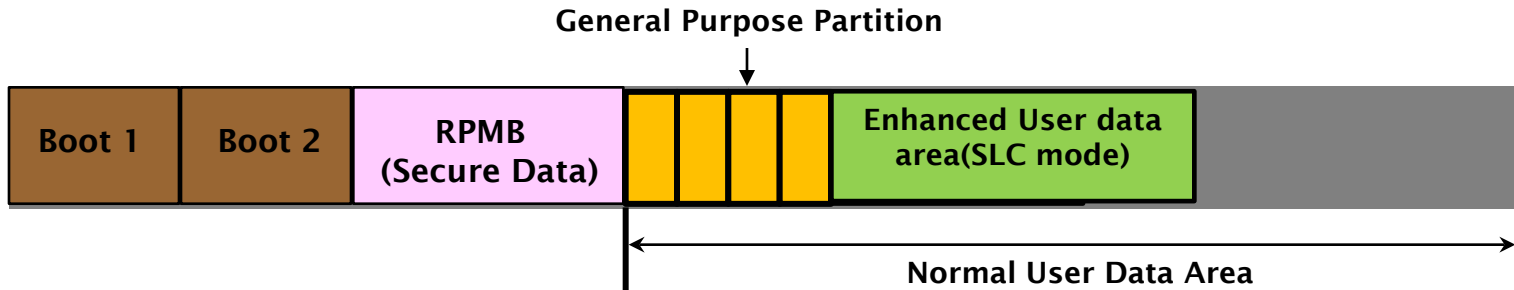
		code	System Meta	Swap	User
Read Centric	Sequential				✓
	Random	✓	✓	✓	
Write Centric	Sequential				✓
	Random		✓	✓	
Reliability		✓	✓	✓	

New Feature for e-MMC 4.4 : Multi Partition



Flexibility to host to manage eMMC 4.4

- 4 general purpose partitions and enhanced user data area can be set in normal user data area

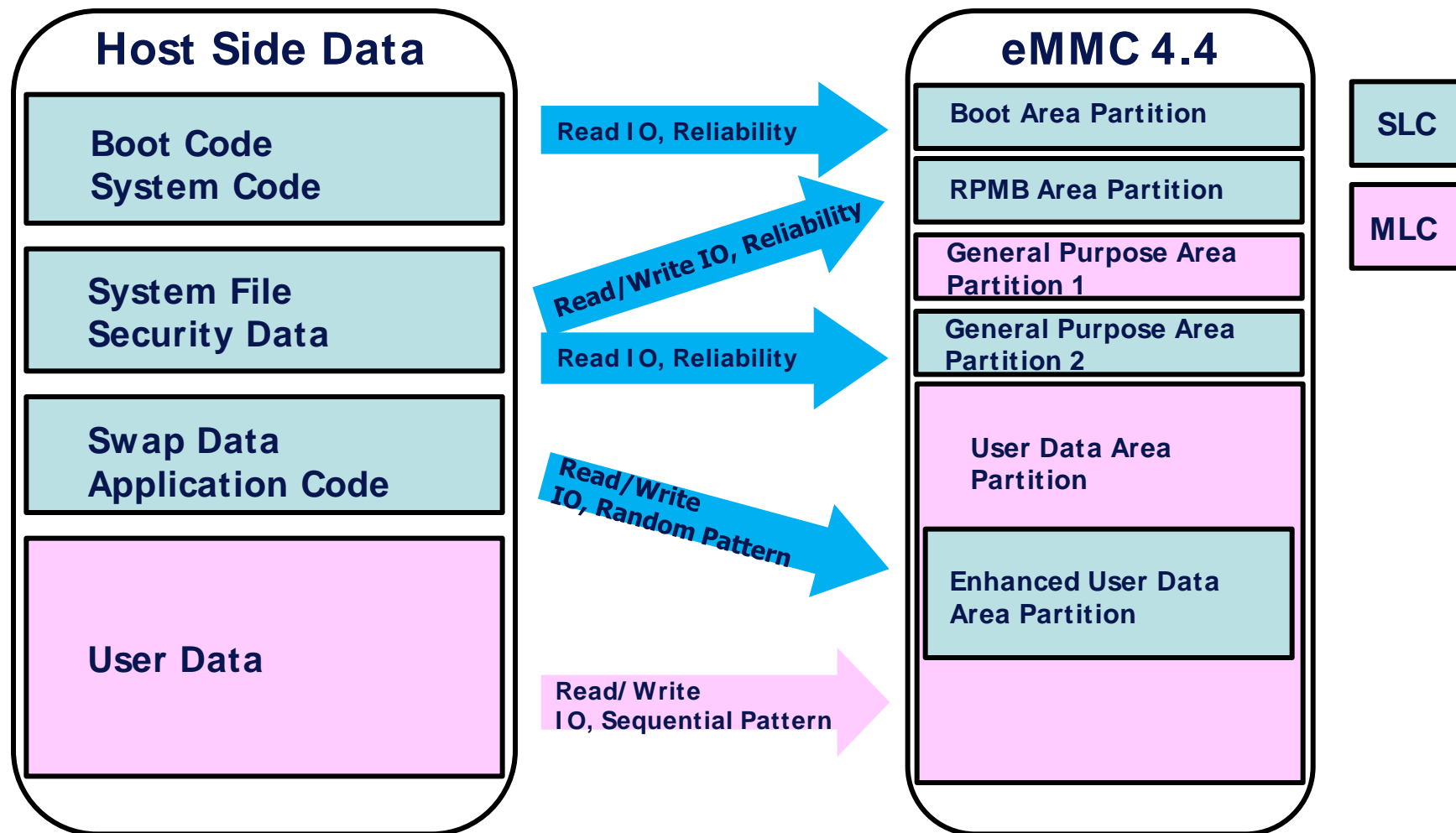


Partitions		NAND type	Default Size	Remarks
Boot Area Partition 1		SLC Mode	128KB	Size as multiple of 128KB (max. 32MB)
Boot Area Partition 2		SLC Mode	128KB	Size as multiple of 128KB (max. 32MB)
RPMB Area Partition		SLC Mode	128KB	Size as multiple of 128KB (max. 32MB)
General Purpose Partitions		MLC "or" Enhanced Area	0KB	Available size can be seen by following: $(EXT_CSD[145] * 8^2 + EXT_CSD[144] * 8^1 + EXT_CSD[143]) * HC_WP_GPR_SIZE * HC_ERASE_GPR_SIZE * 512KB$ byte
User Data Area	Enhanced Area	SLC Mode	0KB	Start address → multiple of Write Protect Group size
	Default Area	MLC	93.1%	

Data Usage Model in eMMC4.4



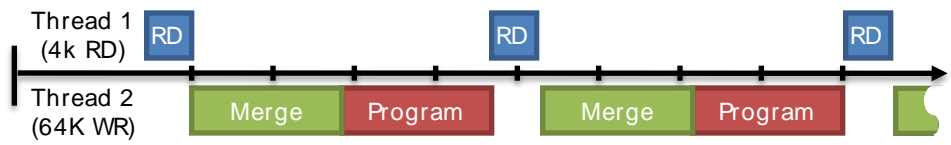
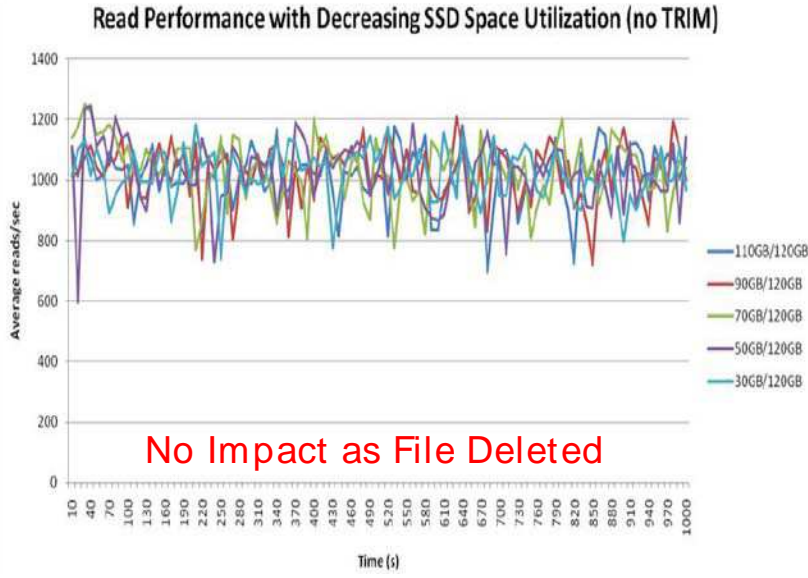
- SLC and MLC partitions to be tailored by use scenarios



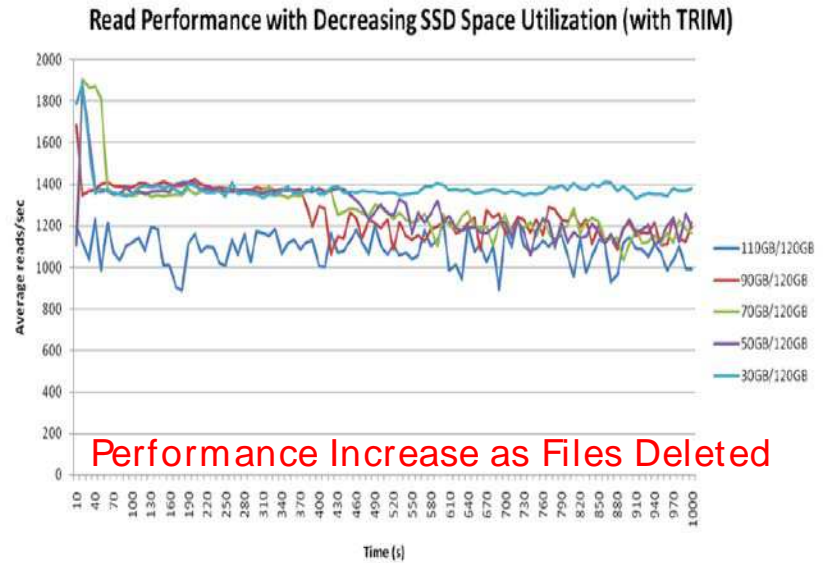
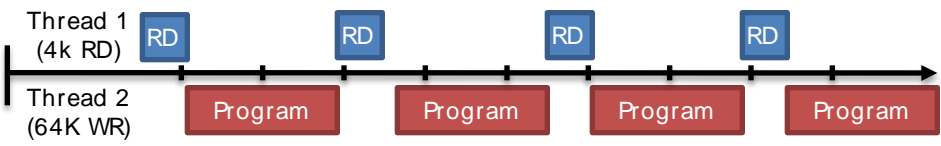


Better Multi-tasking Performance with Trim

■ Trim reduces long write latency and in turn read latency at multi-tasking. Mobile phones very likely having some free space benefit from Trim



(Source : Microsoft)

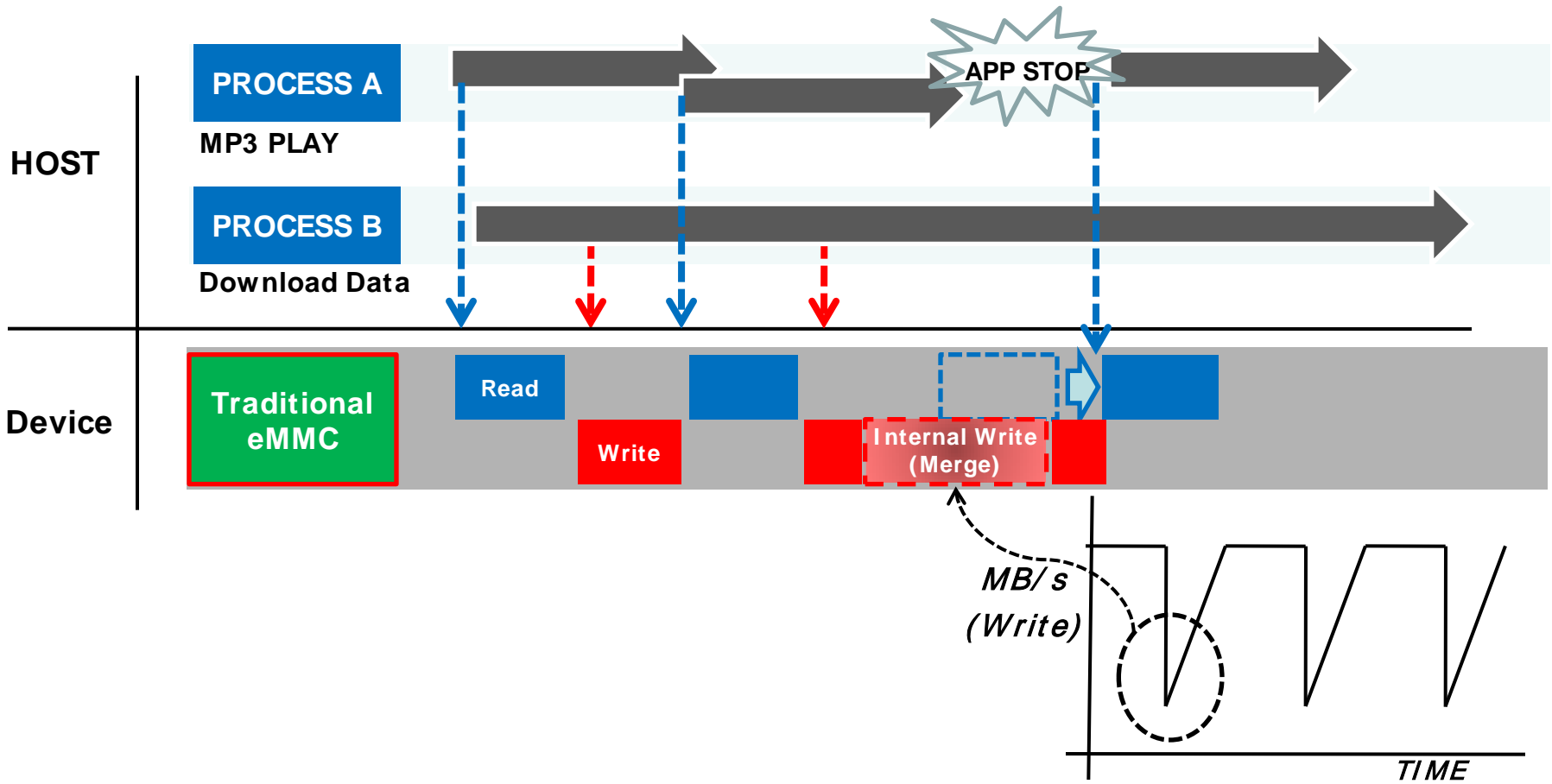


•Conceptual view of increased read thread due to long write thread with Busy



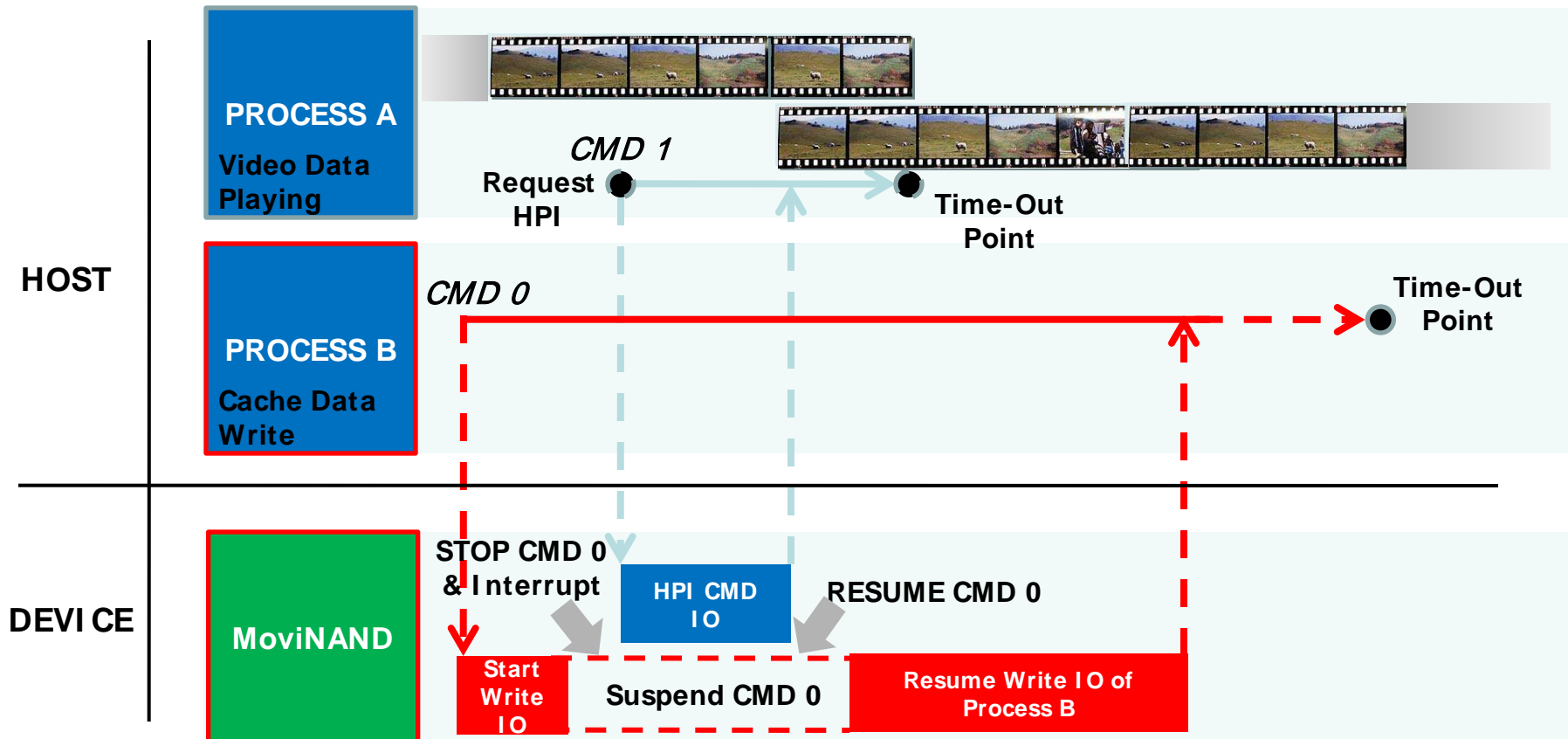
How Write Latency Affects for Multi-tasking

- Multi-tasking needs low write latency in a time-out value for uninterrupted audio/ video play-back



Interrupting Write-Busy Sustains Real-time Task

- Long write-busy of eMMC should be interrupted for high priority real-time task such as audio/ video play-back

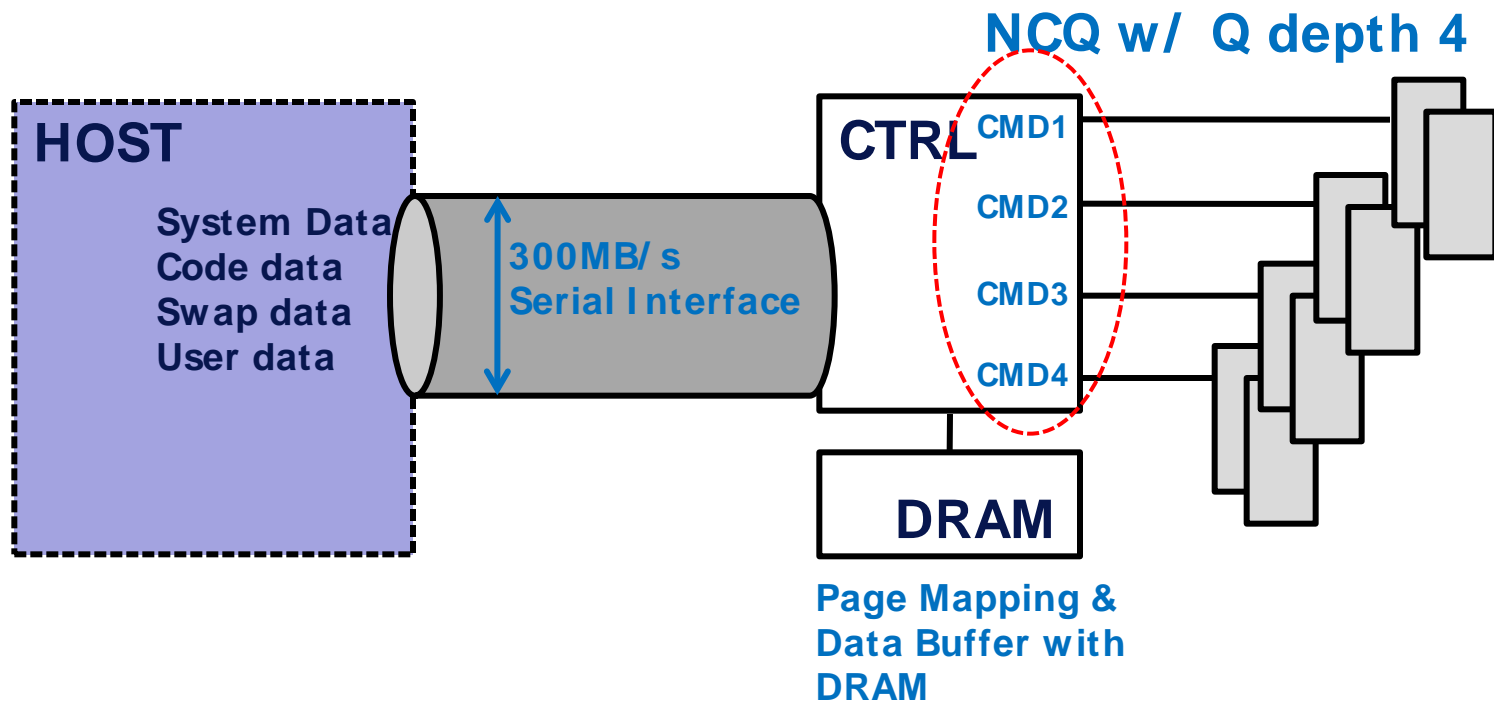


Next Generation Embedded Storage : UFS



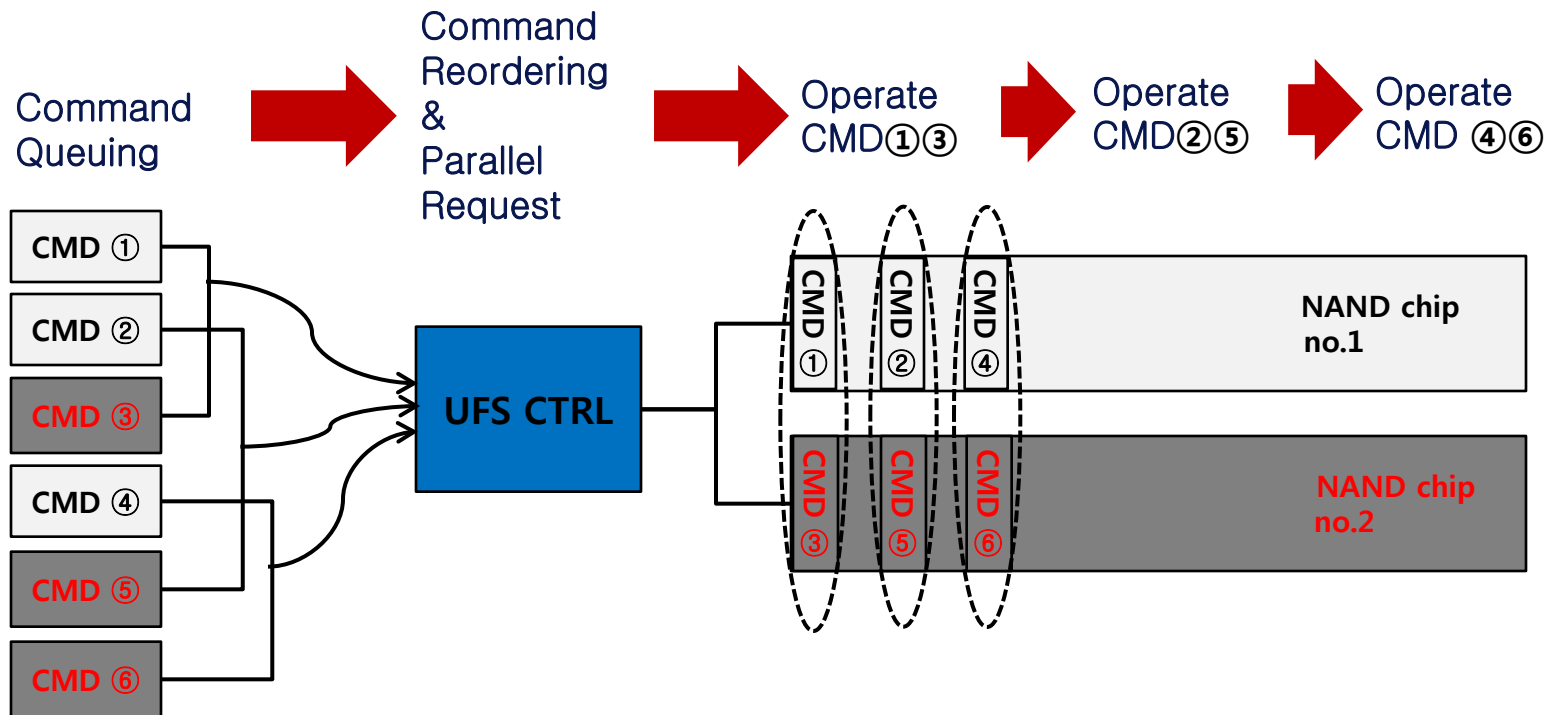
■ Universal Flash Storage based on serial M-Phy

- Serial interface : 300MB/s bandwidth
- Native Command Queuing : Support parallel NAND Flash working for Random/Sequential IO
- Page mapping with DRAM : Reduce internal Merge operation



UFS Command Queuing Enhance Performance

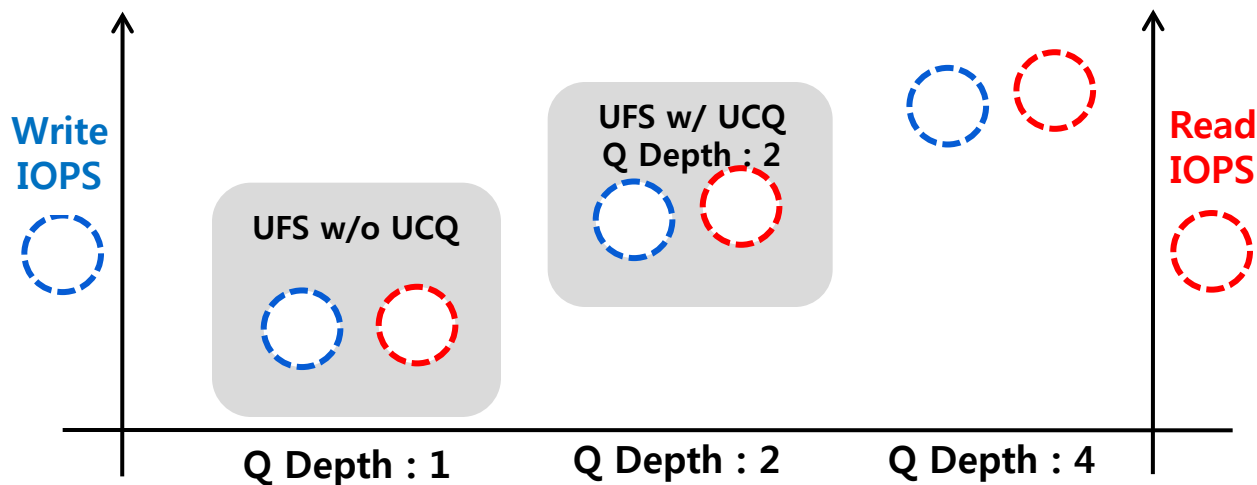
- Normal data buffered and written to different NAND dies in parallel
- Host system should indicate critical data like file-system meta to be written synchronously





Queue Depth & Random I/O Performance

- Random Performance depends on parallel I/O number



* IOPS(Input/Output Operation Per Second) is a common benchmark for computer storage media

- **Optimized Command Queue Depth**

- Depends on the number of NAND channel or way in UFS Device
- Consider Data loss while sudden power off

- **Register for Dynamic setting of Command Queue Depth**

- Add to UCQ mode page



UFS Standardization Roadmap

- UFS version 1.0 is only the baseline
- Additions for performance, low power and reliability will be put into the follow-up spec

	2010		2011			
	Q3	Q4	Q1	Q2	Q3	Q4
UFS Specification	Version 1.0			Version 1.x		Final
e-MMC features ERASE/ TRIM/ Reliable write...	1 st showing (Aug. interim meeting)	2 nd showing				
NCQ	2 nd showing					
Hybrid Power	2 nd showing					



- **Various NAND Flash technology target different markets with different focus in architecture and performance, resulting in wider product portfolio**
- **High performance mobile systems should tailor the use of different NAND Flash technology of single storage with regards to performance and reliability for various application scenarios**
- **NAND-friendly system-level solutions such as Trim, HPI as well as Reliable Write on eMMC will play a critical role in managing high reliability and performance**
- **Command queue architecture on UFS provides a path to future high performance by significantly increasing effective IOPS**