

Storage Controller Technologies for Next Generation Non-Volatile Memories

Hanan Weingarten

Aug. 9, 2016



Preface

- Two parallel trends
 - Non-Volatile memory devices:
 - Proliferation of different types of non-volatile memories
 - Different reliability / performance capabilities
 - Varying prices
 - SSD management is moving to the host
 - Shared resources
 - Predictable performance
 - Application dependent
- Proposal:
 - Single and generic physical layer interface:
 - Support all NAND device characteristics
 - Allow different reliability / performance boosting schemes
 - Interface should be flexible enough to allow different tradeoffs per application

Outline

- Non-volatile memories
- Reliability & Performance boosting methods
- New physical layer interface

Non-volatile memory options (1)

- 1Znm planar NAND:
 - Considered to be the last technology node
 - 1Znm TLC price is still lower than 48L 3D NAND
 - Reliability:
 - TLC required soft-decoding
- 3D NAND:
 - Accommodate TLC and probably QLC
 - Higher performance (more planes / one shot programming)
 - Higher reliability:
 - 3D-TLC equivalent to planar MLC
 - 3D-QLC equivalent to planar TLC: requires soft-decoding

Non-volatile memory options (2)

- Non-NAND Flash
 - MRAM, 3D-Xpoint, ...
 - Not all available today
 - Much faster than NAND (x1000 – x10000)
 - Priced higher, closer to DRAM
 - Applications:
 - Extremely high performance storage
 - DRAM replacement with non-volatile capabilities
 - Non-volatile data buffering / caching
 - Reliability:
 - Extremely good. Very simple ECC required

Non-volatile memory options (3)

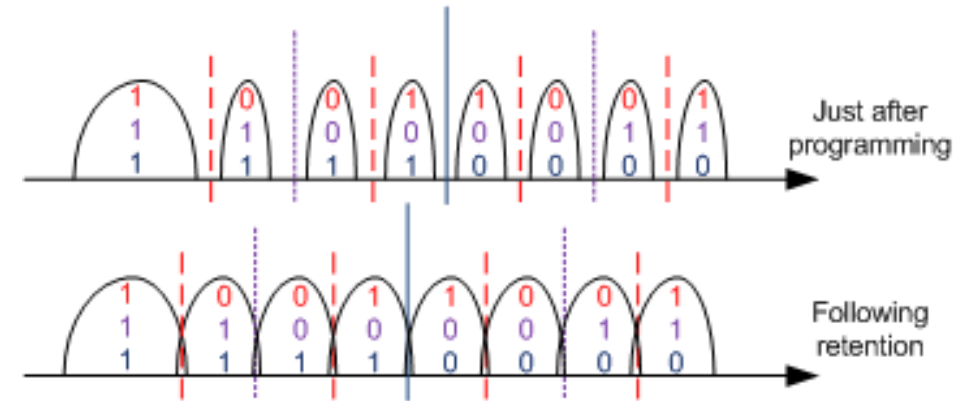
Memory Type	Performance	Reliability	Price
Planar NAND 1Znm MLC	X1	X1	X1
Planar NAND 1Znm TLC	X0.5	X0.1	X0.7
3D NAND MLC	X2	X2-X10	X2
3D NAND TLC	X1	X1	X1.5
3D NAND QLC	X0.5	X0.1	X1
3D Xpoint	X1000	X1000	X20
MRAM	X10000	Infinite	?

Reliability & Performance Boosting Methods (1)

- NAND reliability may be an issue
 - Planar NAND devices are less reliable than 3D-NAND
 - More bits/cell less reliability
 - Several methods may be applied to boost reliability:
 - Read flows
 - Striping
 - Location dependent code-rate adaptation
 - Management adaptation over life-time
- Performance boosting
 - Low Latency Flash Codes may boost performance

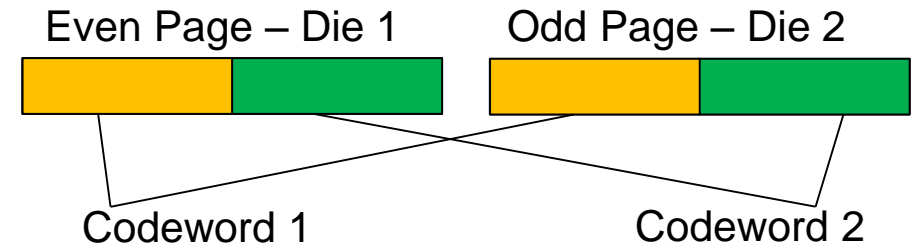
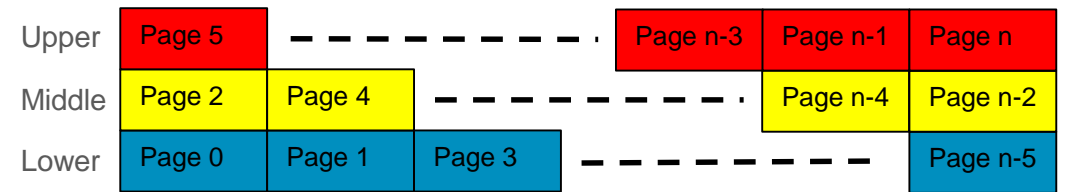
Reliability Boosting Methods (2)

- Read flows that improve reliability:
 - Read Threshold:
 - Optimization of read thresholds, depending on state
 - Soft decode:
 - Support of soft decoding allows attempting more reliable reads after simple, hard decode fails. Performance is traded off
 - Soft sampling algorithms may be used to reduce BER before soft sampling is applied
 - Other flow controls:
 - Modifying read parameters to induce less wear. Initial BER is traded off



Reliability Boosting Methods (3)

- Codeword striping:
 - Different locations within a NAND block exhibit different error rates
 - Example: Planar and 3D TLC NAND
 - Striping codewords between different page types on different dies:
 - Averages error rates
 - Significantly reduces maximum error rates
 - **Improves reliability by X2**
 - Data buffering allows very efficient striping schemes
 - e.g. on planar TLC where data is buffered to SLC and then folded to TLC

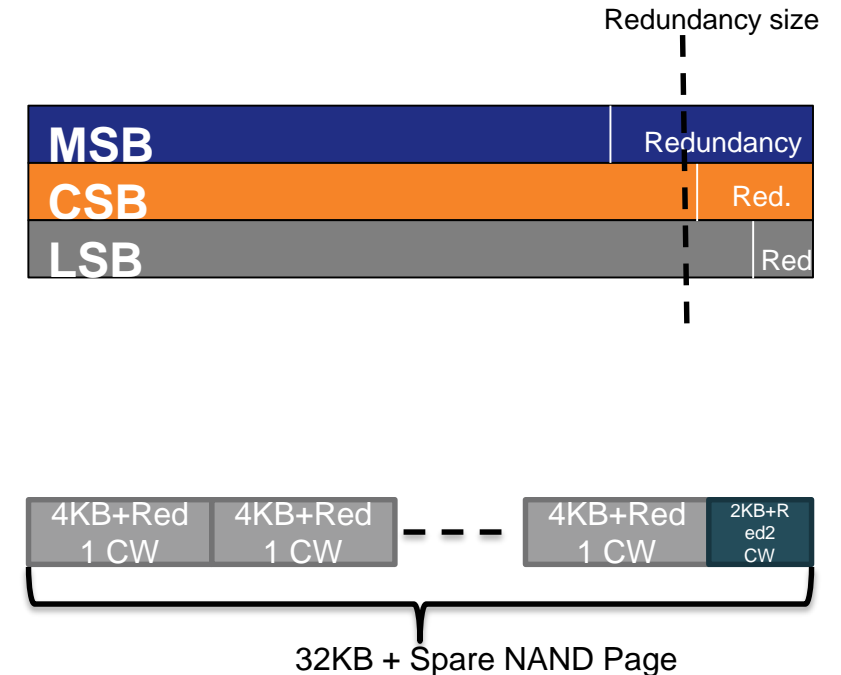


Reliability Boosting Methods (4)

- Codeword striping (cont.):
 - Pros:
 - Improved reliability (upto x2)
 - No throughput loss
 - Robust:
 - Different stresses exhibit different error rates variability across the block
 - Cons:
 - Loss in attainable random IOPs (two dies are occupied in a single read operation)

Reliability Boosting Methods (5)

- Variable rates
 - On the fly code reconfiguration with variable code rates (& payloads) according to bit error rate
 - Up to x2 gain in reliability
 - Semi-aligned codewords to eliminate performance penalty on read IOPs
 - Negligible decode latency penalty
 - Marginal impact on NAND channel utilization



Reliability Boosting Methods (6)

- Variable rates (cont.):
 - Pros:
 - Improved reliability (upto x2)
 - No throughput loss
 - Negligible random performance loss
 - No buffering required for efficient rate allocation optimization
 - Cons:
 - Generality:
 - Different stresses exhibit different error rates variability across the block. Variable rates will typically be optimized to work better under certain types of stresses

Reliability Boosting Methods (7)

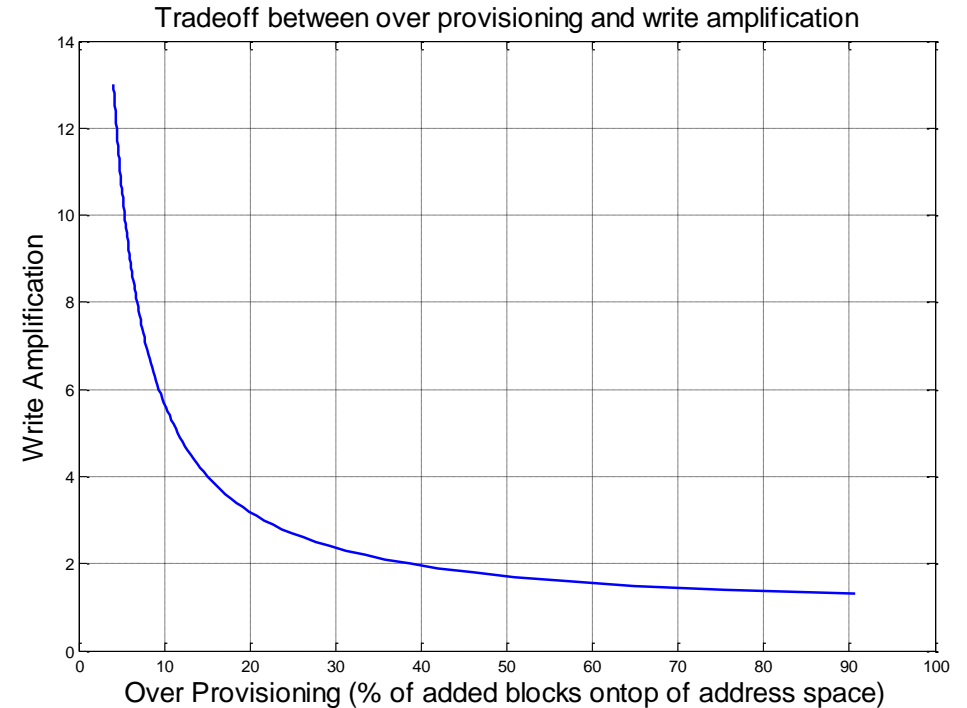
- System life-time adaptation:
 - General problem:
 - Error rates increase with P/E cycles
 - Solutions for handling higher error rates:
 - ECC with lower codes rates:
 - Each NAND block contains less data and more redundancy to support higher error rates
 - Optimized codeword striping to reduce peak error rates
 - Optimized variable rate coding
 - Higher frequency of soft decoding

Reliability Boosting Methods (8)

- System life-time adaptation (cont.):
 - Tradeoff performance to reliability
 - As device ages, change:
 - Non-striping scheme to optimized striping schemes
 - Codeword size increase 2KB → 4KB → 8KB
 - Application use may change with age
 - Random IOPs intensive to sequential intensive
 - Tradeoff capacity to reliability
 - At beginning of life use higher code rates:
 - More NAND space for data payload
 - Allocate larger amount of payload for user data
 - At end of life use lower code rates
 - User space is smaller
 - Capacity variable storage may be handled by application sensitive management

Reliability Boosting Methods (9)

- System life-time adaptation (cont.):
 - Tradeoff code rates to system over provisioning
 - At beginning of life use higher code rates:
 - More NAND space for data payload
 - Allocate constant amount of payload for user data
 - Allocate more of payload for over provisioning → reduce write amplification
 - NAND device P/E cycles increase at a lower pace due to increased over provisioning
 - At end of life use target code rates to support worst error rates
 - Allocate target over-provisioning as defined by:
 - Performance SPEC
 - Reliability SPEC
 - Practically allows improving system DWPD by 10%-30%



Performance Boosting Methods (1)

- Low Latency Flash codes:

- Goal: Enable reading all pages in an MLC/TLC NAND with SLC like latency

- NAND page read times

- Standard MLC NAND requires comparing 3 thresholds to read out 2 pages:

- Lower page: 1 threshold comparison
- Upper page: 2 threshold comparisons

- TLC NAND requires comparing 7 thresholds to read out 3 pages. Scheme is dependent on NAND:

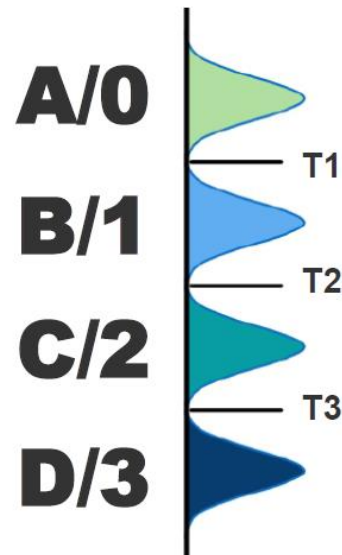
- Lower page: 1 or 2 thresholds
- Middle page: 2 or 3 thresholds
- Upper page: 4 or 2 thresholds

- Read time depends on the number of threshold comparisons



Performance Boosting Methods (2)

- Latency performance may be improved:
 - MLC and TLC rows used to store 3 and 7 pages
 - A single threshold comparison is used to read each page (an SLC read)
 - LLFC is used to enable 2 bit/cell and 3 bit/cell row division into 3 pages and 7 pages
- 2-bit MLC / 3 page example:



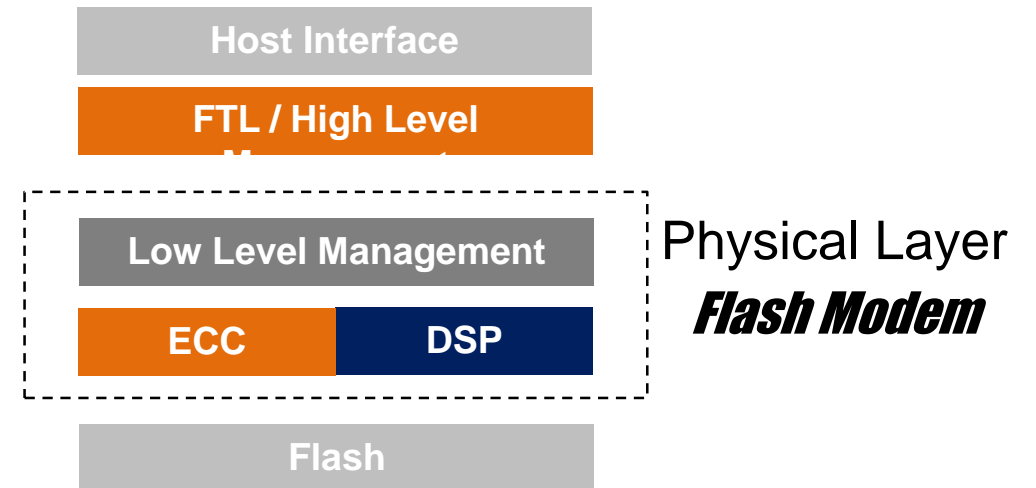
P1	P2	P3	Media	T1	T2	T3
0	0	0	CB	00	01	11
0	0	1	DB	00	01	01
0	1	0	BC	00	10	11
0	1	1	BD	00	10	10
1	0	0	CA	01	01	11
1	0	1	DA	01	01	01
1	1	0	AC	10	10	11
1	1	1	AD	10	10	10

Performance Boosting Methods (3)

- LLFC:
 - Pros:
 - Lower QD1 latency
 - Cons:
 - Loss in capacity
 - LLFC requires redundancy

Physical Layer Interface (1)

- FTL is divided:
 - Physical layer:
 - Reliability / ECC
 - Optional: Bad Block management
 - High level:
 - Data mapping
 - Garbage collection
 - Host interface
 - Bad block management
 -



Physical Layer Interface (2)

- Current state:
 - Even with separation to higher level and physical level, higher level is still tightly coupled to NAND topology and usage
 - FTL Higher level has to be aware of NAND device topology
 - There are many different types of NAND devices:
 - Blocks can have mixed TLC / MLC / SLC pages
 - No one standard NAND topology
 - New NAND devices may impact higher level
 - FTL has to be aware of NAND usage:
 - striping
 - Variable rates within a block
 - Modifying NAND usage for reliability / performance will lead to higher level FTL changes

Physical Layer Interface (3)

- Proposal:
 - Virtualization:
 - Physical layer represents to the higher level a virtual NAND device
 - Higher level picks a block type from a list of possible block types
 - Fixed interface resolution
 - Variable Block capacity
 - Striping / Variable rates are taken care of in physical layer
 - Chosen as part of block type
 - Determines performance / reliability
 - SLC / MLC / TLC block
 - For each block type there is definition of:
 - Block capacity
 - Reliability
 - Sequential / Random performance
 - Basic FTL role does not change. Higher level FTL can now support fixed or life-time variable management with much simpler adaptation to new NAND devices

Physical Layer Interface (4)

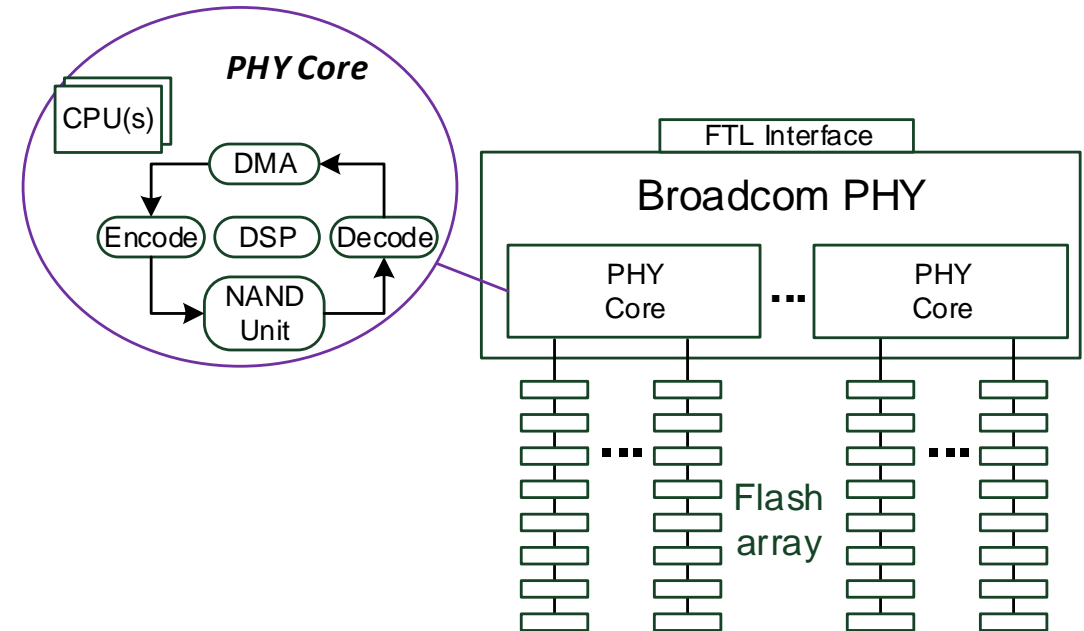
- **Broadcom PHY**

- **Building Blocks:**

- Decoder, Encoder, DSP, NAND Unit
- CPU(s)
- Memory
- FTL Interface

- **Customize per Customer:**

- PHY does not limit performance
 - Limited by Flash device and/or host
- No changes in common building blocks
- Customization: # of channels, # of devices per channel, # of CPUs, memory, frequency





BROADCOM®

connecting everything®