ASIC/Merchant Chip-Based Flash Controllers

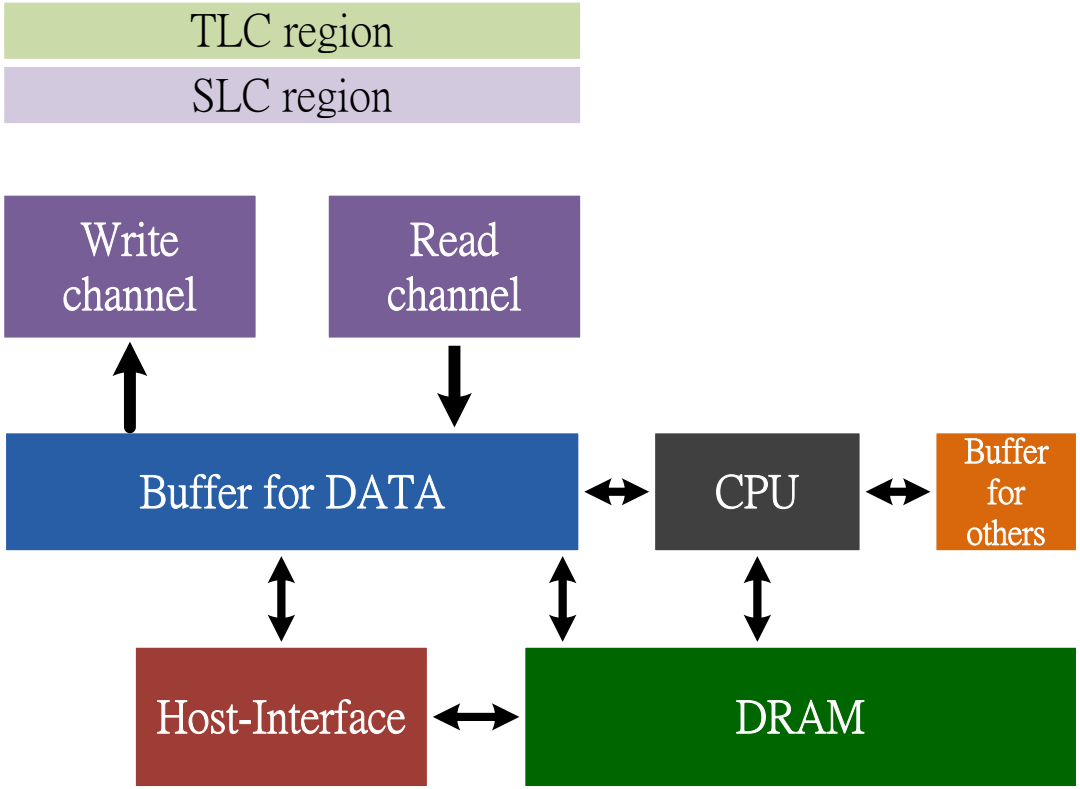
Jeff Yang

Siliconmotion

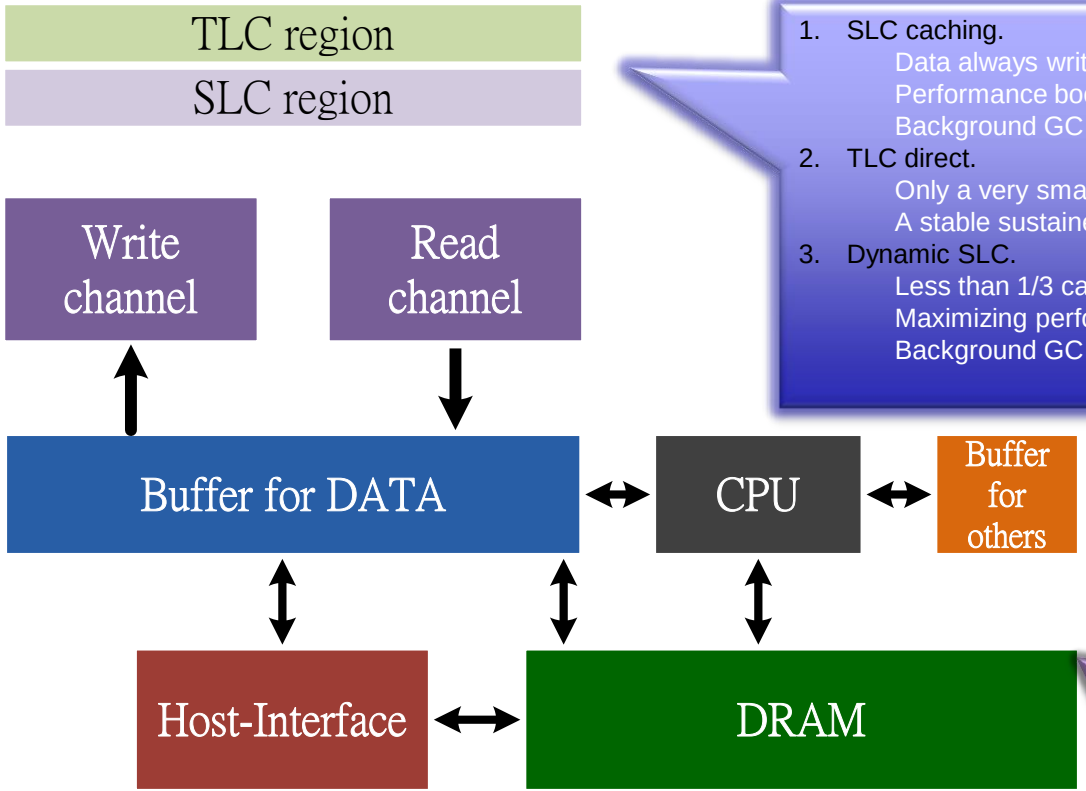
Outline

- Basic controller architecture.
- The challenge on the Merchant Chip-based flash controller.
- The Flash selection combination and the performance requirement.
- Flash write channel and the read channel throughput analysis.
- Hard-decoding only BCH based controller for SATA application throughput requirement.
- From the hard-decoding only to soft-decoding controller.
- Correction capability.
- 3D vs. 2D's NAND architecture.
- Vth tracking and the Data-retention issue.
- RAID

Basic architecture



Application combinations on TLC



1. SLC caching.
   - Data always write into SLC. A Fixed portion of SLC regard as a cache buffer.
   - Performance boost on SLC caching.
   - Background GC to TLC block.
2. TLC direct.
   - Only a very small portion for system usage and small random write data.
   - A stable sustained write performance.
3. Dynamic SLC.
   - Less than 1/3 capacity threshold, using SLC.
   - Maximizing performance boosting period.
   - Background GC to TLC block.

1. Full size DRAM
   - For host data write caching. Full lookup table.
2. Non-DRAM
   - Extremely low cost.
   - Optimize for user experience.
   - More system info access from SLC blocks.
3. Small DRAM.
   - Full lookup table on external buffer
   - No host data buffer.

Challenge: Support all combinations and cost efficiency

2D/3D

- One-pass
- Two-pass
- Multi-pass

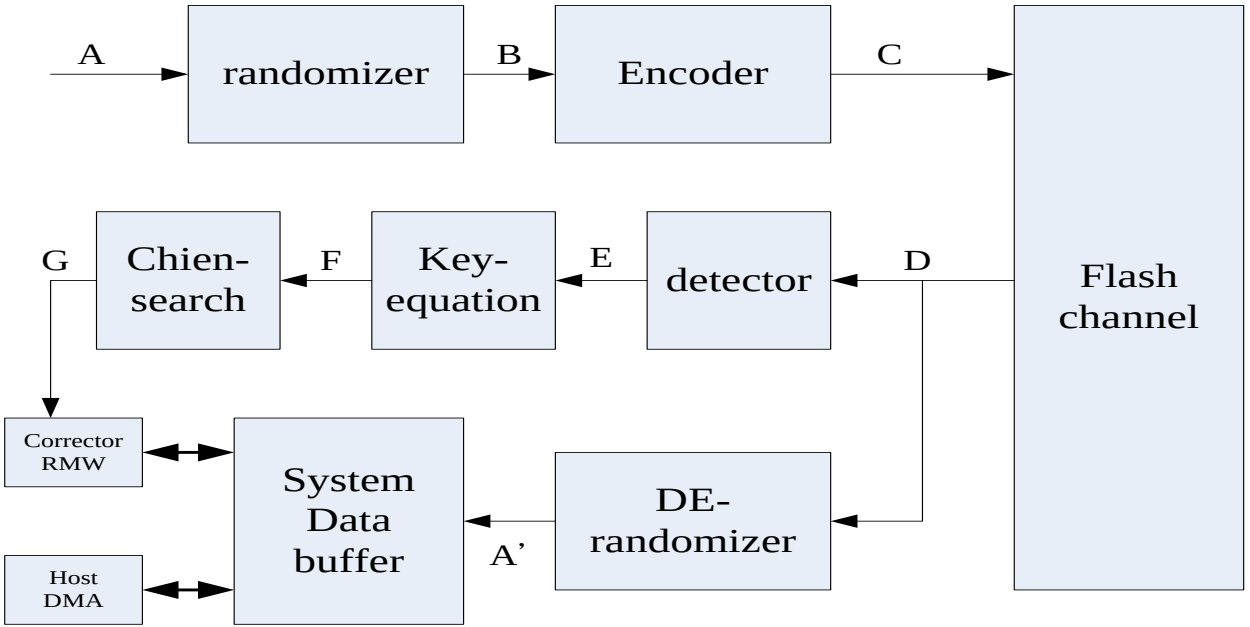
SLC usage

- SLC caching
- TLC direct
- SLC/TLC dynamic

External buffer

- Full DRAM
- Non-DRAM
- Partial DRAM

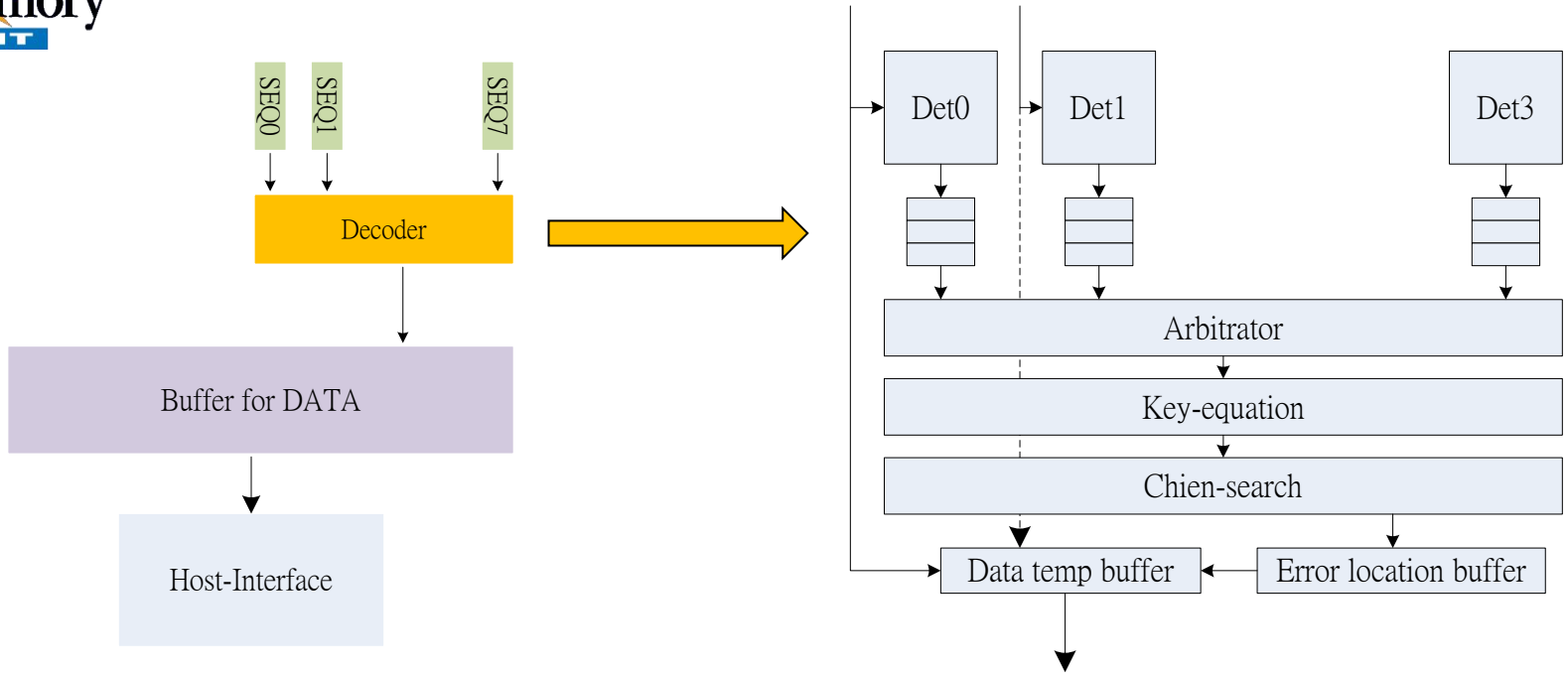
Traditional Write/Read channel with BCH



- A: 1024B
- B: 1024B randomized.
- C: 1024B + 126B-parity
- D: C +error from flash

- A': 1024B with error bit.
- E: syndrome (126B)
- F: error-polynomial (128B)
- G: error location and err-mag

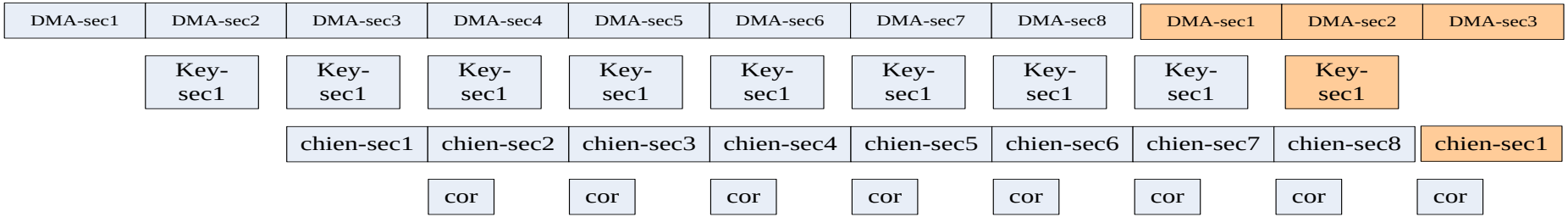
Multi-channel Read in SATA controller



- Share the decoder's hardware with multi-channel.
- Each channel will not encode and decoder at the same time. Share the encoder with Detector.
- The decoder's output should satisfy the host maximum read throughput.

4-stage pipe-line BCH

single key-quation with 4 stage pipeline
T=72bit mode and error bit=72



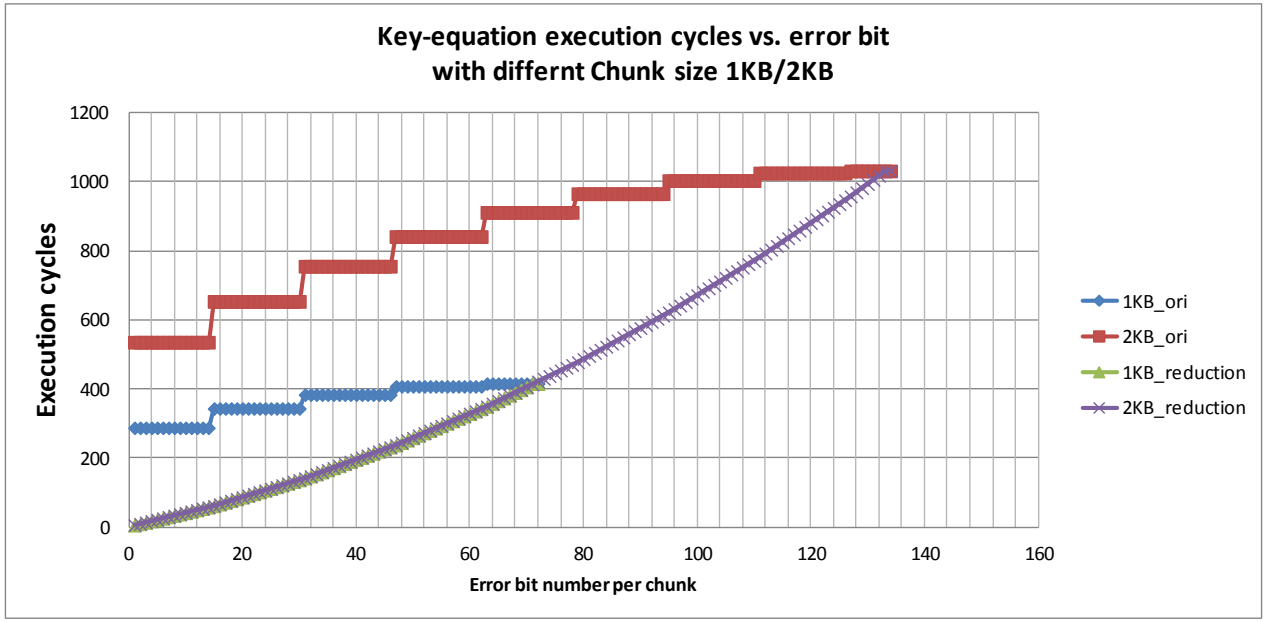
Correction the data in memory.
Diff clock domain handshake

- BCH 72bit mode, 72bit error, chunks size is 1024B + 126B = 1150B
- DMA is 100MHz parallel 16 ➔ 576cycles (200MB/sec per channel)
- Chien-search is operated at 330MHz with parallel16 circuit. ➔ 576cycles.
- Chien-search throughput is 1024/(576x3ns) = 592MB/sec.
- Key-equation cycle is proportional to error bit. (throughput, power consumption bottleneck)
- Key-equation's execution cycle should under 576 cycles
  - It will need a very high parallelism Key-equation on its hardware.

Key-equation operation efficiency.

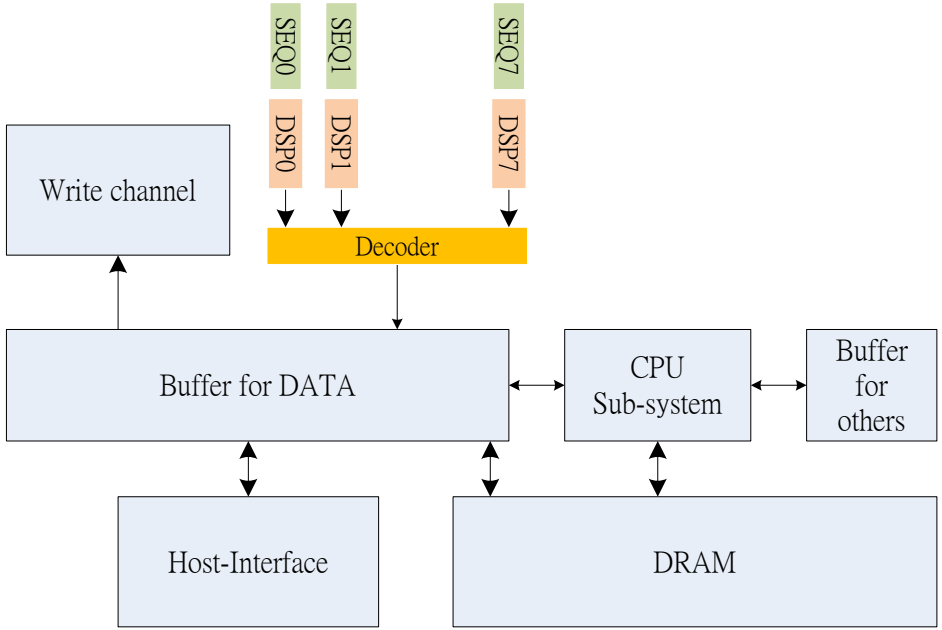
1KB + 126B with 72bit protection.
Cover range to UBER~1e-15
RBER = 3.1e-3
Average error bit = 28bits per chunk

2KB + 252B with 134bit protection
Cover range to UBER ~1e-15
RBER = 3.9e-3
Average error bit = 71bits per chunk



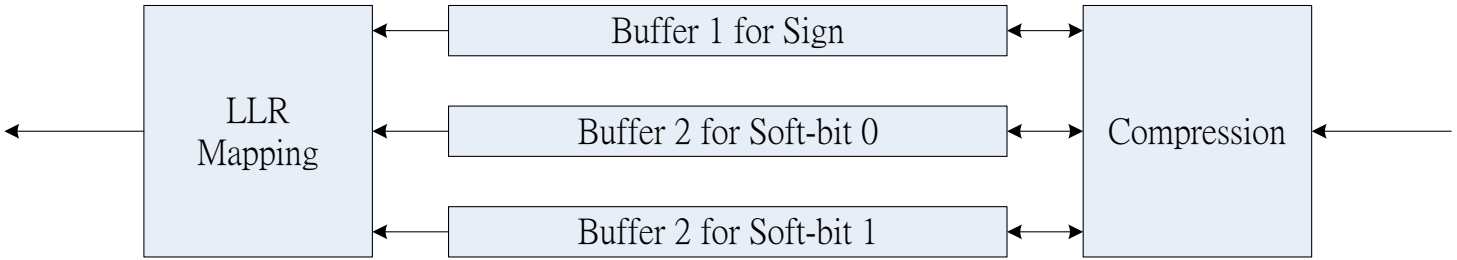
- An very efficiency BM, simplified and inversion free algorithm has been used as an original.
- The further reduction provide much better efficiency.
  - 1KB 10bits error, 288 ➔ 42 cycles. ~85% improvement. (BOL)
  - 2KB 20bits error, 654 ➔ 87 cycles. ~87% improvement. (BOL)
  - 1KB 28bits error, 200 ➔ 127cycles. ~55% improvement. (EOL)
  - 2KB 71bits error, 912 ➔ 414 cycles. ~55% improvement. (EOL)

Soft-information interface



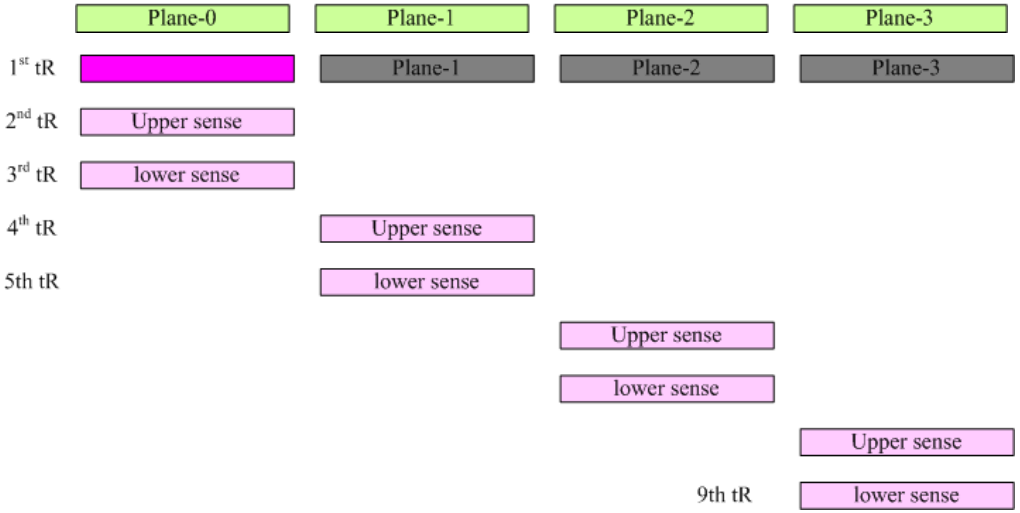
- In order to provide better decoder's correction capability, using the soft-info to get more reliability bits.
- NAND interface support .
  - Traditional read/retry interface.
  - Direct soft-info interface.

DSP engine's buffer size



- The buffer size is the capability to contain the number of chunks soft-bit.
- Access addition soft-info from NAND may need additional read busy time.
- Read the soft-bit under the same busy time will have higher efficiency, but buffer size requirement is huge.

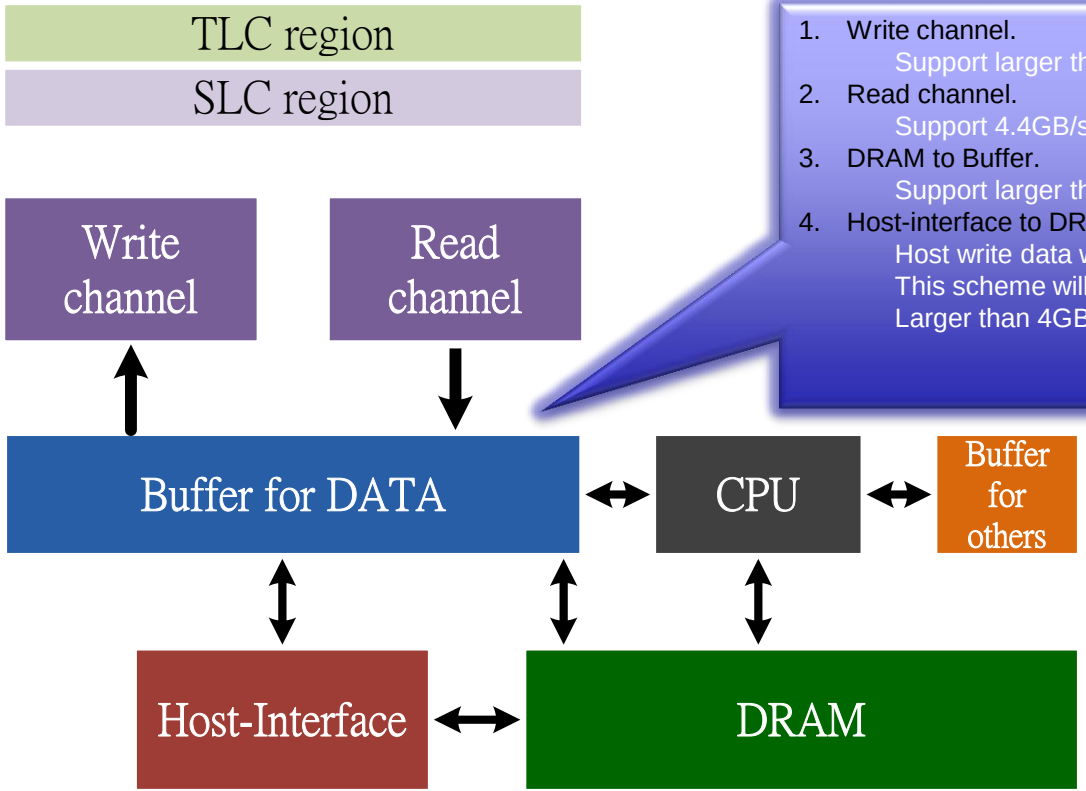
Soft-decoding throughput limitation



- One Transfer time = 2.5ns/1B x 18432B = 46us (400MTs)
- Assume DSP-buffer size 16KB.
  - 9 tR time + 12 transfer-time = 9x(100us) + 12 x (46us) = 1452us
  - Throughput = 64KB/1452us = 44MB/sec
- Assume DSP-buffer size 64KB.
  - 3tR time + 12 transfer-time = 3x(100us) + 8 x(46us) = 668us.
  - Throughput = 64KB/668us = 95MB/sec

In Client SSD applications,
Soft-decoding will regard as the ERROR-Recovery flow.
We will not ask the throughput under recovery mode.
But we will take care the recovery mode trigger rate.

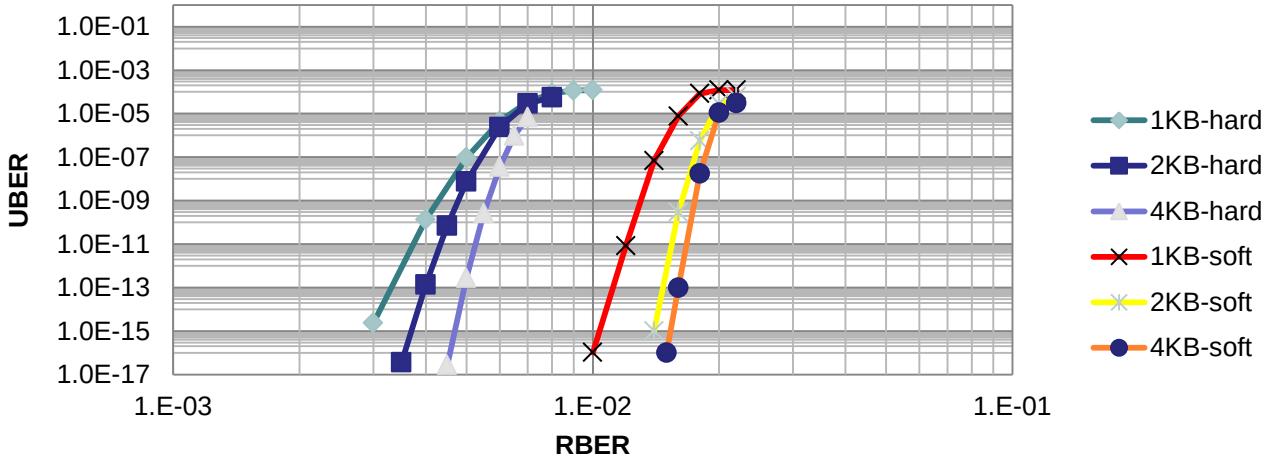
High throughput SOC bottleneck PCIE3x4



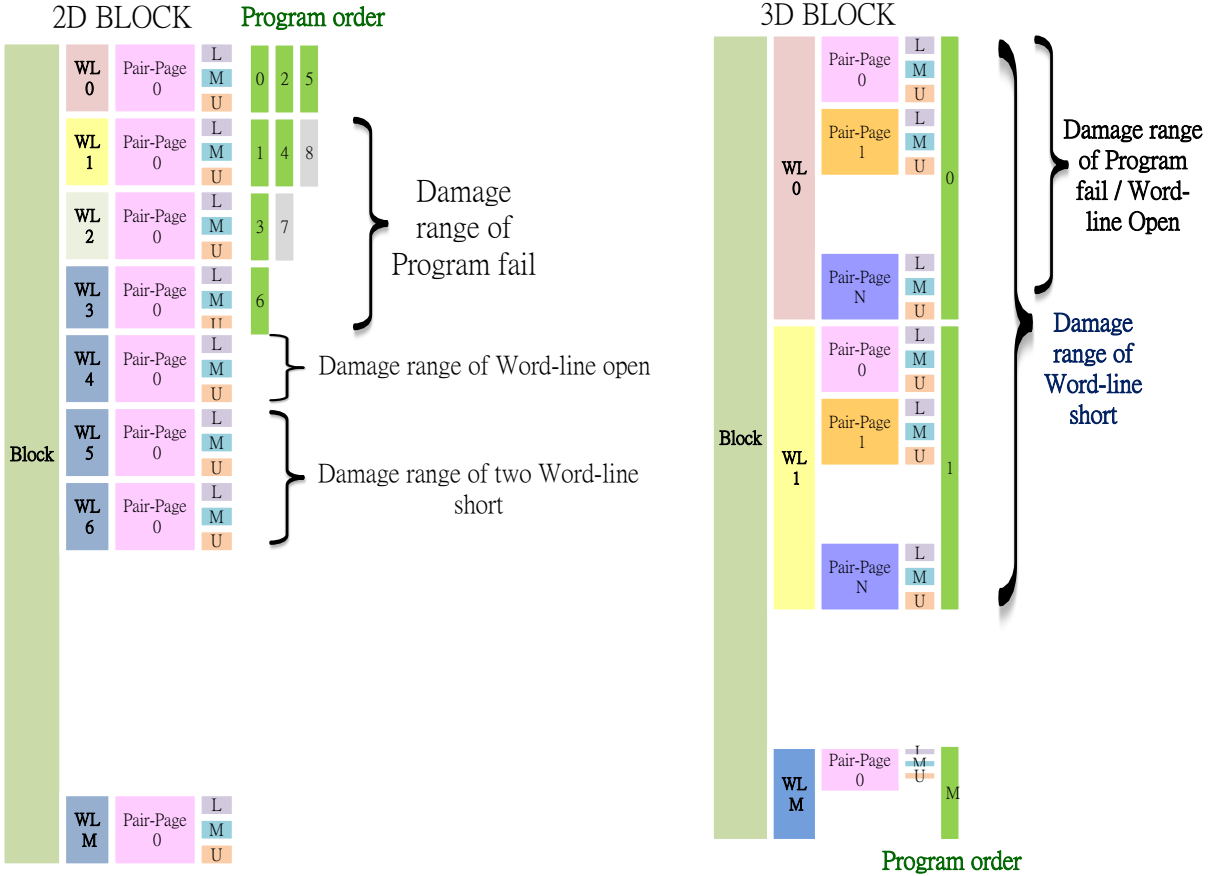
1. Write channel.
   Support larger than 4GB/sec throughput. (5.28GB/sec)
2. Read channel.
   Support 4.4GB/sec throughput. .
3. DRAM to Buffer.
   Support larger than 4GB/sec throughput.
4. Host-interface to DRAM
   Host write data will store in DRAM first.
   This scheme will serve the host behavior better.
   Larger than 4GB/sec throughput.

ECC Chunk

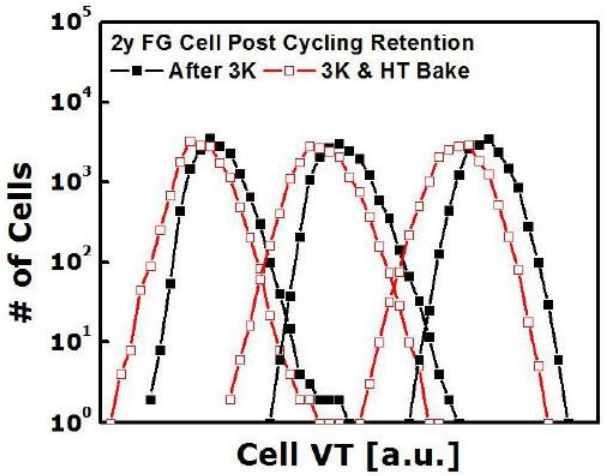
- Fixed code rate: around 0.9, ECC chunk size: 1KB/ 2KB/ 4KB
- Hard-decoding is based on BCH, and soft-decoding is based on LDPC with less than 3-bit channel reliability values.
  - Correction Performance: 4KB better than 1KB
  - Decoding Latency: 1KB better than 4KB

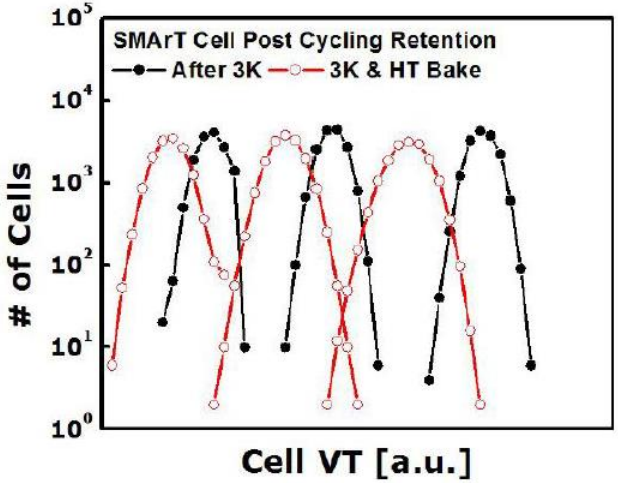


Failure range from 2D to 3D.



Retention issue on 2D/3D

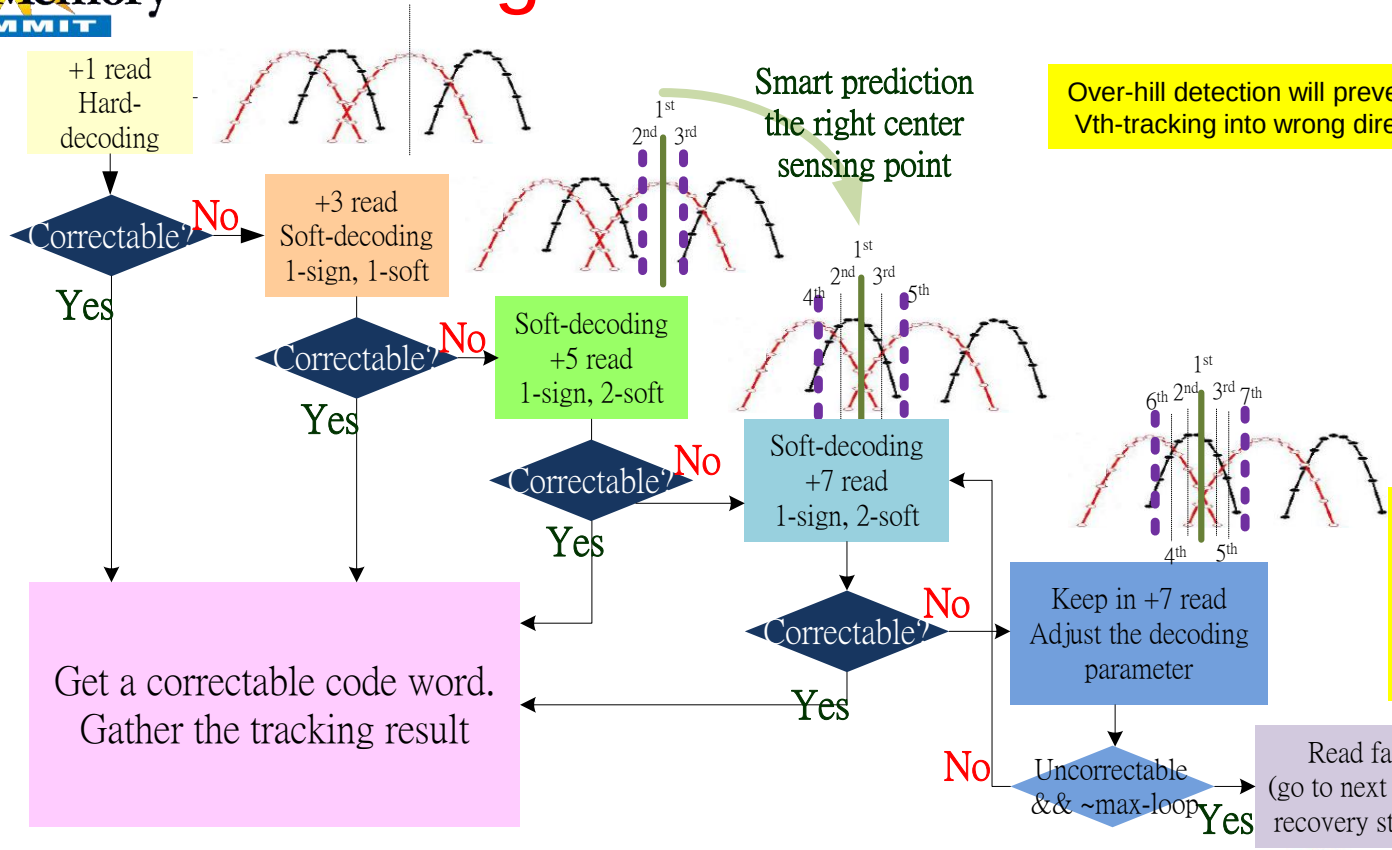




- Both the 2D and 3D will have the data retention problem.
  - 1Znm MLC need 6~10 read-retry tables, But TLC need 40~45 tables with less endurance and retention.
  - 3D will have more severe Data retention issue.

[ref]: E.S. Choi, S.K. Park, “Device Considerations for High Density and Highly reliable 3D NAND Flash Cell in Near Future”. IEDM 2012

DSP algorithm for the Vth-tracking



Over-hill detection will prevent the Vth-tracking into wrong direction.

The parameter include:

1. Sensing step size.
2. Decoder internal coef.
3. Dynamic tuning the LLR
4. ……

The flash trend and ECC correction

Flash Memory SUMMIT

The 3D flash is good!! What are we waiting for?
COST, COST, COST!!!

	2D	3D
EDURANCE	After cycling: Keep the same Error distribution in LOW RBER	After cycling: Keep the same Error distribution in LOW RBER
Data RETENTION	The RBER become worse, the Vth also shifting	Only the Vth-shifting, but RBER is still good.

HDD Trend: RS ➔ LDPC ➔ NB-LDPC
SSD Trend: HM ➔ RS ➔ BCH ➔ LDPC

We Always Need A Stronger ECC

WHY target RBER= 3e-3?
BCH 72bit/1KB will provide UBER< 1e-15 with RBER = 3e-3

	Target	RBER requirement
Normal operation (Base-line)	Extreme low power. Keep the host throughput	RBER = 3e-3
Reliability extension	Vth-tracking to lowering the RBER. Soft-info to have stronger correction	RBER = 1e-2 ~ 1.2e-2

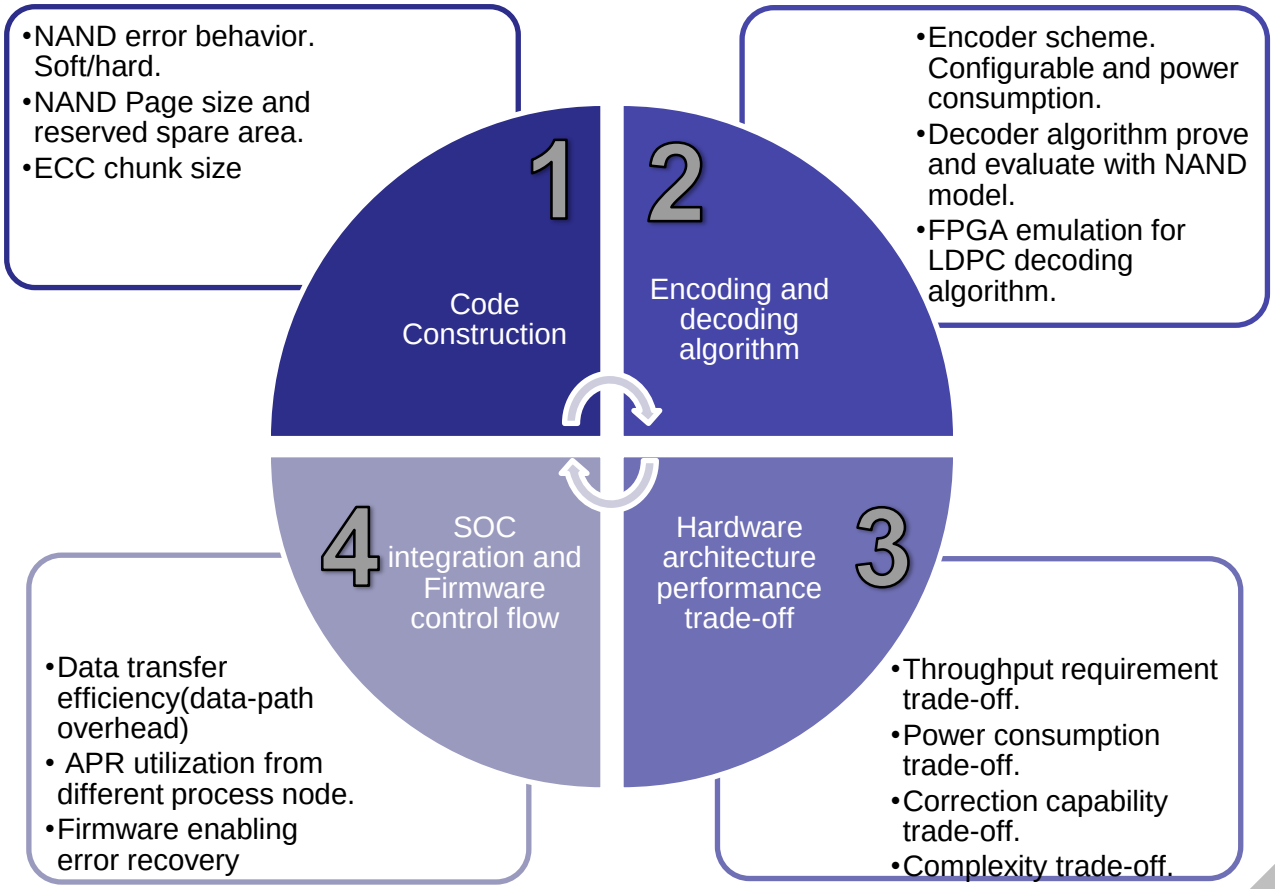
Need a ECC stronger then BCH

Provide the most-cost efficient way to satisfy the reliability

SiliconMotion

ECC design loop related to NAND characteristics.

- We already have 6th generation LDPC decoder.
- Keep improving the LDPC performance.
- For higher throughput ~8GB/sec, we may go back to step1.
- After 28nm process, the design iteration depth will from code-construction to trial APR.
- EX: Find the Routing congestion issue in step 4, it may need to solve from step1.



Before the RAID protect flow…..

DRAM/
DRAM-less/
Small DRAM

SLC-first/
TLC-direct write/
Dynamic SLC

One-pass /
Multi-pass/
Pair-page
mapping

WL to WL short
Failure range

All the issues combine together!

Capacity
(RAID overhead)
Binary/arbitrary

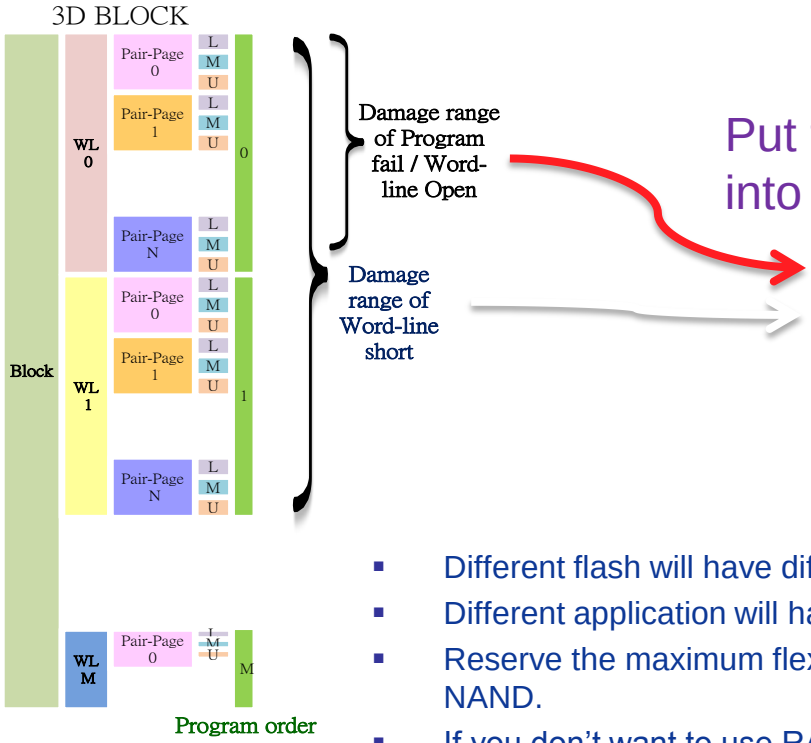
Program failure
DRAM-backup/
Flash-cache/
RAID recover/

WL open
Failure range

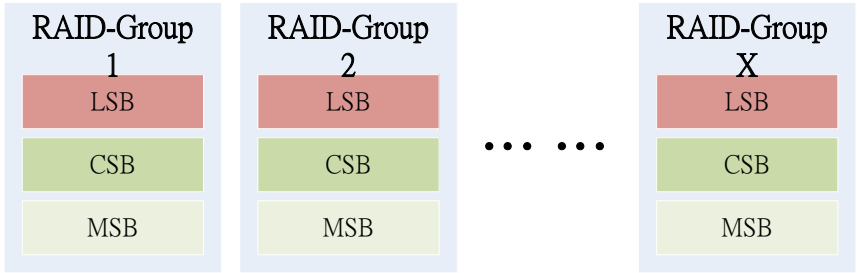
Recovery latency

BUT THE SAME CONCEPT IS…….

Don’t put eggs in the same basket.



Put the pages in the same failure range into different RAID-Group



- Different flash will have different failure range.
- Different application will have different RAID encoding flow.
- Reserve the maximum flexibility for all controller to support all kinds of 2D/3D NAND.
- If you don’t want to use RAID, What alternative you still have?
  - Read-back check after program.