

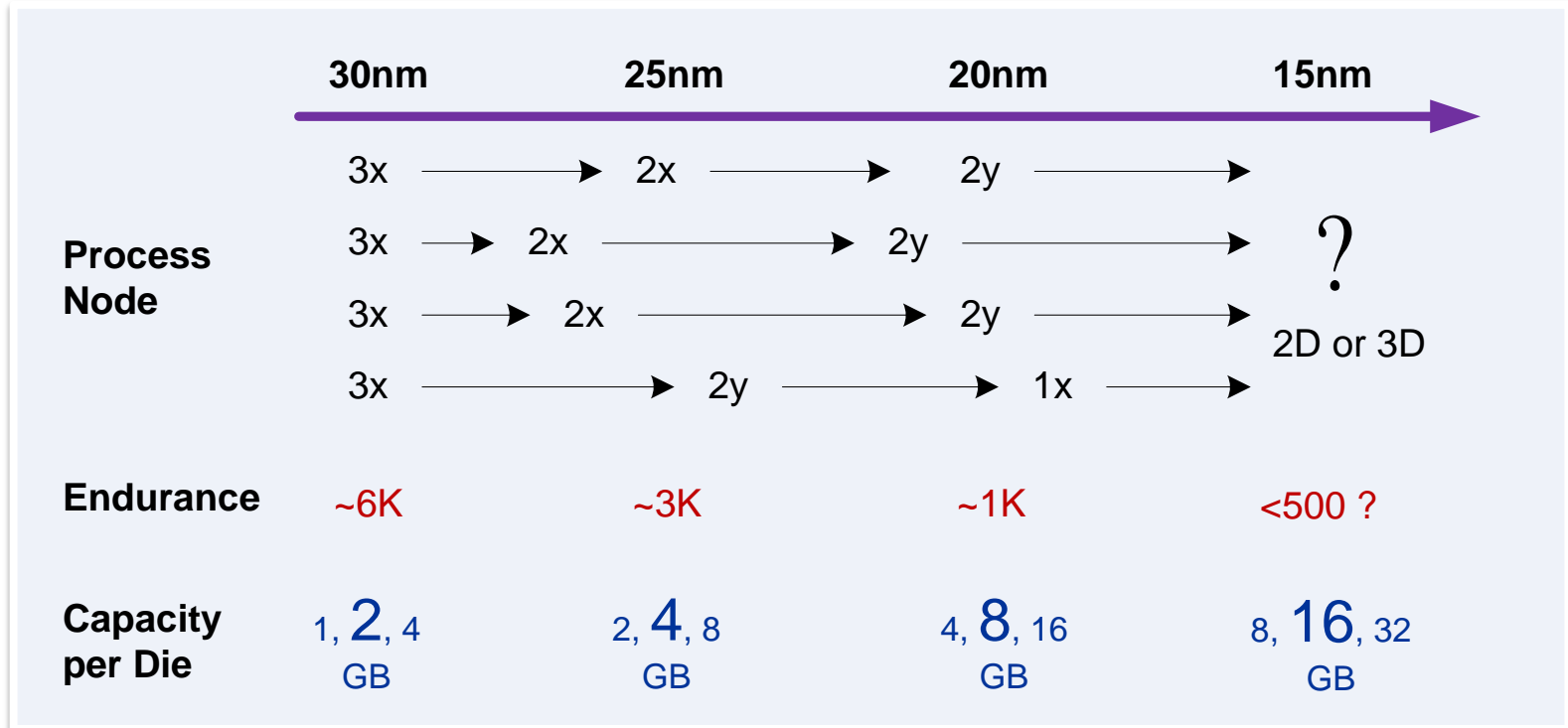
Novel ECC Architecture Enhances Embedded Storage System Reliability

Jeff Yang

Principle Engineer
Silicon Motion

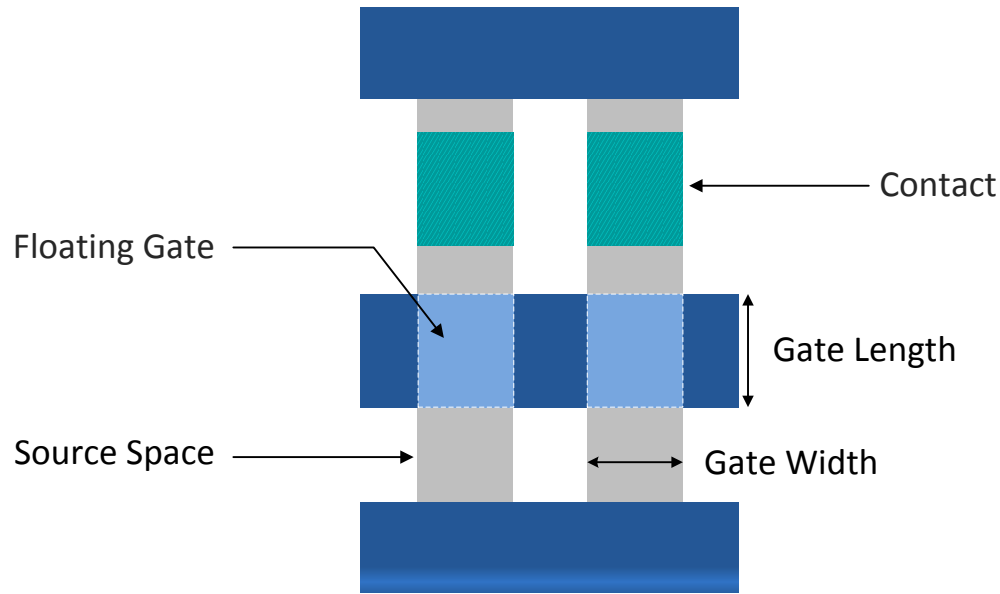
NAND Evolution and Effects

- New NAND Flash
 - Advanced smaller process nodes
 - Double capacity, but with half endurance capability



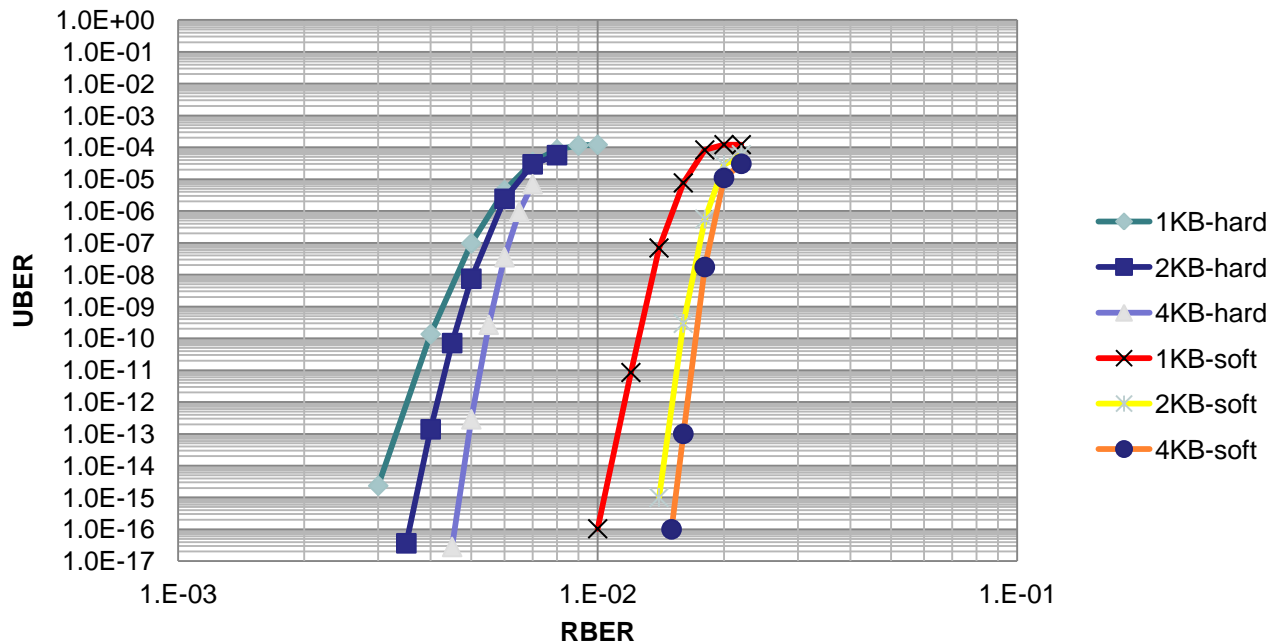
How to Keep the Same Reliability

- Floating gate area = Gate length x Gate width
- Smaller gate area → Lower cost + Worse reliability
- ECC is increasingly important to NAND flash.



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



Endurance vs. Retention with Hard-Decoding

1KB BCH Protection

84 hrs	○	X	X	X	X	X
74 hrs	○	X	X	X	X	X
68 hrs	○	X	X	X	X	X
60 hrs	○	X	X	X	X	X
52 hrs	○	X	X	X	X	X
44 hrs	○	X	X	X	X	X
36 hrs	○	X	X	X	X	X
28 hrs	○	X	X	X	X	X
24 hrs	○	X	X	X	X	X
20 hrs	○	X	X	X	X	X
16 hrs	○	X	X	X	X	X
12 hrs	○	○	X	X	X	X
8 hrs	○	○	X	X	X	X
4 hrs	○	○	○	X	X	X
Endurance	○	○	○	○	○	○
P/E Cycles	600	1200	1800	2400	3000	3600



- All data sectors are correctable
- At least one data sector is uncorrectable

- 2ynm TLC
- 1KB-based BCH
- Room Temp. Burn-in
- 120 °C Bake for Different Durations
- Data retention is much more important because NAND flash is NV-memory.
- High endurance and poor data retention

Endurance vs. Retention with Soft-Decoding

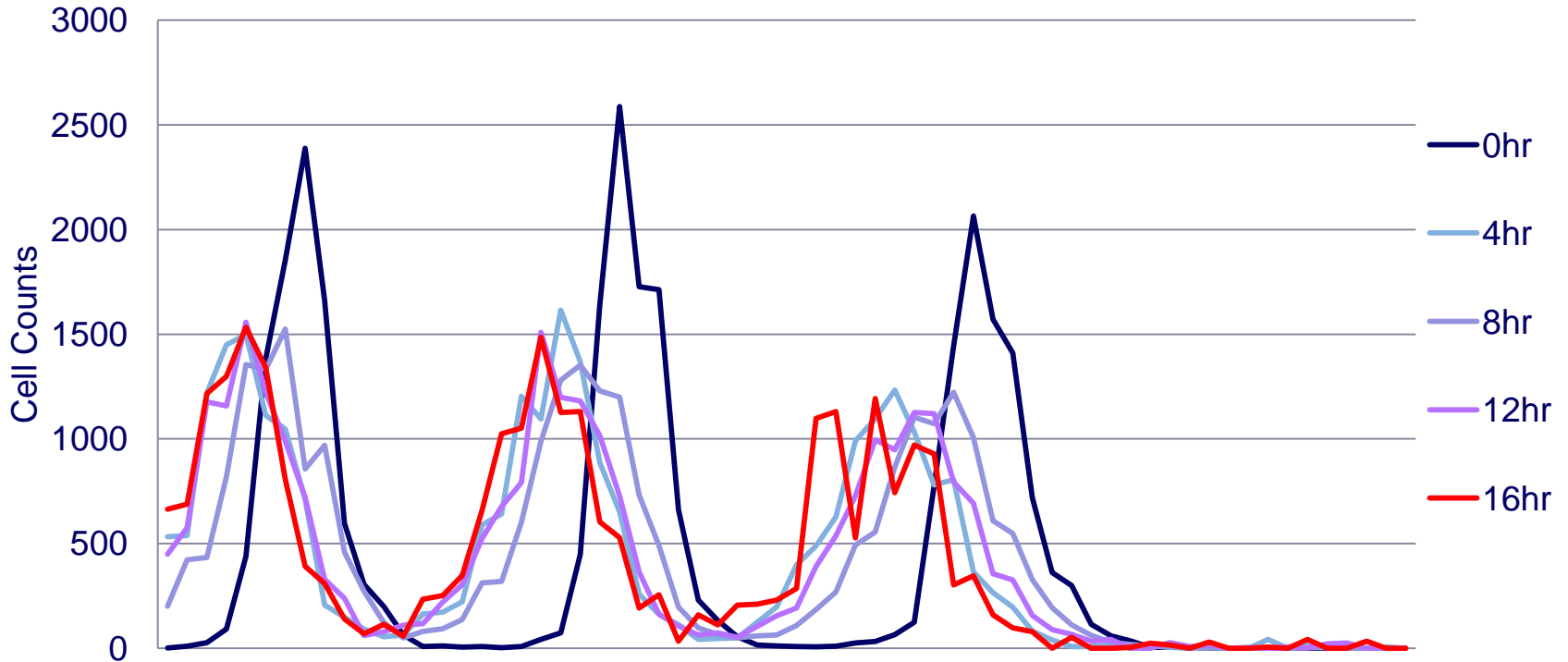
1KB LDPC Protection

84 hrs	o	o	o	x	x	x
74 hrs	o	o	o	x	x	x
68 hrs	o	o	o	x	x	x
60 hrs	o	o	o	o	x	x
52 hrs	o	o	o	o	x	x
44 hrs	o	o	o	o	x	x
36 hrs	o	o	o	o	x	x
28 hrs	o	o	o	o	o	x
24 hrs	o	o	o	o	o	x
20 hrs	o	o	o	o	o	x
16 hrs	o	o	o	o	o	x
12 hrs	o	o	o	o	o	x
8 hrs	o	o	o	o	o	x
4 hrs	o	o	o	o	o	x
Endurance	o	o	o	o	o	o
P/E Cycles	600	1200	1800	2400	3000	3600

-  All data sectors are correctable
-  At least one data sector is uncorrectable

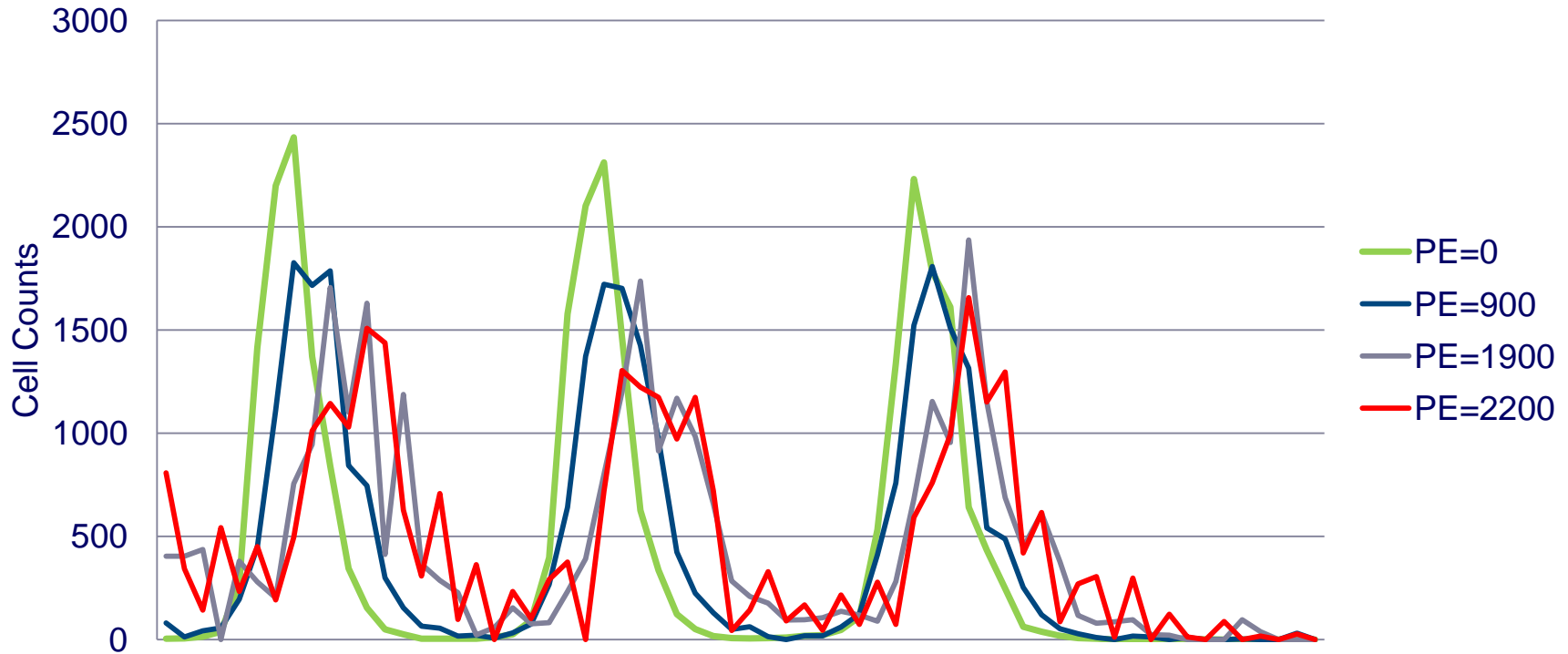
- 2ynm TLC
- 1KB-based LDPC
- Room Temp. Burn-in
- 120°C Bake for Different Durations
- The endurance is larger than 3.6K. It is around two times compared to BCH.
- Under the same data retention condition, LDPC offers three times greater protection than BCH.

Data Retention at 120°C Bake



- The Vth distribution shifts down with the noise variance increasing.
- The curve is more Gaussian-like, bell-shaped.

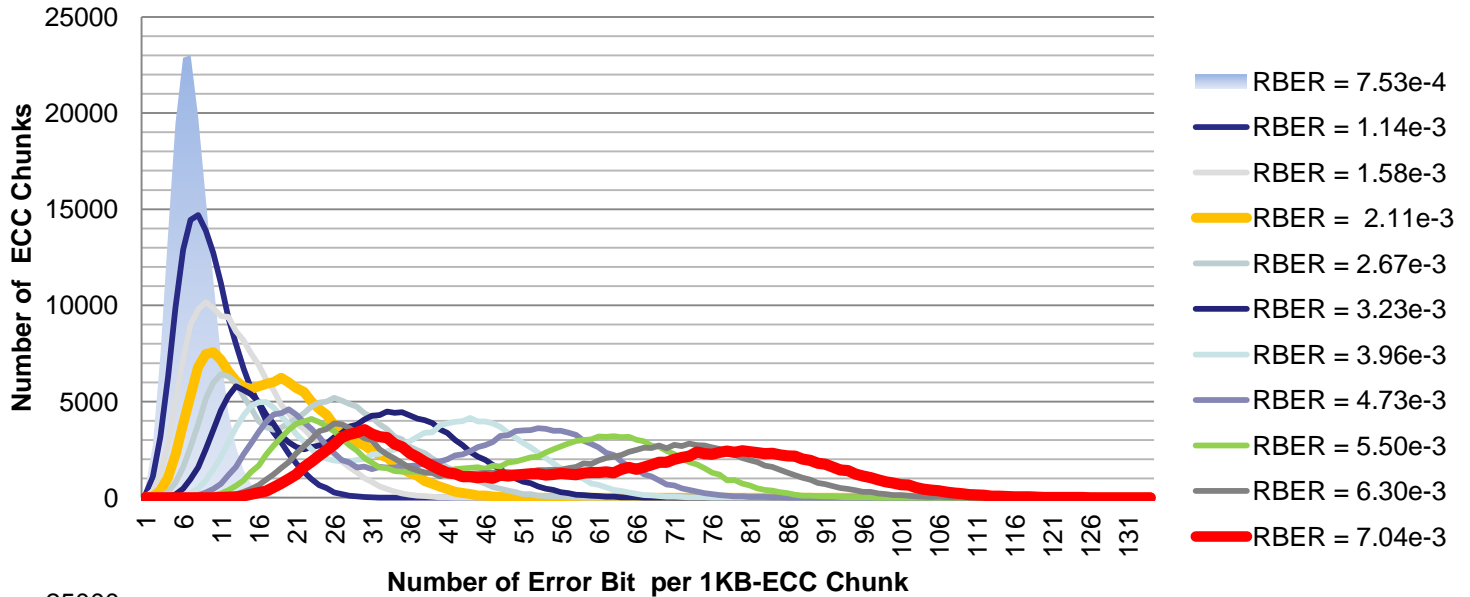
Endurance Test - P/E Cycling



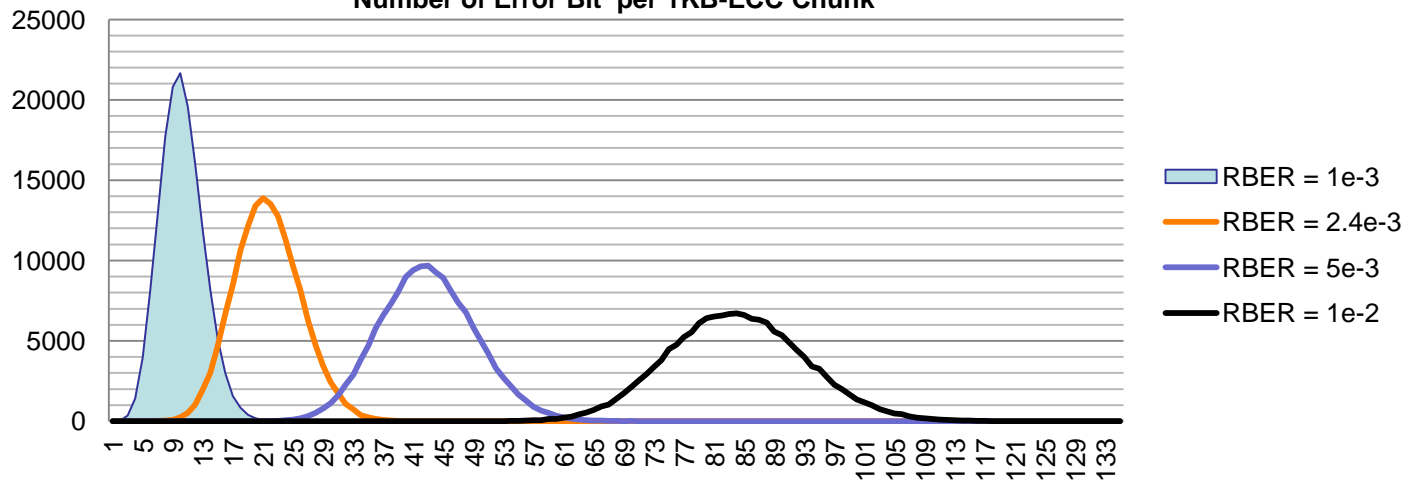
- The Vth distribution shifts to the right.
- The noise variance is also increasing.
- The non-Gaussian parts are generated by endurance disturbance.

Comparison of Error Histogram

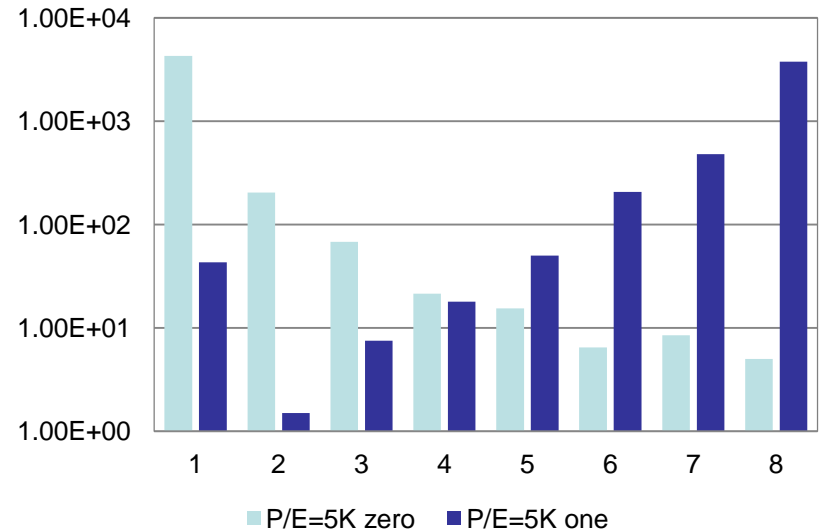
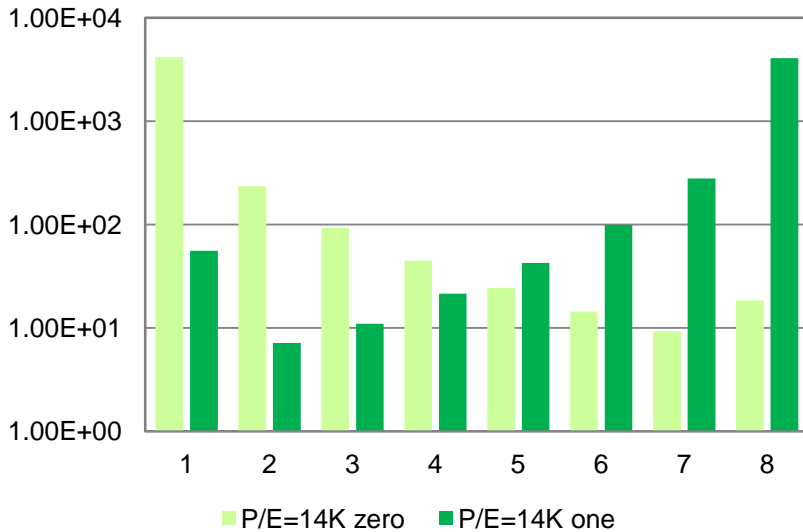
- Measurement on Real 2ynm TLC (after Burn-in Test)



- Measurement on Pure-Gaussian Noise Model



Bit Error Profile in Endurance Test



LLR Value	1	2	3	4	5	6	7	8
AWGN	-4.08764	-2.47504	-1.48491	-0.49495	0.49495	1.484909	2.475038	4.087639
Real NAND (P/E = 14K)	-1.87216	-1.51755	-0.92549	-0.3199	0.241576	0.837346	1.47941	2.343726

- Based on 1KB ECC chunk as an observation space.
- The endurance increases the occurrence of strong errors.
- When P/E ≥ 5K, all the uncorrectable codewords have the same noise problem.

Analysis of Error Profile

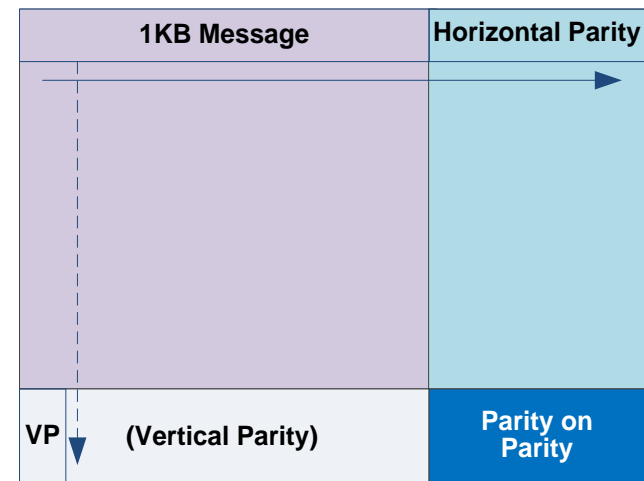
- Strong errors (long and flat tail) in the endurance test
 - When P/E is around 5K, the RBER in one block is $2.6e-3$. But the worst-case error profile will cause the first uncorrectable codeword.
- If one uses soft-decoding to extend the endurance, the strong errors may be overcome.
 - Extend the ECC chunk size (2KB, 4KB)
 - Decoding latency can be increased.
 - Reduce the code rate
 - Storage capacity can be decreased.

Raid-like Protection Within One Block

- An example on TLC
 - 174 word-lines, 522 logical pages
- The last 10 pages for vertical parity, 512 pages for real user data.
 - Each bit column provides 1-bit correction and detects the mis-correction.
- The small-sized ECC chunks will be concatenated into a super long code.
 - Overcome partly non-AWGN noises
- Only extra 2% protection area will increase the correction capability. ($10/522 = 2\%$)
 - Decoding latency will be extended
- Dynamically change the extra protection area.
 - Protection area: from 2% to 4% or 8%
 - Reduce the decoding latency, and increase the correction capability.



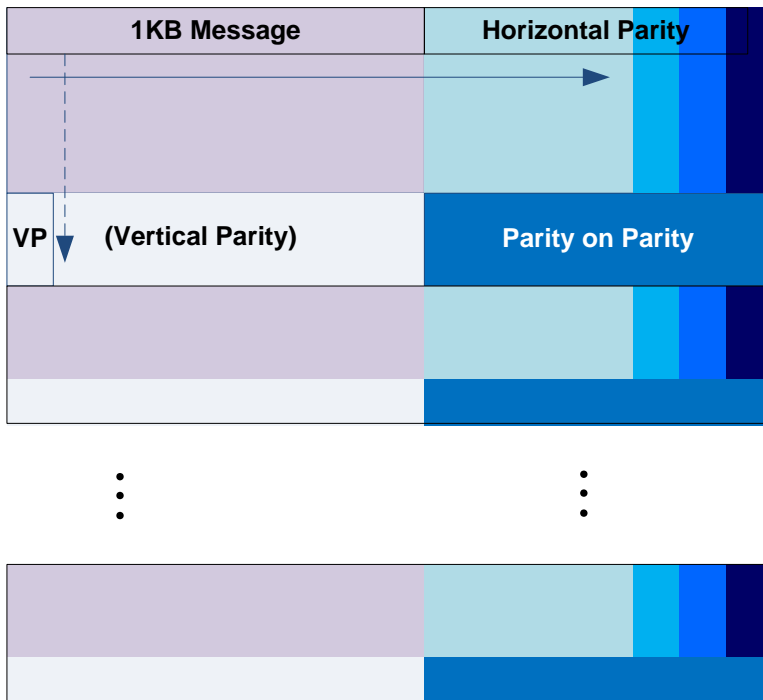
8 Column Groups in One Block



Test Result of 2ynm TLC under extra 1% Protection:

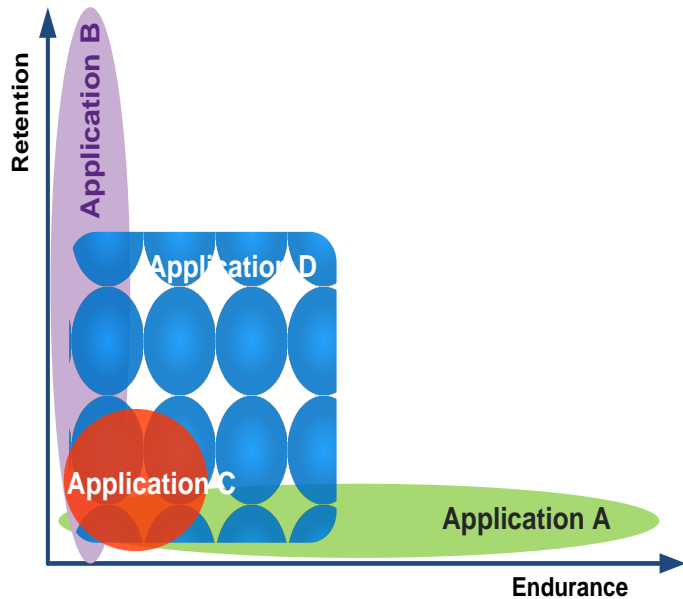
- BCH horizontal hard-decoding
 - Endurance P/E = n
- LDPC horizontal soft-decoding
 - Endurance P/E = 1.5~2n
- LDPC with Rail-like protection
 - Endurance P/E = ~4n

Flexible Protection in 2 Directions



- The vertical protection is programmable.
- The horizontal LDPC code rate is also programmable.
- Accordingly the different kinds of disturbance on various NAND types can be covered and overcome in this protection scheme for all kinds of applications.

Strategies Applicable for Different Applications



- App A: Cache SSD
 - Temporary data storage
 - Data retention capability not important.
 - Lower code rate + higher decoding efficiency
- App B: Read-Only Storage
 - Data retention as the first priority
 - Detect the read-disturbance
 - Detect the retention-disturbance
- App C: General Consumer (uSD, USB)
 - Capacity is important
 - BCH may be enough, but high reliability still relies on LDPC. Cannot use complicated protection.
- App D: Consumer SSD (eMMC)
 - Care the data-retention and the endurance both
 - Need LDPC to provide higher reliability
 - Long decoding latency not acceptable



THANK YOU!

Q & A

Disclaimer Notice

Although efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided “as is” as of the date of this document and always subject to change.