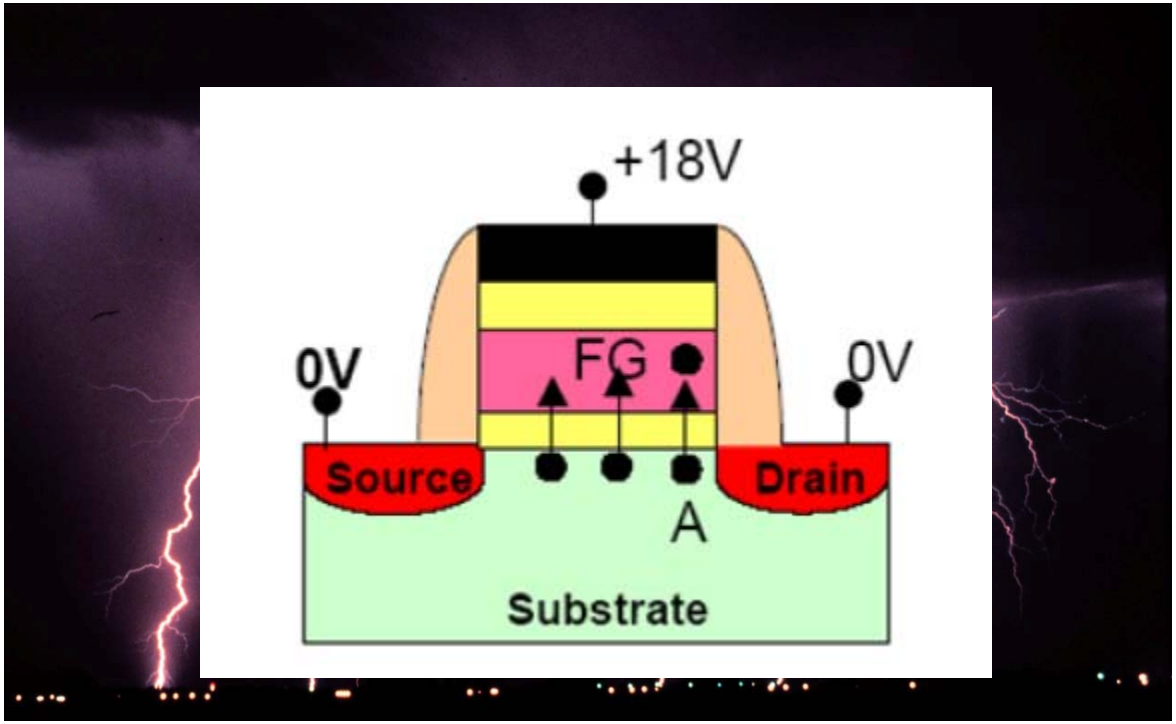
Quantifying Reliability of Solid-State Storage from Multiple Aspects

Hairong Sun, Pete Grayson, and Bob Wood
Fusion-IO
hsun@fusionio.com

May 25, 2011



Tens of volts over nanometer-scale geometries

Outline

- More than 6000 ioDrives are tracked for factory bad block and infant mortality analysis.
- The total capacity tested exceeds 10 TB. About 60,000 erase blocks for each type of flash devices were exercised and worn out during the tests.
- This study discloses a comprehensive set of test data with NAND flash parts from a variety of vendors and different NAND technologies (3x nm, 4x nm, and 5x nm).

Outline

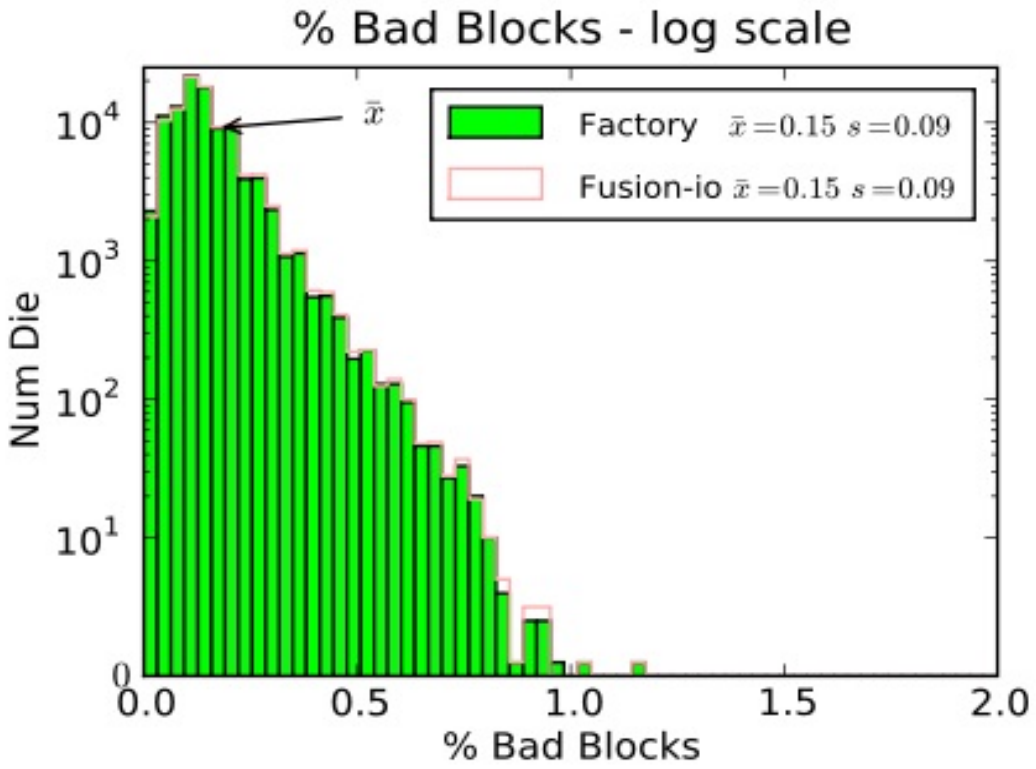
- MAT and Infant Mortality
- Endurance
- Read disturb
- Retention

MAT

We track all the FBB (Factory Bad Blocks) identified by NAND Flash vendors and run MAT (Manufacturing Acceptance Test) at system level for 24~48 hours.

The MAT goes through a variety of test cases and runs about 10~100 P/E cycles with all the failures and their root causes being logged.

MAT



- There typically exists a small increase in the number of bad blocks during MAT. These are “infant mortality” failures or vendor test escapes
- The distribution of bad blocks appears to be Normal at first glance, however a closer look shows linearity at logarithmic scale when the number of bad blocks is greater than 10
- There is a heavy tail in the bad block count distribution with some die having more than 100 bad blocks

Endurance - *NAND flash failure modes*:

Erase failure: As the trapped charges being build up in the tunnel oxide, valid Erased states (a logical “1”) cannot be reached.

Program failure: Hard Program Failures are typically not an endurance phenomena; rather an Inherent Failure.

Read failure: The charge loss or disturb from other cells causes cell charge level to shift so severely that there are excessive bit errors in ECC code words.

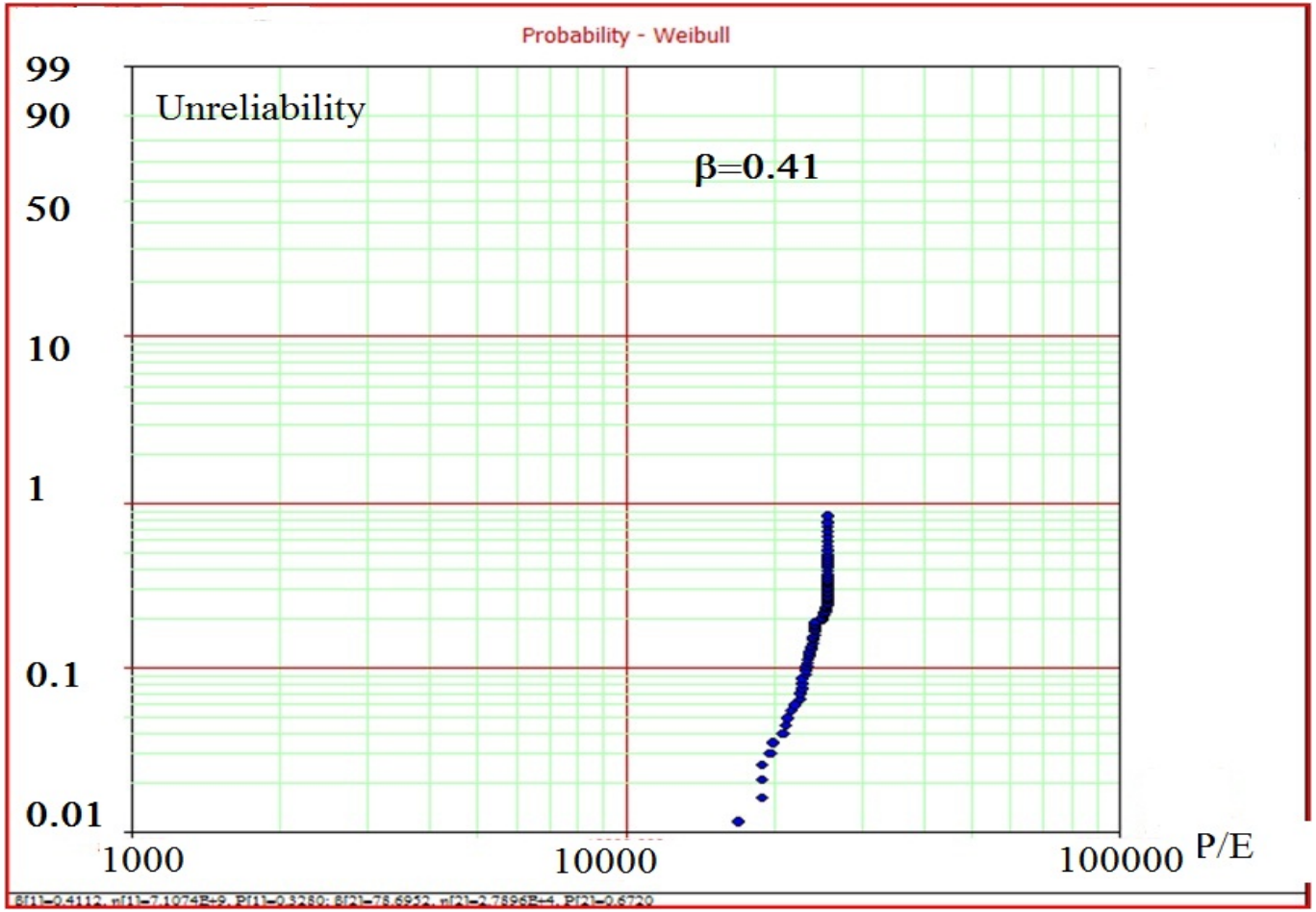
Endurance – *Testing Regime*

Type I Wear-out: Program, Read/Verify, Erase, Program, Read/Verify, Erase,…, till N% of Blocks have failed

Type II Wear-out: Based on Type I results, multiple regions on the ioDrive are worn to varying levels to assess "curvature"

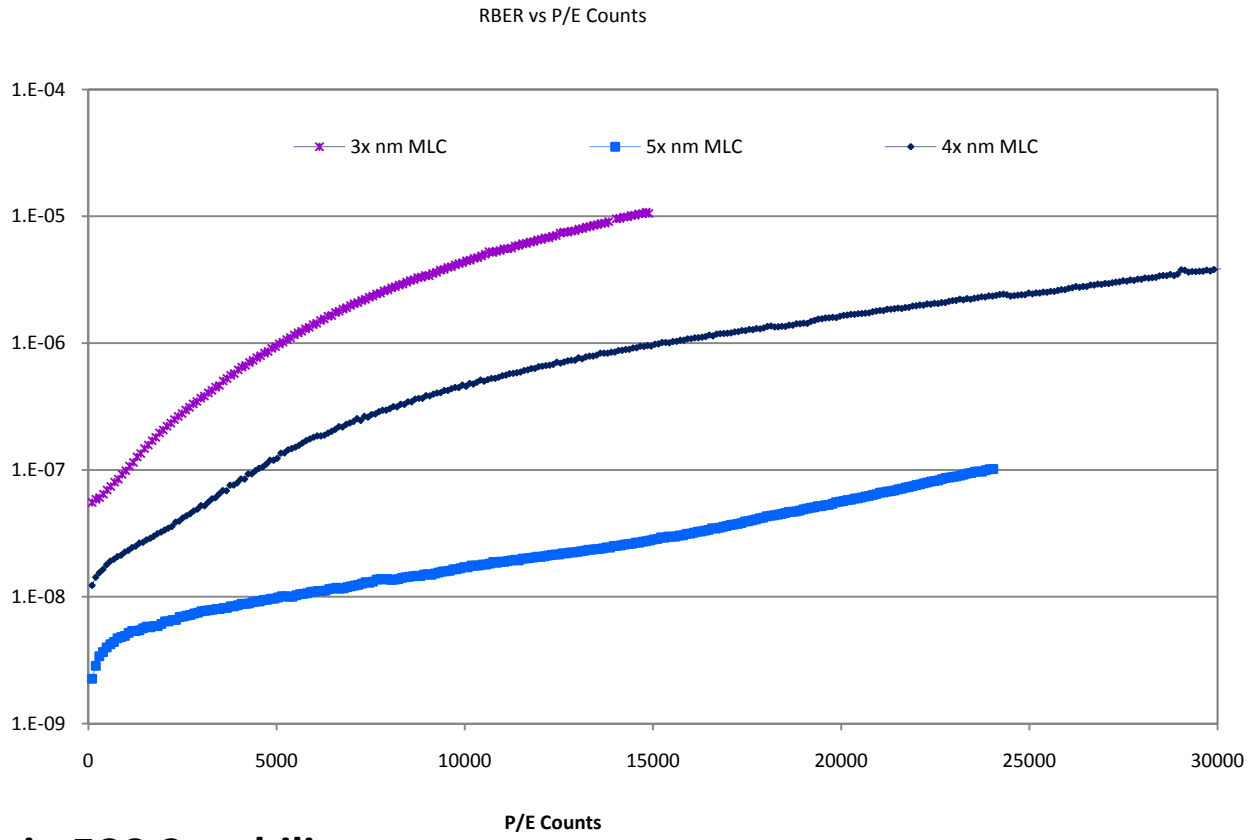
- Block Time To Failure
- RBER vs P/E

Endurance



5x nm MLC: Erase operations generate “fail to erase” errors at 23k ~ 32k P/E cycles.

Endurance



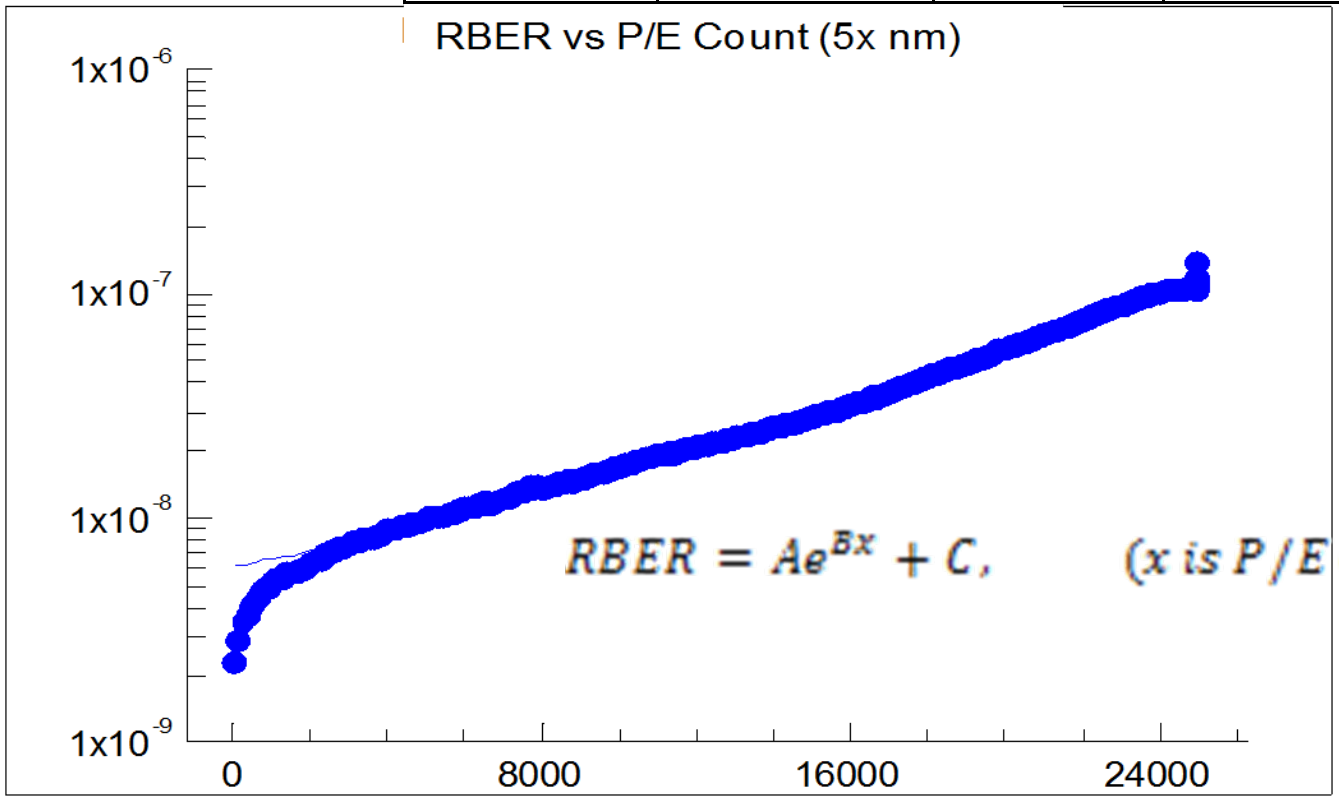
Fusion-io ECC Capability:

Gen-1: correct 11 bit error in 240 Bytes: 6E-5 (RBER)->E-20 (UBER)

Gen-1.5: correct 39 bit error in 960 Bytes: 7.5E-4 (RBER)->E-20 (UBER)

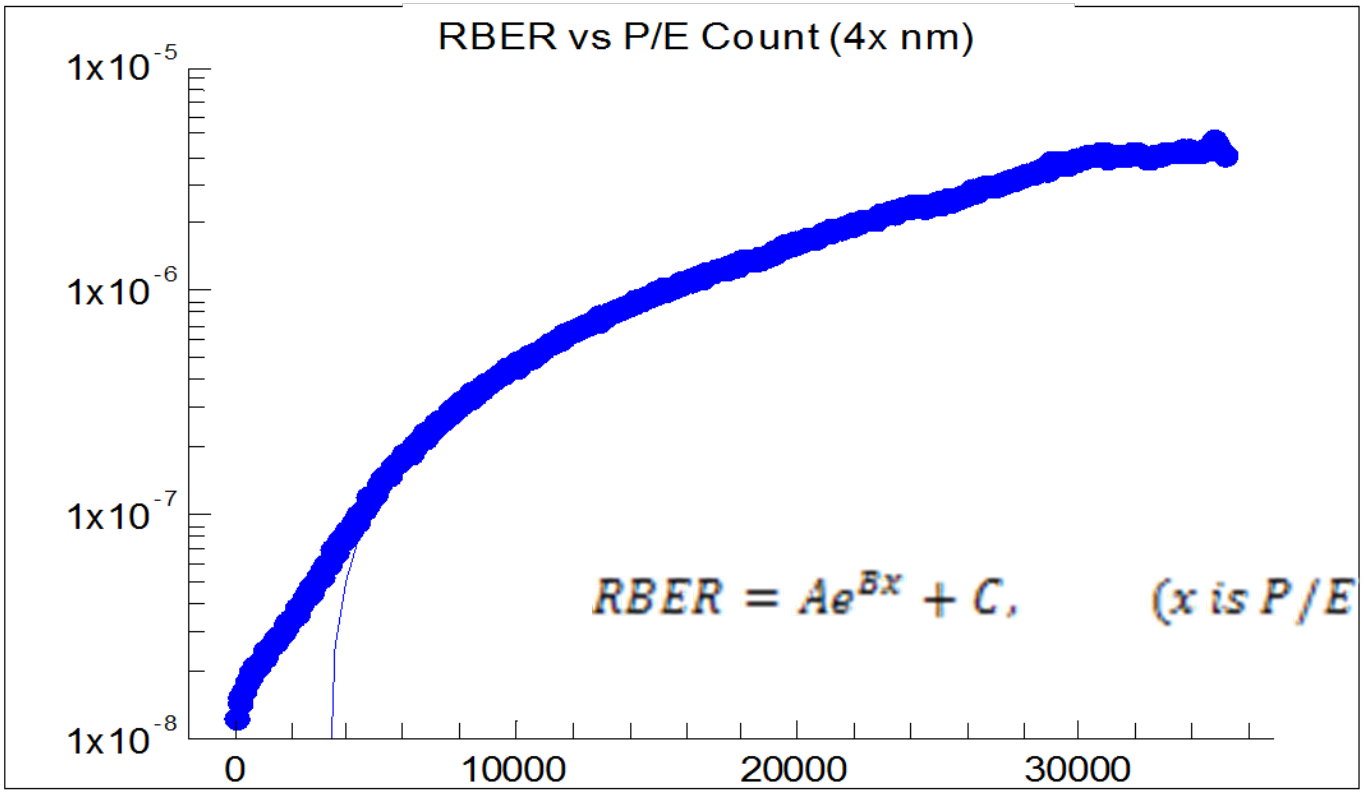
Endurance

5x nm MLC	Value	Std Err	95% Confidence
A	2.6953E-09	1.97E-10	2.31E-09
B	0.0001608	4.12E-06	0.000153
C	5.4685E-09	7.77E-10	3.94E-09



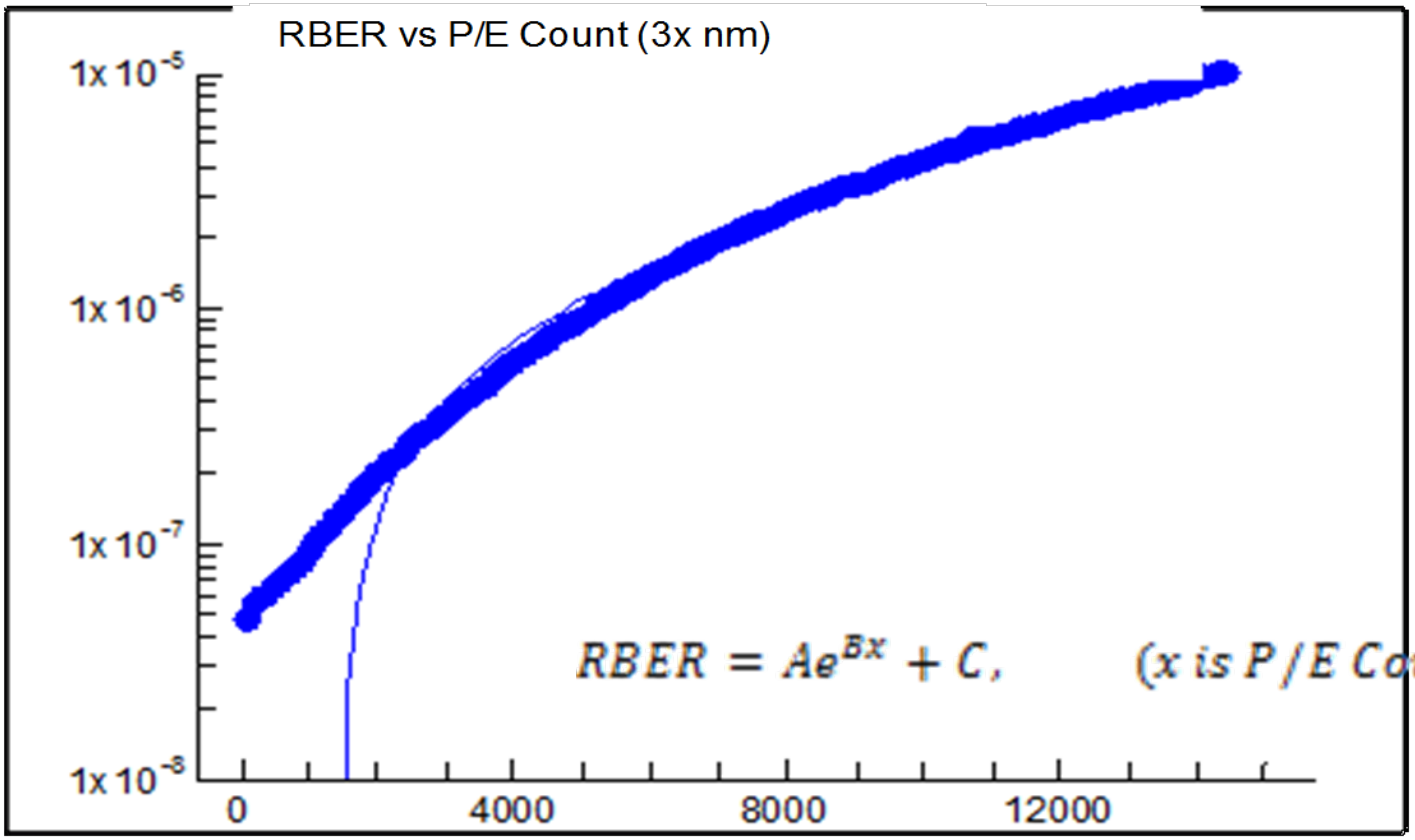
Endurance

4x nm MLC	Value	Std Err	95% Confidence
A	1.3631E-06	8.6E-08	1.19E-06
B	4.6896E-05	1.58E-06	4.38E-05
C	-1.4805E-06	1.01E-07	-1.7E-06



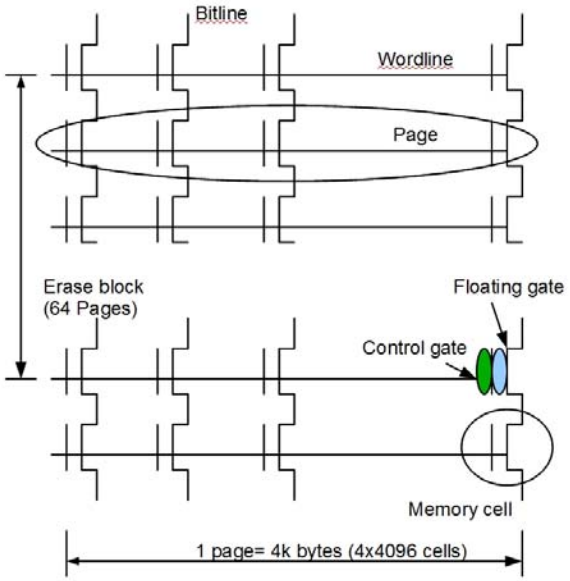
Endurance

3x nm MLC	Value	Std Err	95% Confidence
A	1.1831E-06	5.37E-08	1.08E-06
B	0.0001543	2.85E-06	0.000149
C	-1.4696E-06	8.01E-08	-1.6E-06



Read Disturb

Due to the layout of flash cells, some cells not being read will receive elevated voltage stress. Disturb occurs when charge collects on the floating gate, causing the cell to appear weakly programmed.

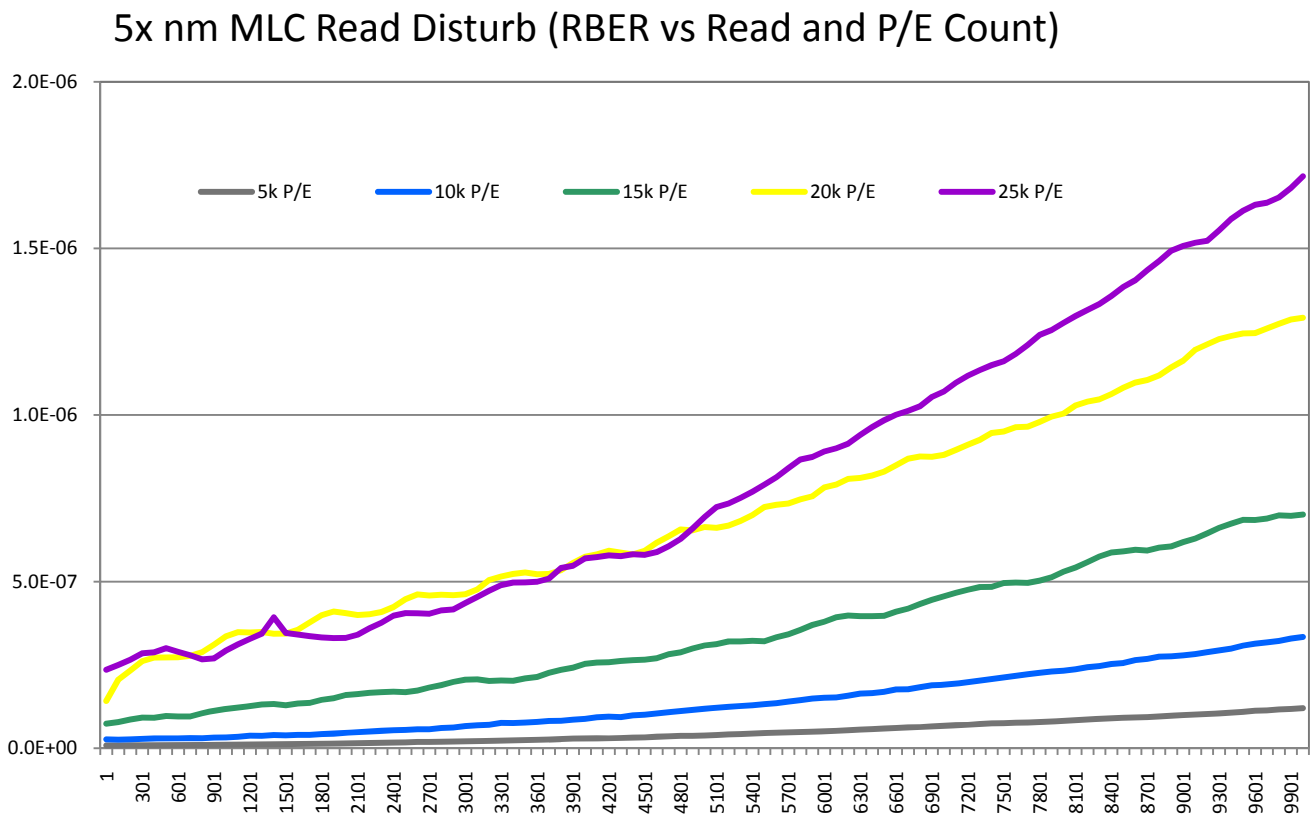


Read Disturb: *Test Regime*

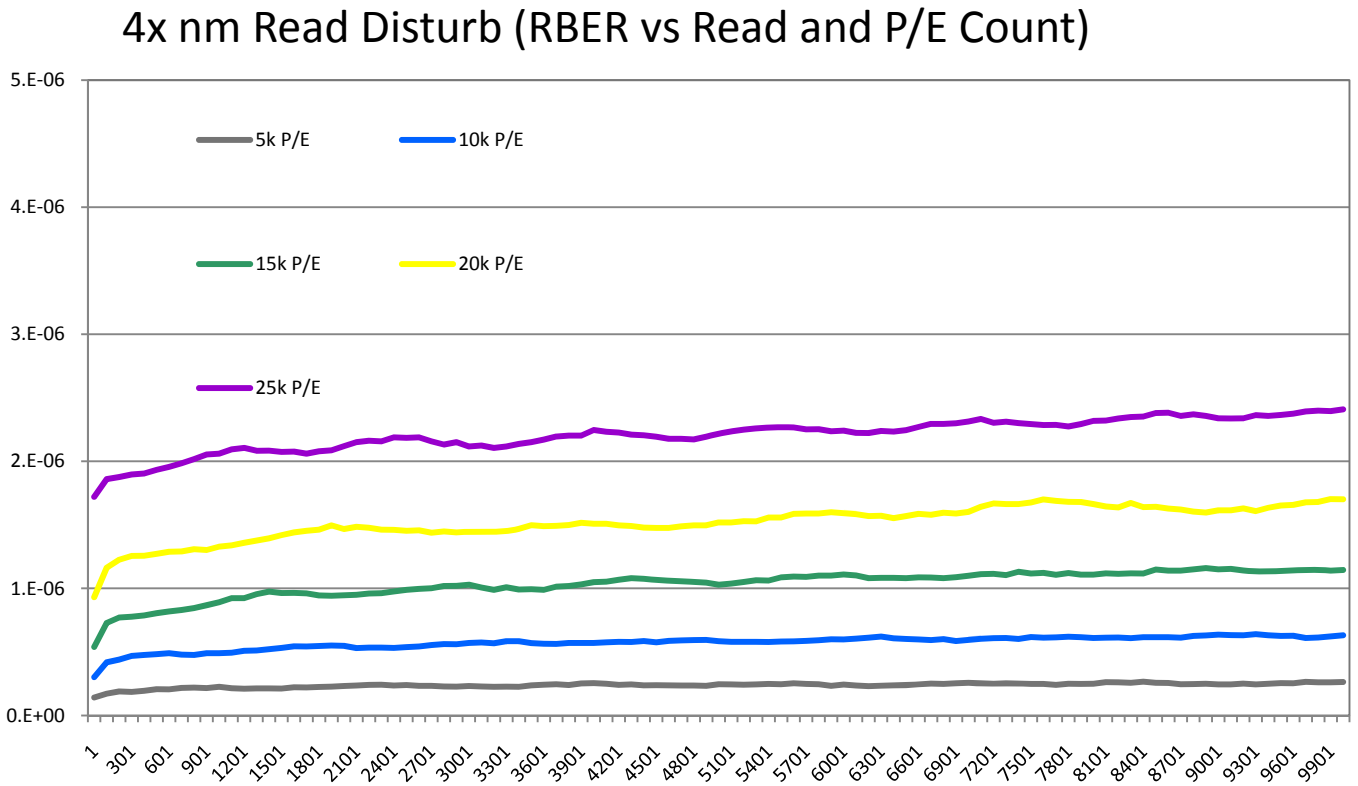
In Type II worn-out ranges: lay out random data pattern, read/verify

- ECC is disabled, flash die are accessed in “raw mode”
- RBER is evaluated *versus* Read and P/E Count

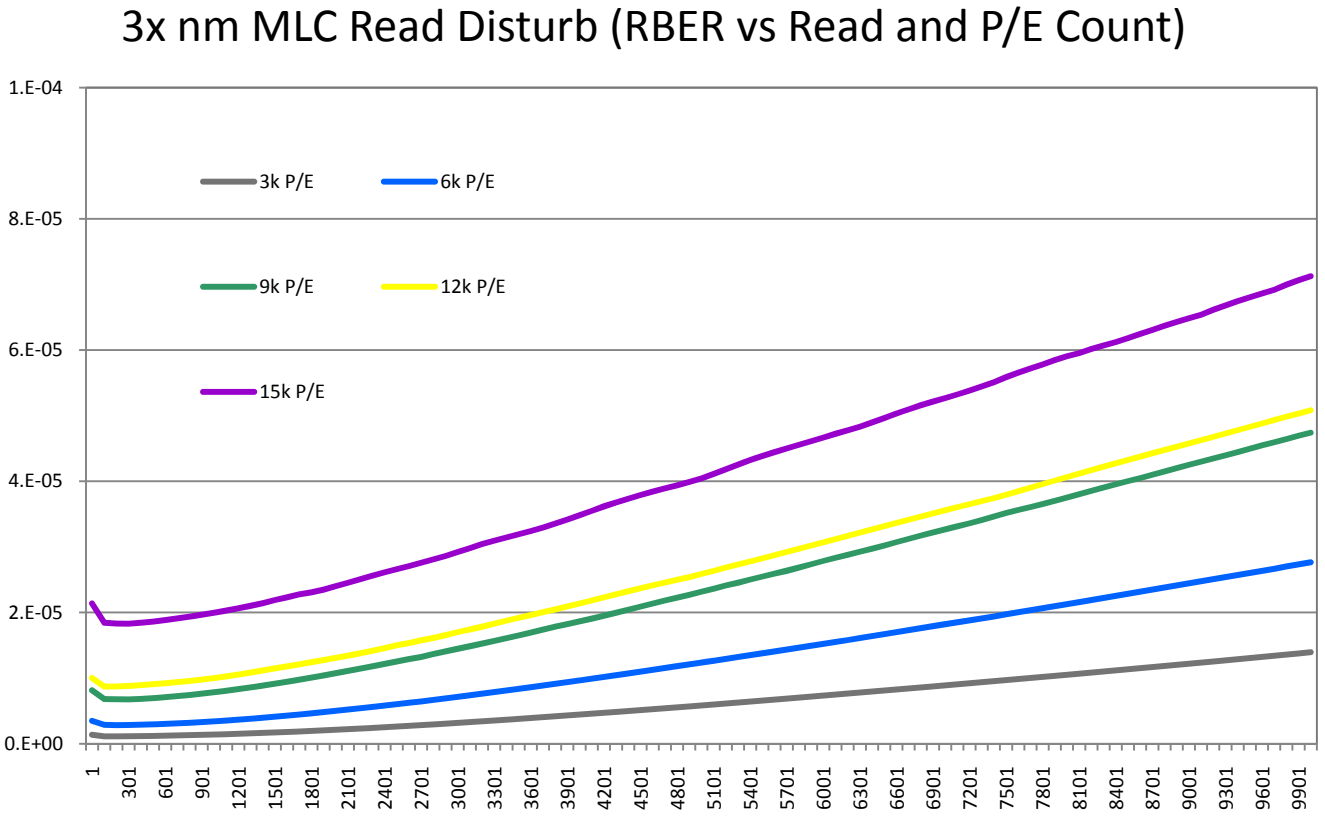
Read Disturb



Read Disturb



Read Disturb



Retention

Charge loss/gain occurring on the floating gate over time will lead to bit flipping. As a flash memory cell is repeatedly programmed and erased, the tunnel oxide and other insulating layers becomes weak leading to Stress-Induced Leakage Currents (SILC) and other charge losses.

De-trapping of trapped charges (surrogate for Programmed charge) also leads to change in effective stored signal.

Retention: *Test Regime*

In Type II worn-out ranges: lay out random data pattern, read/verify, bake at 70 C for pre-determined periods, read/verify

- Acceleration factor: 70 C/40C, 35:1, 21 hrs →1 month

Determine RBER vs Retention Time and P/E Count

- 3 months is a typical enterprise storage retention requirement

Retention

Data retention for 5x nm MLC (RBER)

P/E	0 month	1 month	2 month	3 month
5000	8.035E-09	4.512E-08	6.415E-08	7.67E-08
10000	1.621E-08	7.366E-08	1.280E-07	1.779E-07
15000	4.789E-08	1.479E-07	2.989E-07	4.891E-07
20000	1.122E-07	2.402E-07	5.320E-07	9.76E-07
25000	2.759E-07	5.008E-07	1.484E-06	3.139E-06

Data retention for 3x nm MLC (RBER)

P/E	0 month	1 month	2 month	3 month
3000	3.217E-07	4.077E-06	5.744E-06	7.268E-06
6000	9.555E-07	1.009E-05	1.571E-05	2.106E-05
9000	2.427E-06	2.169E-05	3.483E-05	4.78E-05
12000	4.386E-06	3.369E-05	5.676E-05	7.935E-05
15000	7.958E-06	5.88E-05	1.044E-04	1.506E-04

Retention

P/E cycling → weakness of cell dielectrics → retention loss

Data retention ability is heavily dependent on the applied ECC methods, RBER, flash age in terms of P/E cycles and disturb phenomena. So when we specify data retention, we also need to specify the context:

- What UBER we are targeting;
- What the ECC scheme capability is;
- How many P/E cycles we want to specify for the devices; and
- How many reads we will allow before we refresh data

Summary

Very large sample size study on production systems as part of Fusion-IO Reliability Qualification Test

Encompasses MAT, endurance, read disturb, retention-specific test suites

There exist tradeoffs between endurance, retention and system (FTL) design as pertains to fault handling, ECC methods, Flash management policy and component selection.