



System Modeling for NAND Flash Endurance

**Niles Yang, Sergey Musin*, Yichao Huang, Alexandra Bauche, Oleg Kragel,
Gautham Reddy, Michael Zhu, Gautam Dusija, Uday Chandrasekhar, Jianmin
Huang, Chris Avila, Steven Sprouse**

SanDisk Corp., Milpitas, California, USA.

***Softeq Development Corp., Minsk, Belarus**

Forward Looking Statement

During our meeting today we will be making forward-looking statements.

Any statement that refers to expectations, projections or other characterizations of future events or circumstances is a forward-looking statement, including those relating to industry trends, expenses, future memory technology, production capacity and technology transitions and future products.

Actual results may differ materially from those expressed in these forward-looking statements due to the factors detailed under the caption “Risk Factors” and elsewhere in the documents we file from time-to-time with the SEC, including our annual and quarterly reports.

We undertake no obligation to update these forward-looking statements, which speak only as of the date hereof.

Abstract

- *The physical limits of sub-20nm NAND flash memory require the SSD / eMMC system design to compensate for the memory deficiencies, including the endurance and disturbance.*
- *In order to achieve the optimized system design, a comprehensive and detailed understanding of the NAND endurance behavior upon the data transmission between the system and the host is required.*
- *However, it is often not feasible to obtain such detailed information from the system hardware. In this presentation, a system modeling environment from SanDisk is described and examples are given. This environment is used to model the NAND flash characteristics upon usage traces sent from the host to the SSD / eMMC system.*
- *The characteristics includes the system and NAND W/E and read endurance information, and the NAND flash read access patterns onto the memory at the wordline level.*
- *This information is useful for better understanding of the vulnerability of the NAND flash memory used in the system, and in turn, both system and NAND design can be optimized accordingly. As a result, higher product quality and better customer experience can be achieved.*

Outline

- Motivations
- NAND endurance degradation
- System modeling requirements and overview
- Modeling results: W/E Endurance
- Modeling results: Read endurance
- System Co-optimization with NAND
- Conclusion

Outline

- **Motivations**
- NAND endurance degradation
- System modeling requirements and overview
- Modeling results: W/E Endurance
- Modeling results: Read endurance
- System Co-optimization with NAND
- Conclusion

Motivations

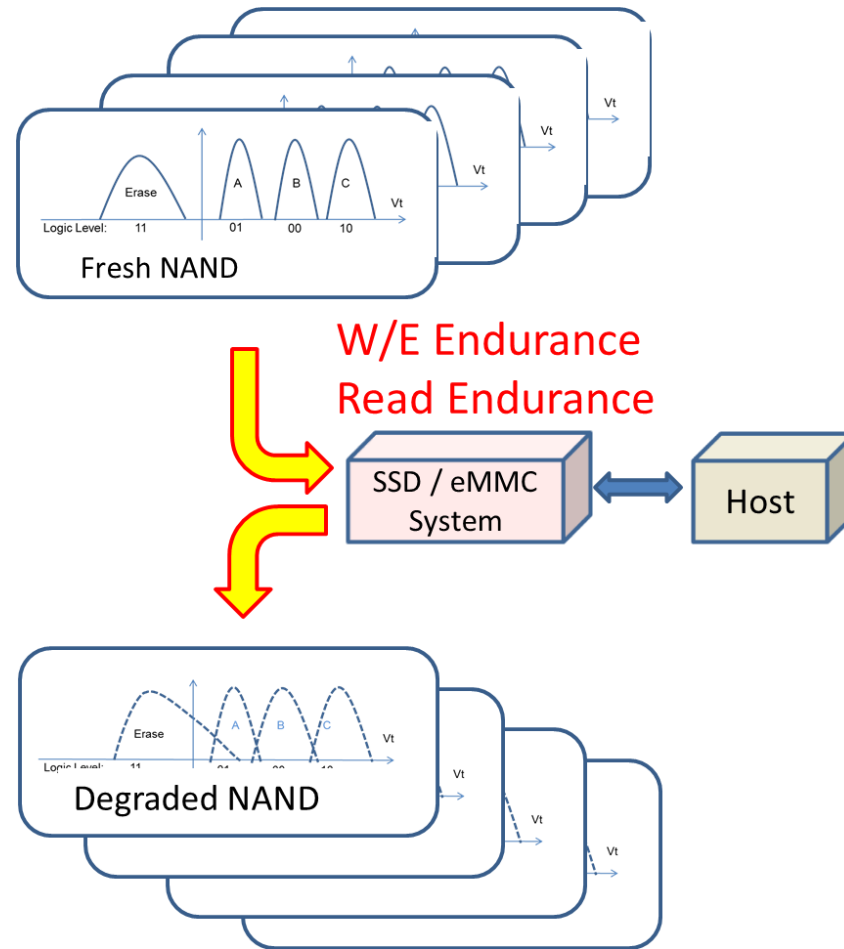
- Help thoroughly characterize and understand the endurance issues related to the SSD / eMMC products with the NAND flash memory for ultimate reliability and quality
- Provide a sufficient and fast modeling environment on the SSD / eMMC system for write and read endurance
- Better understand the system endurance requirements in order to optimize the system design and conduct the NAND endurance characterization effectively

Outline

- Motivations
- **NAND endurance degradation**
- System modeling requirements and overview
- Modeling results: W/E Endurance
- Modeling results: Read endurance
- System Co-optimization with NAND
- Conclusion

Degradation of the NAND flash at sub-20nm regime

- Worse endurance and disturbance than previous generations
- The system design needs to fully utilize the limited NAND endurance budget
- Systems with the sub-20nm NAND require interactive design optimization



Outline

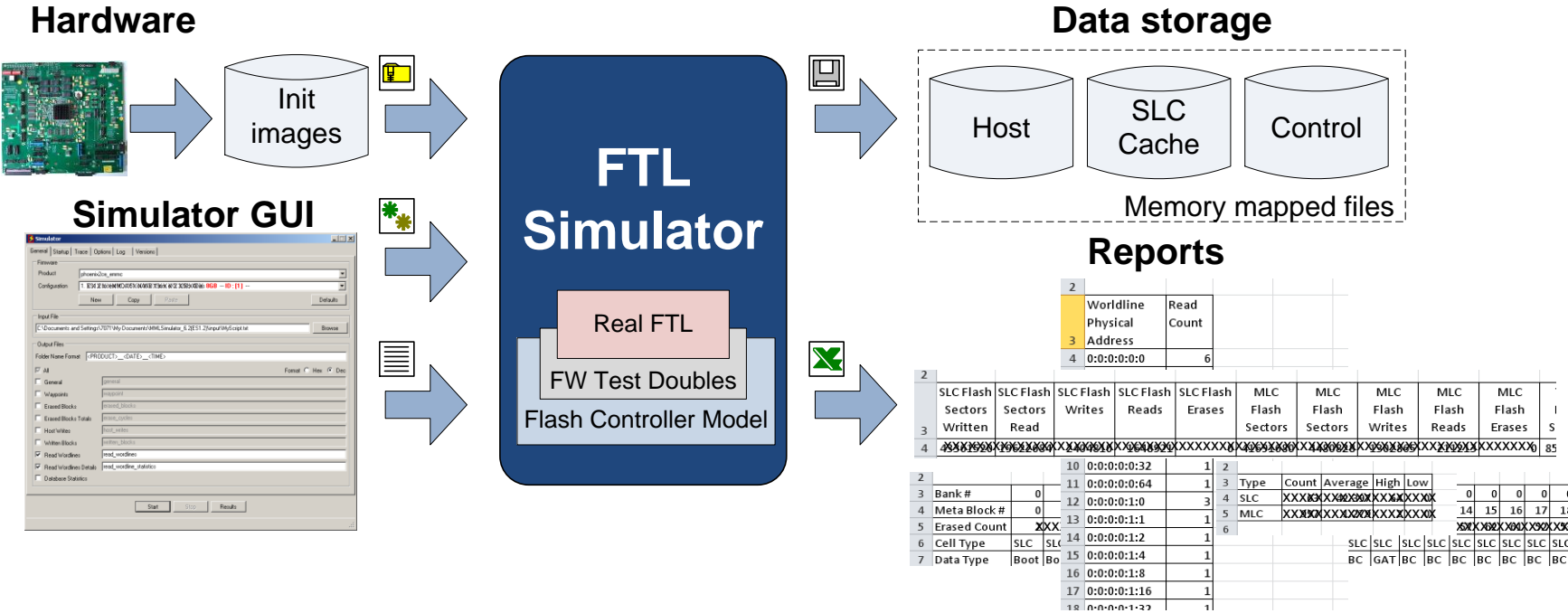
- Motivations
- NAND endurance degradation
- **System modeling requirements and overview**
- Modeling results: W/E Endurance
- Modeling results: Read endurance
- System Co-optimization with NAND
- Conclusion

System modeling environment (1): Requirements

- As close to the real products as possible while being highly efficient and informational
 - Integrated with advanced flash memory model
 - Utilizes production firmware images to model the product features
 - Based on the user usage traces
- Focusing on the Flash Translation Layer (FTL) simulation.
- Provide statistic reports to track the key system parameters.
- Easily deployed for large scale and multi-configuration simulations
- Friendly user interface and data format for post analysis

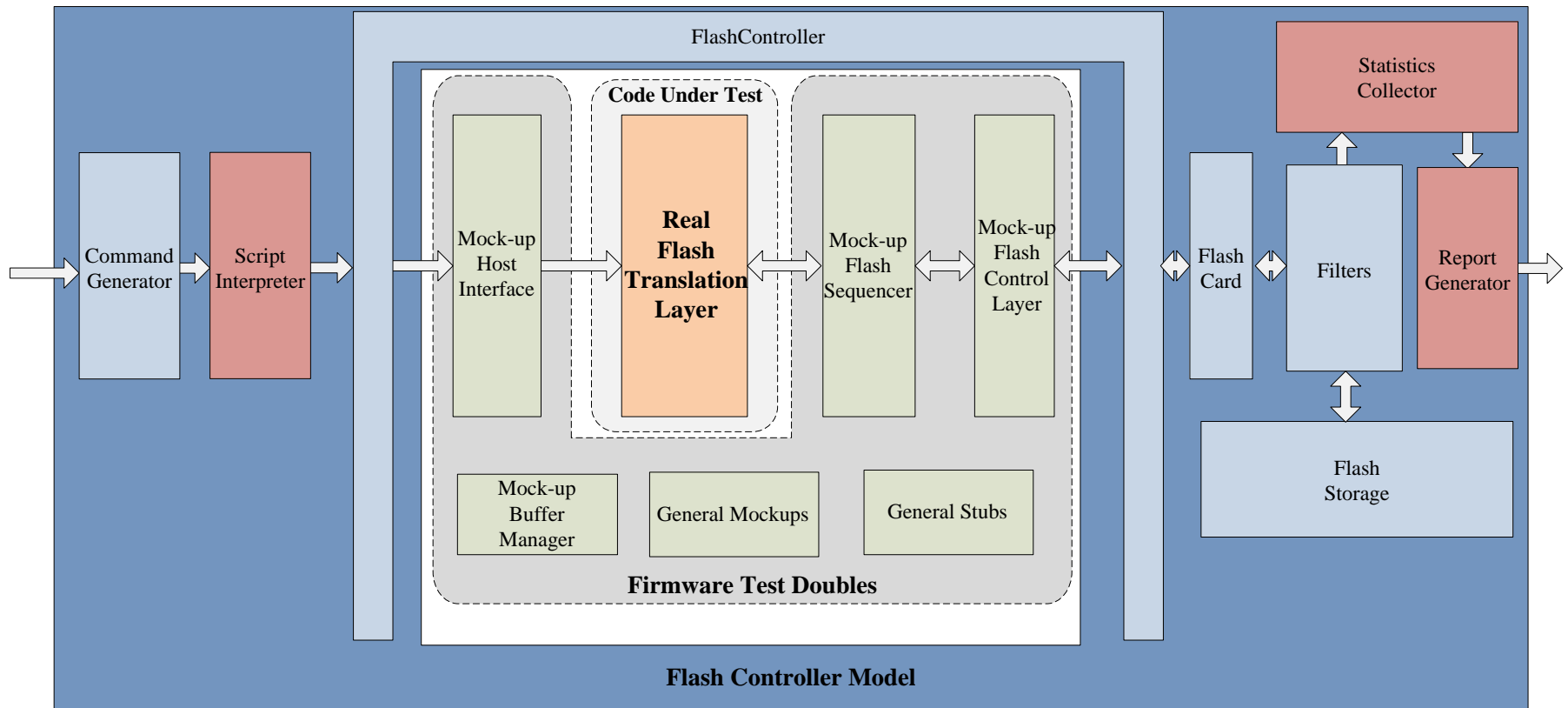
System modeling environment (2): Overview

- Hardware images created for the simulator
- Real FTL inside the simulator with system components.
- Input traces are applied and the statistical reports are generated from the modeling



System modeling environment (3): FTL Simulator implementation

- Firmware Test Doubles (FTD) used as the simulation layer with the Real FTL
- Mockups in the FTD are used for fast modeling speed
- Flash controller model include all the key components of the system



Outline

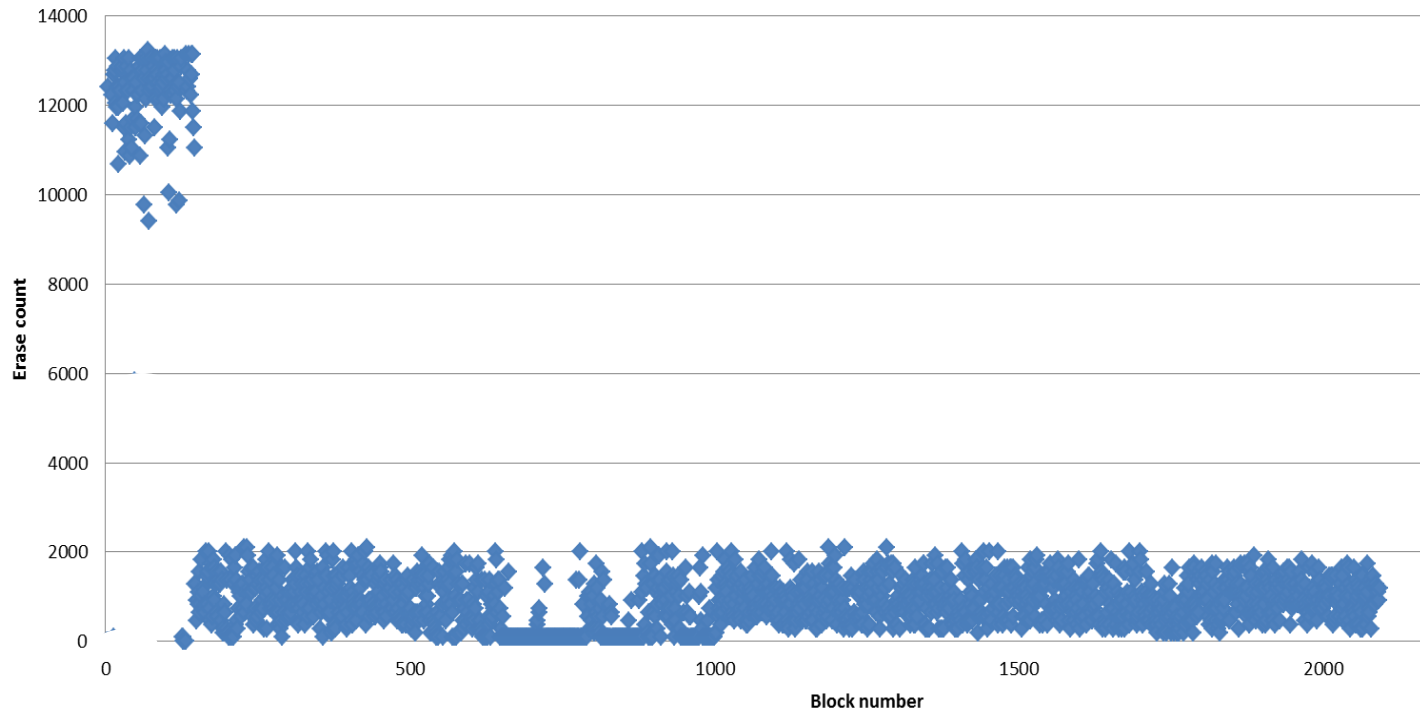
- Motivations
- NAND endurance degradation
- System modeling requirements and overview
- **Modeling results: W/E Endurance**
- Modeling results: Read endurance
- System Co-optimization with NAND
- Conclusion

Modeling results

- 32GB prototype eMMC system (see following slides):
 - W/E Endurance
 - Read endurance
- Hardware comparison
 - FTL Modeling achieves 20~30x faster speed than using the identically configured hardwares for endurance characterizations with sufficient accuracy

W/E Endurance Analysis

- SLC and MLC memory blocks show different W/E counts
- All write operations are captured
 - Host write operation
 - Internal system maintenance write operations
 - Wear-leveling
 - Write amplification factor (WAF): 1~2

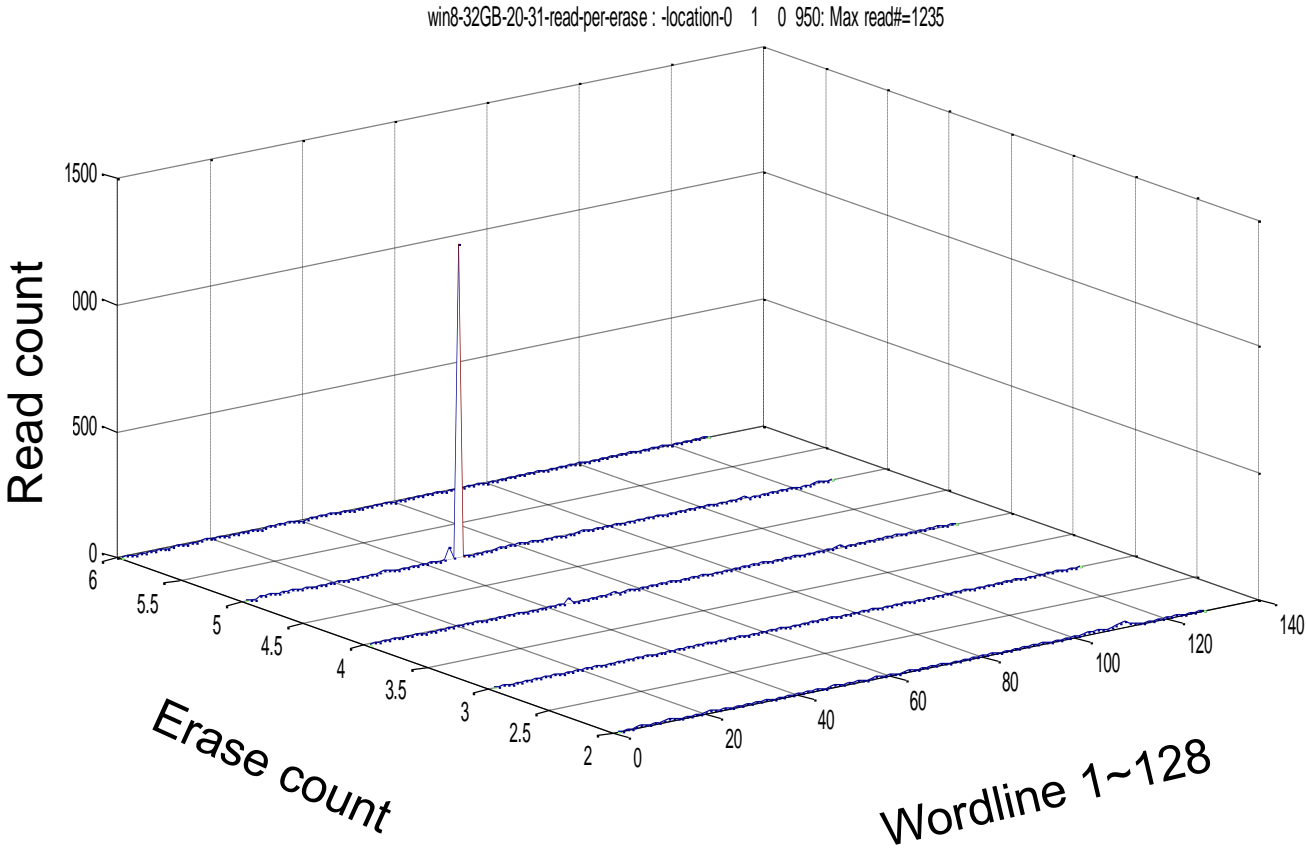


Outline

- Motivations
- NAND endurance degradation
- System modeling requirements and overview
- Modeling results: W/E Endurance
- **Modeling results: Read endurance**
- System Co-optimization with NAND
- Conclusion

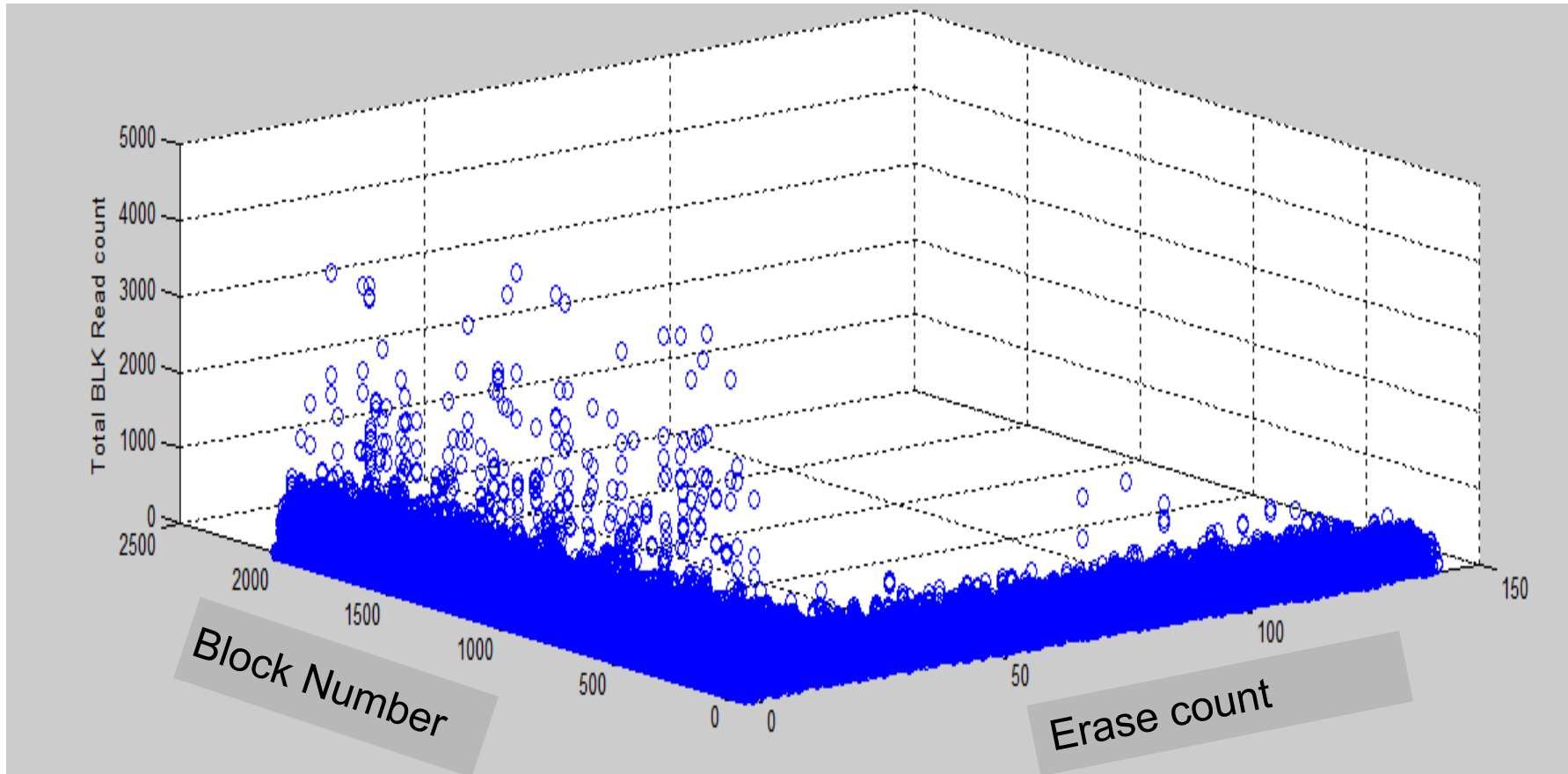
Read endurance with respect to the erase count within a block

- Within a block, read accesses to the wordline between erase counts can be significantly different from one erase to the next
- Wordline read pattern appears to be heavily single wordlines based
 - Read accesses to other wordlines are significantly lower



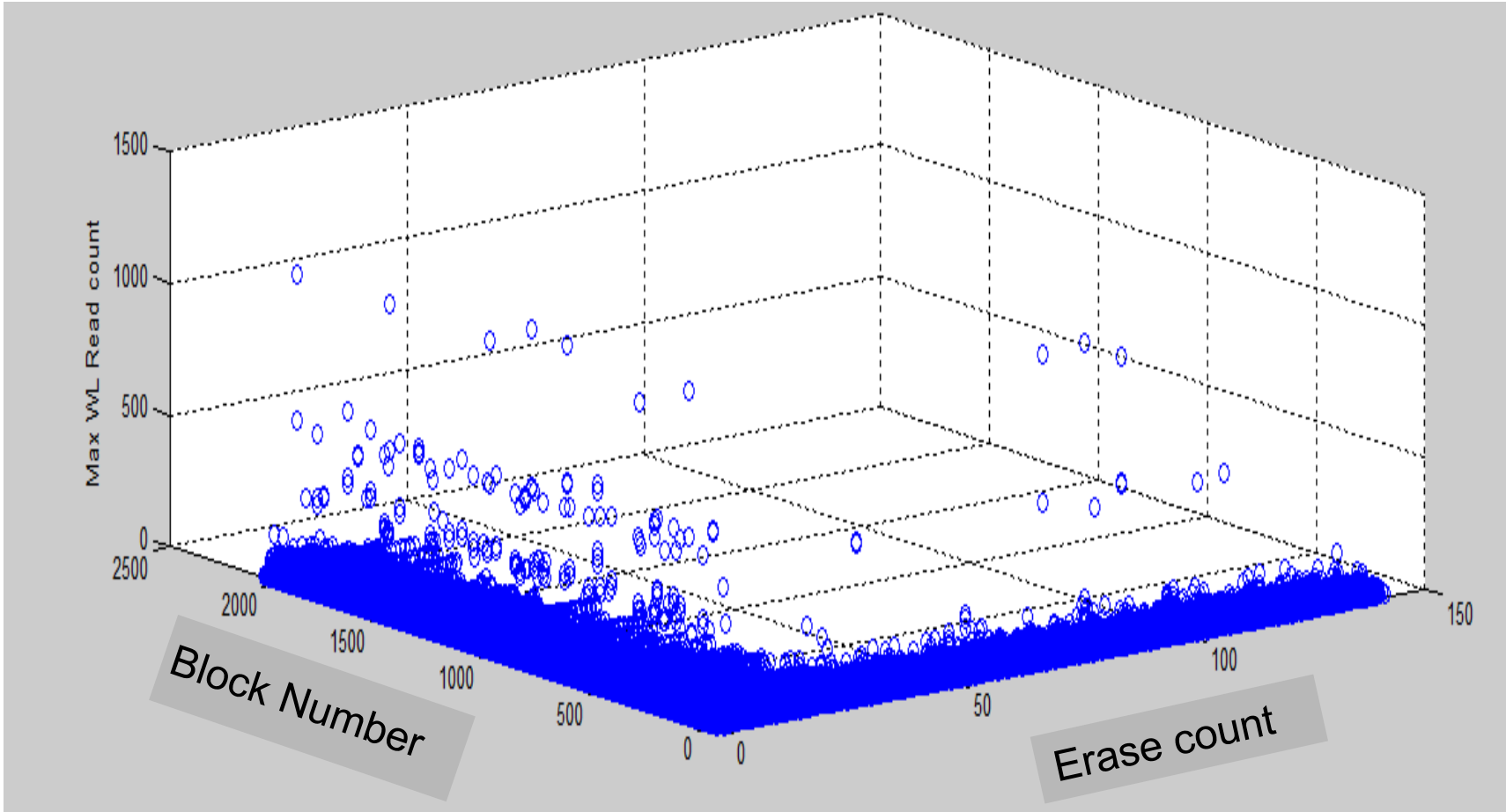
Read endurance with respect to the W/E cycle and block number (1)

- Block level read counts of all blocks at various erase counts
 - 3-D illustration of the block read counts as a function of block number and erase count
 - At any particular erase number, the combined read access to all the wordlines of a block is recorded



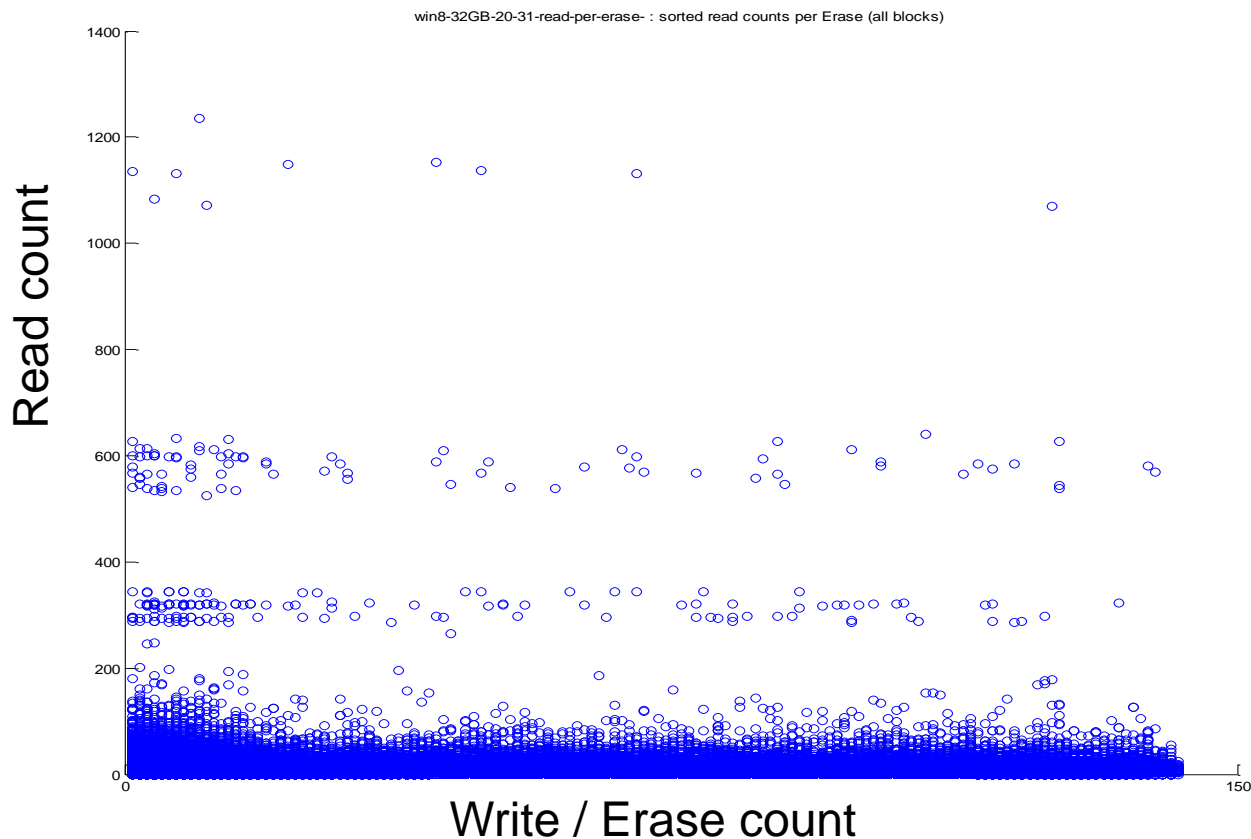
Read endurance with respect to the W/E cycle and block number (2)

- Wordline level read counts of all blocks at various erase counts
 - 3-D illustration of the maximum wordline read counts as a function of block number and erase count
 - At a particular erase number, the max read access to the wordlines of a block is sorted and recorded



Read endurance with respect to the system W/E cycles

- Read counts per wordline at various erase count
- Data combines read counts on all the wordlines of all the blocks at particular erase count
 - 40 million data points captured
 - Both host and internal read operations are included. Read amplification factor (RAF): ~2.



Summary of the endurance modeling

- W/E endurance includes all erase operation on the memory block.
- Read endurance study requires:
 - The statistics of each block in the memory to be recorded and tracked at every erase cycle
 - The statistics of each wordline in the memory to be recorded and tracked at every erase cycle
- Amplification effects:
 - Flash write and read includes both host and internal write and read
 - Modeling shows the complete statistics for both host and internal write and read operations
 - RAF can be obtained
- The most erased block is not necessarily the most read, vice versa.
- Implication of the connection between the write and read endurance

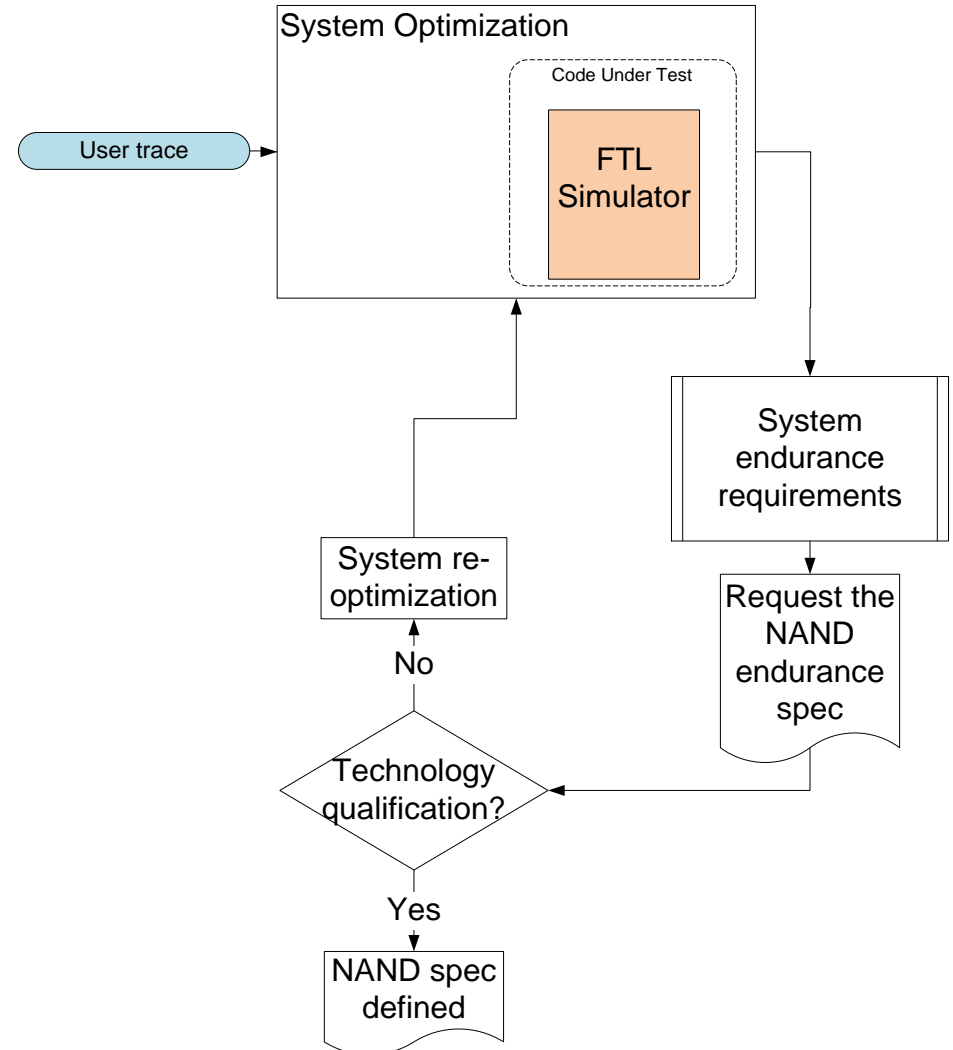
Outline

- Motivations
- NAND endurance degradation
- System modeling requirements and overview
- Modeling results: W/E Endurance
- Modeling results: Read endurance
- **System Co-optimization with NAND**
- Conclusion

System co-optimization with NAND

Product optimization requires:

- Endurance information for both W/E and read
- Optimum system reliability that requires NAND optimization
- System maintenance carefully designed to maximize the use of NAND endurance
- NAND endurance defined based on the system application
- Fully statistical FTL modeling used for accurate system evaluation with 25~30x faster evaluation time than the hardware
- Multi-configuration considerations



Outline

- Motivations
- NAND endurance degradation
- System modeling requirements and overview
- Modeling results: W/E Endurance
- Modeling results: Read endurance
- System Co-optimization with NAND
- **Conclusion**

Conclusion

- A SSD / eMMC system modeling environment is presented for endurance study.
- System endurance statistics can be obtained through modeling where the hardware may be slow and insufficient
- Both W/E and read endurances can be obtained down to the block and wordline level of the NAND memory
- Optimized product design could be achieved with the help of the FTL simulator to track the key system endurance parameters that are important for the NAND endurance optimization.

Acknowledgement

This work has been supported by both System Design Engineering and Firmware Engineering at SanDisk

©2013 SanDisk Corporation. All rights reserved. SanDisk is a trademark of SanDisk Corporation, registered in the U.S. and other countries.

**** Thank you ****