Reliability of 3D NAND flash memory with a focus on read voltage calibration from a system aspect

Nikolaos Papandreou, Nikolas Ioannou, Thomas Parnell, Roman Pletka, Milos Stanisavljevic,
Radu Stoica, Sasa Tomic, Haralampos Pozidis
IBM Research – Zurich
Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland
email: npo@zurich.ibm.com

Abstract—**This paper discusses the reliability challenges of 3D NAND flash memory and their impact on flash management for enterprise storage applications. Emphasis is given to the read voltage calibration and its critical role in achieving low error-rates and low latency read performance, as well as in enabling accurate block health estimation. We present experimental results that demonstrate the improvements in endurance, retention and read-disturb from different read voltage calibration schemes, and we address their requirements from a system perspective, i.e., the accuracy vs. complexity trade-off. We discuss the above aspects for state-of-the-art 3D TLC and QLC NAND flash memory.**

Keywords-3D NAND flash; quad-level cell (QLC); triple-level cell (TLC); endurance; data retention; read voltage calibration

I. INTRODUCTION

3D NAND flash memory has made inroads into the enterprise storage space and the data center, creating a new tier between fast, volatile main memory and slow, non-volatile hard-disk drives. Moreover, the technological advancements in vertical integration, cell design and manufacturing, and the improvements in the read and write algorithms, have enabled triple-level cell (TLC) and quad-level cell (QLC) NAND flash with enterprise-level reliability, achieving further capacity increase and cost-per-bit reduction [1]-[3].

In terms of reliability, 3D NAND flash has shown superior performance compared to the latest generations of 2D NAND flash in various aspects, e.g., in terms of endurance and data retention capabilities, and susceptibility to cell-to-cell interference [4], [5]. However, 3D NAND flash exhibits its own reliability issues, such as fast and abrupt raw bit-error rate (RBER) changes under retention and read-disturb stress, as well as increased page and layer variability as a result of process-induced variations [6]-[10]. As the number of bits per cell increases from 3 bits in TLC to 4 bits in QLC, the reliability challenges are further exacerbated. To meet the endurance requirements of enterprise systems, the NAND flash memory controller needs to accommodate advanced flash management and error-handling algorithms that are able to perform under various types of workloads [11], [12]. Read voltage calibration is a key component of the NAND flash memory controller that aims to track and adjust the read voltages according to the specifics of the workload stress. For example, read- or write-intensive workloads have a different impact on the error characteristics of the NAND flash blocks and on the underlying threshold voltage (V_{TH}) distributions.

In this paper, we discuss the reliability issues of 3D NAND flash with a focus on read voltage calibration from a system perspective. Section II compares the relative endurance and read latency between the latest 2D and newest 3D NAND technologies and presents some of the flash management algorithms of modern NAND flash controllers. Section III focuses on the read voltage calibration and discusses the complexity aspects for TLC and QLC NAND. Section IV first discusses the challenges that arise with the increase in characterization data with modern TLC and QLC NAND flash and then presents experimental measurements from state-of-the-art 3D TLC NAND flash devices. The characterization results demonstrate the reliability challenges of 3D TLC NAND and motivate the need for effective and efficient read voltage calibration. To address the calibration requirements from a system perspective, we present different read voltage calibration strategies and discuss their trade-offs in terms of accuracy and complexity. Finally, Section V summarizes the results of this paper.

II. FLASH MANAGEMENT FOR ENTERPRISE STORAGE

Fig. 1 compares the relative endurance and read latency of memory devices from different NAND flash generations. The data points are normalized to the 1x nm planar MLC NAND technology. The endurance numbers were obtained from large-scale device characterization and denote the average block endurance assuming that the same error correction strength is used in all cases and that read voltage calibration is employed, i.e., the read voltages are optimal before any page read. The read latency numbers correspond to the average latency among the different page types of each technology, e.g., lower and upper pages for MLC NAND.

The technological advancements with 3D NAND flash have enabled TLC and QLC with endurance similar to or better than the latest 2D MLC NAND. Further, QLC NAND offers a 33% capacity increase per cell compared to TLC, however, it shows lower endurance and higher read latency. The lower

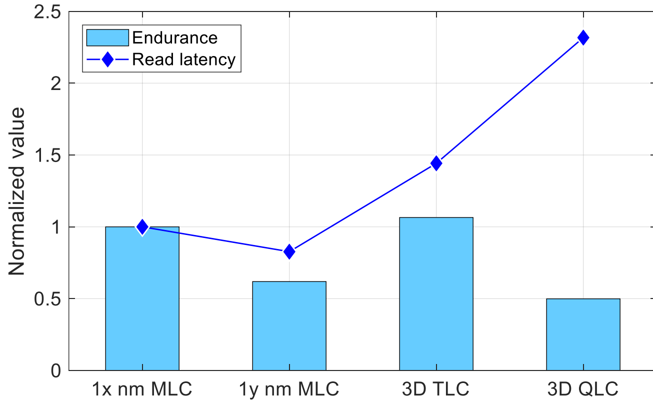

Figure 1. Comparison of relative endurance (in P/E cycles) and read latency (in µs) between latest 2D and recent 3D NAND flash technologies. All data points are normalized to the 1x nm MLC NAND technology.

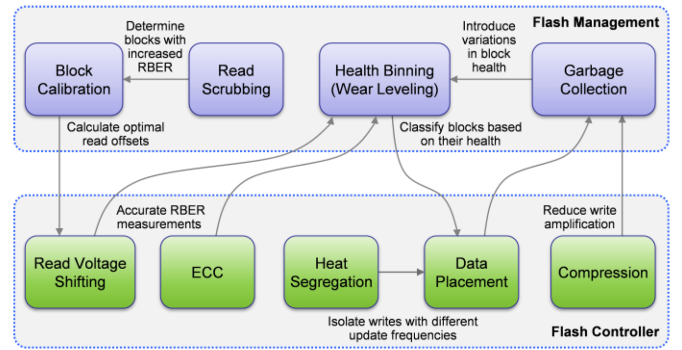

Figure 2. Block diagram illustrating flash management functions and synergies between firmware and hardware in a state-of-the-art NAND flash memory controller. Read voltage calibration is a key component to enable low error-rates and low latency read performance, as well as accurate block health estimation and grading.

endurance is attributed to the increase of the V_{TH} levels and the subsequent reduction of the available margin between the levels, which can lead to significant overlapping and thus an increase of bit-errors. The higher read latency is expected since more read voltages need to be applied in QLC to extract the 4 bits of information from each cell. One approach to improve the endurance and reduce the read latency experienced by the user is the adoption of a hybrid SLC/QLC controller architecture with a fast SLC cache of variable size depending on parameters such as the device utilization and workload properties [12], [13].

Fig. 2 gives an overview of NAND flash management algorithms and their synergies in a state-of-the-art NAND flash controller [11], [12]. In addition to error count information obtained directly from the host reads, background read scrubbing is used to identify blocks and pages with increased error rates. This enables timely block calibrations such that the optimal read voltages are available in advance, i.e., before the host attempts to read the data, thereby leading to a significant reduction of read retries. Moreover, the use of optimal read voltages is essential in order to get accurate error count measurements for estimating and monitoring the health of a block, e.g., detect permanent error-rate degradation of a block due to wear caused by program/erase (P/E) cycling effects. This information is typically reported by the error-correction code (ECC) decoder and can be leveraged by the controller's flash management to perform health binning.

Health binning is a method of wear leveling that classifies flash blocks into a range of health grades, so that the data placement unit can efficiently place write-hot data into healthier blocks and write-cold data into less healthy blocks. It was shown in [14] that health binning is capable of improving the endurance by up to 80%. Combined with accurate write-heat separation in the data placement process, significant reductions in the write amplification can be further achieved as demonstrated in [11]. As the garbage collection process selects blocks based on the amount of invalid data in the block, it naturally introduces variations in the block health as some blocks are garbage collected significantly earlier than others. This variation is used for grading blocks in health binning. Further, data compression helps to reduce write amplification since less pages have to be written in NAND flash for the same amount of host data.

III. READ VOLTAGE CALIBRATION

Due to the changes in the V_{TH} distributions as a result of retention time, read-disturb, program-disturb and other effects, it is apparent that the read voltages need to be re-adjusted, either periodically or on-demand. Read voltage calibration aims to determine a set of voltage offset values to adjust (correct) the read voltages by taking into account the direction and amount of the V_{TH} changes, so that the number of bit errors during a page read is minimized. Fig. 3 illustrates the effect of negative shift of the programmed V_{TH} distributions due to charge loss over time in a TLC NAND with 8 V_{TH} levels (L0, …, L7). As a result, the vertical solid lines corresponding to the read voltages (V1, …, V7) after programming need to be adjusted by a set of offset values (ΔV1, …, ΔV7) that account for the retention effects. In general, the amount of shift of each V_{TH} distribution, and thus the corresponding read voltage offset value, may be different for the various levels.

Typically, the memory vendors provide specific commands that allow the NAND flash controller to instruct the chip to

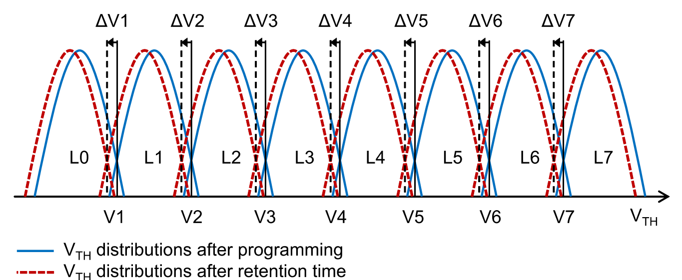

Figure 3. Illustration of read voltage adjustment due to retention errors in TLC NAND flash with 8 V_{TH} levels.

apply a set of offset values on the default read voltages prior to a page read. Those offset values can be selected from a predefined list of available options that each is suitable for a specific type of stress, e.g., X number of program/erase cycles, or Y number of read-disturb cycles, etc., or can be calculated on the fly based on an algorithm that uses information from the read data, e.g., error count or bit flips [15]-[18].

Read retry is a common dynamic read voltage adjustment method in case of ECC failure after a page read. In particular, the controller attempts to retry reading the page by applying a corrective set of predefined read voltage offsets with the goal of reducing the error count and thus passing the ECC decoder. This process, however, may need to be applied several times by trying different offset options, which results in an increase in the overall read latency.

On the other hand, background read voltage calibration refers to algorithms that run periodically in the controller as a background process. The goal of such algorithms is to pro-actively calculate and store the optimal read voltage offsets for the different pages and blocks, such that those offset values can be directly applied prior to a host read, thus resulting in an ECC pass and minimum latency. Fig. 4 illustrates an example of an algorithmic process that aims to find the new optimal offset value for a particular read voltage by searching towards the direction (positive or negative update) that minimizes the bit errors. The algorithm may try a fixed or variable number of offset values based on constant or adaptive step size.

Read voltage calibration has become a challenging task for TLC and QLC NAND due to the increased number of read voltages (7 for TLC, 15 for QLC) and increased number of pages and layers per block in modern 3D NAND devices. This has a direct impact on the amount of metadata required to keep track of the read voltage offsets for different pages and blocks, as well as on the background read overhead to perform the calibration. In the next section, we discuss methods to reduce this overhead.

IV. 3D NAND Flash Reliability

A. Challenges with 3D NAND Characterization Data

With the increase of block capacity in modern 3D NAND flash, there is a significant increase in the amount of data that are collected during device testing and characterization. As an example, we consider the case where the blocks under test are selected from multi-die packages of 64-layer 3D TLC NAND. Given a typical page size of 16kB, a block size of 1k to 2k pages (the block size typically varies across different vendors), a total of 16 blocks under test per package, 10 characterization readouts (e.g., 5 cycling data points followed by 5 retention data points), a sweep of 9 read voltages (e.g., 4 positive, 4 negative, and default–zero offset values), a total of 45GB of data are collected per run. This amount should be further multiplied by the number of packages tested in parallel.

This substantial increase in experimental data poses significant challenges in analyzing the results but at the same

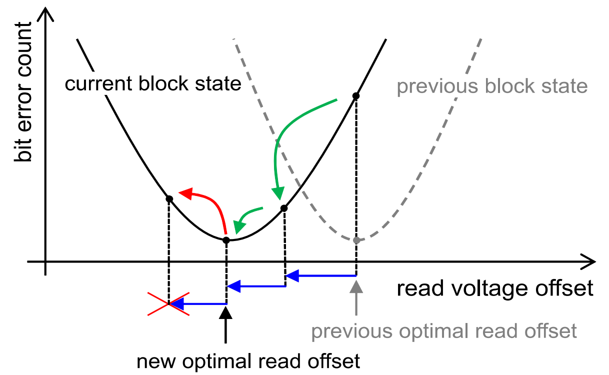

Figure 4. Illustration of the read voltage calibration principle. The read voltage is adjusted by a corrective offset that corresponds to the changes of the V_{TH} distributions in the current state of the block.

time offers new opportunities. Recently, various efforts have been made to utilize machine learning (ML) to gain valuable insights from raw NAND flash characterization data [19]-[22]. With the advances in modern ML algorithms in both accuracy and expressiveness, ML-based data analytics can be a valuable tool for characterization, as well as for real-time flash management applications in 3D NAND flash-based systems.

B. 3D TLC NAND Characterization Measurements

Fig. 5 shows characterization measurements of 64-layer 3D TLC NAND blocks where each block is subjected to successive phases of 3k P/E cycles followed by (a) 3 weeks retention time or (b) 6k read-disturb cycles. In both cases, the maximum page-RBER is reported under default (blue), semi-optimized (green) and optimized read voltages (red). The *optimized* method refers to a calibration scheme that finds the optimal offset value for each page in the block, where each read voltage is adjusted independently. The *semi-optimized* method corresponds to a version with less complexity and lower overhead, where all the read voltages of each page type are adjusted by the same offset. The results of Fig. 5 clearly highlight the importance of read voltage calibration in maintaining a low RBER under different device stress conditions. The RBER improvements, under both retention and read-disturb stress, are substantial and, in particular, they enable an extension of the lifetime of the NAND blocks compared to using the default read voltages. In [9], the analysis of the V_{TH} distributions in 3D TLC NAND presented the tracking and updating characteristics of the optimal offsets for each read voltage and page type under different types of device stress. Moreover, it was shown that the different pages have different read offset requirements.

One approach to reduce the amount of metadata and background read overhead is to organize pages in groups with similar characteristics, so that a set of offset values can be maintained for all pages of the same group. In this case, the RBER curves for both *optimized* and *semi-optimized* read voltages will deteriorate depending on the size of the groups. With the increase in the number of pages per block in QLC NAND, the amount of calibration metadata increases as well.

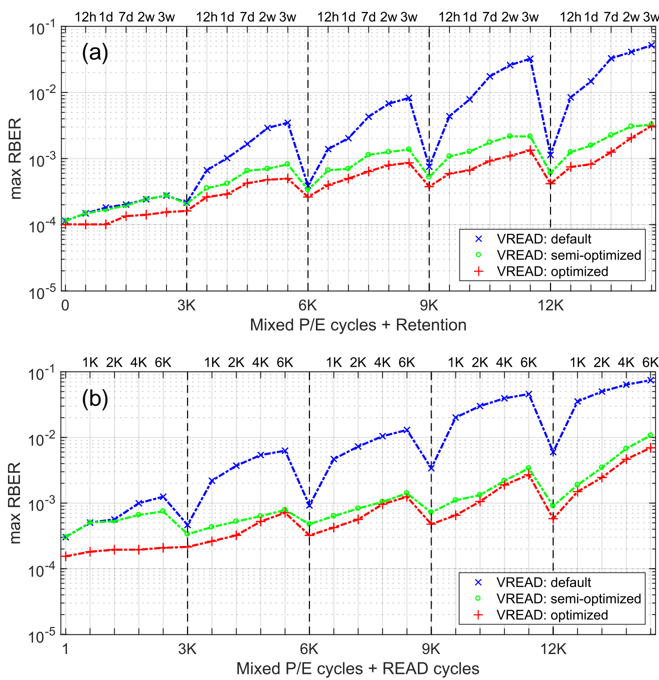

Figure 5. Characterization measurements of 3D TLC NAND blocks corresponding to successive phases of 3k P/E cycles followed by (a) 3 weeks retention or (b) 6k block read-only cycles.

Therefore, a trade-off between complexity and accuracy is expected depending on the calibration scheme. In [6]-[9], the abrupt RBER increase during early retention and read-disturb stress was characterized for MLC and TLC 3D NAND. These results emphasize the need for accurate and timely read voltage calibration in modern 3D NAND flash controllers.

V. Conclusions

In this paper, we discussed the importance of read voltage calibration to improve the RBER and extend the lifetime of the blocks in modern 3D TLC and QLC NAND flash controllers. Due to the increasing number of layers and pages per block, in addition to the higher number of read voltages from TLC to QLC, calibration becomes a challenging task in terms of metadata and background read overhead. We presented various approaches in terms of algorithms and grouping of pages to reduce the computational burden of calibration from a system perspective. Characterization results based on 3D TLC NAND demonstrated the effectiveness of the presented methods.

Acknowledgment

The authors wish to thank their colleagues from the IBM FlashSystem development and characterization teams.

References

[1] H. Kim, S. Ahn, Y. G. Shin, K. Lee, and E. Jung, “Evolution of NAND Flash Memory: From 2D to 3D as a Storage Market Leader,” IEEE International Memory Workshop (IMW), 2017.

[2] T. Griffin, P. Breen, G. Tressler, and N. Papandreou, “Can Next-Generation 3D TLC NAND Extend Enterprise Applications?,” Flash Memory Summit, 2017.

[3] S. Venkatesan and M. Aoulaiche, “Overview of 3D NAND Technologies and Outlook,” 18th Non-Volatile Memory Technology Symposium (NVMTS), 2018.

[4] N. Righetti and G. Puzzilli, “2D vs 3D NAND technology: Reliability benchmark,” IEEE International Integrated Reliability Workshop (IIRW), 2017.

[5] K. Mizoguchi, T. Takahashi, S. Aritome, and K. Takeuchi, “Data-Retention Characteristics Comparison of 2D and 3D TLC NAND Flash Memories,” IEEE International Memory Workshop (IMW), 2017.

[6] P. Breen, T. Griffin, N. Papandreou, T. Parnell, and G. Tressler, “3D NAND Assessment for Next Generation Flash Applications,” Flash Memory Summit, 2016.

[7] Y. Luo, S. Ghose, Y. Cai, E. F. Haratsch, and O. Mutlu, “Improving 3D NAND Flash Memory Lifetime by Tolerating Early Retention Loss and Process Variation,” ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems, 2018.

[8] Q. Xiong et al., “Characterizing 3D Floating Gate NAND Flash,” ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems, 2017.

[9] N. Papandreou, H. Pozidis, T. Parnell, N. Ioannou, R. Pletka, S. Tomic, P. Breen, G. Tressler, A. Fry, and T. Fisher, “'Characterization and Analysis of Bit Errors in 3D TLC NAND Flash Memory,” IEEE International Reliability Physics Symposium (IRPS), 2019.

[10] C. Zambelli, R. Micheloni, and P. Olivo, “Reliability challenges in 3D NAND Flash memories,” IEEE International Memory Workshop (IMW), 2019.

[11] R. Pletka, I. Koltsidas, N. Ioannou, S. Tomic, N. Papandreou, T. Parnell, H. Pozidis, A. Fry, and T. Fisher, “Management of Next-Generation NAND Flash to Achieve Enterprise-Level Endurance and Latency Targets,” ACM Transactions on Storage 14(4), 33:1-33:25, 2018.

[12] R. Pletka, R. Stoica, N. Ioannou, S. Tomic, N. Papandreou, and H. Pozidis, “Designing Enterprise Controllers with QLC 3D NAND,” Flash Memory Summit, 2018.

[13] R. Stoica, R. Pletka, N. Ioannou, N. Papandreou, S. Tomic, and H. Pozidis, “Understanding the design trade-offs of hybrid flash controllers,” IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), 2019.

[14] R. Pletka and S. Tomic, “Health-Binning: Maximizing the Performance and the Endurance of Consumer-Level NAND Flash,” 9th ACM International Systems and Storage Conference (SYSTOR), 2016.

[15] N. Papandreou, T. Parnell, H. Pozidis, T. Mittelholzer, E. Eleftheriou, C. Camp, T. Griffin, G. Tressler, and A. Walls, “Enhancing the reliability of MLC NAND flash memory systems by read channel optimization,” ACM Trans. on Design Automation of Electronic Systems, 20 (4), 2015.

[16] Y. Luo, S. Ghose, Y. Cai, E. F. Haratsch, and O. Mutlu, “HeatWatch: Improving 3D NAND flash memory device reliability by exploiting self-recovery and temperature awareness,” IEEE International Symposium on High Performance Computer Architecture (HPCA), 2018.

[17] Q. Li, M. Ye, Y. Cui, L. Shi, X. Li, and C. J. Xue, “Sentinel cells enabled fast read for NAND flash,” 11th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage), 2019.

[18] K. Ho, P. Fang, H. Li, C. Wang, and H. Chang, “A 45nm 6b/cell charge-trapping flash memory using LDPC-based ECC and drift-immune soft-sensing engine,” IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), pp. 222–223, 2013.

[19] B. Fitzgerald, D. Hogan, C. Ryan, and J. Sullivan, “Endurance prediction and error reduction in NAND flash using machine learning,” 17th Non-Volatile Memory Technology Symposium (NVMTS), 2017.

[20] Y. Nakamura, T. Iwasaki, and K. Takeuchi, “Machine Learning-Based Proactive Data Retention Error Screening in 1Xnm TLC NAND Flash,” IEEE International Reliability Physics Symposium (IRPS), 2016.

[21] C. Zambelli et al., “Characterization of TLC 3D-NAND Flash Endurance through Machine Learning for LDPC Code Rate Optimization”, IEEE International Memory Workshop (IMW), 2017.

[22] Y. Liao, C. Huang, C. Zeng, and H. Chang, “Data Analysis and Prediction for NAND Flash Decoding Status”, IEEE International Memory Workshop (IMW), 2017.