

Characterizing 3D Charge Trap NAND Flash: Observations, Analyses and Applications

Fei Wu[†], Yue Zhu[†], Qin Xiong[†], Zhonghai Lu[‡], You Zhou^{¶†}, Weizhen Kong[§], Changsheng Xie[†]

[†]Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, China

[‡]School of Information and Communication Technology, KTH Royal Institute of Technology, Stockholm, Sweden

[¶]School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China

[§]Huawei Technologies, Shenzhen, China

{wufei, yuezhu}@hust.edu.cn, {qinxiong, zhonghai}@kth.se, zhouyou@hust.edu.cn,

kongweizhen@huawei.com, Cs_xie@hust.edu.cn

Abstract—In the 3D era, the Charge Trap (CT) NAND flash is employed by mainstream products, thus having a deep understanding of its characteristics is becoming increasingly crucial for designing flash-based systems. In this paper, to enable such understanding, we implement comprehensive experiments on advanced 3D CT NAND flash chips by developing an ARM- and FPGA-based evaluation platform. Based on the experimental results, we first make distinct observations on the characteristics of 3D CT NAND flash, including its performance and reliability features. Then we give analyses of the observations from physical and circuit aspects. Finally, based on the unique characteristics of 3D CT NAND flash, suggestions to optimize the flash management algorithms in real applications are presented.

Index Terms—3D CT NAND flash, performance, reliability

I. INTRODUCTION

Over the past years, NAND flash has become the mainstream storage medium since it provides non-volatility and high performance. To satisfy the sustained demand growth in the storage market, NAND flash memory vendors continuously shrink the feature size according to the Moore’s Law, improving the density and meanwhile lowering the per-bit price. However, due to the technical complexity, the scaling down becomes increasingly challenging and costly, especially when the feature size is below 30nm [1], which poses greater urgency of innovative flash architecture. 3D NAND flash represents a very promising solution to solve the constraints of planar devices. By enabling scaling in Z direction rather than X and Y directions, tremendous improvements in storage density can be achieved, while vendors can employ larger feature sizes to guarantee the reliability. In the 3D era, there are two noteworthy technology trends:

- *Charge Trap is becoming a mainstream technology.* Unlike planar NAND flash products, which are based on *Floating Gate (FG)* [2] technology, all vendors except Micron/Intel joint venture have chosen *Charge Trap (CT)* [3] in their 3D NAND products. The main difference between these two technologies is that an FG cell employs conductive polycrystalline silicon as the medium to store electrons while a CT cell adopts insulating charge trapping layer, which is typically made of silicon nitride. Compared with

This research is sponsored by the National Natural Science Foundation of China under Grant No 61872413 , No. U1709220, No. 61821003, and Wuhan Science and technology Project No. 2017010201010108, and the Fundamental Research Funds for the Central Universities No.2016YXMS019, and the 111 Project (No. B07038). This work is also supported by Key Laboratory of Data Storage System, Ministry of Education.

*Corresponding Author: Changsheng Xie, Cs_xie@hust.edu.cn

FG NAND flash, CT NAND flash has better scalability and less coupling effects among cells [4].

- *TLC dominates.* *Triple-level cell (TLC)* NAND flash, which stores 3 bits per cell, only accounts for a small fraction of the planar NAND flash market due to the poor reliability compared with *single-level cell (SLC)* (1 bit per cell) and *multi-level cell (MLC)* (2 bits per cell). Due to the employment of the rolled back process technology node, the emerging 3D TLC is able to achieve the technical specifications of the last planar MLC flash, and has accounted for over 50% of the industry bits by the end of 2017, which is predicted to exceed 80% in 2019 [5].

In NAND flash, once a page is programmed, it cannot be over-programmed until being erased. However, program and erase operations are performed at different granularities (program→page and erase→block). In order to simulate NAND flash as a block device in general-purpose file systems by hiding the out-of-place update and block-erase properties, flash translation layer (FTL) is employed between file systems and NAND flash. Due to the unique characteristics of NAND flash, FTL needs to be optimized by utilizing them to improve the performance and reliability of NAND flash-based storage devices. For planar NAND flash, the characteristics have been comprehensively and deeply investigated, and high-efficiency NAND flash management methods have been proposed. Nevertheless, only a few prior studies have been conducted on the characteristics of 3D NAND flash. Xiong et al. have studied the characteristics of 3D FG NAND flash and given some implications on system designs [6], [7]. However, due to the differences in structures and materials between FG and CT NAND flash, those implications of 3D FG NAND flash cannot be directly applied to 3D CT NAND flash. Hence, it is important and urgent to have a profound understanding of the characteristics of 3D CT NAND flash.

In this paper, we conduct comprehensive experiments to characterize the performance and reliability features of 3D CT NAND flash and make meticulous analyses. We also summarize important observations and provide insights into how they can be utilized to optimize the flash management algorithms for 3D CT NAND flash.

The rest of this paper is organized as follows. Section 2 introduces the background of the work. Section 3 shows experimental setups. Evaluation results, analyses and applications are presented in Section 4. Section 5 demonstrates related works and Section 6 concludes this paper.

II. BACKGROUND AND PRELIMINARIES

A. 3D Charge Trap NAND Flash

NAND Flash Cell. A NAND flash cell stores 1-bit or several bits of data by dividing the threshold voltage (V_{th}) into multiple regions. The region in which the V_{th} of a cell locates represents the current value of data. As illustrated by Fig. 1, seven *read reference voltages* (V_{ref}), V_{ref1} to V_{ref7} , divide a TLC cell into eight states, E and S_1 to S_7 , which can be decoded into 111, 110, 100, 000, 010, 011, 001 and 101, respectively. In this paper, we use the format *ABC* to represent the 3-bit data stored in a TLC cell, where A , B and C denote the least significant bit (LSB), the center significant bit (CSB) and the most significant bit (MSB), respectively.

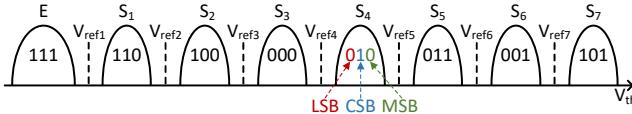


Fig. 1: Threshold voltage distribution of TLC NAND flash (in the format of *ABC*, A , B and C denote LSB, CSB and MSB, respectively).

3D CT NAND Flash Organization. Compared with planar NAND flash, which arranges NAND flash cells in 2D arrays (as demonstrated in [8]), 3D NAND flash is formed by those in a 3D structure, as shown in Fig. 2. The medium responsible for data storage is made up of an array of word lines (WLs), which are continuously connected from top to bottom along the channel side, and are connected to the source line and bitlines through source line selectors (SLSs) and bitline selectors (BLSs), respectively. A WL is the basic unit of program operations and each WL is composed of 3 pages: lower page, middle page and upper page, which correspond to LSB, CSB and MSB, respectively. The 4 WLs in the same X-Y plane form a *tier* and their control gates are connected together (to clearly illustrate WLs, the WLs in the same tier are separated in Fig. 2), and the WLs in the same X-Z plane form a *string*.

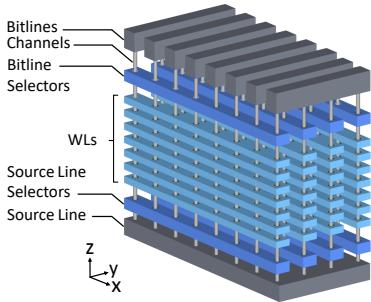


Fig. 2: Bird's eye view of 3D CT NAND flash array (In the evaluated flash chip, an array actually contains 48 tiers. For convenient illustration, only 7 tiers are plotted in Fig. 2).

As with planar NAND flash, there are three basic operations for 3D CT NAND flash: *read*, *program* and *erase*.

Read. A read operation obtains data stored in the target page. When a target page is read, SLS and the corresponding BLS are turned on, and all the tiers exclusive of the target page are applied a *read pass voltage* (V_{pass}) so that data in the target

page can be properly propagated to the output. Meanwhile, a sequence of V_{ref} s are applied to the tier containing the target page in order to read out the data. According to the type of the target page, read reference voltages are divided into 3 sets: lower page $\rightarrow \{V_{ref3}, V_{ref7}\}$, middle page $\rightarrow \{V_{ref2}, V_{ref4}, V_{ref6}\}$, and upper page $\rightarrow \{V_{ref1}, V_{ref5}\}$. During a read operation, the read reference voltages in the corresponding set are sequentially applied.

Program. Program operations store specific data into the target WL. In 3D CT NAND flash, program operations are performed in a *one-shot* behavior, which means that *the three pages in the selected WL are programmed simultaneously*. Similar to the read operations, during a program operation, SLS and the BLS corresponding to the target WL are turned on, and all the tiers except for the tier containing the WL are applied a V_{pass} , so that the target WL can be selected. Meanwhile, a series of incremental staircase program pulses (ISPP) and program verify (PV) processes are performed so that electrons can be charged into the storage layers of the selected WL, completing the program operation.

Erase. An erase operation wipes all the data in the selected block. Similar to program operations, the erase operation involves a sequence of erase pulses and hard erase verify (HEV) processes. The erase pulse is achieved by applying a high voltage, V_{erase} , between channels and control gates, which move the thresholds of all cells in a block towards the E state. After each erase pulse, an HEV process follows. The HEV process checks if there are still some cells that have threshold voltages higher than V_{ref1} , in which case V_{erase} is incremented by ΔV_E and another erase pulse needs to be applied. Usually, a maximum number (E_{max}) of erase pulses is set so that erase failure occurs if E_{max} is reached.

B. Metrics

For data storage, performance and reliability are always the two most important technical indicators. The main metrics of NAND flash characteristics are briefly introduced in this section.

1) **Performance:** Nowadays, with the rapid development of NAND flash interface protocols, the latest interfaces could provide up to 800MT/s and 400MT/s when they run on Open NAND Flash Interface (ONFI) 4.0 [9] and Toggle 2.0 [10], respectively. Unfortunately, the performance of NAND flash chips is greatly limited by internal latencies. **Erase latency**, **program latency** and **read latency** refer to the time consumed for wiping out all data stored in a block, for programming data stored in data register to the NAND array, and for copying data from NAND array to cache register and enabling data output from the cache register to host, respectively.

2) **Reliability:** During the lifetime, NAND flash can fail due to a wide range of reasons, which jointly threaten the data safety. After repeated erase and program operations, cells become unreliable due to trap creation in tunnel oxide and interfacial damages until erase or program failures arise. **Endurance** denotes the number of program/erase (P/E) cycles that NAND flash can withstand before a failure appears. Besides, data can be corrupted by four other main factors, **retention**, **fast detrapping**, **read disturb** and **program disturb**, all of which do not damage cells permanently. Retention and fast detrapping errors result from charge leakage over time. Read disturb is a phenomenon that a read operation causes a weak programming effect on the other (unread) WLs in the same block. When programming a

WL, electrons can be *unintentionally* injected into the cells of other WLs in the same block, caused by parasitic capacitance or a *weak programming effect*, called program disturb.

III. EXPERIMENTAL SETUP

A. Experimental Platform

In order to accurately measure the performance and reliability of 3D CT NAND flash, we build an ARM- and FPGA-based NAND flash experimental platform, *General Storage Tester (GST)*, that enables us to directly control raw NAND flash chips without ECC, as shown in Fig. 3. The *ARM* runs a stripped-down Linux operating system, which is responsible for 1) receiving commands/data from a personal computer (PC) via 1Gbps Ethernet ports or RS-232 ports; 2) parsing user commands into atomic commands supported by NAND flash chips; 3) transferring atomic commands/data to the FPGA; and 4) receiving returned data from the FPGA and sending them back to the PC. A custom flash controller is implemented in the *FPGA*, which transforms the commands/data from the *ARM* into the corresponding signals to control NAND flash chips and reads data from NAND flash chips to the *ARM*. By monitoring the ready/busy (R/B) signal, the *FPGA* supports measuring the latency of each atomic command executed by NAND flash chips with the precision of $1\mu s$. An experimental platform can connect to up to 4 daughter boards, each of which supports at most 8 NAND flash chips.

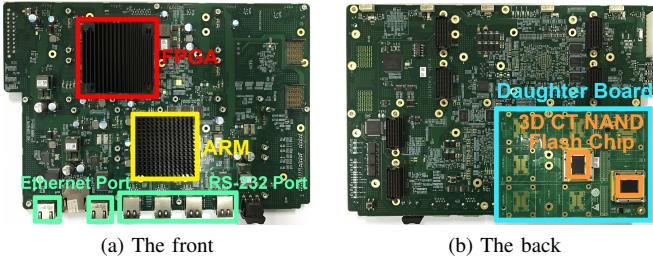


Fig. 3: 3D NAND flash experimental platform.

B. Experimental Acceleration

We measure NAND flash chips over various P/E cycles and different retention ages at room temperature ($25^\circ C$) to imitate the real environment. In order to characterize reliability after long retention ages (e.g., 1 year), we accelerate retention error tests under high temperature according to the Arrhenius Law [11]. Table I shows the accelerated retention ages (t_{acce}) at various high temperatures (T_{acce}) to achieve the corresponding equivalent retention ages (t_{room}) at room temperature (T_{room}) used in our experiments. For example, by putting the tested 3D CT NAND flash for 12.9 hours at $85^\circ C$, it suffers a 1-year equivalent retention age at $25^\circ C$. *In this paper, experiments except for retention errors are all completed at room temperature.*

TABLE I: Accelerated and equivalent retention ages.

t_{room} (T_{room})	t_{acce} (T_{acce})
1 month ($25^\circ C$)	2.7h ($75^\circ C$)
6 months ($25^\circ C$)	16.2h ($75^\circ C$)
1 year ($25^\circ C$)	12.9h ($85^\circ C$)
3 years ($25^\circ C$)	24.8h ($90^\circ C$)
5 year ($25^\circ C$)	26.8h ($95^\circ C$)

C. Experimental Object

In our experiments, we use a representative 3D CT NAND flash product, *BiCS2¹ TLC from Toshiba*. It uses Silicon-Oxide-Nitride-Oxide-Silicon (SONOS) structure and its charge traps are similar to 3D CT NAND flash from SK Hynix and Samsung. The main parameters of our measured chips are listed in Table II.

TABLE II: Parameters of 3D CT NAND flash chips.

Parameter	Value
Capacity	2.48Tb (310GB)
Page size	(16384 + 1952) bytes
Block size	576 pages
Plane size	1972 blocks
Die size	2 planes
Target size	2 dies
Chip size	4 targets

D. Experimental Methodology

Random Data. In actual products, data randomization mechanism is widely employed in flash controller or integrated into flash chips to reduce the raw bit error rates (RBERs) and prolong the lifetime. Therefore, we implement a pseudo-random binary sequence generator to produce random data and use a scrambler designed for 3D NAND flash to further randomize the data. The randomized and scrambled data² are adopted in each program operation to mimic real data.

P/E Cycling. Blocks are repeatedly erased and programmed with random data to simulate the real process of usage. For all experiments, a dwell time of 10s, which is the idle time between an erase operation and a program operation to the same block, is employed at room temperature.

IV. EVALUATION RESULTS

A. Performance

In order to characterize the performance of 3D CT NAND flash, we employ random data to repeatedly program and erase fresh blocks and record the latencies of each erase, program and read operation for every 100 P/E cycles until they fail. Fig. 4 shows the latencies of erase (a), program (b) and read (c) operations within the lifetime.

Observation 1. *The performances of erase and program operations vary predictably with the increase of P/E cycle.*

As shown in Fig. 4a and Fig. 4b, with the continuous wear of NAND flash, the erase latency exhibits a ladder-shaped growth, forming several steps and fluctuates near the joint of each two steps, while the program latency declines gradually. In 3D CT NAND flash, program and erase operations are achieved by Fowler-Nordheim (FN) tunneling effect [12], which charges and discharges electrons into and from the storage layer, respectively. However, as electrons are repeatedly tunneled through the tunnel layer, defects, which unintentionally trap electrons and cause the Trap-Assisted Tunneling (TAT) effect, are formed. The TAT effect reduces the insulativity of the tunnel layer [13], making

¹Since the later generations, BiCS3 and BiCS4, only have *engineering samples* available and in order to provide guidance for the design of real flash-based storage systems through our characterization, we use the mature *customer samples* of BiCS2 as our experimental object.

²No data randomization mechanism is integrated into Toshiba BiCS2 chips. The generated data are directly programmed to flash cells.

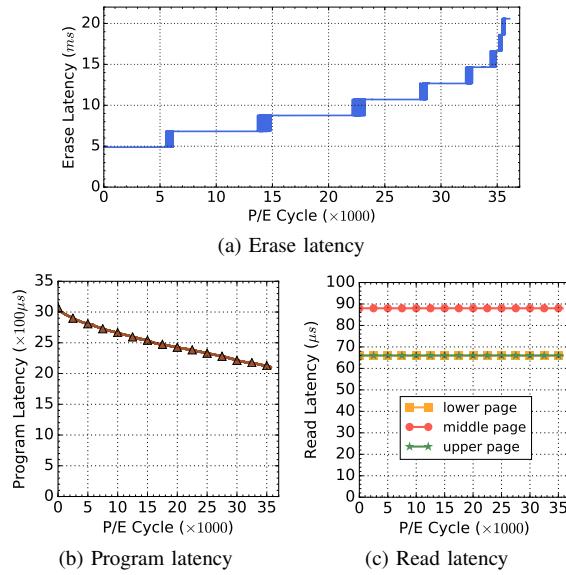


Fig. 4: NAND flash operation performances over P/E cycles. Since in 3D CT NAND flash, program operations are performed in *one-shot* behavior, the program latencies of lower, middle and upper pages are not separated.

electrons easier to be tunneled through. For erase operations, although the TAT effect accelerates the detrapping process, electrons trapped in the defects form an opposite electric field against the erase pulse and make electrons in the storage layer harder to be tunneled out. Therefore, with more defects accumulating in the tunnel layer, a larger opposite electric field is formed, thus more time is needed to erase the cells to a voltage beneath V_{ref1} . Since erase operations are composed of continuous erase-verify processes, the harder electrons can be tunneled out, the more erase-verify processes are performed, thus the erase latency is increased in a ladder shape. For program operations, the TAT effect accelerates the programming process. In addition, as the threshold voltage of a cell is shifted to a higher value, fewer electrons are needed to be programmed into the storage layer, hence reducing the program time. Since the step length of the pulse of program operations are much shorter than erase operations, the decreasing of the program latency exhibits a more smooth curve.

Application 1. The sudden failure of storage medium may cause irretrievable data loss, which has always been a problem in the maintenance of storage systems. Nevertheless, according to our evaluation results, this problem can be mitigated by utilizing the characteristics of 3D CT NAND flash. As shown in Fig. 4, the erase and program latencies show negative and positive correlations with the remaining lifetime, respectively, indicating the health conditions of NAND flash. Therefore, by choosing proper machine learning algorithms and using such performance information as indicators, a *lifetime prediction model* can be trained to predict the remaining lifetime of NAND flash. The lifetime prediction model can be implemented in FTL at different granularities (block-level, chip-level, etc.) according to users' demands, so that data in the storage medium can be migrated before failures occur.

Observation 2. *Read performance remains constant during the entire lifetime of NAND flash.*

entire lifetime of NAND flash.

As shown in Fig. 4c, the read latencies of all types of pages remain unchanged during the entire lifetime, with reading lower and upper pages taking 66 μs , and middle pages costing 88 μs . Unlike erase or program operations which involve moving electrons through the tunnel layer, the read operation is not affected by the wear degree of NAND flash, thus read latency stays constant regardless of the P/E cycle. The variance of read latencies can be explained by Fig. 1. When reading a lower/upper page, only two read reference voltages (lower page $\rightarrow \{V_{ref3}, V_{ref7}\}$, upper page $\rightarrow \{V_{ref1}, V_{ref5}\}$) need to be applied in sequence, while reading a middle page involves applying three read reference voltages ($\{V_{ref2}, V_{ref4}, V_{ref6}\}$), resulting in a longer read time.

Application 2. In 3D CT NAND flash, the performance variance of reading different types of pages can be exploited to improve the read performance of flash-based storage systems. According to Fig. 4c, the read latencies of lower and upper pages are 25% smaller than that of middle pages. Therefore, we can program read-hot data to lower/upper pages and read-cold data to middle pages, so that the overall read performance of the storage systems can be optimized.

B. Reliability

1) Endurance: The endurance of a NAND flash block represents its lifetime. In order to quantitatively characterize the endurance of 3D CT NAND flash, we randomly choose 100 blocks from the flash chips, then repeatedly program and erase the blocks until they fail, and record the RBERs of each page and block for every 100 P/E cycles.

Observation 3. *3D CT TLC NAND flash achieves similar endurance of planar MLC NAND flash.*

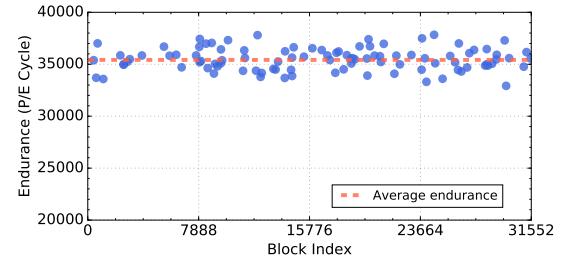


Fig. 5: Endurance distributions among 100 blocks.

In our experiment, all blocks end up with erase failures. As shown in Fig. 5, the endurance of 3D CT TLC NAND flash blocks exhibits an average of 35,417 P/E cycles, showing a significant longer lifetime than planar TLC NAND flash (typical value: 1,000 P/E cycles [14]), and is comparable with planar MLC NAND flash (typical value: 20,000 P/E cycles [14]). The outstanding endurance benefits from the larger feature size, the new materials and structure. Among these selected blocks, the maximum endurance reaches 37,883 P/E cycles, and the minimum is 32,921 P/E cycles, resulting in a standard deviation of 1,064 P/E cycles.

Observation 4. *RBERs show a near exponential growing trend with the increase of P/E cycles, and exhibit large variations among and inside blocks.*

Fig. 6a exhibits the growing trend of RBERs with P/E cycles for each block from the beginning of the test until they fail, with

different colors representing different blocks. We can observe from Fig. 6a that block RBERs (the average RBERs of a block) grow near-exponentially with P/E cycles, and exhibit various growing speeds among different blocks. We further analyze the RBERs of each page within a block, and the results are shown in Fig. 6b. We can observe that page RBERs exhibit a dispersive distribution, indicating enormous process variations among pages. Moreover, the average RBERs of upper and middle pages grow faster than that of the lower pages. According to Fig. 1, the probability that threshold voltages of CSBs shift to adjacent states are higher than that of LSBs and MSBs, resulting in more bit errors. Moreover, since P/E operations shift threshold voltages to higher values, states with lower threshold voltages tend to shift more. Therefore, the MSBs, which have states with lower threshold voltages compared with LSBs, are more likely to suffer bit errors, resulting in various growing speeds of RBERs among different types of pages.

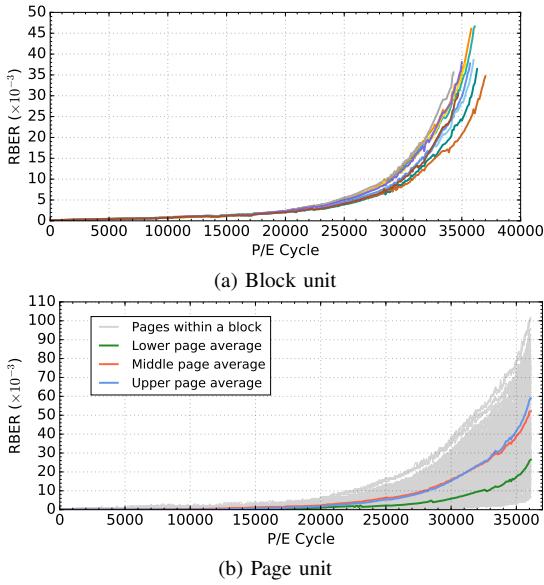


Fig. 6: RBERs variations with P/E cycles.

Application 3. In traditional FTL strategies, a block is identified as a bad block and then discarded if one of its pages fails. However, according to our evaluation results, flash blocks are fault-tolerant, which means that when a page fails, the other pages in the block might still work in healthy states. Therefore, even if a page fails in a block, we can continue to use the remaining healthy pages in the block to improve the storage utilization and NAND flash lifetime.

Application 4. The reliability variance between pages in a block can also be leveraged to reduce the ECC overhead, such as decoding latency and redundant storage capacity. That is, we can employ a strong ECC to ensure the reliability of weak pages (pages with lower reliabilities), and a weak ECC for strong pages (pages with higher reliabilities) to reduce unnecessary ECC decoding expenses.

2) **Retention error:** Retention errors usually occur in systems storing cold data. In order to characterize retention errors, we randomly choose 100 fresh blocks and divide them into 4 groups, then employ random data to repeatedly program and erase each group to a specific P/E cycle. After that, we place the

chips under room temperature for 1 day and 1 week, then under high temperature to imitate retention for 1 month, 6 months, 1 year, 3 years and 5 years, as shown in Table I. At the end of each retention period, we read the selected blocks under room temperature to get the average RBERs in each group.

Observation 5. RBERs show a near-logarithmic growth with the increase of retention time.

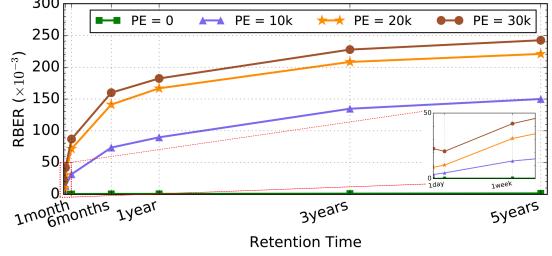


Fig. 7: RBERs variations with retention time.

As shown in Fig. 7, RBERs grow sharply within the first 6 months of retention time, and increase mildly afterward. When NAND flash chips are placed for a long time without operations, electrons gradually escape from the storage layer since they form an electric field, whose intensity is proportional to the number of electrons inside the storage layer. Therefore, as electrons gradually escape with the increase of retention time, the electric field intensity is reduced, slowing down the RBERs growing speed. We can also observe from the figure that a larger P/E cycle corresponds to higher RBERs, which is because the tunnel oxide wears more seriously with a larger P/E cycle, thus electrons are easier to escape. Moreover, an interesting phenomenon can be observed from the figure that the RBERs experience a slight drop after 1 day's retention under the P/E cycle of 30,000. Actually, this slight drop happens to all P/E cycles at different retention stages (not shown in the figure due to the discrete sampling in retention time). Since the P/E operations and the retention process shift threshold voltages to different directions, at some point, the retention process partly "repairs" the errors caused by P/E operations.

Application 5. Traditional FTL algorithms employ a *refresh* policy, which periodically migrates cold data, to avoid high retention errors. However, a fixed refresh frequency may not suit all blocks in the system, since according to our evaluation results, retention rates differ for blocks at different wear levels. Therefore, given the maximum tolerable error rate of the system, we can calculate and apply different refresh intervals for blocks at different P/E cycles so that unnecessary migration overheads can be alleviated.

3) **Fast detrapping:** Fast detrapping, which results in a sharp rise of RBERs, happens immediately after every program operations. In order to characterize the fast detrapping phenomenon, we choose blocks that are programmed and erased to different P/E cycles, and for each block, perform the following operations: 1) program a WL (instead of a whole block in order to observe the RBERs variations immediately after programming) with random data and record the programmed data; 2) read the three pages in the programmed WL every 100 milliseconds for 300 seconds. Finally, we compare the programmed data and the data read out to obtain the variations of page RBERs within the first 300 seconds after programming, and a representative fast detrapping phenomenon is shown in Fig. 8.

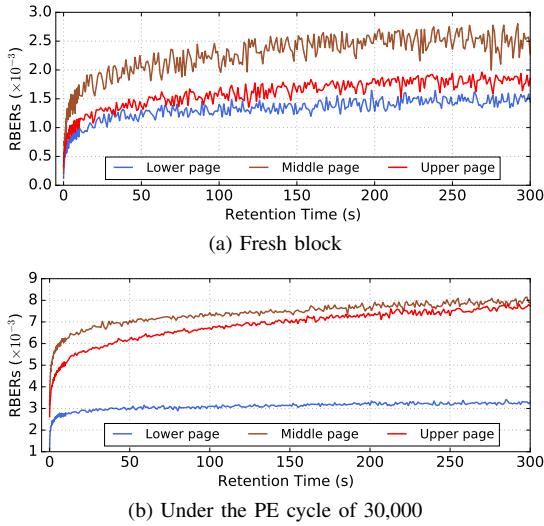


Fig. 8: Fast detrapping phenomenon.

Observation 6. *RBERs rise sharply immediately (within ten seconds) and then remain nearly constant after programming.*

Fig. 8 exhibits the growing trends of page RBERs shortly after a WL has been programmed under the P/E cycle of 0 and 30,000. We can observe from the figure that RBERs of the three pages increase rapidly within the first 10 seconds after programming, with lower, middle and upper page reaching $8.2\times$, $7.2\times$ and $3.4\times$ of the initial RBERs in fresh blocks, respectively. Under the wear of 30,000 P/E cycles, the three values are 2.6 , 2 and 1.9 , respectively. During program operations, a portion of electrons are trapped in shallow traps, thus are less stable compared with electrons trapped in deep traps. Therefore, they can easily escape from the storage layer through defects in oxide immediately after programming, shifting threshold distributions towards lower values, thus resulting in the sharp rise of RBERs in the initial stage. After the majority of the unstable electrons escape from the storage layer, charge loss becomes slow and steady. In addition, with the increase in P/E cycles, threshold distributions shift to higher values after program operations, which is in the opposite shift direction of fast detrapping and weakens it. Thus, the rising rates under 30,000 P/E cycles are lower than those in fresh blocks.

Application 6. Fast detrapping is a unique phenomenon in 3D CT NAND flash, which leads to multiple times of increase in RBERs due to its charge trap structure, hence this characteristic should be considered when designing flash management techniques for systems using 3D CT NAND flash. Since fast detrapping is caused by electrons trapped in shallow traps during program operations, we can use a *re-program* approach that implements an additional program operation to enhance electrons into deep traps, so that the sharp rise of RBERs caused by fast detrapping can be effectively reduced. By carefully designing the FTL algorithm and hiding the additional program operation in the background (implement the second program when flash chips are in idle states), the overheads caused by the re-program operation can be alleviated or even eliminated.

4) Read disturb error: Read disturb errors occur when pages in a block are repeatedly read without any erase operations. In order to characterize the read disturb error in 3D CT NAND

flash, we choose blocks at different P/E cycles, equally divide them into two groups (group A and group B), and for each group we perform the following operations: 1) erase and program an entire block with random data; 2) read a specific page (for group A we choose Page 3, Page 4 and Page 5, for group B we choose Page 279, Page 280 and Page 281 to research on the influence of location and types of pages on read disturb errors) for 2,000,000 times and record data of the entire block every 100,000 times to get the distributions of RBERs.

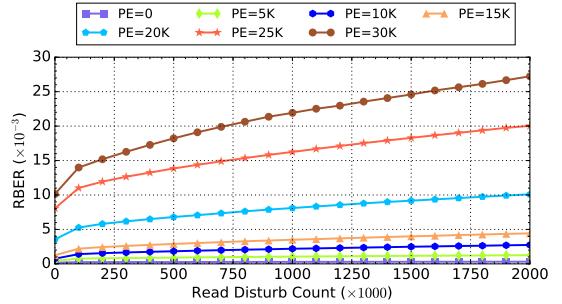


Fig. 9: Variation of read disturb induced RBERs with different read disturb counts and P/E cycles.

Observation 7. *Read disturb induced RBERs grow near-linearly with increased read disturb counts, with the earlier stage growing much faster.*

According to our evaluation results, reading different types of pages produce similar levels of read disturb errors, hence in Fig. 9 we only present RBERs variations when reading a middle page as a representative case. we can observe from the figure that with the increase of read disturb counts, the RBERs rise sharply within the first 10,000 read disturb counts, and grows relatively slowly subsequently. Moreover, blocks under higher P/E cycles are more vulnerable to read disturb and exhibits higher increase rates. When reading a page, a pass through voltage, V_{pass} , is applied to tiers not being read in the block. The continuous exertions of V_{pass} generate *weak programming effects* that gradually shift the threshold voltages to higher values, resulting in the rise of RBERs. In addition, blocks under higher P/E cycles have more defects in the tunneling oxide, thus electrons are much easier to be injected into the storage layer through read disturb.

Observation 8. *Read disturb exerts non-uniform influences on different tiers inside a block, with neighboring tiers suffering more seriously.*

Fig. 10a and Fig. 10b exhibit cases in which a page at the edge (Page 4) and a page in the middle (Page 280) of the block are selected as the target page to be read 2 million times, respectively. Since all types of pages generate similar results, Fig. 10 gives the general situations. The cross-tier variations of RBERs before read disturb at 30,000 P/E cycles are due to the different voltages applied to different tiers during the continuous P/E operations. We observe from the figure that when a target page is repeatedly read, RBERs of the tier containing the target page experience only a slight increase, while RBERs of all the other tiers in the block rise dramatically. Moreover, neighboring tiers of the target page suffer relatively more read disturb than the other tiers, since they are applied a voltage (V_{passH}) higher than V_{pass} to reduce coupling effects.

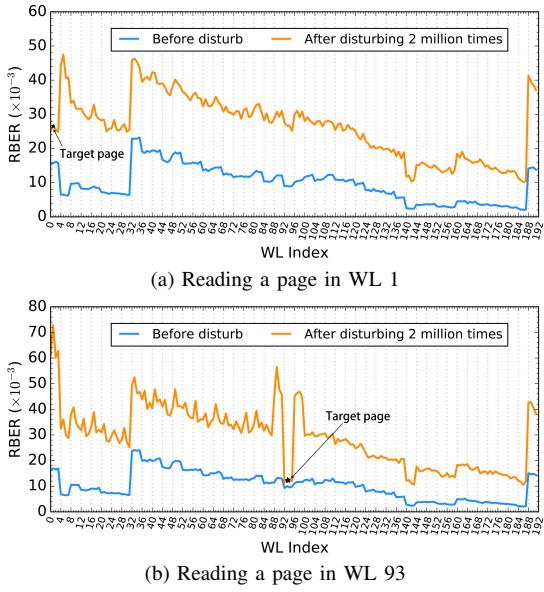


Fig. 10: Distribution of read disturb induced RBERs under the P/E cycle of 30,000.

Application 7. Based on the evaluation results, we suggest two approaches to alleviate the read disturb problem. One approach is to maintain several copies of read-hot data in different flash chips. Thus, the read disturb can be amortized and meanwhile the read performance can also be improved (we can avoid a read request being blocked by directing it to an idle flash chip that contains one copy of the requested data). Another approach is to refresh a block when its pages have been massively read, in this way, errors caused by read disturb can be corrected.

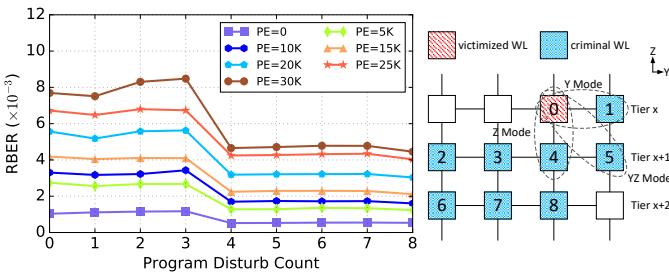


Fig. 11: Variations of RBERs caused by program disturb with various P/E cycles and disturb counts.

5) **Program disturb error:** As shown in Fig. 11b, in 3D NAND flash, there are three spatial relationships between a criminal WL (the WL being programmed) and a victimized WL (the WL suffering program disturb): 1) Y mode (the two WLs are in the same tier); 2) Z mode (the two WLs are in the same strings); and 3) YZ mode (the two WLs are in neither the same tier nor the same strings). In order to explore the characteristics of program disturb in 3D CT NAND flash, we 1) choose blocks at different P/E cycles; 2) randomly program a WL (not the last WL in a tier) as the victimized WL in each block and place those blocks for 30 minutes to eliminate the influence caused by fast detrapping; and 3) for each block, read the victimized WL,

program the next 8 WLs of the victimized WL³ (e.g., victimized WL index: n , criminal WL indexes: $n + 1, n + 2, \dots, n + 8$), and read the victimized WL after each program operation. Since each tier contains 4 WLs, Y-, Z- and YZ-modes program disturb can be covered by the above scheme.

Observation 9. *Z-mode program disturb has the most significant impact on the reliability.*

As shown in Fig. 11, the RBERs vary gradually when the program disturb is in Y mode or YZ mode (the program disturb counts are 1, 2, 3, 5, 6 and 7). But for Z-mode program disturb, the RBERs are strongly and weakly affected by the 4th and the 8th program disturb, respectively. When performing a program operation on a criminal WL, a victimized WL can be disturbed by weak programming effect and coupling effect. The coupling effect mainly exists between WLs in the same strings, sharply weakens with the increase of distance, and is much stronger than the programming effect. Thus, programming the WL $n + 4$ affects the reliability most, and the other program operations have weaker effects.

Just like read disturb, program disturb also unintentionally injects electrons into the storage layer. However, as shown in Fig. 11, the RBERs decrease after being program disturbed. In order to investigate this phenomenon, we analyze the state shifts before and after program disturb, as shown in Fig. 12. Since more than 99.9% of the shifts occur among adjacent states, Fig. 12 only shows the adjacent state shifts. If a cell shifts from a lower state to a higher one, it is a positive state shift (e.g., $S_6 \rightarrow S_7$), otherwise, a negative state shift (e.g., $S_7 \rightarrow S_6$).

Observation 10. *Program disturb can eliminate fast detrapping errors to some extent.*

As mentioned in Section IV-B3, electrons sitting in shallow traps can easily escape from the storage layer, shifting threshold voltages to lower values, causing even negative state shifts. As a result, the injected electrons caused by program disturb may make the cells, which have shifted to lower states, back to the original states. As shown in Fig. 12, the 4th program disturb dramatically reduces the negative state shift counts, and slightly increases the positive state shift counts, resulting in the lower RBERs.

Application 8. WLs in 3D CT NAND flash are programmed in a fixed order and the effect of program disturb from a criminal WL to a victimized WL can be quantified. Moreover, program disturb and fast detrapping shift threshold voltages to opposite directions. Therefore, both program disturb and fast detrapping errors can be reduced by carefully controlling the program pulses so that a majority of the threshold voltage shifts can be counteracted.

V. RELATED WORK

During the past few years, plenty of research works have experimentally characterized planar NAND flash [2], [8], [15]–[21] and flash-based devices [22]. Cai et al. conducted the major series of investigations on various reliability issues, including error patterns [18], threshold voltage distribution [19], and three main sources of errors (retention [20], read disturb [21] and program disturb [8]). These studies increase designers’ knowledge about the inherent features of NAND flash, and motivate them to design high-efficiency NAND flash management algorithms based on the features.

³The WL-order programming scheme is suggested by the manufacturers and employed in practice.

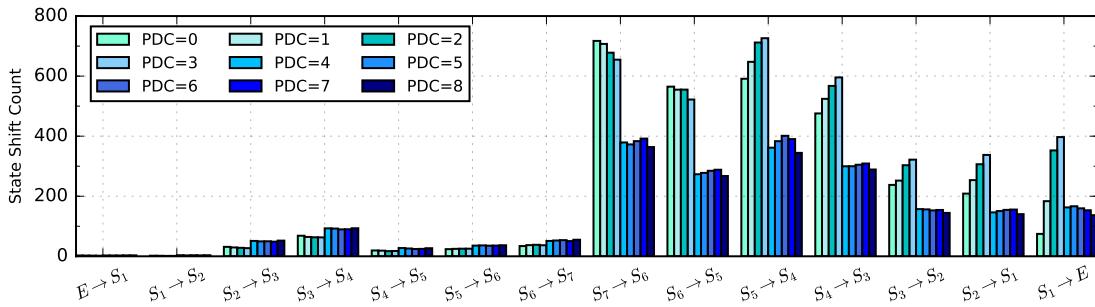


Fig. 12: State shifts with different program disturb counts (PDC) under the P/E cycles of 30,000.

Due to the scaling limitation of planar NAND flash, the 3D structure has been proposed and is attracting increasing attention. Most research works about 3D NAND flash are at simulation levels or integrated circuit levels, which usually use self-designed structures and custom-fabricated chips [4], [23]–[25]. Xiong et al. comprehensively evaluated the commercial 3D FG NAND flash product for the first time [6], [7]. Based on their observations, Zhu et al. built a read disturb error model and proposed a location-aware redistribution method [26] to improve 3D FG NAND flash reliability. However, in the 3D era, the CT technology has replaced the FG technology as the mainstream, and the characteristics of commercial 3D CT NAND flash products have not been comprehensively investigated. The lack of research in this area makes NAND flash management algorithms inefficient since those algorithms cannot utilize the characteristics of 3D CT NAND flash. To our knowledge, this is the first study that comprehensively characterizes advanced commercial 3D CT NAND flash products.

VI. CONCLUSION

In this paper, we implement comprehensive characterizations of advanced 3D CT NAND flash from performance and reliability aspects. According to the experimental data measured on the real platform, we make multiple distinct observations and give detailed analyses. 3D CT NAND flash exhibits outstanding endurance, diverse degradation speed among pages and blocks, fast detrapping, cross-tier variations of errors, and slight program disturb, which differ from those of planar and 3D FG NAND flash. Based on the observations, we discuss some possible approaches that utilize these characteristics to optimize flash management techniques of 3D CT NAND flash in real applications. We believe this work can give researchers and designers deeper understandings of the characteristics of 3D CT NAND flash and help improve the efficiency of NAND flash management algorithms.

In the future, we will further characterize 3D CT NAND flash from other manufacturers, research on the similar and different features among various types (planar, 3D FG and 3D CT) of NAND flash, and deeply explore the applications of the characteristics of 3D CT NAND flash in flash management techniques.

REFERENCES

- [1] K. Prall, “Scaling non-volatile memory below 30nm,” in *Proc. of NVSMW*, 2007, pp. 5–10.
- [2] R. Micheloni, L. Crippa, and A. Marelli, *Inside NAND Flash Memories*. Springer Science & Business Media, 2010.
- [3] C.-H. Lee, J. Choi, and C. K. et al., “Multi-level NAND flash memory with 63 nm-node TANOS (Si-Oxide-SiN-Al2O3-TaN) cell structure,” in *Tech. Dig. of VLSI Technol.*, 2006, pp. 21–22.
- [4] R. Micheloni, *3D Flash Memories*. Springer, 2016.
- [5] J. Yoon, R. Godse, and G. T. et al., “3D-NAND scaling & 3D-SCM implications to enterprise storage,” *Flash Memory Summit*, Aug. 2017.
- [6] Q. Xiong, F. Wu, and Z. L. et al., “Characterizing 3D floating gate NAND flash,” in *Proc. of SigMetrics*, 2017, pp. 32–33.
- [7] ———, “Characterizing 3D floating gate NAND flash: Observations, analyses, and implications,” *ACM Tos*, May 2018.
- [8] Y. Cai, O. Mutlu, and E. F. H. et al., “Program interference in MLC NAND flash memory: Characterization, modeling, and mitigation,” in *Proc. of ICCD*, 2013, pp. 123–130.
- [9] Intel, Micron et al., “Open nand flash interface specification,” *Technical Report ONFI*, 2014.
- [10] T. Grunze, “400 MT/s NAND interface solutions,” *Flash Memory Summit*, Aug. 2011.
- [11] D. A. Baglee, “Characteristics & reliability of 100A oxides,” in *Proc. of IRPS*, 1984, pp. 152–155.
- [12] R. H. Fowler and L. Nordheim, “Electron emission in intense electric fields,” vol. 119, no. 781, pp. 173–181, 1928.
- [13] R. Degraeve, F. Schuler, and B. K. et al., “Analytical percolation model for predicting anomalous charge loss in flash memories,” *IEEE TED*, vol. 51, no. 9, pp. 1392–1400, 2004.
- [14] Samsung Semiconductors, “3D TLC NAND to beat MLC as top flash storage,” *EETimes*, 2015.
- [15] L. M. Grupp, A. M. Caulfield, and J. C. et al., “Characterizing flash memory: Anomalies, observations, and applications,” in *Proc. of MICRO*, 2009, pp. 24–33.
- [16] P. Desnoyers, “Empirical evaluation of NAND flash memory performance,” *ACM SIGOPS Operating Syst. Rev.*, vol. 44, no. 1, pp. 50–54, 2010.
- [17] S. Boboila and P. Desnoyers, “Write endurance in flash drives: Measurements and analysis,” in *Proc. of FAST*, 2010, pp. 115–128.
- [18] Y. Cai, E. F. Haratsch, and O. M. et al., “Error patterns in MLC NAND flash memory: Measurement, characterization, and analysis,” in *Proc. of DATE*, 2012, pp. 521–526.
- [19] ———, “Threshold voltage distribution in MLC NAND flash memory: Characterization, analysis, and modeling,” in *Proc. of DATE*, 2013, pp. 1285–1290.
- [20] Y. Cai, Y. Luo, and E. F. H. et al., “Data retention in MLC NAND flash memory: Characterization, optimization, and recovery,” in *Proceedings of HPCA*, 2015, pp. 551–563.
- [21] Y. Cai, Y. Luo, and S. G. et al., “Read disturb errors in MLC NAND flash memory: Characterization, mitigation, and recovery,” in *Proc. of DSN*, 2015, pp. 438–449.
- [22] A. Tavakkol, M. Sadrosadati, and S. G. et al., “MQSim: A framework for enabling realistic studies of modern multi-queue SSD devices,” in *Proc. of FAST*, 2018, pp. 49–65.
- [23] S. Aritome, Y. Noh, and H. Y. et al., “Advanced DC-SF cell technology for 3D NAND flash,” *IEEE TED*, vol. 60, no. 4, pp. 1327–1333, 2013.
- [24] J. Wu, D. Han, and W. Y. et al., “Comprehensive investigations on charge diffusion physics in SiN-based 3D NAND flash memory through systematical Ab initio calculations,” in *Proc. of IEDM*, 2017, pp. 1–4.
- [25] N. Righetti and G. Puzzilli, “2D vs 3D NAND technology: Reliability benchmark,” in *Proc. of IIRW*, 2017, pp. 1–6.
- [26] Y. Zhu, F. Wu, and Q. X. et al., “ALARM: A location-aware redistribution method to improve 3D FG NAND flash reliability,” in *Proc. of NAS*, 2017, pp. 1–10.