Fortifying the NAND Flash Supply Chain with Innovative Security Primitives

Matchima Buddhanoy
Colorado State University
Fort Collins, CO, USA
matchima.buddhanoy@colostate.edu

Biswajit Ray
Colorado State University
Fort Collins, CO, USA
biswajit.ray@colostate.edu

ABSTRACT

The infiltration of counterfeit electronics into the global supply chain is a growing concern that poses significant challenges for both manufacturers and consumers. This paper introduces Flash-Odometer, an innovative technique designed to estimate the age and usage statistics of NAND flash memory blocks. Flash-Odometer works by analyzing key characteristics of NAND arrays, which are highly dependent on the defect density in the gate dielectric of the flash memory cells. Our experimental evaluation, conducted on state-of-the-art 3D NAND chips from a leading manufacturer, demonstrates that the Flash-Odometer accurately predicts the program-erase (PE) cycle count of NAND memory blocks. This provides a unique and reliable method for estimating chip usage, which is instrumental in detecting recycled memory chips.

CCS CONCEPTS

• **Hardware → External storage**.

KEYWORDS

3D NAND Flash Memory, Counterfeit, Endurance, Odometer

ACM Reference Format:
Matchima Buddhanoy and Biswajit Ray. 2024. Fortifying the NAND Flash Supply Chain with Innovative Security Primitives. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD '24), October 27–31, 2024, New York, NY, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3676536.3689917

1 INTRODUCTION

NAND flash memory technology is widely adopted in industry as a leading non-volatile data storage solution due to its superior bit density. However, this widespread use has made flash memory chips a prime target for counterfeiters [2, 15, 26], who infiltrate various stages of the NAND supply chain. Figure 1 illustrates the major players and the complexity of the supply chain for NAND flash memory chips. There are a few flash memory chip manufacturers, such as Micron, Samsung, Toshiba, and SK-Hynix. However, the number of companies that make solid-state drives (SSDs) is orders of magnitude larger, and the number of companies that integrate

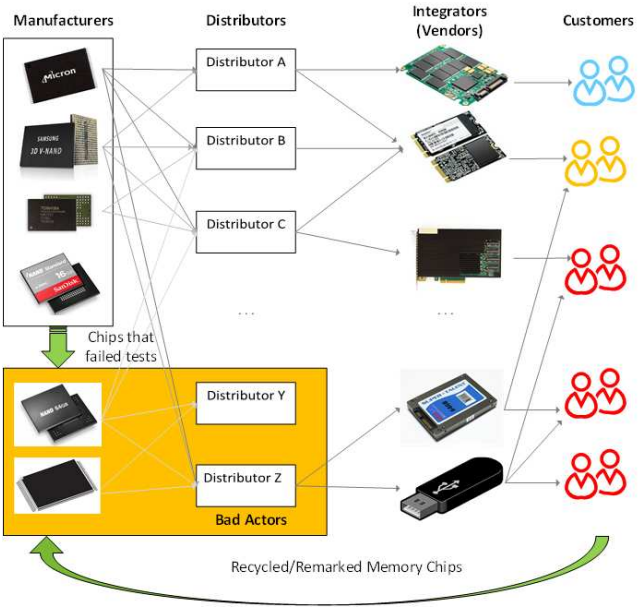

Figure 1: System view of the NAND flash memory supply chain and the involvement of counterfeiters

NAND flash memories in their products is many orders of magnitude larger. These electronic system integrators buy flash memory chips from numerous chip distributors, scattered all over the world. Some of these distributors are authorized and buy chips from the original component manufacturers (OCMs). Unfortunately, several other unauthorized distributors may open the door for counterfeit flash memory chips to enter the market. The result could be that, for example, SSD manufacturers unknowingly build their products using counterfeit chips bought from these distributors. Thus, there have been several media reports that SSDs made by Kingfast contained counterfeit NAND chips [2]. Similarly, 1,500 flash memory chips bought by Raytheon for missile systems were discovered to be counterfeits [26]. According to a report from eBay [15], fake flash memory cards and SSDs usually possess NAND chips that have less than half of their labeled capacity and have slower access speed.

There are multiple pathways through which counterfeit flash memory chips can infiltrate the supply chain:

- **Recycling Used Chips**: Flash chips are often used as mass-storage media in many electronic products that may have limited lifetimes, including smartphones, SSDs, USB drives, and a plethora of emerging Internet-of-Things (IoT) devices.

ICCAD '24, October 27–31, 2024, New York, NY, USA

ACM ISBN 979-8-4007-1077-3/24/10
https://doi.org/10.1145/3676536.3689917

However, the flash memory chips inside these electronic gadgets, in most cases, remain functional even after the end of the product's lifetime. Counterfeiters exploit this by retrieving used flash memory chips from printed circuit boards and reselling them as new at higher prices.

- **Rejected Dies and Fall-Out Chips**: Rejected memory dies that fail post-fabrication tests can re-enter the supply chain through counterfeiters with access to chip packaging sites located in various countries. Even though flash foundries mark some dies as rejected during die-sort testing, these dies can still be repackaged and sold by counterfeiters.
- **Re-branding Inferior Chips**: Counterfeiters may purchase inferior flash chips from less reputed manufacturers and re-brand them to sell at higher prices, misleading consumers about their quality and origin.
- **Cloning Chips**: Counterfeiters with access to foundry facilities can manufacture cloned flash memory chips, making them indistinguishable from genuine ones.

The use of inferior or defective counterfeit non-volatile flash memories results in significant economic losses for the original chip manufacturers, and it can lead to the failure of end-user applications. This ranges from data loss and premature end-of-life of devices to more serious catastrophic events. Developing robust methods to trace the origins of flash memory chips is crucial to mitigate these risks.

Existing approaches for tracing the origins of flash memory chips can be easily circumvented by determining counterfeiters. For example, manufacturers typically store identity information (such as lot and wafer number, manufacturer ID, and date) in a dedicated memory block of the die. However, counterfeiters can erase and reprogram this information once they gain physical access to the chip. Consequently, several research groups have proposed flash memory-based Physical Unclonable Functions (PUFs) [7, 9, 10, 17, 23, 27] for tracking memory origins. Unfortunately, PUF-based chip authentication requires detailed characterizations of individual chips and the maintenance of large databases, which is cumbersome and not commonly practiced by NAND manufacturers. Thus, developing a cost-effective anti-counterfeiting technique for flash memory chips remains a significant challenge.

This paper introduces Flash-Odometer, a technique for accurately estimating the chip usage which is typically quantified by the program-erase cycle (PEC) count (or N_{PE}) of its constituent memory blocks. Flash-Odometer leverages key NAND array characteristics—such as block erase time, intrinsic bit error rate (BER), and high-temperature data retention (HTDR)—which are critically influenced by defect density in the gate-dielectric layer of memory cells. Given the finite endurance of flash memory, estimating PEC count from these intrinsic characteristics offers far-reaching implications, including the detection of recycled memory chips.

The rest of the paper is organized as follows. Section 2 gives the background of 3D NAND flash memory. Section 3 discusses related work while Section 4 describes the Flash-Odometer technique in detail. The experimental setup and experimental procedure are explained in Section 5. Section 6 shows the results and discussion. Finally, Section 7 concludes the paper.

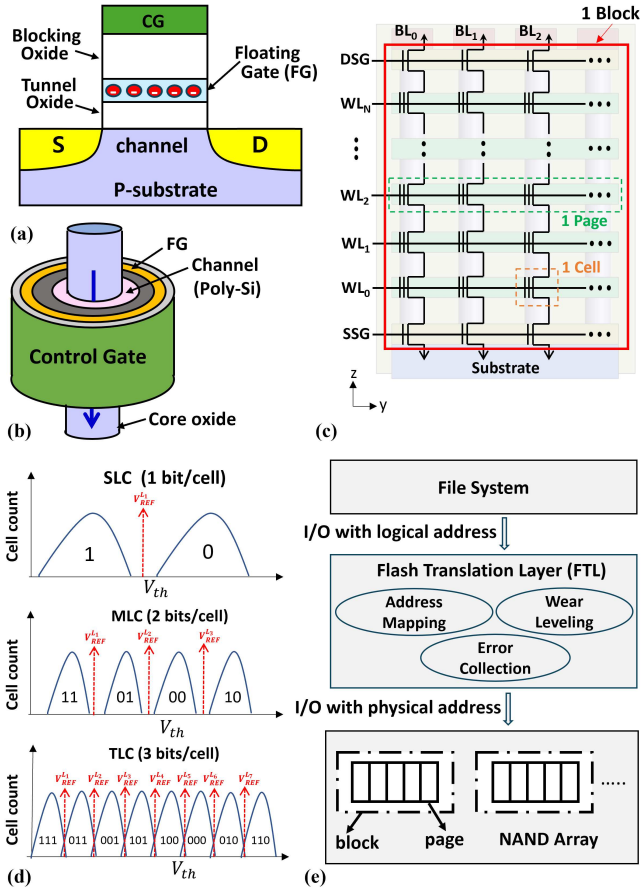

Figure 2: (a) Cross section of a single 3D flash memory cell. (b) Structure of a single 3D NAND flash memory cell in Gate-All-Around (GAA) geometry. (c) Circuit diagram of a flash memory block. (d) Cell V_{th} distribution of SLC, MLC, and TLC storage modes. (e) Schematic of a flash storage system.

2 BACKGROUND

Flash memory cells: Figure 2(a) illustrates the device structure of a flash memory cell made using planar fabrication technology, while the structure of a 3D NAND memory cell, which posses a Gate-All-Around (GAA) geometry, is shown in Figure 2(b). NAND flash memory cell is a floating-gate (FG) metal-oxide-semiconductor field-effect transistor (MOSFET). The presence of trapped negative charge on FG effectively increases the transistor's threshold voltage (V_{th}) relative to the case when there is no charge on the FG. Thus, a flash memory cell is a charge-based analog memory. The program operation charges the FG with electrons via Fowler-Nordheim tunneling, whereas the erase operation removes the charges from the FG. A flash memory cell read operation involves applying a read voltage on the control gate (V_{REF}) and sensing the cell V_{th}. An erased cell conducts the current, and that is sensed as a logic-1, whereas a programmed cell does not conduct the current, and that is sensed as a logic-0.

Flash memory array: Figure 2(c) shows the organization of a NAND flash memory block. Cells in a row constitute a page, and their control gates are connected to a shared word line (WL). The page size varies from 2-16 kilobytes depending on the manufacturer. A collection of pages forms a flash block. A flash memory chip typically includes multiple blocks. The cells in a vertical column are connected to a metal bit line (BL) at one end and to the ground at the other end. Thus, a BL can be pulled down to the ground only if all FG transistors in a column are active (resembling the operation of the NAND gate). The NAND architecture means that data are read or programmed at the page level, whereas erase operations are performed at the level of entire memory blocks. Any flash cell that is set to a logic-0 by a program operation can only be reset to a logic-1 by erasing the entire block.

Traditional flash memory cells store one bit of information (SLC – single-level cell), which requires two different V_{th} states as illustrated in Figure 2(d). The (V_{REF}) is set in between the erased state and programmed state distributions so that there is enough noise margin to correctly identify the cell states as shown in Figure 2(d). Recent advances in controlling and sensing different levels of charge on the FG enabled modern flash memory cells that can store two bits of information (MLC – multi-level cell), three bits (TLC – triple-level cell), or even four bits (QLC – quad-level cell). Figure 2(d) shows the analog V_{th} distribution of the MLC and TLC storage modes. Since the voltage margin between successive V_{th}-states reduces with a higher number of bits per cell, the corresponding cell reliability and endurance are lower for higher bits per cell storage.

Flash memory system: The simplified schematic of a flash storage system, which illustrates the interaction mechanism between the controller and the flash memory chip, is shown in Figure 2(e). The memory controller sends commands to the flash chip to perform storage operations such as write/read, and the flash chip responds by storing/sending digital data from/to the controller. Generally, the memory controller comes with many useful functions. One of the most important functions is error correction, where some failed bits can be corrected after reading from the memory cells. This digital abstraction hides the exact physical properties of the underlying memory bits. Hence, the controller functions are mostly algorithmic, which are agnostic to the exact physical properties of memory bits.

3 RELATED WORK

Over the past few years, researchers have proposed several methods for detecting counterfeit integrated circuits (ICs) [5, 8, 14, 21, 28]. In the following, we summarize the generic as well as flash-specific anti-counterfeiting techniques [4].

Physical/electrical tests: Most counterfeit IC detection methods rely on advanced physical and electrical inspection techniques, including high-tech imaging solutions like X-ray, SEM, and TEM, as well as electrical parametric and functionality tests [5]. While these methods can identify internal and external defects or anomalies associated with counterfeit chips, they face significant challenges. These include prolonged testing times, high costs, the destructive nature of some tests, limited effectiveness, and a lack of automation.

Use of anti-fuse memory: A common approach for tracing the history of a chip is using electronic chip IDs (ECIDs) [6], stored in anti-fuse memory which cannot be modified after the initial write operation. Unfortunately, none of the flash manufacturers includes any on-chip anti-fuse memory. Instead, they store this ECID information in the flash memory itself, which can be easily modified through fault injection by a counterfeiter.

Flash-PUF [7, 9, 10, 12, 17, 23, 27]: Several researchers have proposed the use of PUFs for defying counterfeit problems for flash memory chips. In principle, a PUF-based approach is very attractive, but it requires lengthy PUF extraction from each chip as well as maintenance of large databases for every manufactured chip. In addition, it requires a means for contacting the chip manufacturer to verify the authenticity of each chip, which may place an additional burden on system integrators.

Flash-Watermark [16, 24]: Flash-Watermark is a recently proposed anti-counterfeiting technique specifically for NAND and NOR flash memory chips. It allows permanently imprinting manufacturer watermarks and the chip's usage history that cannot be reversed. The technique utilizes repeated program-erase stressing in order to selectively control the physical properties of the flash cells and hence imprint watermark information into the flash media in an irreversible manner.

Flash-DNA [25]: Flash-DNA is another anti-counterfeiting technique designed for NAND flash memory to identify its manufacturer or origin. The method utilizes the systematic variation between memory word lines which are consistent and unique for a given process but completely different for different manufacturing processes.

Flash Timing Characteristics [11, 22]: It has been suggested that monitoring flash timing characteristics—such as block erase time, page program time, and page read time—can effectively detect recycled flash memory chips. Among these, block erase time has been shown to be the most reliable indicator, as it significantly increases with the program-erase cycle (PEC) count, making it a strong metric for identifying recycled memory.

4 PROPOSED METHOD

The proposed technique for Flash-Odometer relies on extracting NAND array characteristics that are highly sensitive to the PEC count (N_{PE}) of the memory block. We identify three key characteristics for Flash-Odometer namely block erase time, intrinsic BER, and HTDR characteristics. Although all three characteristics are sensitive to the PEC count of a memory block, measurement results show that the HTDR characteristics offer a fine-grained estimation technique of the PEC count with a higher confidence level.

Figure 3 illustrates the general approach for identifying recycled memory chips. Figure 3(a) shows the PDF of the identified NAND array characteristic as a function of its usage condition or PEC count. The distribution width in the array characteristics reflects the block-to-block or chip-to-chip process variation. If process variation in the identified characteristics is high, there will be an overlap between the distribution of fresh and used conditions. In that case, identification of the recycled memory cannot be done with 100% confidence. However, if there is no overlap, the identification can be done with 100% confidence. Figure 3(b) shows the corresponding cumulative distribution function of the specific memory characteristic. The overlap point between the fresh and

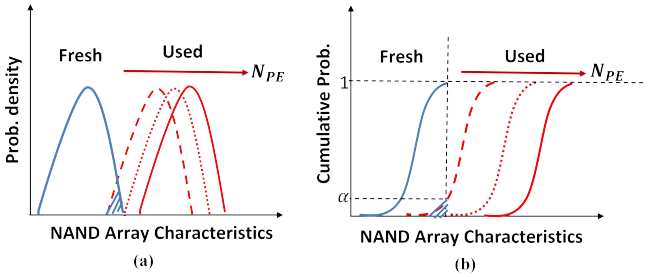

Figure 3: (a) Probability density function (PDF) and (b) the corresponding cumulative distribution function (CDF) of the NAND array characteristics used for usage estimation.

used (dashed line) distribution is denoted by α. The confidence level for detecting recycled memory with this specific usage condition can be calculated by the following equation [11]:

$$Confidence = (1 - \alpha) \times 100\% \quad (1)$$

5 EXPERIMENTAL EVALUATION

5.1 Experimental set-up

The experimental evaluation is performed on COTS 3D 64-layer FG TLC NAND flash memory chips from a well-known vendor. Each memory chip has 256 gigabits of storage capacity while containing 1,008 flash blocks. Each flash block consists of 2,304 flash pages of size 16 kilobytes. Figure 4(a) illustrates a custom hardware-designed board used for interfacing the raw NAND chip with the computer. The board includes a Ball Grid Array (BGA) socket to insert the NAND flash memory chip and an FT2232H mini module from Future Technology Devices International (FTDI) to interface the memory chip with a computer through a Universal Serial Bus (USB) connection. We follow the command sets defined by the Open NAND Flash Interface (ONFI) to perform basic memory operations such as read, write, and erase. This custom-designed hardware setup allows us to access the raw memory bits without any error corrections.

5.2 Experimental procedure

Figure 4(b) illustrates the experimental procedure. We first perform the endurance test on the memory chips by repeatedly performing erase and program operations on several memory blocks with different number of PE cycles (i.e., N_{PE}=1, 100, 200, 300, …, 1,000). Next, we write a random data pattern on all the blocks. Then, we evaluate the HTDR characteristics of the memory chip by baking the chip at 120°C for 5 minutes. Note that using the Arrhenius model with an assumption that the activation energy (E_a) for 3D NAND is 1 eV [1, 13], we find that 5 minutes bake at 120°C is equivalent to ~42 days of data retention at room temperature. After the HTDR test, we cool down the memory chip to room temperature and then read all memory pages to calculate the raw BER.

6 RESULTS AND DISCUSSION

We evaluate our proposed Flash-Odometer approach using 64-layer 3D NAND flash chips in TLC configuration.

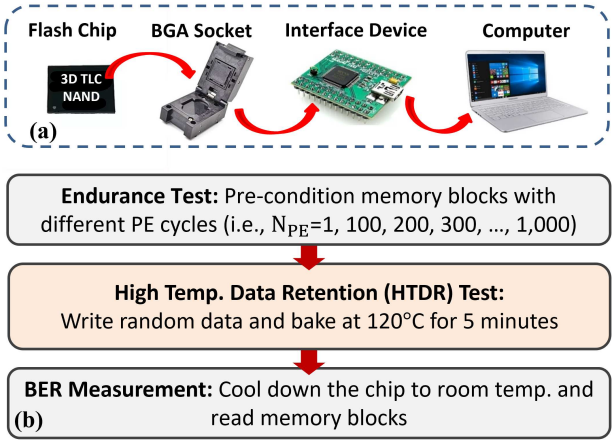

Figure 4: (a) Experimental setup. (b) Experimental procedure.

6.1 Block Erase Time

Block erase time usually increases monotonically with PEC count (N_{PE}) and hence it was previously identified as a good indicator for recycled NAND flash memory detection [22]. However, the previous evaluation was performed on 2D NAND memory chips. In this work, we extend the erase time evaluation on state-of-the-art 3D NAND memory chips. Generally, the block erase time can be measured through the ready/busy pin status as illustrated in the yellow box in Figure 5. At the beginning of the erase operation, the ready/busy pin status changes from 1 to 0 and then switches back to 1 once the operation is finished. More detail of the block erase time measurement can be found in our previous work [19]. Figure 5 shows the evaluation result. We find that the erase time of a NAND memory block increases with higher N_{PE}, as reported previously. One issue with the erase time-based chip usage identification is the discrete nature of the erase time increase with N_{PE}. For example, Figure 5 shows that erase time increases from its nominal value for N_{PE} >50 but it remains constant in the following range: 50< N_{PE} <500. Thus, an increase in erase time may correctly detect a recycled memory chip but it cannot be used for accurate estimation of block usage in terms of PEC count.

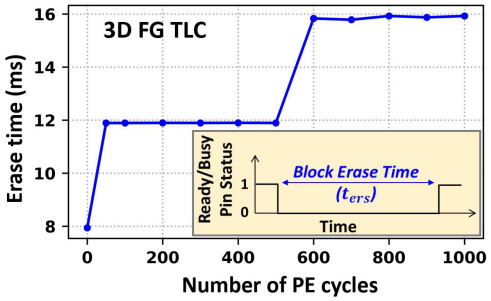

Figure 5: Erase time versus number of PE cycles.

6.2 Intrinsic BER

High-density NAND flash memory shows raw bit errors just after write operation. The percentage of the write-error is coined as intrinsic BER. The intrinsic BER is significantly lower than the standard error correction code (ECC) limit and these errors are corrected by the NAND controller during page read operation. Hence, intrinsic BER is not visible to the end users. Usually, the intrinsic BER increases with block usage and hence it can be considered as another candidate for Flash-Odometer.

Figure 6 summarizes our experimental characterization results of intrinsic BER as a function of N_{PE}. Because each cell stores three bits of information in TLC memory technology, every physical memory layer or WL consists of three logical pages sharing the same WL. The most significant bits (MSB) of the logic states of all the memory cells connected to a given word line form the logical MSB page. Similarly, the least significant bits (LSB) and the central significant bits (CSB) of the logic state of each cell form the LSB and CSB page respectively. The number of memory cells belonging to a given WL determines the logical page size which is 16 kilobytes for the chip under test.

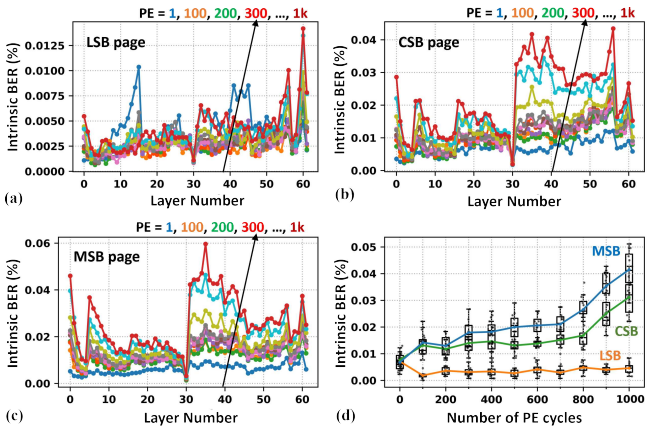

Figure 6: Plots of intrinsic BER as a function of layer number for (a) LSB, (b) CSB, and (c) MSB pages, respectively. (d) The trend of the intrinsic BER with N_{PE}.

Figure 6 shows the intrinsic BER corresponding to LSB, CSB, and MSB pages separately in (a), (b), and (c) respectively. The x-axis represents the vertical layer number, where lower numbers stand for memory layers located at the bottom of the 3D stack. The y-axis represents the average value of intrinsic BER measured in that memory layer. Different colors represent different PEC counts of the memory block. Even though Figure 6(d) shows an increasing trend of the intrinsic BER with N_{PE}, the BER value is significantly lower at the initial N_{PE} values making it a less distinguishable metric for exact usage estimation.

A possible explanation behind the less sensitivity of intrinsic BER with NAND usage is the Incremental Step Pulse Programming (ISPP) scheme. All NAND flash memory technology employs the ISPP scheme during page write operation. ISPP scheme involves multiple program-verify cycles to dynamically adjust the cell V_{th} values irrespective of their physical condition. Thus, the V_{th} distribution of both fresh and worn-out memory appear quite similar just after the write operation producing comparable intrinsic BER.

6.3 HTDR Characteristics

Although the intrinsic BER immediately after the write operation is low, the BER observed after the HTDR test is significantly higher. The detailed measurement steps of the HTDR test are discussed in Section 5.2. Figure 7 summarizes the measurement results, showing BER values across different vertical layers for memory blocks with different usage conditions characterized by N_{PE} numbers. Despite the BER varying significantly depending on the vertical layer location and logical page type, Figure 7(a)-(c) demonstrates that the BER across all vertical layers increases monotonically with N_{PE} for all three logical pages.

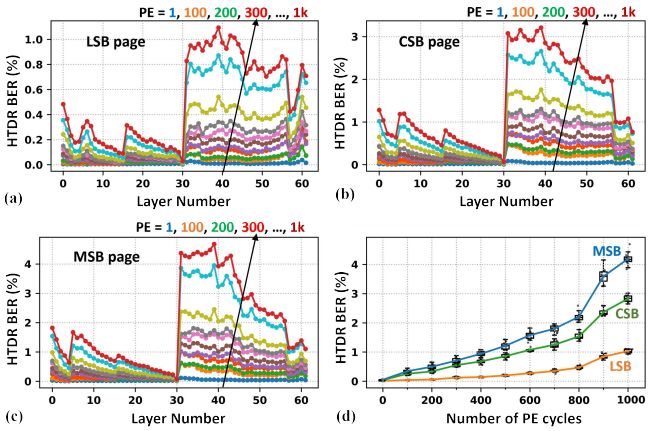

Figure 7: Plots of HTDR BER as a function of layer number for (a) LSB, (b) CSB, and (c) MSB pages, respectively. (d) The trend of the HTDR BER with N_{PE}.

This trend is further highlighted in Figure 7(d), where BER values are plotted for one vertical memory layer (layer 42) under different PE cycle conditions. Each data point in this plot represents measurements from 12 different memory blocks with identical PE cycle conditions, and the spread in BER values for each condition is depicted using a box plot. We observe that the BER increases almost linearly until N_{PE} reaches 800, after which the trend becomes non-linear. This suggests that the HTDR-BER can be effectively used in a Flash-Odometer technique, where block usage can be estimated through linear interpolation of pre-characterized data, as long as N_{PE} remains below 800. For more extensive use, a sophisticated non-linear model could be developed using the pre-characterized BER vs. N_{PE} data to accurately predict block usage with this technique.

Since the accuracy of the Flash-Odometer technique diminishes with increased variability in the identified flash parameters, we performed a detailed analysis of HTDR-induced BER variability. It's important to note that BER values vary significantly depending on both the logical page type and the vertical layer location. For instance, as shown in Figure 7(d), the LSB pages exhibit the smallest increase in BER after the HTDR test, while the MSB pages display the largest increase. This page-type dependent variability

is attributed to the logical encoding of the cell's V_{th} states, which influences BER based on the logical page type. Typically, higher V_{th} states experience more significant shifts after the HTDR test compared to lower V_{th} states. Since shifts in high V_{th} states primarily cause errors in the MSB page, these pages are more prone to errors following the HTDR test. Similarly, layer-dependent variability of BER can be explained using the unique array architecture of 3D NAND memory [18, 20]. For example, the abrupt transition in the BER trend around layer number 32 is due to the two-tier (or double deck) process involved in the 3D NAND fabrication process [3]. In addition, the reactive ion etching process involved in the monolithic fabrication of 3D NAND gives rise to a unique layer-dependent endurance variation [18, 20].

In order to minimize the variability and improve the accuracy of the Flash-Odometer, we propose to choose BER values from a single memory layer with a given logical page type as a representative metric per memory block. For example, we chose MSB pages from memory layer 42 as our representative candidate to illustrate the block-to-block variability and accuracy of the proposed technique. Figure 8 summarizes our accuracy analysis, where we show the cumulative distribution plot of the HTDR-induced BER from the chosen page. Each distribution plot consists of 100 different data points obtained from 100 different memory blocks. We find that even though there is a small overlap between the distributions N_{PE} >100, there is no or negligible overlap between the distribution corresponding to fresh condition and N_{PE} >100. Thus, HTDR characteristics offer detection of recycled memory blocks with 100% confidence for N_{PE} >100.

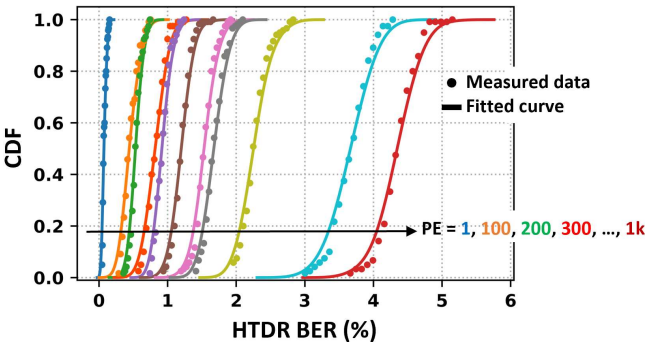

Figure 8: CDF of HTDR-induced BER from a chosen page under different PE cycles.

6.4 Physics behind HTDR based Flash-Odometer

Figure 9 illustrates the underlying concept of the proposed HTDR-based Flash-Odometer technique. In Figure 9(a), a memory cell is shown in its fresh state. During write operation, electrons are stored exclusively in the FG because the tunnel oxide is defect-free. The surrounding oxide layers ensure that the electrons have longer retention, as depicted in the corresponding energy band diagram. In contrast, Figure 9(b) shows the worn-out condition of the same cell after undergoing a certain number of PE cycles. Over time, the tunnel oxide layer develops defect states, where

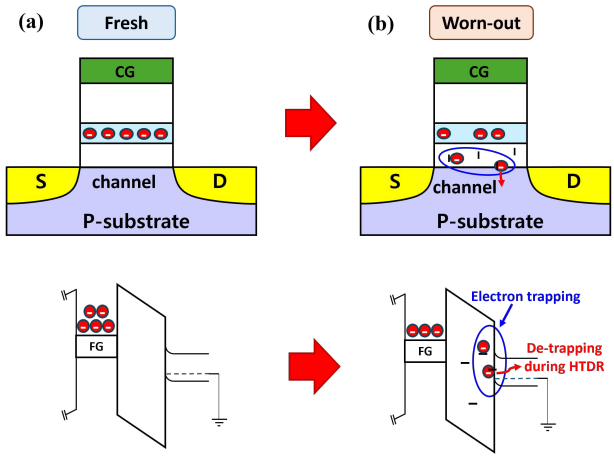

Figure 9: Cartoons and energy band diagrams of (a) fresh and (b) worn-out flash memory cells right after write operation. The worn-out memory cell reveals the defect states along with the trapped electrons.

electrons can become trapped during the write operation. Since both cells store five electrons in this illustrative example, their V_{th} values will be nearly identical. Thus, the ISPP scheme will halt the programming process for the worn-out cell of Figure 9(b), despite it having only three electrons on the FG. Since the location of the stored electrons critically determines the detention characteristics of the flash memory cell, the HTDR characteristics of the worn-out cell will differ significantly from the fresh condition. The trapped electrons near the interface cause rapid de-trapping, leading to poor HTDR performance in the worn-out cell.

7 CONCLUSION

In this paper, we propose a technique called Flash-Odometer, where the age or the usage of the flash memory blocks can be estimated. Our proposed technique leverages key NAND array characteristics including the block erase time, the intrinsic BER, and the HTDR BER. We find that the HTDR BER can deliver an estimated N_{PE} of the memory block with high accuracy as the HTDR BER of flash memory cells truly reflects the defect density in the oxide layer. The experimental evaluation is performed on COTS 3D 64-layer TLC NAND flash memory chips. The experimental results reveal that we are able to detect the aged blocks with 100% confidence for N_{PE} > 100. Hence, this technique is a promising technique to detect counterfeit flash memory chips on the market.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant 2403540, 2317563 and 2346853.

REFERENCES

[1] Matchima Buddhanoy, Preeti Kumari, Umeshwarnath Surendranathan, Maryla Wasiolek, Khalid Hattar, and Biswajit Ray. 2022. Total Ionizing Dose Effects on Long-Term Data Retention Characteristics of Commercial 3-D NAND Memories. *IEEE Transactions on Nuclear Science* 69, 3 (2022), 390–396.
[2] eTeknix. 2019. Kingfast unknowingly sent counterfeit SSDs with mislabelled Flash NAND for review. *eTeknix* (August 2 2019). https://www.eteknix.com/kingfast-

unknowingly-sent-counterfeit-ssds-with-mislabelled-flash-nand-for-review/ Retrieved August 2, 2019.
[3] Akira Goda and Krishna Parat. 2012. Scaling directions for 2D and 3D NAND cells. In *2012 International Electron Devices Meeting (IEDM)*.
[4] Holden Gordon, Jack Edmonds, Soroor Ghandali, Wei Yan, Nima Karimian, and Fatemeh Tehranipoor. 2021. Flash-Based Security Primitives: Evolution, Challenges and Future Directions. *Cryptography* 5, 1 (2021), 7.
[5] Ujjwal Guin, Ke Huang, Daniel DiMase, John M. Carulli, Mark Mohammad Tehranipoor, and Yiorgos Makris. 2014. Counterfeit Integrated Circuits: A Rising Threat in the Global Semiconductor Supply Chain. *Proc. IEEE* 102 (2014), 1207–1228.
[6] Ujjwal Guin, Xuehui Zhang, Domenic Forte, and Mohammad Tehranipoor. 2014. Low-cost On-Chip Structures for Combating Die and IC Recycling. In *Proceedings of the 51st Annual Design Automation Conference* (San Francisco, CA, USA) *(DAC '14)*. Association for Computing Machinery, New York, NY, USA, 1–6.
[7] Zimu Guo, Xiaolin Xu, Mark M. Tehranipoor, and Domenic Forte. 2017. FFD: A Framework for Fake Flash Detection. In *Proceedings of the 54th Annual Design Automation Conference 2017* (Austin, TX, USA) *(DAC '17)*. Association for Computing Machinery, New York, NY, USA, Article 8, 6 pages.
[8] Ke Huang, John M Carulli, and Yiorgos Makris. 2012. Parametric counterfeit IC detection via Support Vector Machines. In *2012 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)*. 7–12.
[9] Shijie Jia, Luning Xia, Zhan Wang, Jingqiang Lin, Guozhu Zhang, and Yafei Ji. 2015. Extracting Robust Keys from NAND Flash Physical Unclonable Functions. In *Proceedings of the 18th International Conference on Information Security - Volume 9290* (Trondheim, Norway) *(ISC 2015)*. Springer-Verlag, Berlin, Heidelberg, 437–454.
[10] Moon-Seok Kim, Dong-Il Moon, Sang-Kyung Yoo, Sang-Han Lee, and Yang-Kyu Choi. 2015. Investigation of Physically Unclonable Functions Using Flash Memory for Integrated Circuit Authentication. *IEEE Transactions on Nanotechnology* 14, 2 (2015), 384–389.
[11] Preeti Kumari, B. M. S. Bahar Talukder, Sadman Sakib, Biswajit Ray, and Md Tauhidur Rahman. 2018. Independent detection of recycled flash memory: Challenges and solutions. In *2018 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*. 89–95.
[12] Shabnam Larimian, Mohammad Reza Mahmoodi, and Dmitri B. Strukov. 2020. Lightweight Integrated Design of PUF and TRNG Security Primitives Based on eFlash Memory in 55-nm CMOS. *IEEE Transactions on Electron Devices* 67, 4 (2020), 1586–1592.
[13] Yixin Luo, Saugata Ghose, Yu Cai, Erich F. Haratsch, and Onur Mutlu. 2018. HeatWatch: Improving 3D NAND Flash Memory Device Reliability by Exploiting Self-Recovery and Temperature Awareness. In *2018 IEEE International Symposium on High Performance Computer Architecture (HPCA)*. 504–517.
[14] Rashmi Moudgil, Dinesh Ganta, Leyla Nazhandali, Michael Hsiao, Chao Wang, and Simin Hall. 2013. A novel statistical and circuit-based technique for counterfeit detection in existing ICs. In *Proceedings of the 23rd ACM International Conference on Great Lakes Symposium on VLSI* (Paris, France) *(GLSVLSI '13)*. Association for Computing Machinery, New York, NY, USA, 1–6.
[15] Fake Flash News. 2018. Fake Flash News - Internet & eBay Fraud. https://fakeflashnews.wordpress.com/ Retrieved August 5, 2018.
[16] Prawar Poudel, Biswajit Ray, and Aleksandar Milenkovic. 2020. Flashmark: watermarking of NOR flash memories for counterfeit detection. In *Proceedings of the 57th ACM/EDAC/IEEE Design Automation Conference* (Virtual Event, USA) *(DAC '20)*. IEEE Press, Article 8, 6 pages.
[17] Pravin Prabhu, Ameen Akel, Laura M. Grupp, Wing-Kei S. Yu, G. Edward Suh, Edwin Kan, and Steven Swanson. 2011. Extracting device fingerprints from flash memory by exploiting physical variations. In *Proceedings of the 4th International Conference on Trust and Trustworthy Computing* (Pittsburgh, PA) *(TRUST'11)*. Springer-Verlag, Berlin, Heidelberg, 188–201.
[18] Md Raquibuzzaman, Md Mehedi Hasan, Aleksandar Milenkovic, and Biswajit Ray. 2022. Layer-to-Layer Endurance Variation of 3D NAND Flash Memory. In *2022 IEEE International Reliability Physics Symposium (IRPS)*. 1–5.
[19] Md Raquibuzzaman, Aleksandar Milenkovic, and Biswajit Ray. 2022. EXPRESS: Exploiting Energy–Accuracy Tradeoffs in 3D NAND Flash Memory for Energy-Efficient Storage. *Electronics* 11, 3 (2022), 424.
[20] Md Raquibuzzaman, Aleksandar Milenkovic, and Biswajit Ray. 2023. Intrablock Wear Leveling to Counter Layer-to-Layer Endurance Variation of 3-D NAND Flash Memory. *IEEE Transactions on Electron Devices* 70, 1 (2023), 70–75.
[21] Daniele Rossi, Vasileios Tenentes, Saqib Khursheed, and Sudhakar M. Reddy. 2018. Recycled IC detection through aging sensor. In *2018 IEEE 23rd European Test Symposium (ETS)*. 1–2.
[22] Sadman Sakib, Preeti Kumari, Bashir M. Sabquat Bahar Talukder, Md. Tauhidur Rahman, and Biswajit Ray. 2018. Non-Invasive Detection Method for Recycled Flash Memory Using Timing Characteristics †. *Cryptogr.* 2 (2018), 17.
[23] Sadman Sakib, Aleksandar Milenković, Md Tauhidur Rahman, and Biswajit Ray. 2020. An Aging-Resistant NAND Flash Memory Physical Unclonable Function. *IEEE Transactions on Electron Devices* 67, 3 (2020), 937–943.
[24] Sadman Sakib, Aleksandar Milenković, and Biswajit Ray. 2020. Flash Watermark: An Anticounterfeiting Technique for NAND Flash Memories. *IEEE Transactions on Electron Devices* 67, 10 (2020), 4172–4177.
[25] Sadman Sakib, Aleksandar Milenković, and Biswajit Ray. 2021. Flash-DNA: Identifying NAND Flash Memory Origins Using Intrinsic Array Properties. *IEEE Transactions on Electron Devices* 68, 8 (2021), 3794–3800.
[26] VentureBeat. 2011. Feds close the books on a huge chip counterfeiting scheme. *VentureBeat* (September 25 2011). https://venturebeat.com/2011/09/25/feds-close-the-books-on-a-huge-chip-counterfeiting-scheme/ Retrieved August 6, 2018.
[27] Yinglei Wang, Wing-kei Yu, Shuo Wu, Greg Malysa, G. Edward Suh, and Edwin C. Kan. 2012. Flash Memory for Ubiquitous Hardware Security Functions: True Random Number Generation and Device Fingerprints. In *2012 IEEE Symposium on Security and Privacy*. 33–47.
[28] Zhichao Xu, Aijiao Cui, and Gang Qu. 2020. A New Aging Sensor for the Detection of Recycled ICs. In *Proceedings of the 2020 on Great Lakes Symposium on VLSI* (Virtual Event, China) *(GLSVLSI '20)*. Association for Computing Machinery, New York, NY, USA, 223–228.