Toward Instantaneous Sanitization through Disturbance-induced Errors and Recycling Programming over 3D Flash Memory

Wei-Chen Wang, Ping-Hsien Lin, Yung-Chun Li, Chien-Chung Ho, Yu-Ming Chang, Yuan-Hao Chang

November, 2019

Outline

• Introduction
• Background and Motivation
• Instantaneous Sanitization and Recycling Programming Designs
• Experimental Results
• Conclusion
Why Data Security Becomes More and More Critical?

- **Data security** has risen to be one of the most critical concerns nowadays
  - Enterprise services start to distribute their data on the cloud and data centers
  - Users gradually change their habits to access data from the remote locations
- Applications/systems usually cache sensitive user data on the many different local storage devices for having the better I/O performance
  - Urgent needs are raised to protect sensitive user data from unauthorized uses
- **Sanitization (or secure deletion)** is proposed to meet the needs of completely removing the sensitive data form the storage devices physically

Source: Dell Inc., Dell Data Security Survey 2016-03 [1]
The Deficiency of Existing Sanitization Techniques

- Comparison of existing sanitization techniques

<table>
<thead>
<tr>
<th>Item to delete</th>
<th>Erasure-based</th>
<th>Encryption-based</th>
<th>Overwriting-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item to delete</td>
<td>The corresponding flash block</td>
<td>The encryption key</td>
<td>The data itself</td>
</tr>
<tr>
<td>Advantages</td>
<td>True data secure deletion</td>
<td>Efficiency</td>
<td>True data secure deletion</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Inefficiency</td>
<td>Key management, data loss risk</td>
<td>Disturbance issue</td>
</tr>
</tbody>
</table>

- There is no previous work takes flash’s error and voltage characteristics into considerations on the design of sanitization techniques
  - Disturbance-induced errors
  - Available window of threshold voltage

This work aims at realizing the efficiently fine-grained sanitization design by conversely exploiting the adverse effects, i.e., program disturbance, on flash
Outline

• Introduction
• Background and Motivation
• Instantaneous Sanitization and Recycling Programming Designs
• Experimental Results
• Conclusion
Background – NAND Flash Memory Basis

- NAND flash memory, a kind of non-volatile memory, persistently stores data through injecting electrons into each memory cell in a page.
- **Program**
  - Electron is moved into the floating gate (FG), and the threshold voltage is raised.
- **Erase**
  - Electron is removed from the floating gate (FG), and the threshold voltage is thus lowered.
- **V\text{read}**
  - The Voltage used to apply on the wordline where the page resides in
    - Cells’ Vt value < V\text{Read} : return 1
    - Cells’ Vt value > V\text{Read} : return 0
- **Error bits**
  - Cells whose Vt status are not in their target Vt window, and they can lead to wrong data value while they are applied with V\text{Read}
Background – Program Disturbance

- Incremental step pulse programming (ISPP) method
  - Programmed pages are programmed and verified in turn for multiple iterations
  - Memory cells’ threshold voltage ($V_t$) distribution are gradually pushed to pass through their target program verify voltage

- Program Disturbance
  - When enough electrons are forced into a neighboring cell during a program, that cell will appear as weakly programmed in the next page read of that cell’s page.

Background – Program Disturbance on 3D Flash Memory

• The disturbance problem is getting worse as the scaling of memory cells keeps advancing for the cost consideration.
• 3D flash memory presents a good opportunity to further reduce the bit cost, it also introduces a new sort of disturbance from the vertical direction ($x$, $y$, $z$ disturb directions).
Research Observation

• In the past, lot of studies and works tell us to keep the Vt distribution of cells with the different data status as wide as possible for improving the reliability
  – Less error bits & program disturbances

• However, there exists the possibility to realize the efficient sanitization if we the properly control and utilize cells’ Vt distribution and disturbance properties
  – The error bits arisen after squeezing two states of an SLC flash memory into a narrower window of Vt can still be correctable if the value of Vt shift is properly selected.
Research Observation

- In the past, lot of studies and works tell us to keep the Vt distribution of cells with the different data status as wide as possible for improving the reliability
  - Less error bits & program disturbances
- However, there exists the possibility to realize the efficient sanitization if we the properly control and utilize cells’ Vt distribution and disturbance properties
  - The error bits arisen after squeezing two states of an SLC flash memory into a narrower window of Vt can still be correctable if the value of Vt shift is properly selected.
  - Under this case, programming adjacent pages can help in sanitizing written data in some SLC pages with a narrow window of Vt
Motivations and Objectives

• The two key observations motivate us to seek for an innovative design to realize a sanitization feature by means of skillfully exploiting the negative effect, i.e., program disturbance

• We aim at pursuing the optimized solution to efficiently achieve the highest degree of data sanitization
  – Each old version of data shall be destroyed whenever a new version of the data arrives

• Technical Challenges will fall on
  – How to realize data sanitization through programming of the adjacent pages
  – How to reuse the window of Vt after the rearrangement of Vt states for an SLC page so as to further reduce the performance overheads
  – How to redesign the FTL management at system level with integrating the proposed programming strategy
Outline

• Introduction
• Background and Motivation
• Instantaneous Sanitization and Recycling Programming Designs
  – Instantaneous sanitization
  – Recycling Programming
  – System Architecture and Management Process
• Experimental Results
• Conclusion
Design Concept – Disturbance-induced Error

- 3D flash memory suffers from more directions of program disturbance which could lead to more data error
  - For some cells of P17, their $V_t$ distribution might be pushed to pass through the read voltage $V_{\text{read}}$, and they will return the wrong $V_t$ state
  - Disturbance-induced error refers to the data error caused by the program disturbance
- If there are too many disturbance-induced errors in a flash page, it fails on ECC decoding
Instantaneous Sanitization

• In our **instantaneous sanitization** design, we propose to deliberately create a certain number of disturbance-induced errors on a flash page in the very beginning stage while it is programmed.
  – To **speed up the accumulation of the disturbance-induced errors** on a flash page.
• It purposely lets Vt distribution of cells whose desired data state is “1” and that of cells whose desired data state is “0” have a certain level of overlap.
• With the purposely generated overlapping, the page data can be instantaneously sanitized with being exactly once disturbed.

Original state “1” and “0” under Conventional ISPP

Correctable Error Bits

![Diagram showing correctable error bits]

Uncorrectable Error Bits

![Diagram showing uncorrectable error bits]
Recycling Programming

- A **recycling programming design** is also proposed to reduce the needs of frequently invoking GC process by reusing the sanitized pages.
- The main idea is to squeeze two states of an SLC flash memory into a narrower window of Vt so as to leave a wide range of unused Vt window
  - A sanitized page can be used to contain the newly written data without being erased, as if it is recycled and reused.
- Note that, there exists a limited number on applying the recycling programming design on a flash page.
System Architecture

• Most of the FTL components do not need to be changed or modified since the proposed ISRP design works by altering the Vt distribution of each flash page
  – It only needs to modify the number of program steps and program voltage used in flash programming stage
• Only the address allocator needs to be slightly modified for preventing valid page from being disturbed
  – Every two pages are grouped together to form a page region
Interleaving Region Allocation Strategy

- The main function of IRA strategy are twofold:
  - Assigning and arranging page locations for storing the data which exist the risk of being sanitized in the nearly future
  - Preventing the pages which keep the valid data from being disturbed while the instantaneous sanitization is applied on their adjacent pages.
Outline

• Introduction
• Background and Motivation
• Proposed mechanism
• Experimental Results
• Conclusion
Experimental Setup

• The proposed sanitization scheme is implemented on a real 3D NAND flash memory
  – Device-level experiment
    • Feasibility
  – System-level experiment
    • Capability

A device-under-testing (DUT) board with our flash memory

A pin-control board to support designing the proposed instantaneous sanitization programming
Experimental Results – Device-Level Results

• The experimental results of device-level show the capability and feasibility of instantaneous sanitization with proposed mechanism.
Experimental Results – System-Level Results (Performance)

- When the recycling programming design is adopted and each page can be programmed 4 times (i.e., ISRP×4), the write response time improves excessively by 65.81% to 86.91% under the eight investigated traces.

![Normalized write latency graph]

![Normalized read/write latency graph]
Experimental Results – System-Level Results (Erase count)

- The number of P/E cycle of the proposed ISRP approach without adopting the recycling programming design (i.e., ISRP\textsubscript{1}) is as same as that of the scrubbing approach under each trace.
- However, the proposed ISRP scheme can further reduces the number of block erase count ranging from 69.41% to 88.82% while adopting the 4-time recycling programming design
Outline

• Introduction
• Background and Motivation
• Proposed mechanism
• Experimental Results
• Conclusion
Summary

- In this work, we propose a instantaneous sanitization design to provide the highest level of data security with zero live-data-copy for SLC flash memory.

- A recycling programming design is also proposed to reduce the number of GC invocation so that the performance and lifetime can be further improved.

- The proposed scheme can deliver the optimal performance and alleviate the program disturbance.

- The evaluation on proposed scheme is based on real flash chip. The device-level evaluation shows the write response time improves excessively by 65.81% to 86.91% under the eight investigated traces.
Question & Answer