Work-in-Progress: Mitigating Write Disturbance in Phase Change Memory Architectures

Chao-Hsuan Huang, Ishan G Thakkar
Department of Electrical and Computer Engineering, University of Kentucky, Lexington, KY, U.S.A.
{chu276, igthakkar}@uky.edu

ABSTRACT

Phase Change Memory (PCM) is seen as a potential candidate that can replace DRAM as main memory, due to its better scalability. However, writing '0s' in PCM cells requires high-temperature RESET operations, which induce write disturbance errors in neighboring idle PCM cells due to excessive heat dissipation. This paper introduces low-temperature partial-RESET operations for writing '0s' in PCM cells. Compared to traditional RESET operations, partial-RESET operations dissipate negligible heat, and therefore, do not cause disturbance errors in neighboring cells during PCM writes.

KEYWORDS : Phase Change Memory (PCM), Write Disturbance, Reset Operation, Reliability.

1 BACKGROUND AND MOTIVATION

A Phase Change Memory (PCM) cell embeds a resistive heater and a small volume of chalcogenide material $Ge_2Sb_2Te_5$ (GST) between two electrodes [1], as shown in Fig. 1(a). The GST volume can be programmed into two different states (i.e., crystalline and amorphous) with dramatically different electrical resistance. The amorphous high-resistance (usually in the MΩ range) state represents a "0", while the crystalline low-resistance (usually in the KΩ range) state represents a "1". To write a PCM cell, two basic operations, RESET and SET, are needed. The SET operation that writes a "1" into the PCM cell requires a long-duration and low-amplitude current pulse I_{SET} (Fig. 1(c)). On the other hand, the RESET operation that writes a "0" into the PCM cell requires a short-duration and high-amplitude current pulse I_{RESET} (Fig. 1(c)), to program the GST volume into the amorphous state. While a SET operation does not typically cause any disturbance, a reliable RESET operation can heat up the GST volume to ~1226.85°C [2] locally, which can dissipate excessive amount of heat into the neighboring PCM cells. *This dissipated heat can cause write disturbance (WD) errors in the neighboring PCM cells, which can potentially change their information content.*

1.1 Write Disturbance (WD) in PCM Cell-Array

Fig. 1(b) shows a PCM cell-array, where PCM cells are arranged in multiple wordlines (WLs) and bitlines (BLs). The "aggressor", shown in Fig. 1(b) at the center, is a PCM cell that is undergoing a RESET operation. The excessive heat generated by this aggressor cell dissipates into the neighboring PCM cells, in the same wordline as well as in the adjacent wordlines, increasing their temperature. If these affected neighboring cells are in the RESET state (storing "0"s), their amorphous GST volumes can become partially crystalline due to the increase in their temperature, making them the "victims" of the aggressor cell. The "0"s stored in these victim cells can be erroneously read as "1"s, if the partial crystallization of their GST volumes significantly reduces their resistance. The probability (F) of such readout error, which is referred to as write disturbance (WD) error henceforth, to occur in a victim cell is given by Eq. (1) [2].

$$F = 1 - \exp(-(\frac{t_{fail}}{t_0})^{\beta}) \qquad (1)$$

Here, β and t_0 are constants that depend on the PCM technology node and temperature of the victim cell [2].

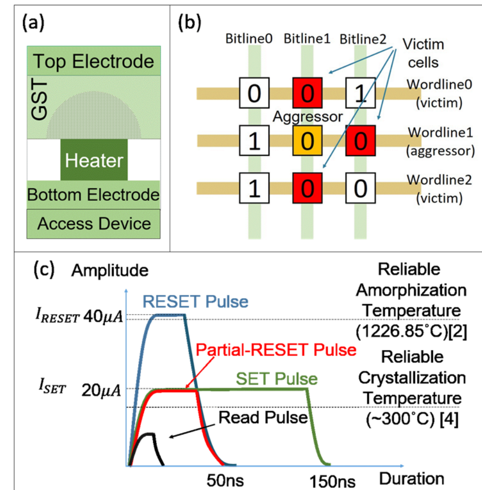

Fig. 1: (a) Basic structure of a PCM cell; (b) A schematic of PCM cell-array; (c) PCM programming pulses from [5].

1.2 Related Work

To mitigate WD, conventional designs (e.g., DIN [2], SD-PCM [3], and ADAM [4]) focus on reducing the number of aggressor and victim cells per PCM write of a cacheline, and also employ Verify and Restore (VnR) mechanism to iteratively rewrite the affected cachelines, until all WD errors are recovered. DIN and ADAM employ frequent data compression to reduce number of "0"s in the compressed data, to ultimately reduce the number of aggressor and victim cells. SD-PCM uses data compression to reduce victim cells, and additionally, it employs error-correcting codes to recover from inflicted WD errors. The VnR mechanism used in these techniques incurs very high energy and latency overhead. To mitigate this overhead, ADAM retrieves the 'victim' cachelines from the last level cache, relaxing the need of using VnR [4].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CASES'19 Companion, October 13–18, 2019, New York, NY, USA
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6925-1/19/10…$15.00

2 PARTIAL-RESET

To mitigate WD, we propose to use partial-RESET operations instead of full-RESET operations to write "0"s. Unlike prior works that focus on reducing the aggressor and victim cells per PCM write operation, partial-RESET focuses on eliminating the root cause of WD, that is, the excessive heat dissipation from the aggressor cells. A partial-RESET operation uses a lower-amplitude short-duration current pulse (red curve in Fig. 1(b)) to write "0"s, instead of the high-amplitude RESET pulse. As shown in Fig. 2, a partial-RESET pulse programs a PCM cell's GST volume in a poly-crystalline state, which renders lower resistance than the amorphous state (the full-RESET state). Nevertheless, the poly-crystalline PCM state (partial-RESET cell) still renders 10-20× higher resistance than the crystalline PCM state (SET cell), which can provide enough readout margin (Fig. 2) to distinguish the PCM cells storing "1"s (SET cells) from PCM cells storing "0"s (partial-RESET cells). Moreover, even if the resistance of a partial-RESET cell drifts (increases) with time due to the atomic restructuring of the GST volume [5] (Fig. 2), the readout margin remains unviolated (Fig. 2), allowing the partial-RESET cell to be unerringly distinguished from the SET cell. We think that the polycrystalline partial-RESET state of a PCM cell (Fig. 2) is deterministically achievable with reasonable repeatability, as it is analogous to an intermediate resistance state of a multi-level (MLC) PCM cell [9], which is traditionally achieved with deterministic repeatability by applying a current pulse with intermediate amplitude and/or width [8].

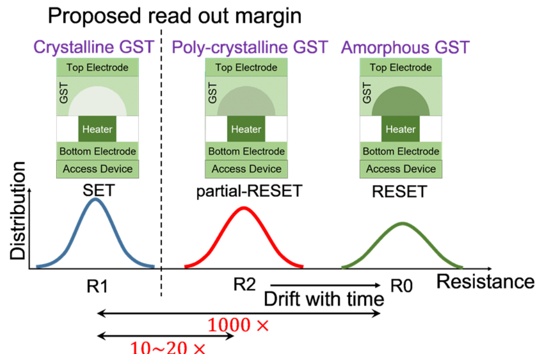

Fig. 2: Illustration of partial-RESET, SET, and RESET PCM cells, and their resistance distributions.

The lower-amplitude current pulse used for the partial-RESET operation hardly increases the temperature of an aggressor cell above the crystallization point (300°C), which results in negligible heat dissipation from the aggressor cell into the victim cells. This in turn increases the β value in Eq. (1) to ~50 [2], rendering a negligible value of ≈ 10^{-12} for the WD error probability F. Thus, with nearly zero WD error probability, partial-RESET operations effectively eliminate the WD problem in PCM architectures. Moreover, the use partial-RESET operations eliminates the need of traditional VnR mechanism [2]-[4], and as a result, saves PCM architectures from the excessive latency and energy overheads related to the VnR mechanism.

3 RESULTS

For our trace-driven evaluations with PARSEC benchmarks [7], we use the PCM configuration and simulation environment from [4]. Moreover, we employ the method from [2] to evaluate per-write energy and latency values for the baseline, ADAM [4], and our proposed partial-RESET PCM architectures, for 20nm PCM technology. Fig. 3(a) shows number of total WD errors (BL+WL WD errors) per-write for three PCM architectures across three benchmarks. As evident, partial-RESET has zero WD errors. Fig. 3(b) and 3(c), respectively, present per-write energy and latency, for three PCM architectures. Our partial-RESET PCM does not require expensive VnR mechanism, which yields the least per-write energy and latency values for partial-RESET, compared the baseline and ADAM.

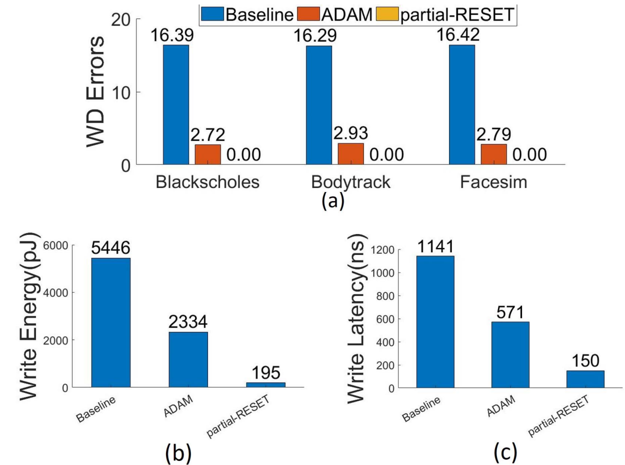

Fig. 2: (a) Number of total WD errors (BL+WL WD errors) across three PARSEC benchmarks, (b) energy, and (c) latency, per-write for the baseline, ADAM, and partial-RESET PCM architectures. The bars for partial-RESET in (a) are not visible due to their ~zero heights.

4 CONCLUSIONS AND FUTURE WORK

This paper shows that the use of partial-RESET operations can eliminate write disturbance (WD) errors, in addition to improving the energy-efficiency and latency of reliable write operations, in SLC PCM architectures. Thus, partial-RESET operations represent a promising solution for mitigating WD in the scaled SLC PCM implementations of the future.

Going forward, we plan to perform a detailed system-level simulation analysis of partial-RESET operations with more number of benchmark applications. Furthermore, we intend to explore the usefulness of partial-RESET operations for mitigating WD errors in MLC PCM architectures as well.

REFERENCES

[1] H. Horii et al., "A novel cell technology using N-doped GeSbTe films for phase change RAM," in Proc. VLSIT, 2003.

[2] L.Jiang et al.," Mitigating write disturbance in super dense phase change memories," in Proc. DSN, 2014.

[3] R.Wang et al., "SD-PCM: Constructing reliable super dense phase change memory under write disturbance," in Proc. ASPLOS, 2015.

[4] S. Shivam et al., "ADAM: Architecture for write disturbance mitigation in scaled phase change memory," in Proc. DATE, 2018.

[5] I. Thakkar et al., "DyPhase: A Dynamic Phase Change Memory Architecture with Symmetric Write Latency and Restorable Endurance," in IEEE TCAD, vol. 37, no. 9, September 2018.

[6] M. Poremba et al., "NVMain: An architectural-level main memory simulator for emerging non-volatile memories," in Proc. ISVLSI, 2012.

[7] C. Bienia, S. Kumar, J. P. Singh, and K. Li, "The PARSEC Benchmark suite: Characterization and architectural implications," in Proc. PACT, 2008.

[8] G. W. Burr et al., "Recent progress in phase-change memory technology," in IEEE JETCAS, vol.6, no.2, June 2016

[9] M. Stanisavljevic et al, "Demonstration of reliable triplelevel-cell (TLC) phase-change memory," in Proc. IMW, 2016.