Enhancing the lifetime of Non-Volatile Memory by hybrid cache replacement policy

Presented by : Sampath Kumar S, 206116021

Guide : Ms B Shameedha Begum

Non-Volatile caches

- Modern computers need large on-chip caches.
- Scalability of SRAM and DRAM is constrained by
  - Low cell density
  - High Leakage power
- Non-Volatile Memory is better alternative for SRAM and DRAM due to High cell density and Low Leakage power.
- NVM can eliminate leakage energy when they are in standby, thereby reducing the total energy consumption.

Problem Statement

- Due to **limited write endurance** and **high write energy** in NVM, Effective write mechanism is required to utilize the full efficiency of Non-Volatile Memory.
- The proposed technique will **improve the lifetime** of the cache sub systems by combining **Wear-leveling** and **Write Minimization techniques**.

Wear Leveling Techniques

- Wear Leveling Techniques are used for pro-longing the service lifetime of erasable computer storage. This technique ensures writes are distributed across the cache sets and blocks.

EqualWrites

- Using Wear-Leveling Technique, EqualWrites reduces the Intra-Set variations by redirecting writes of the hot blocks (Maximum writes) to the cold blocks (Minimum writes).

Write Minimization Techniques

- Write Minimization Techniques are used for reducing the number of writes on the erasable computer storage to pro-long the lifetime of the erasable computer storage with negligible area overhead.

Flip-N-Write

- Write Minimization Technique perform quick bit-by-bit inspection before write. It stores the data in actual format or in flipped format depending on the hamming distance Computed.

Drawbacks of Existing Solution

- Existing Solution does not utilize the full capacity and efficiency of the Non-Volatile Memory.
- The major disadvantage of Non-Volatile Memory is their high error rate during write operation, which is not considered.
- EqualWrites distribute the number of writes on the block level and it does not handle repeated writes on the same cell inside the cache block, which may lead to Security threat.

Proposed Methodology

- The high error rate during write operation and unequal writes can be controlled by designing a **hybrid cache replacement policy** which in turn improves the lifetime of the Cache Subsystems.
- During new data write into memory cell, either **0 to 1** or **1 to 0 transition** possible.
- The error rate during **0 to 1 transition** is 100 times higher than **1 to 0 transition**.

Proposed Methodology

- Before writing a new data into the cache block, a quick bit-by-bit inspection is done to compute,
  - Hamming distance between Existing data and Incoming data
  - number of 0 to 1 transition
- If the hamming distance is more than half of the size of the block, data will be stored in inverted format, which will ensures at-most half number of cells will be involved in write operation.
- The line within the set which has least number of **0 to 1 transition and minimum writes within the set** will be chosen.

Proposed Methodology

- In addition, for ensuring reliability of the data stored within the cache block, Hamming Codes are used.
- Within the cache block, a repeated write operations on the selective cells may shorten the lifetime of the cache block, which is considered as serious Security threat.
- This can be overcome by rotating the bit location within the cache line without fixing the static location.

Work done

Tools

- GEM5 Simulator
- Evaluation using PARSEC benchmark suite
- **Operating System :** Ubuntu 14.04
- **Language :** C++, Python

Partial Implementation

- Installation of GEM5 Simulator
- Integration of GEM5 with PARSEC benchmark suite
- Analyzed the write variation in Intra-Set and Inter-Set
- Implemented EqualWrites Technique

References

[1] Mittal, Sparsh, and Jeffrey S. Vetter. "Equalwrites: Reducing intra-set write variations for enhancing lifetime of non-volatile caches." IEEE Transactions on Very Large Scale Integration (VLSI) Systems 24.1 (2016): 103-114.

[2] Monazzah, Amir Mahdi Hosseini, Hamed Farbeh, and Seyed Ghassem Miremadi. "LER: Least-error-rate replacement algorithm for emerging STT-RAM caches." IEEE Transactions on Device and Materials Reliability 16.2 (2016): 220-226.

[3] Cho, Sangyeun, and Hyunjin Lee. "Flip-N-Write: A simple deterministic technique to improve PRAM write performance, energy and endurance." Microarchitecture, 2009. MICRO-42. 42nd Annual IEEE/ACM International Symposium on. IEEE, 2009.

Thank You