# Reliability Improvement of STT-MRAM Cache Memories in Data Storage Systems

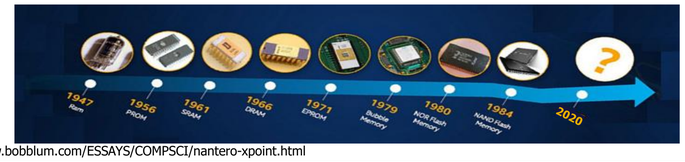
**Elham Cheshmikhani**, Hamed Farbeh, & Hossein Asadi

Data Storage, Networks, & Processing
DSN Lab
http://dsn.ce.sharif.edu/

IEEE/ACM Design, Automation & Test in Europe (DATE), Virtual Conference and Exhibition, February 2021

## Introduction: Cache Memory Technology 1

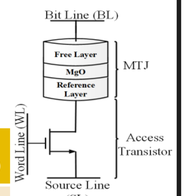
- Recent technology, architecture, and application trends
  - Lead to new requirements
  - Exacerbate old requirements
- Conventional cache memories suffer from low density, high leakage power, low reliability due to soft error, and high sensitivity to process variation
- We need to rethink the cache memory system to
  - Fix DRAM and SRAM issues and enable emerging technologies
  - Satisfy all requirements




https://www.bobblum.com/ESSAYS/COMPSCI/nantero-xpoint.html

## Replacing Conventional Technology 2

- Some alternatives are RRAM, FeRAM, PCM, MRAM, and STT-MRAM
- The most promising technology is Spin Transfer Torque Magnetic RAM
  - MTJ
    - Free layer
    - Reference layer
    - Tunnel layer
  - Access Transistor

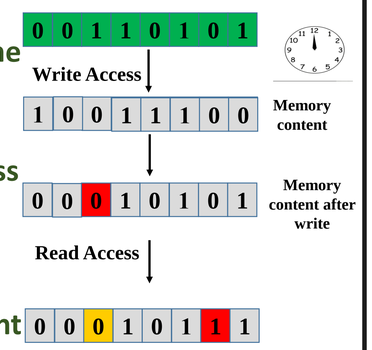



| | Area ($mm^2$) | Read Latency (ns) | Read Energy (nJ) | Leakage Power (mW) | Write Latency(ns) | Write Energy (nJ) |
|---|---|---|---|---|---|---|
| 1 MB SRAM | 0.87 | 1.748 | 0.054 | 33.750 | 1.491 | 0.051 |
| 4MB STT-MRAM | 0.791 | 2.730 | 0.13 | 2.477 | 11.212 | 0.352 |

3 to 4 times denser | Comparable read latency | 1/10 times lower leak | Higher write delay | Some reliability challenges

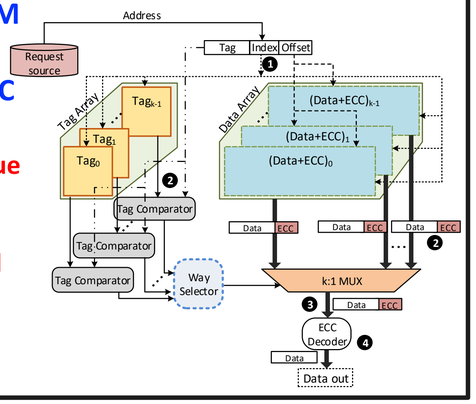
## STT-MRAM: Reliability Challenges 3

- **Retention Failure:** Without applying any current, an idle cache cell may flip stochastically
- **Write Failure:** Stochastic behavior of magnetization process does not switch on a write operation
- **Read Disturbance:** A cache cell may unintentionally flip by current applied for reading a cache block



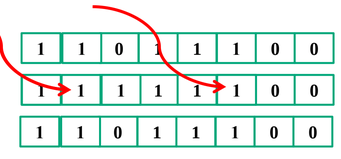

## Conventional Cache: Using ECC 4

- Consider the STT-MRAM conventional cache
- Error is Corrected in ECC Decoder Unit
- Extra reads are imposed due to each line read from Data array
- Writing "1"s is difficult and results in more write failure
- Higher temperature increase all the error rates



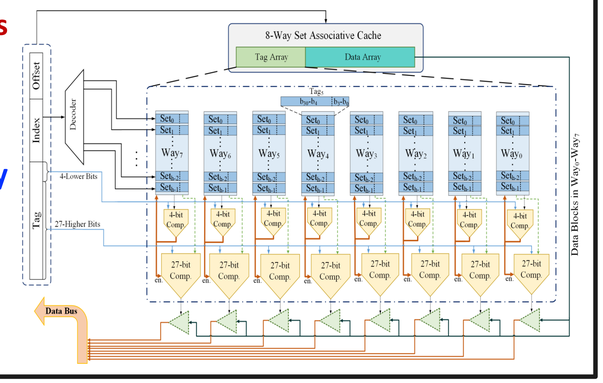

## Proposed Fault Avoidance Method: TA-LRW Replacement Poli 5

- One of the main sources of error is heat
- Write operation is the source of heat in STT-MRAM
- Thermal-Aware Least Recently Written (TA-LRW)
- Distributes write operations
- Distributes temperature
  - Write failure ↓ ☺
  - Read disturbance ↓ ☺
  - Retention failure ↓ ☺



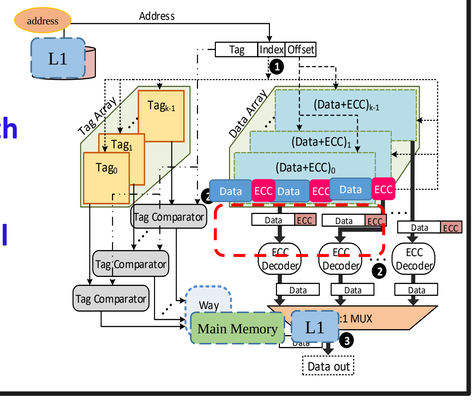

## Proposed Fault Tolerance Architecture: 3RSeT 6

- Decreasing read operations causes read disturbance reduction
- 3RSeT partially read the tags
  - Reduction in number of reads



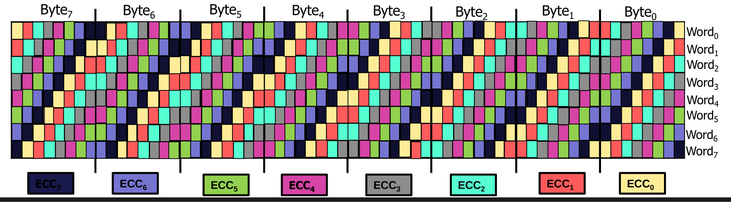

## Read Error Accumulation Preventer (REAP) 7

- REAP Cache
  - Simultaneous Read during tag comparison
  - Swaps ECC decoder with MUX unit
  - Replicates decoders to simultaneously check all blocks
  - Guarantees to perform ECC check for all cache lines




## IncRemental OBlique Interleaved ECC (ROBIN) 8

- Large variation in bit positions
  - Distributing the transitions between codewords
  - Minimizing the effect of workloads variation and data pattern differences
- ROBIN: IncRemental Oblique Interleaved ECC
  - Similar contribution of all parts of block
    - All bytes of each 64-bit dataword
    - All words, Similar bit positions in all bytes




## Reliability Optimized STT-MRAM Memory (ROSTAM) 9

- Combining all the fault tolerance and avoidance methods
- Using TA-LRW, 3RSeT, REAP and ROBIN
- Total error rate → 10x ↓
- MTTF → ↓ 3 orders of magnitude