ASIA SOUTH PACIFIC DESIGN AUTOMATION CONFERENCE

Energy, Latency, and Lifetime Improvements in MLC NVM with Enhanced WOM Code

Huizhang Luo[1], **Liang Shi[1*]**, Qiao Li[1],

Chun Jason Xue[2] and Edwin H.-M. Sha[1]

[1]Chongqing University
[2]City University of Hong Kong

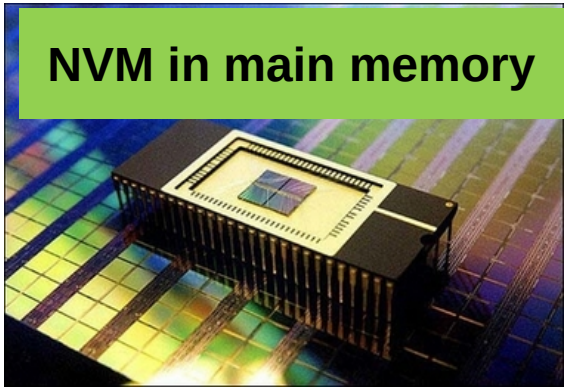
重慶大學 CHONGQING UNIVERSITY

香港城市大學 City University of Hong Kong

Outline

Introduction

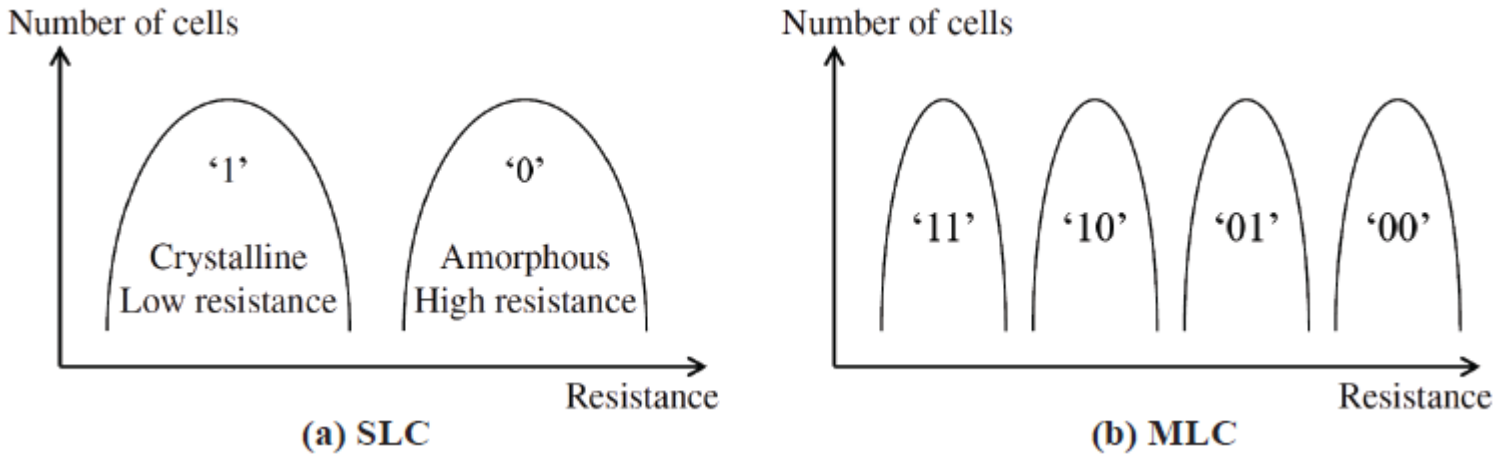
- Increasing memory capacity requirement



Replacing DRAM with NVM

- Challenges of DRAM
  - Limited scalability, High leakage power
- Opportunities of NVM
  - High density, low leakage power, Non-volatility

MLC NVM



➤ Multiple level cell (MLC) NVM

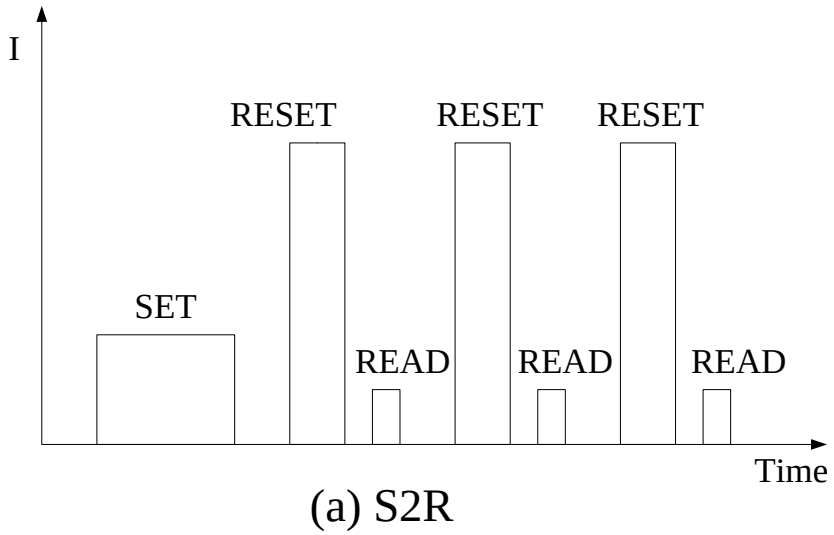
By further dividing the large resistance range
Increases the data density, lower price, lower energy per bit

➤Complex and costly programming

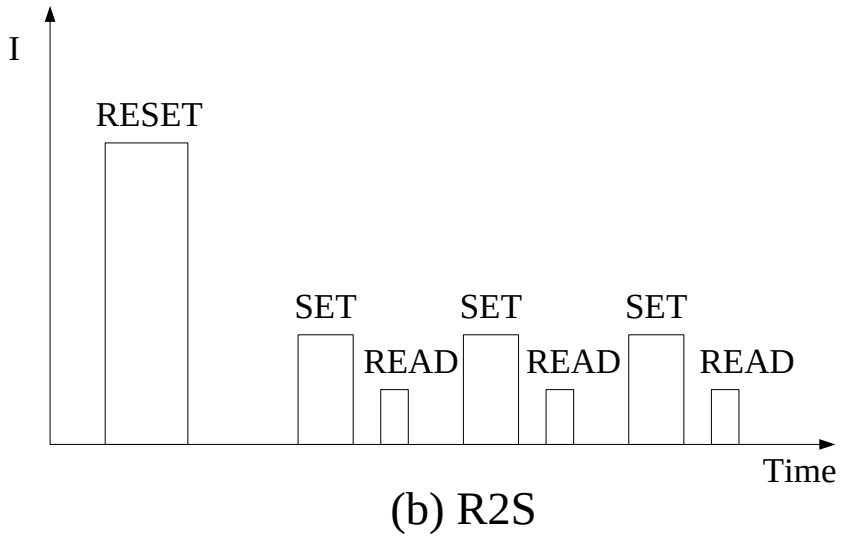
Program-and-verify (P&V), much higher write latency, energy, and write times on indirect cell states
e.g. R01 and R10 for MLC PCM, R2-5 for TLC ReRAM

Outline

Program-and-verify (P&V)



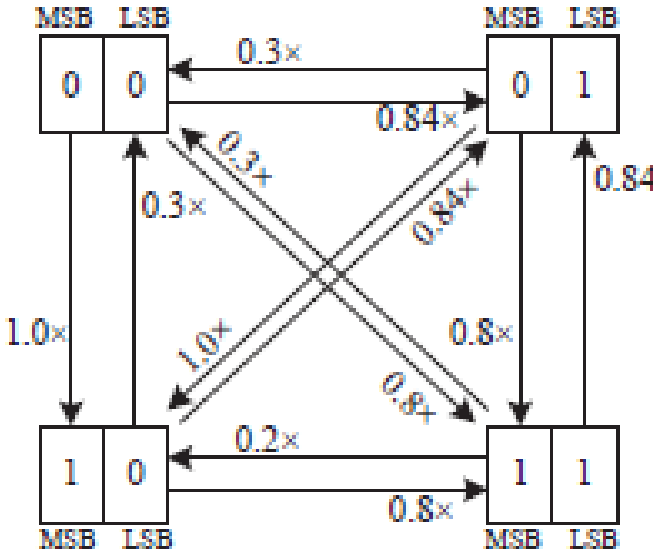
(a) S2R

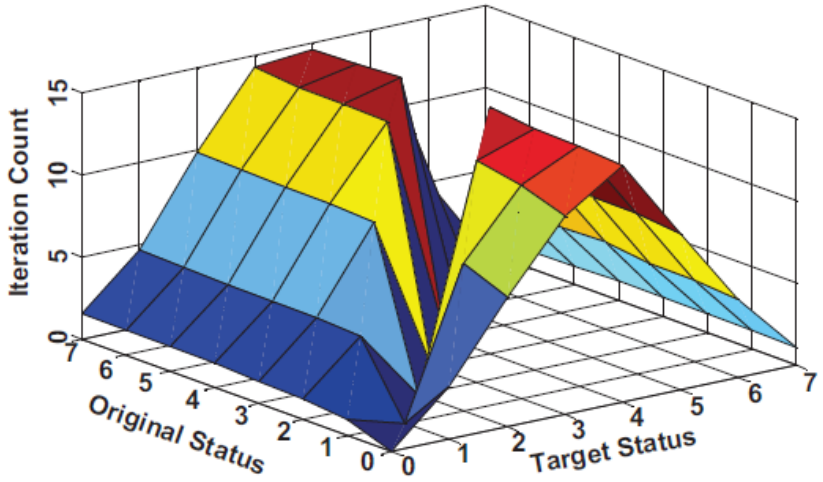


(b) R2S

- Programming of MLC NVM requires precisely writing a cell (P&V)
- The number of iterations is highly dependent on **the original and written data**

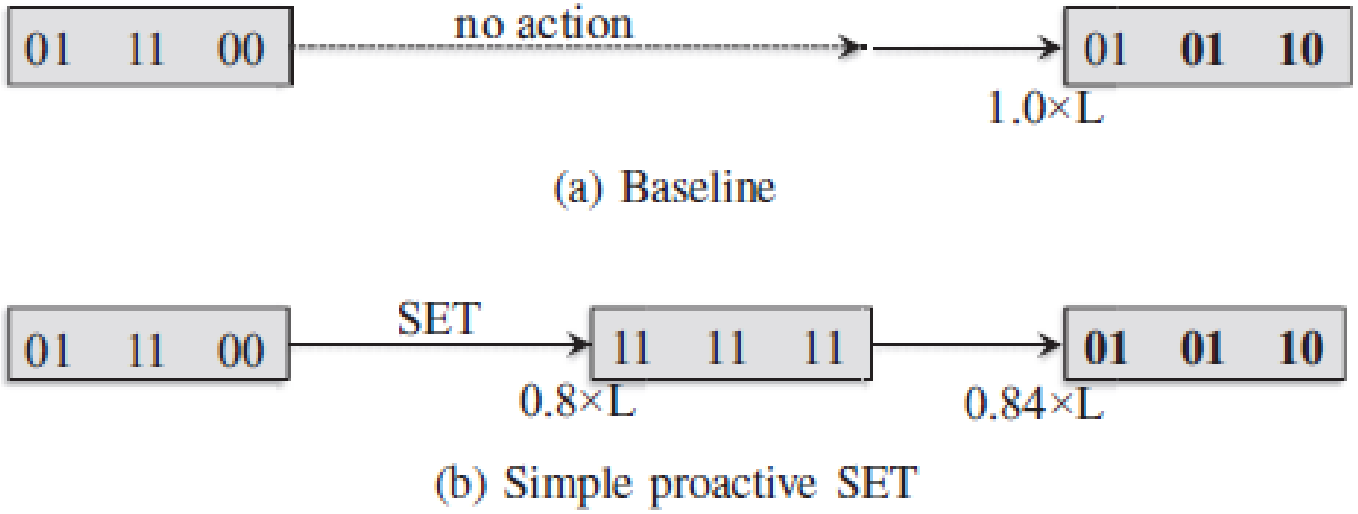
The number of iterations





- MLC NVMs have **asymmetric** in number of P&V iterations
- PreSET [ISCA 12, ICCD 15], WOM (write-once-memory) [DATE 14, ISLPED 13, TC 16] are proposed by taking advantage of the write asymmetric

PreSET

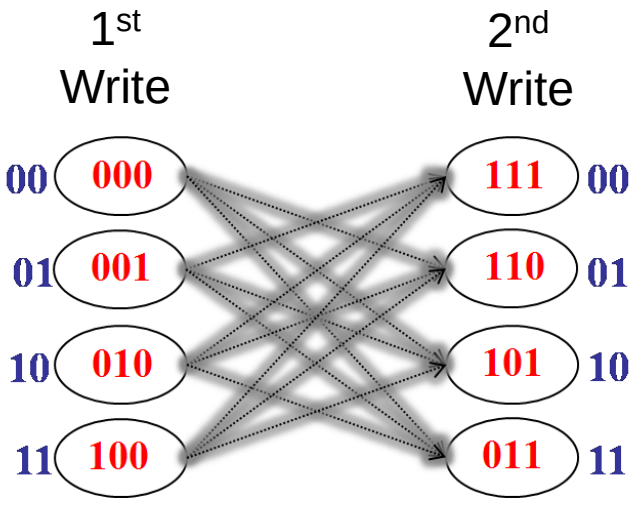


- PreSET exploited this asymmetry by **proactively setting all cells** of a memory line into one full state
- It significantly **increases the write power** due to the proactive write
- It **impairs the lifetime** of the main memory since the writes to memory cells are approximately doubled

Write-Once Memories (WOM)

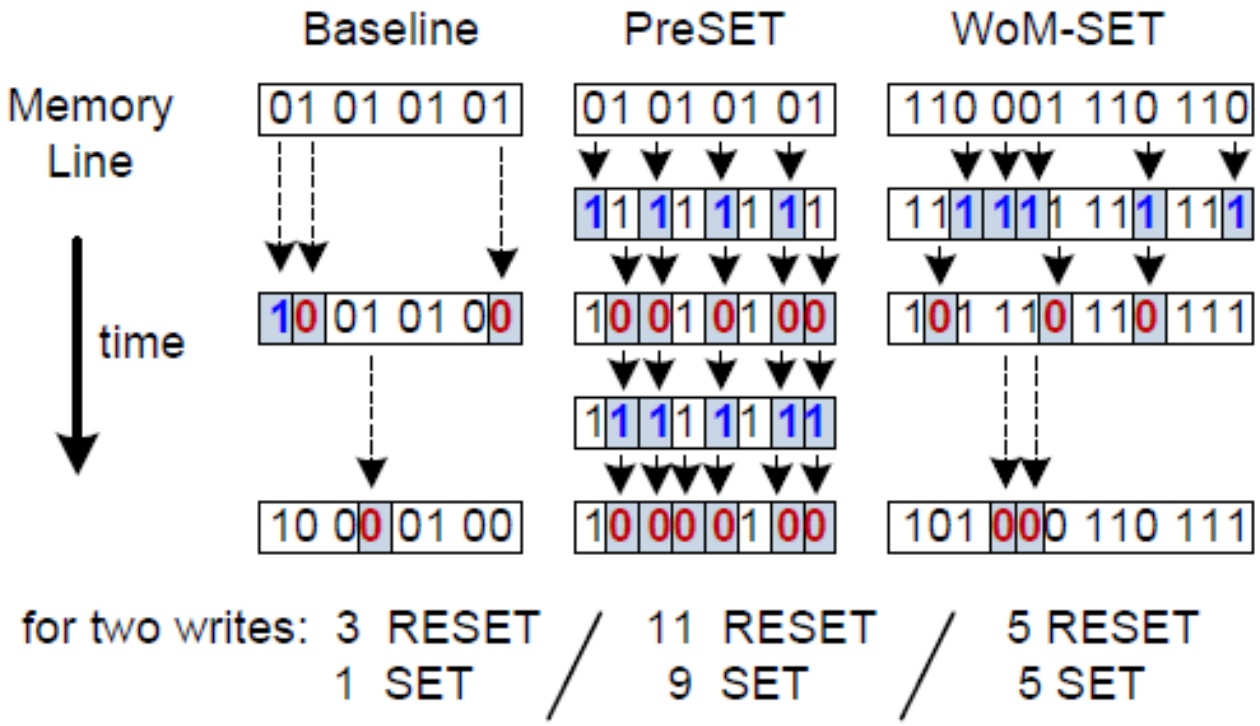
- Introduced by **Rivest and Shamir**, "*How to reuse a write-once memory*", 1982
- The memory elements represent bits (2 levels) and are irreversibly programmed from **0** to **1**

Bits Value	1st Write	2nd Write
00	000	111
01	001	110
10	010	101
11	100	011



- WOM are popular on NVM:
- [MICRO 09] [MSST 10] [ISIT 13] [MSST 14] [ISIT 15] [MSST 15] [FAST 15] [DATE 14] [TC 16]

Write-Once Memories (WOM)



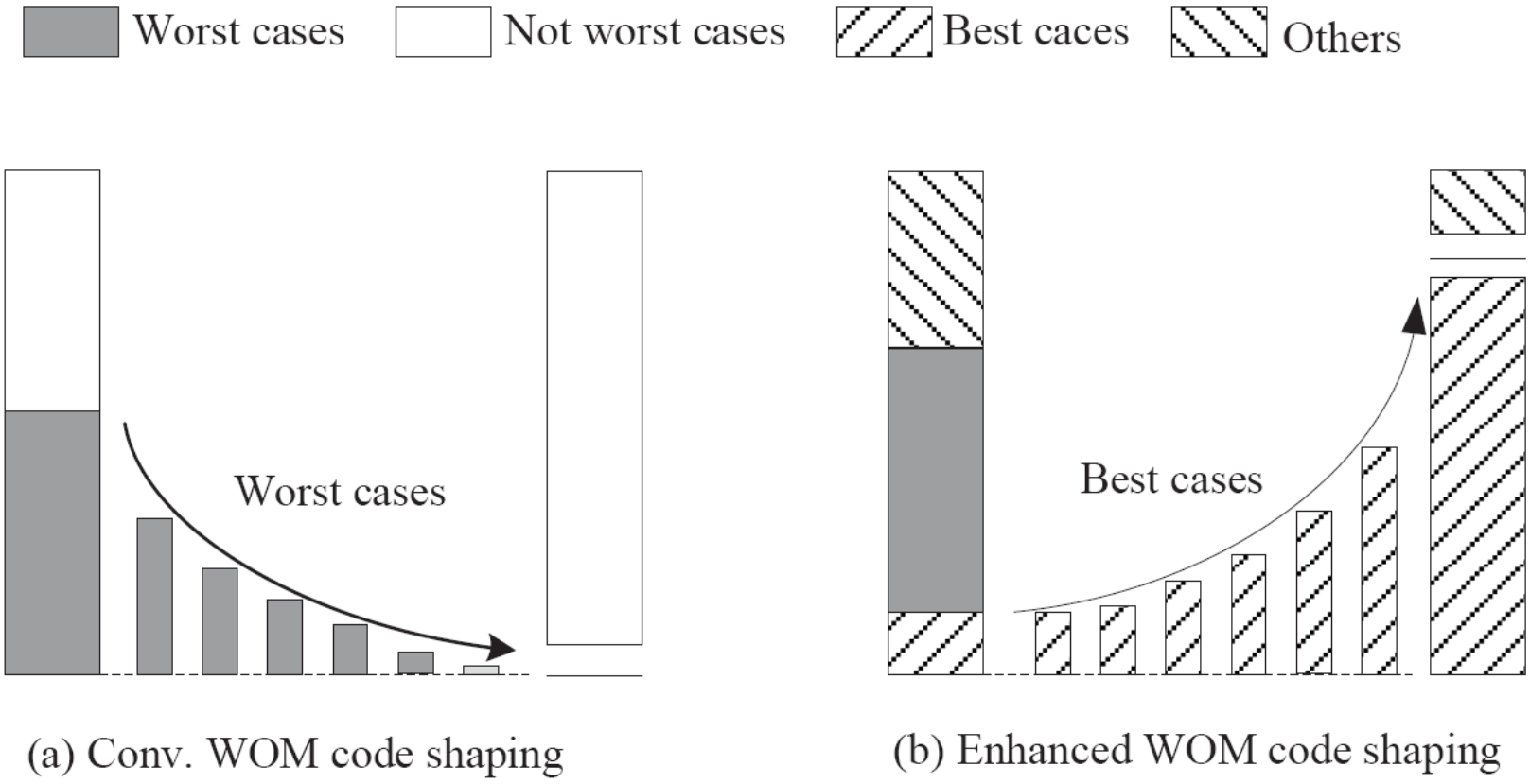
- 50% of extra space
- **Eliminating** the longest (the worst-case) latency writes
- **Without consideration** of the shortest (the best-case) latency writes

Outline

Problem and Goal

- Want to leverage MLC-NVM's strengths
  - Higher density
  - More scalability than existing technologies (DRAM)
- But, also want to mitigate MLC NVM's weaknesses
  - Higher latency/energy, poor endurance
- ***Our goal*** in this work is to propose new optimizations designed to mitigate the weaknesses of MLC NVM

Motivation



(a) Conv. WOM code shaping

(b) Enhanced WOM code shaping

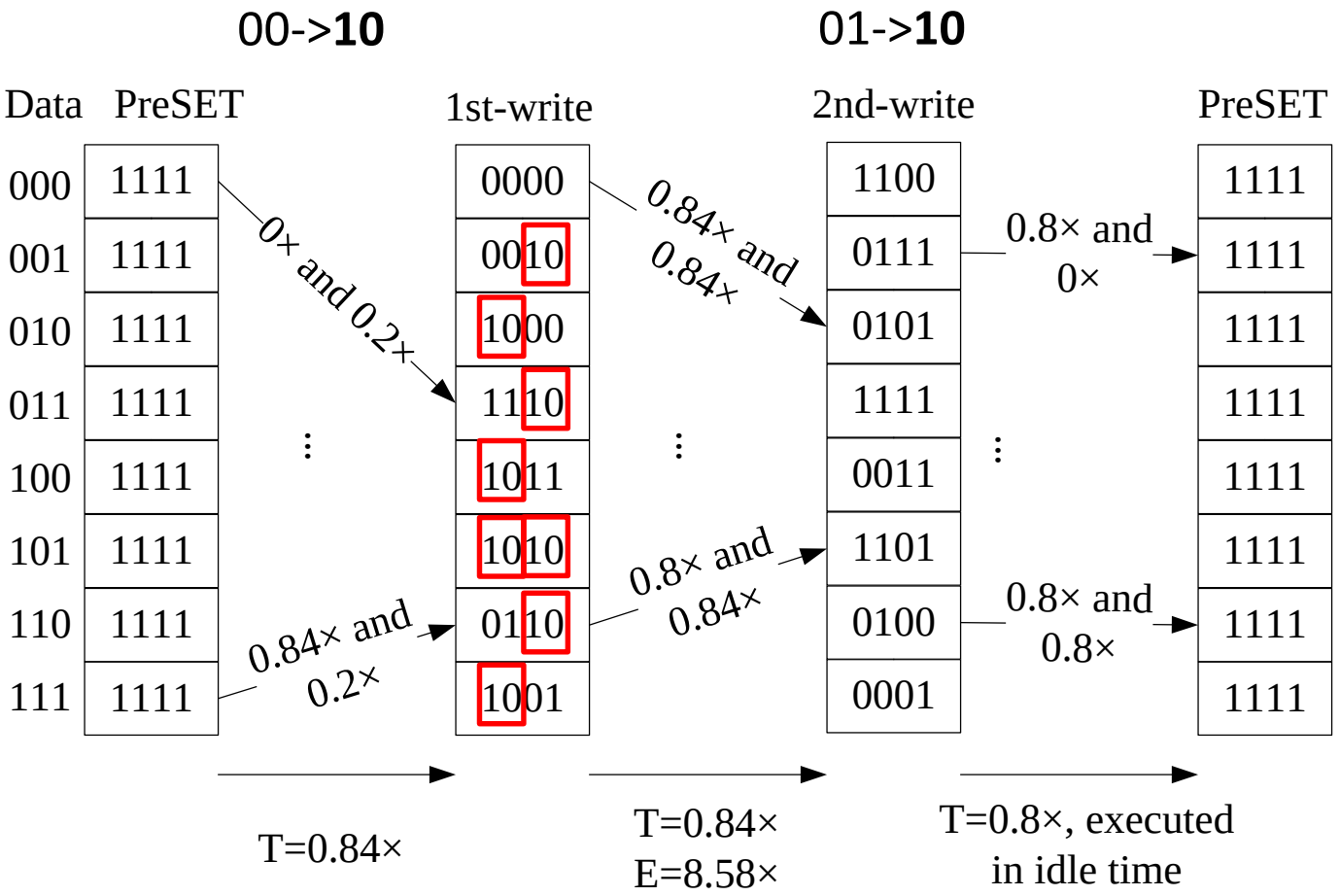
Minimizing the Worst Case

Maximizing the Best Case

Outline

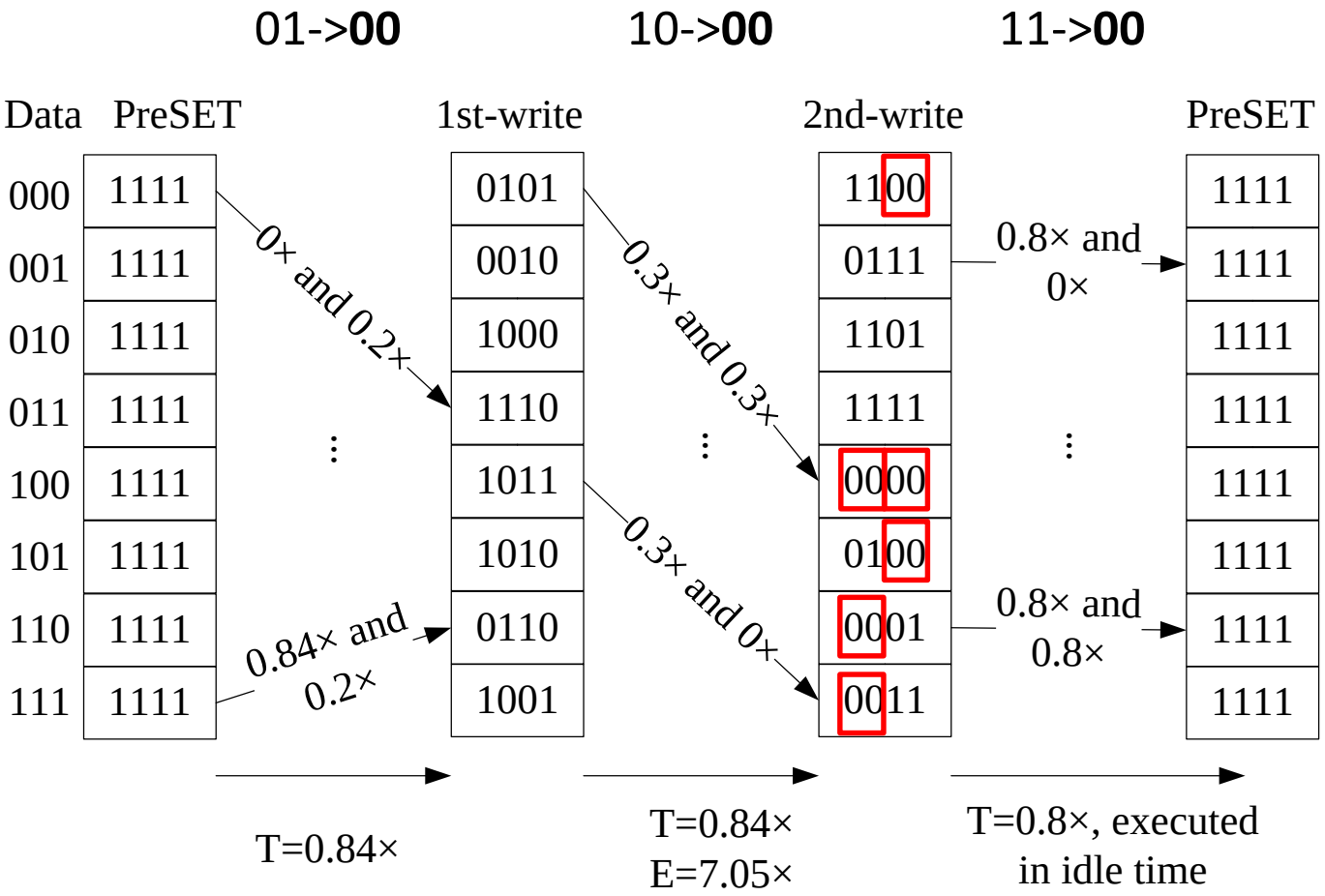
WOM for MLC NVM

- Extend WOM on minimizing the worst case for MLC NVM

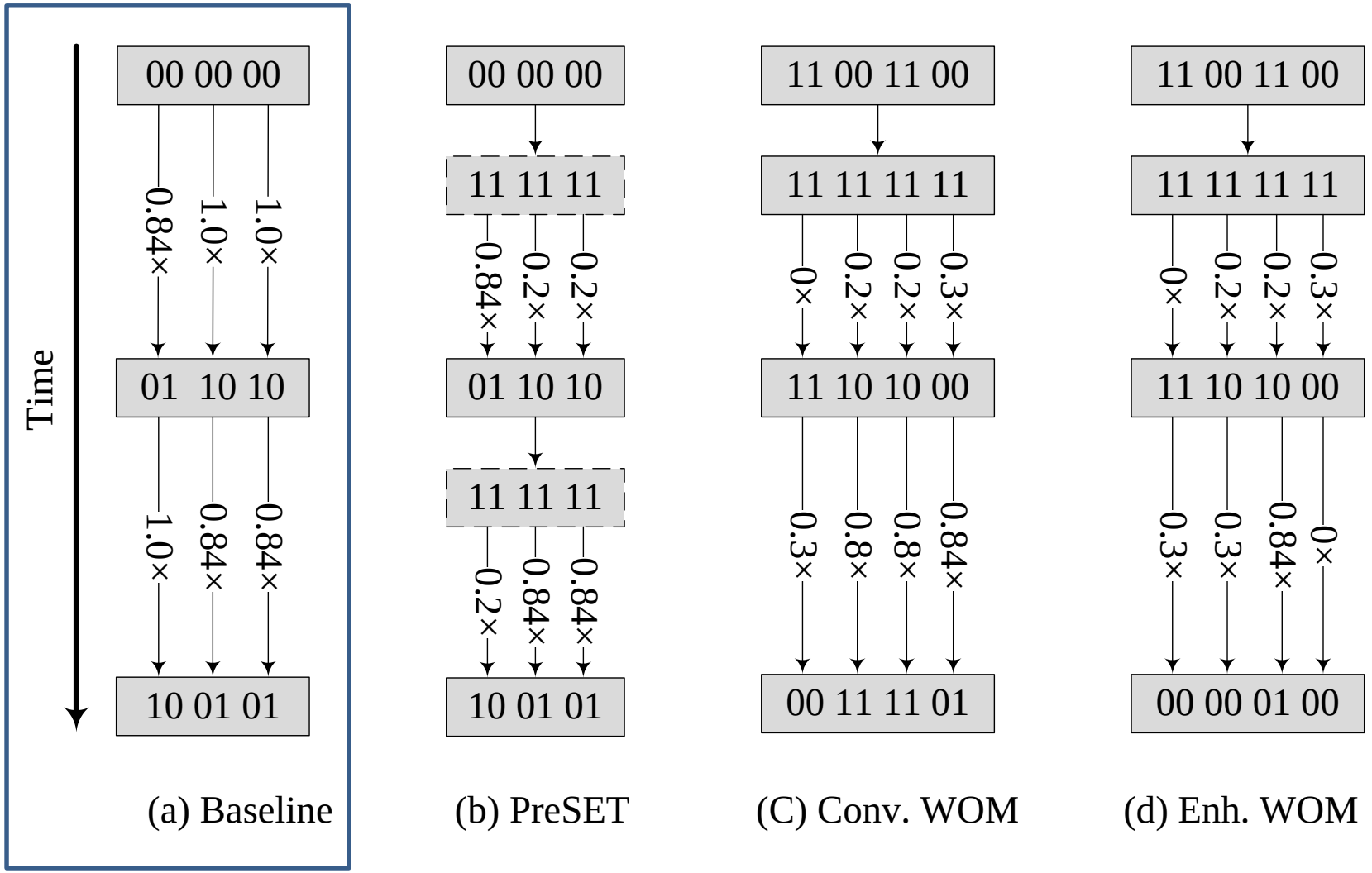


Enhanced WOM

- Enhanced WOM is designed for maximizing the best case



Example


Time
00 00 00
0.84×
1.0×
1.0×
01 10 10
1.0×
0.84×
0.84×
10 01 01
(a) Baseline
00 00 00
11 11 11
0.84×
0.2×
0.2×
01 10 10
11 11 11
0.2×
0.84×
0.84×
10 01 01
(b) PreSET
11 00 11 00
11 11 11 11
0×
0.2×
0.2×
0.3×
11 10 10 00
0.3×
0.8×
0.8×
0.84×
00 11 11 01
(C) Conv. WOM
11 00 11 00
11 11 11 11
0×
0.2×
0.2×
0.3×
11 10 10 00
0.3×
0.3×
0.84×
0×
00 00 01 00
(d) Enh. WOM

Outline

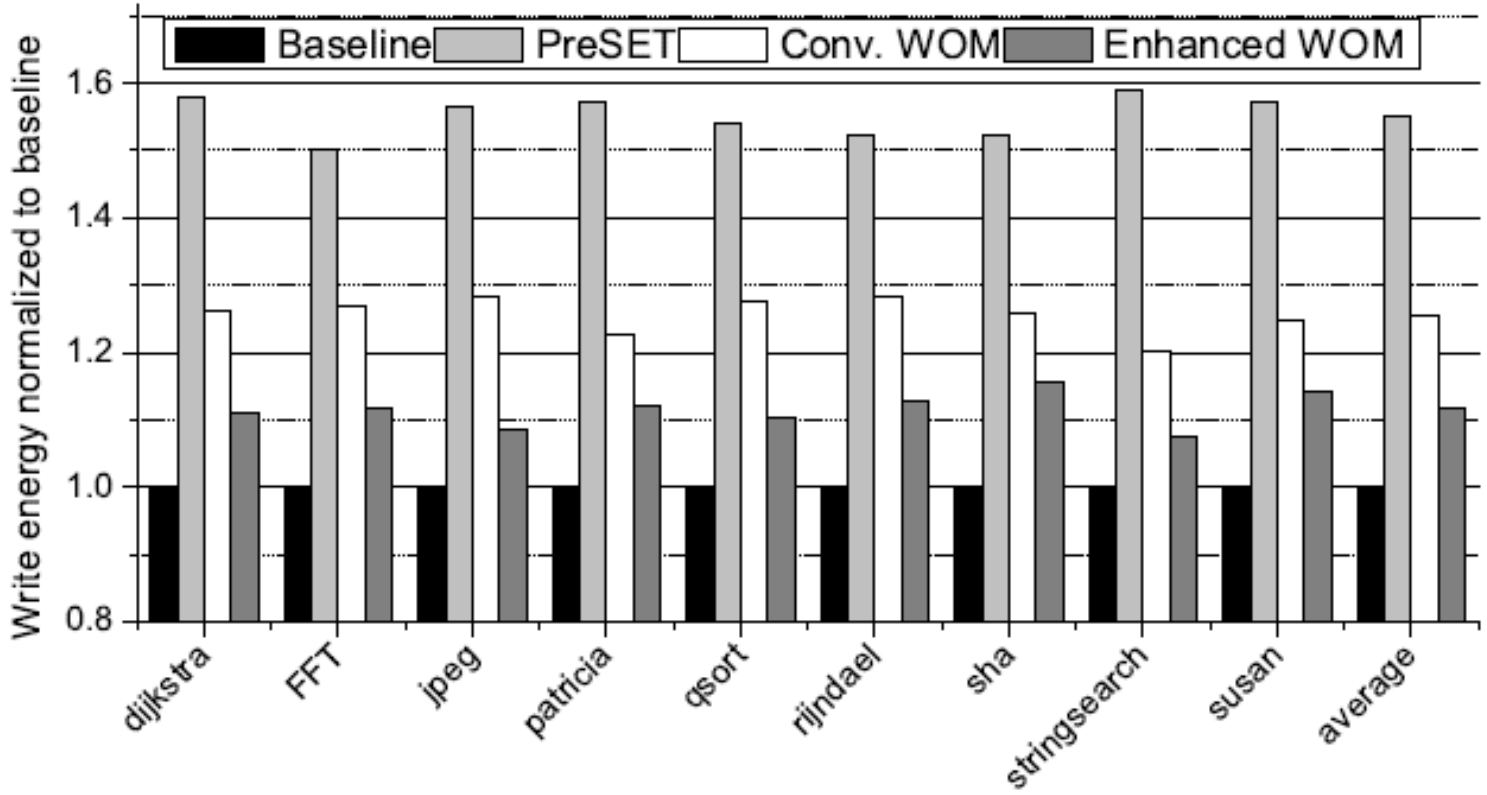
Experimental Setup

- Experimental setup
  - Simulation platform: Gem5
- Benchmarks:
  - Mibench
- PCM configurations:

TABLE IV
PCM CHIP CONFIGURATIONS [14, 10].

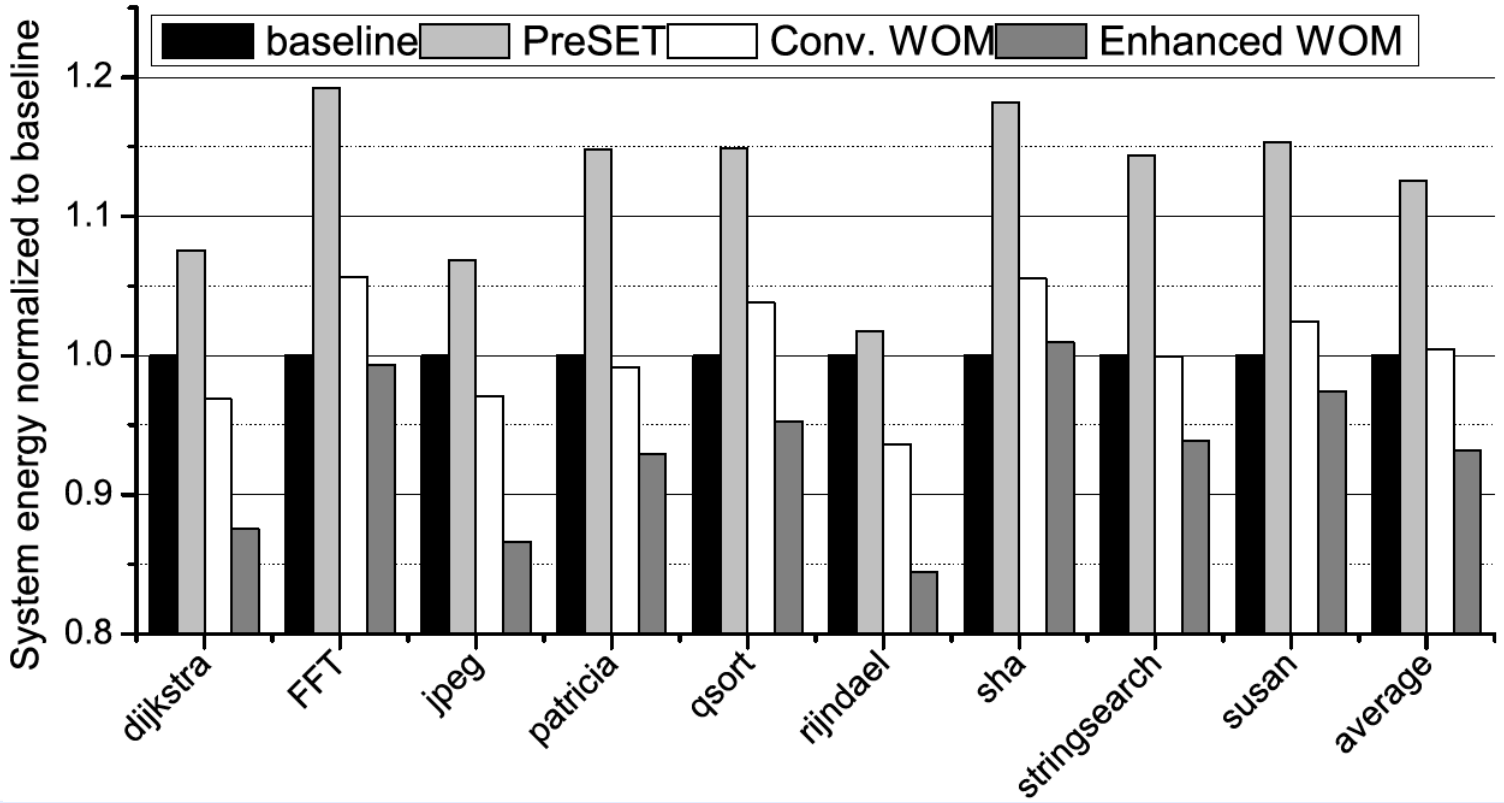
STR		RTS	
SET current	150 uA	RESET current	250 uA
SET timing	150 ns	RESET timing	75 ns
RESET current for '10', '01','00'	200 uA 225 uA 250 uA	SET timing for '01', '10','11' $(pulse \times pulse)$	15*9=135 ns 15*13=195 ns 15*23=345 ns
RESET timing	50 ns	SET current	150 uA

Write Energy on Main Memory



It is note that the enhanced WOM still incurs write energy overhead compared to the baseline, 11.6% on average

Overall Energy



The results show that the write energy overhead of main memory is partly covered by the performance improvement

Performance Results

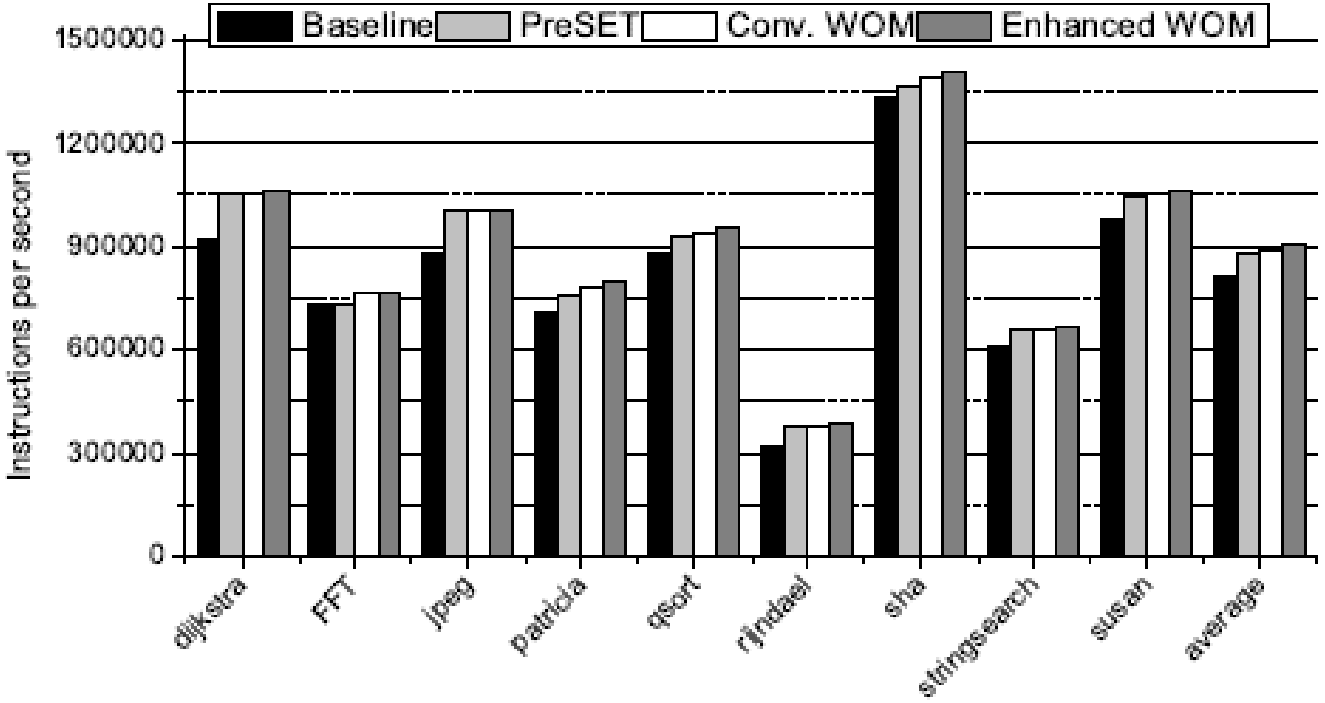


Fig. 10. Instructions per second.

Compared to the baseline, PreSET, conventional WOM code, and enhanced WOM code speedup the system performance by 7.8%, 9.8%, and 10.3%, respectively

Outline

Conclusion

- The adoption of MLC NVMs is limited by their high programming energy and latency as well as the low endurance
- We propose an enhanced WOM code for MLC NVMs, w hich exploits the asymmetric characteristic in MLCNVM c ell state transitions
- Unlike the conventional WOM codes that focus on elimin ating the worst-case latency writes, we propose to enlarg e the best-case latency writes in MLC NVM cell state tra nsitions

ASIA SOUTH PACIFIC DESIGN AUTOMATION CONFERENCE

Energy, Latency, and Lifetime Improvements in MLC NVM with Enhanced WOM Code

Huizhang Luo[1], **Liang Shi[1*]**, Qiao Li[1],

Chun Jason Xue[2] and Edwin H.-M. Sha[1]

[1]Chongqing University
[2]City University of Hong Kong

重慶大學 CHONGQING UNIVERSITY

香港城市大學 City University of Hong Kong