ASPDAC 2016

Peak-to-average Pumping Efficiency Improvement for Charge Pump in Phase Change Memories

Huizhang Luo, Jingtong Hu, Liang Shi, Chun Jason Xue and Qingfeng Zhuge

CHONGQING UNIVERSITY

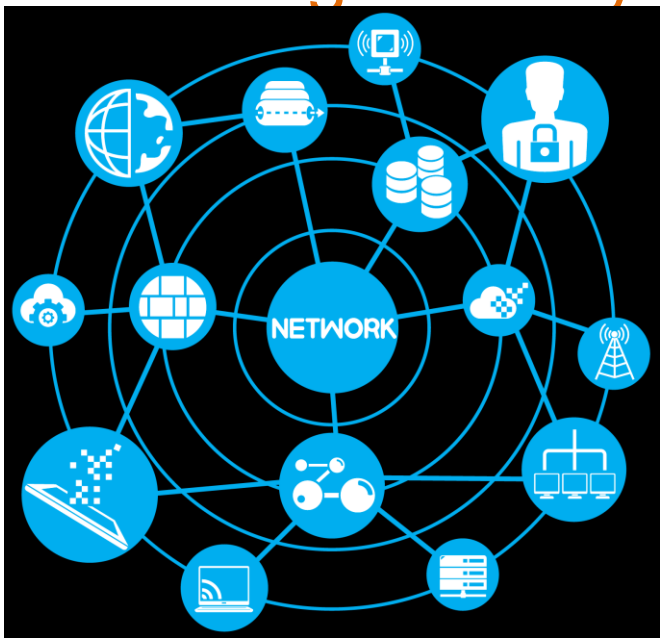
Oklahoma State University

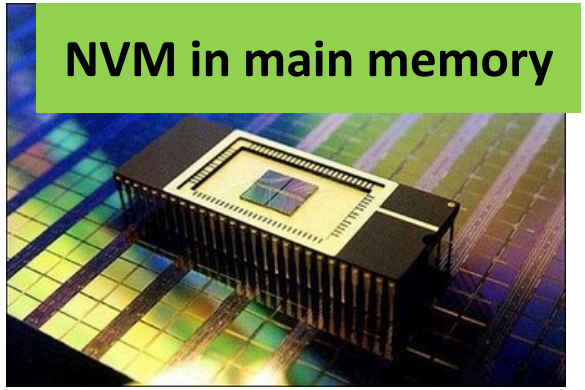
香港城市大學 City University of Hong Kong

Outline

Introduction

- Increasing memory capacity requirement





PCM is one of such NVM

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

PCM in Future [ITRS 2013]

Technology	Cell size (SLC)	Read Latency	Write Latency	Write Energy
DRAM	6 F^2	< 10 ns	< 10 ns	2 pJ/bit
PCM	4 F^2	< 10 ns	< 50 ns	29.7 pJ/bit

➤ Similar access latency with DRAM

Thus, PCM has similar write parallelism with DRAM.

➤ High write energy

Energy consumption of PCM remains a key problem.

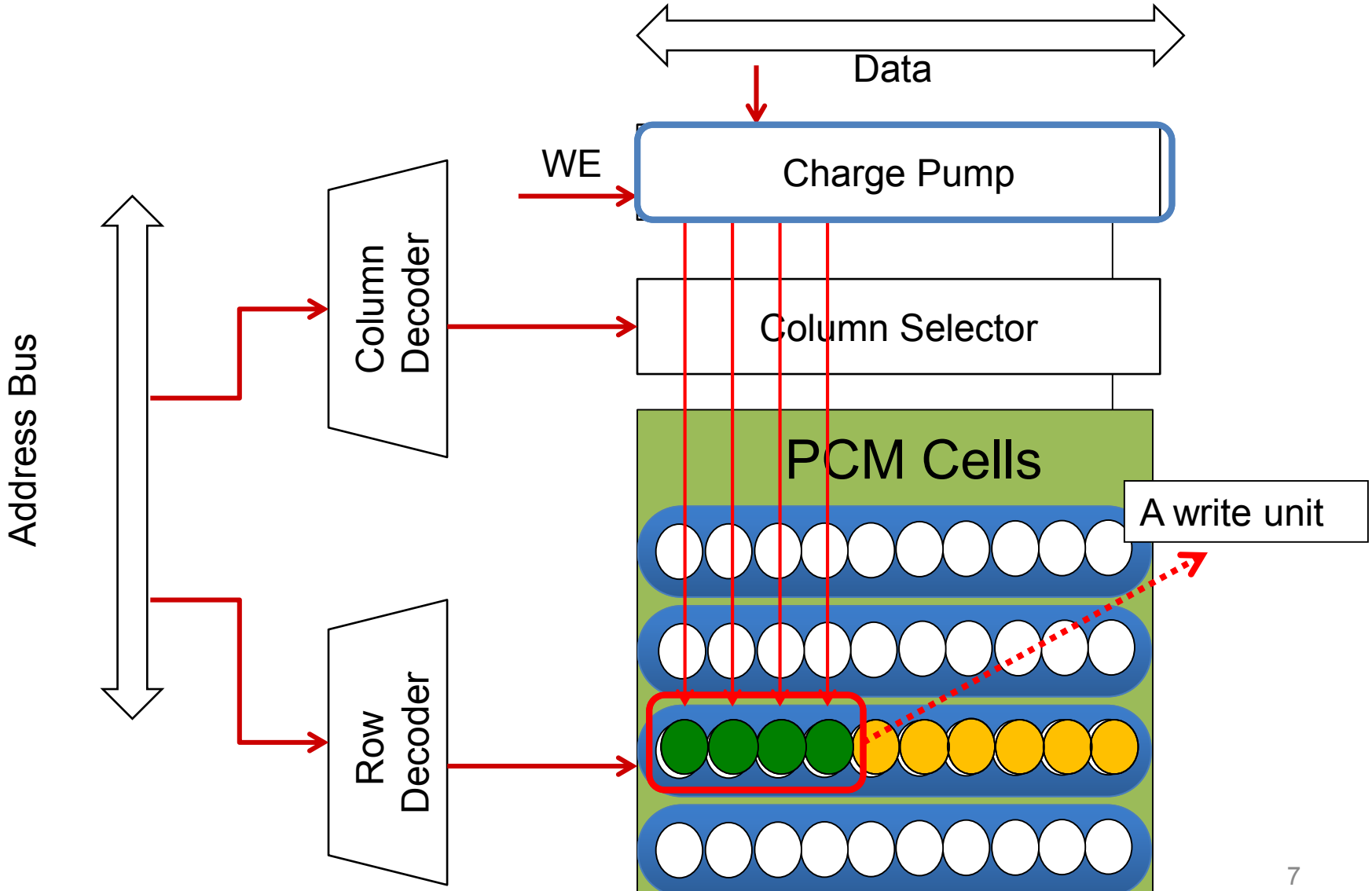
High Write Energy

- Much higher write voltage and current
  - RESET: 5V, 100uA,
  - SET: 3V, 50uA,
  - V_{dd} 1.5V for DRAM, 22uA
- Energy loss in the charge pumps (CPs)
  - Parasitic power, leakage power
  - Lead to low pumping efficiency
- Key idea
  - We improve the pumping efficiency to reduce energy consumption of PCM.

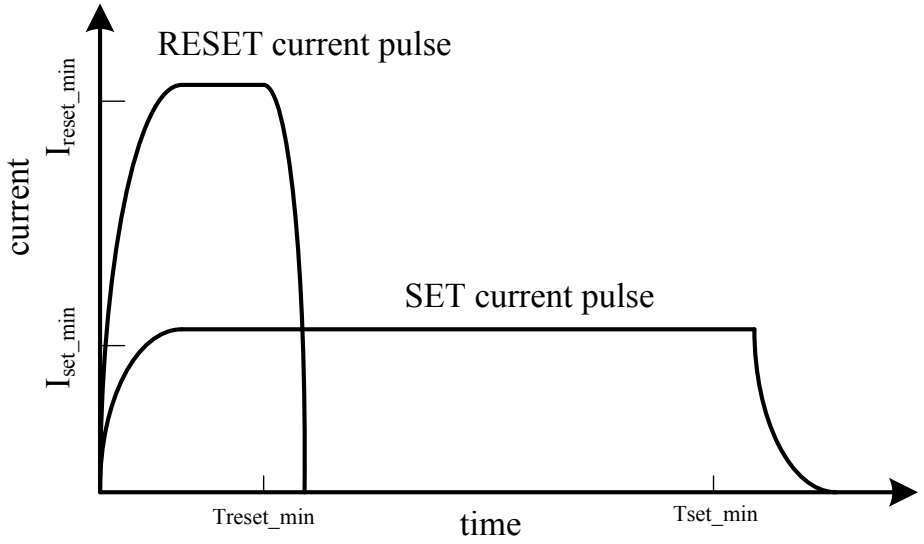
Outline

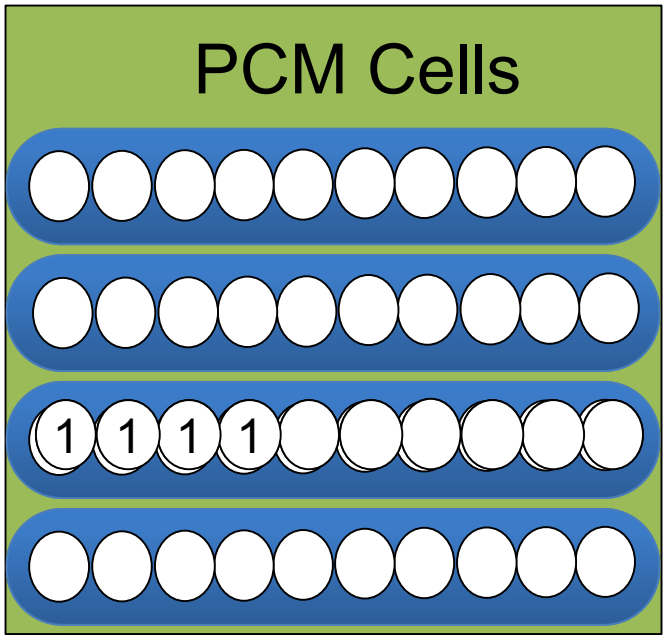
1. Introduction
2. Background
3. Problem Definition
4. Peak To Average Write Scheme
5. Experiment
6. Conclusion

Write Process in PCM



Asymmetric Write in PCM



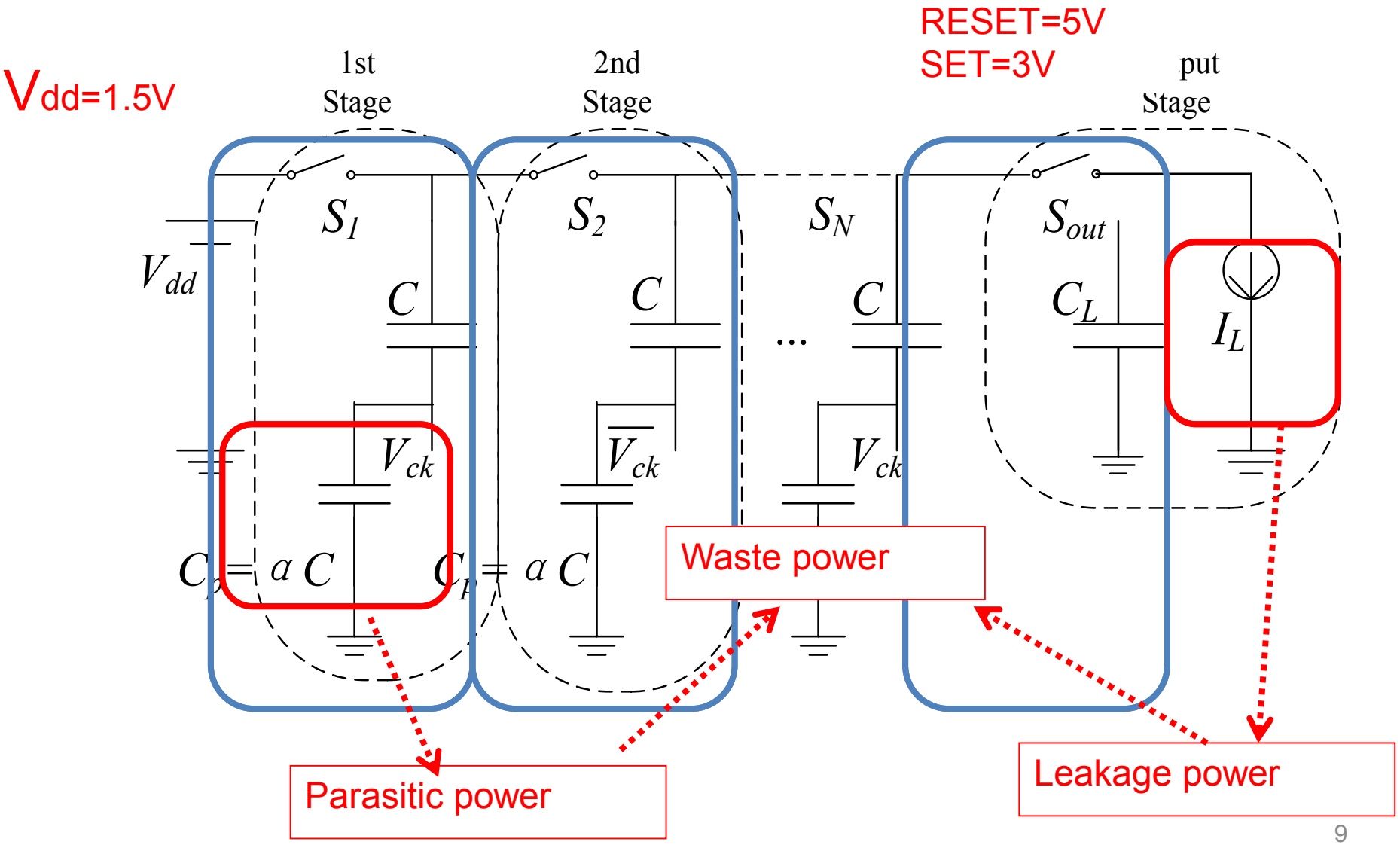


Writing a ‘0’ requires much higher current than writing a ‘1’.

Write request=‘0000’, write current requirement is high.

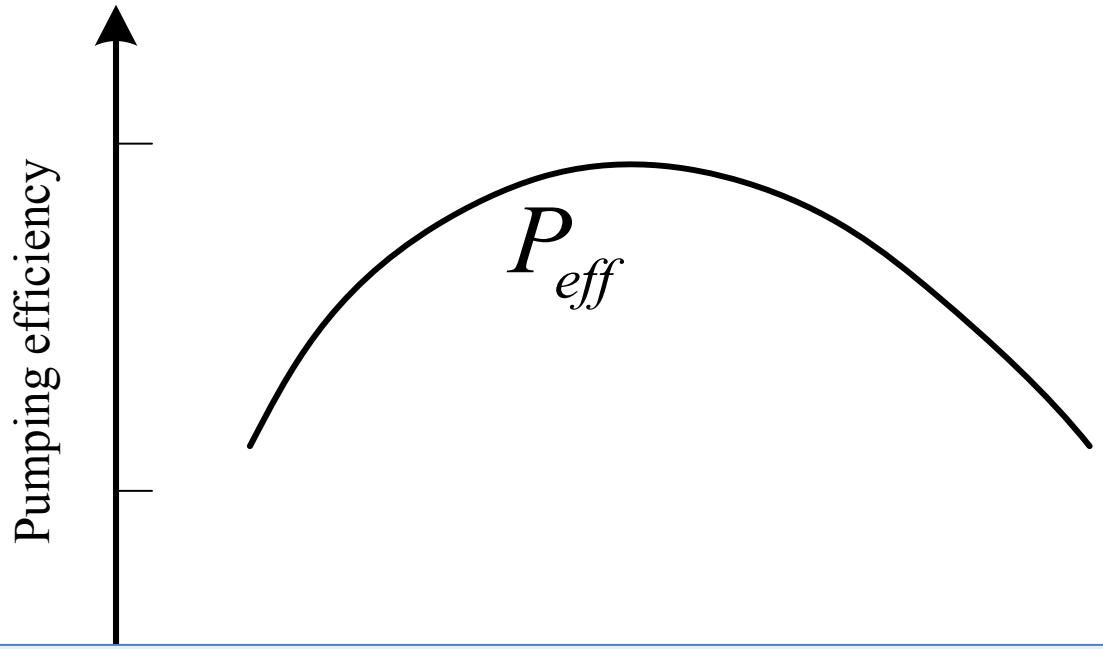
Write request=‘1111’, write current requirement is low.

PCM Charge Pump



Pumping Efficiency [ISCA 14]

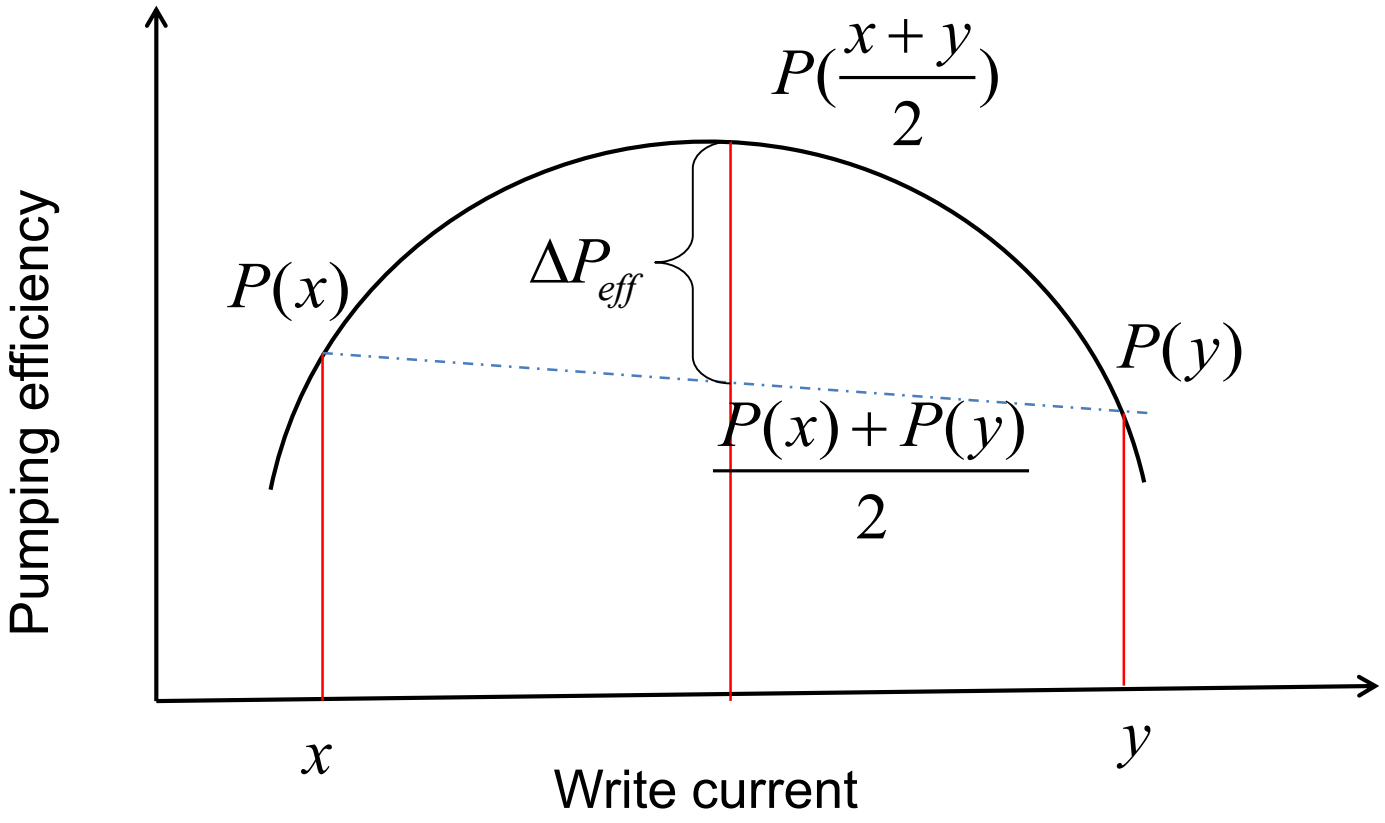
- Waste power leads to low pumping efficiency.



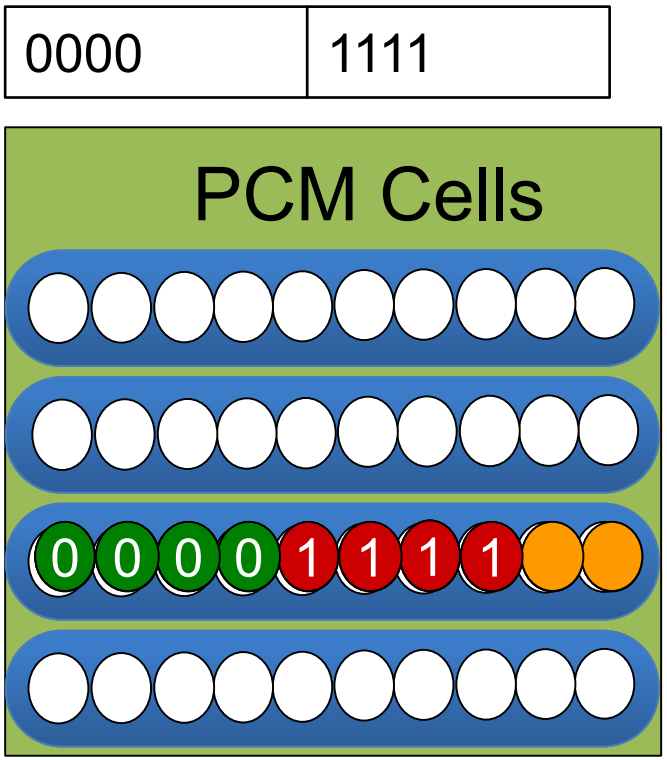
The pumping efficiency is a concave function of the write current [10].

Characteristic of a Concave Function

$$\forall_{x,y}\, 2f\left(\frac{x+y}{2}\right) \geq f(x)+f(y)$$

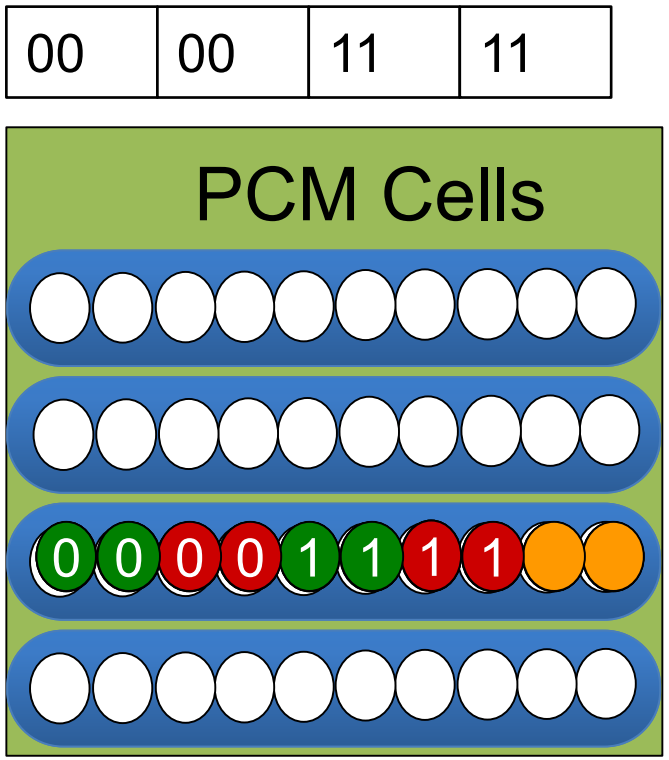


Key Idea



I1=400uA I2=200uA

Conventional scheme



I1=300uA I2=300uA

Uniform write current scheme

Outline

1. Introduction
2. Background
3. Problem Definition
4. Peak To Average Write Scheme
5. Experiment
6. Conclusion

Problem Definition

- The property can be extended to *num serial write current* requirements.

$$f\left(\frac{I_1 + I_2 + \ldots + I_{num}}{num}\right) \geq \frac{f(I_1) + f(I_2) + \ldots + f(I_{num})}{num}$$

Overall pumping efficiency can be improved if the current requirements of the *num* serial write units are uniform.

- Average write current:

$$I_{avg} = \frac{\sum_{i=1}^{num} I_i}{num}$$

- Write variation:

$$WV = \frac{1}{I_{avg}} \sqrt{\frac{\sum_{i=1}^{num} \left(I_i - I_{avg}\right)^2}{num - 1}}$$

Motivation

- The write variation distribution

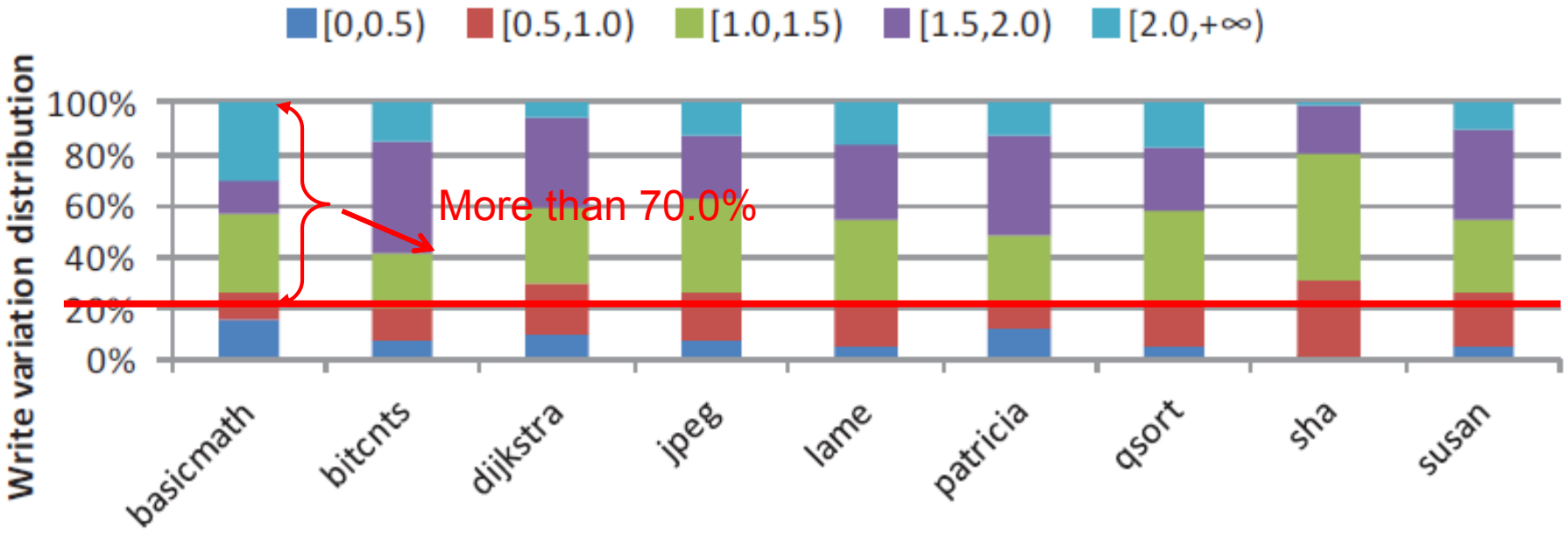


Fig. 5. The write variation distributions for different workloads.

➤ The write variations are high. This motivates us to propose a write scheme to reduce write variation. Such that the pumping efficiency will be improved for energy saving.

Outline

1. Introduction
2. Background
3. Problem Definition
4. Peak To Average Write Scheme
5. Experiment
6. Conclusion

Peak To Average Write Scheme (PTA)

- Key idea
  - When the write units are queued in the memory controller, each write unit is divided into several sub-units.
  - The sub-units are then regrouped as new write units before they are sent to the PCM chip.
- Two challenges
  - 1.how to obtain a minimal write variation during regrouping?
  - 2.What modifications of PCM chips and controller should be made?

Challenge 1

- The problem:
  - Assume that the current requirements of each sub-write unit I_{ij} are known. The problem is to regroup the write units such that the overall write variation is minimized.
  - This problem is NP-hard, which can be proved by reducing from the Subset Sum Problem.
- Subset Sum Problem :
  - Given a set of non-negative integers, and a value *sum*, find the subset of the given set with sum equal to the given *sum*.

IP Formulation (Off-line)

- Objective function
  - The total write variation is the minimal.

$$\min \frac{1}{I_{avg}} \sqrt{\frac{\sum_{i=1}^{num}\left(\sum_{j=1}^{M} X_{i,j}\cdot C_j - I_{avg}\right)^2}{num-1}}$$

- Constraints
  - First, each sub-write must be in one and only one write group.

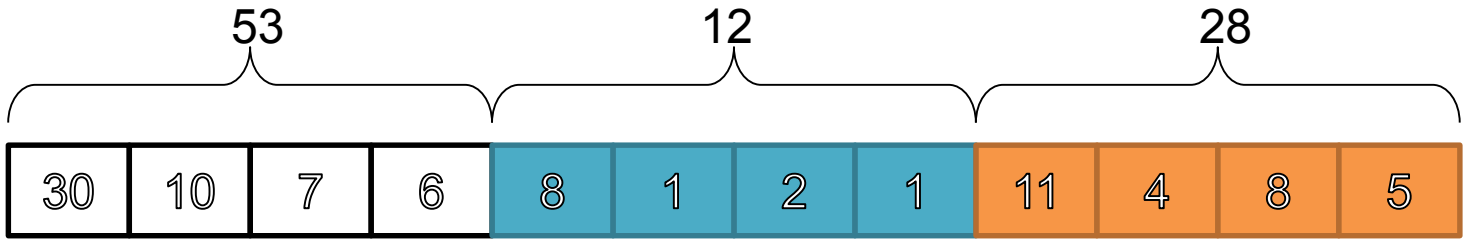
$$\forall_j \sum_{i=1}^{num} X_{i,j} = 1$$

  - Second, each group has *sub* sub-write units.

$$\forall_i \sum_{j=1}^{num\cdot sub} X_{i,j} = sub$$

The time complexity of the IP formulation is exponential.

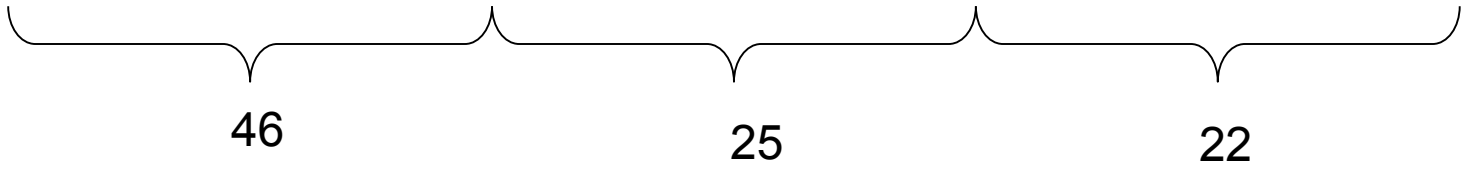
Partition Strategy (PS, Online)



Compared with θ $\theta = 7.75$

After partition:

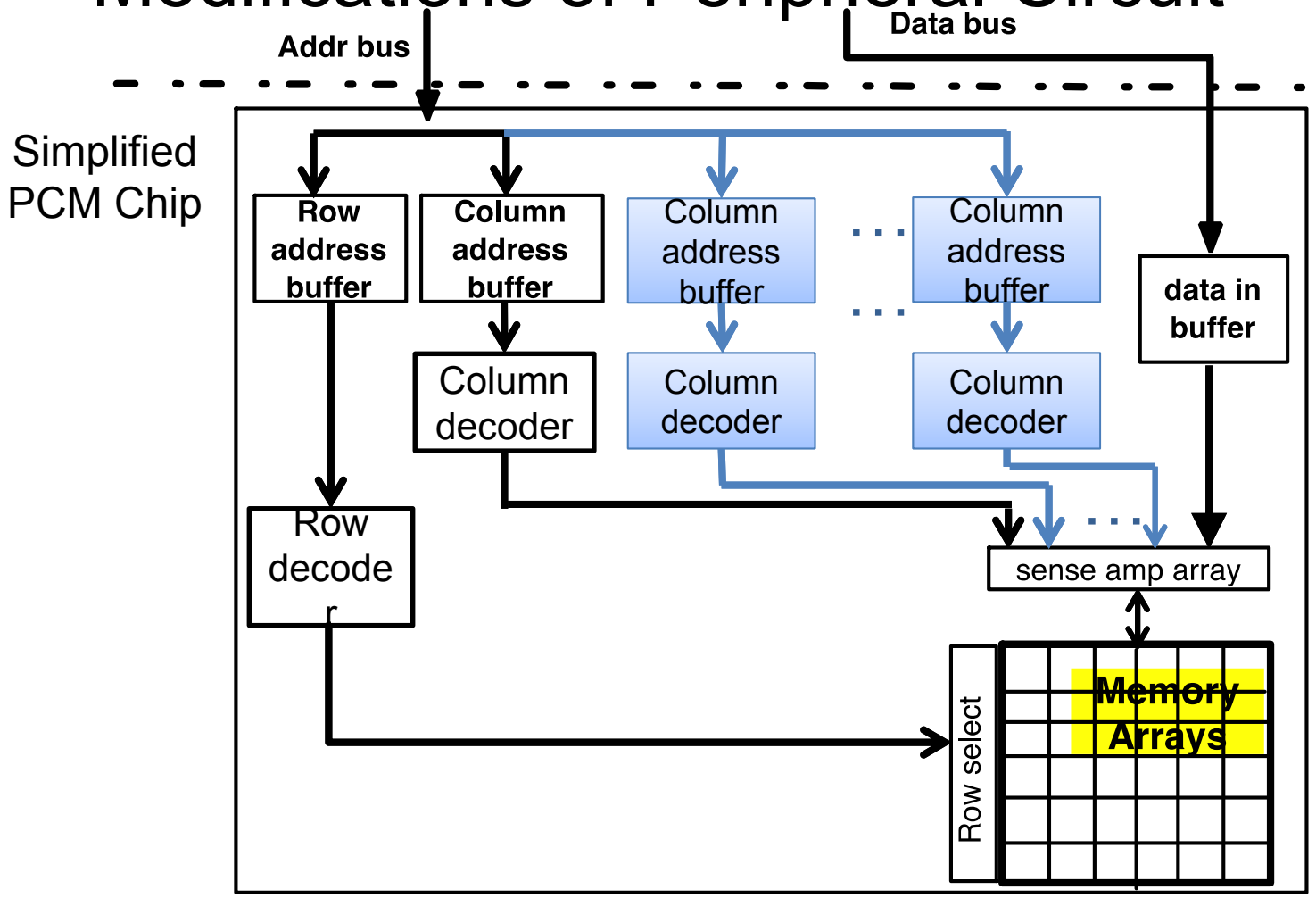
Regrouped:



Time complexity: o(n), space overhead : 1 array.

Challenge 2

- Modifications of Peripheral Circuit



Outline

1. Introduction
2. Background
3. Problem Definition
4. Peak To Average Write Scheme
5. Experiment
6. Conclusion

Experimental Setup

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

Chip	1.8V V_{dd}, 64 concurrent RESET power buget
Charge pump	working frequent: 133MHZ RESET/SET/READ working voltages: 5/3/3V
READ	3V, 8.4 μA, 5.6nJ per line
Write	RESET: 5V, 100 μA, 29.7pJ per bit, 50ns operation latency SET: 3V, 50 μA, 22.5pJ per bit, 150ns operation latency
PCM write unit	size: 64bit, divided into 4 sub-write units

Write Variations

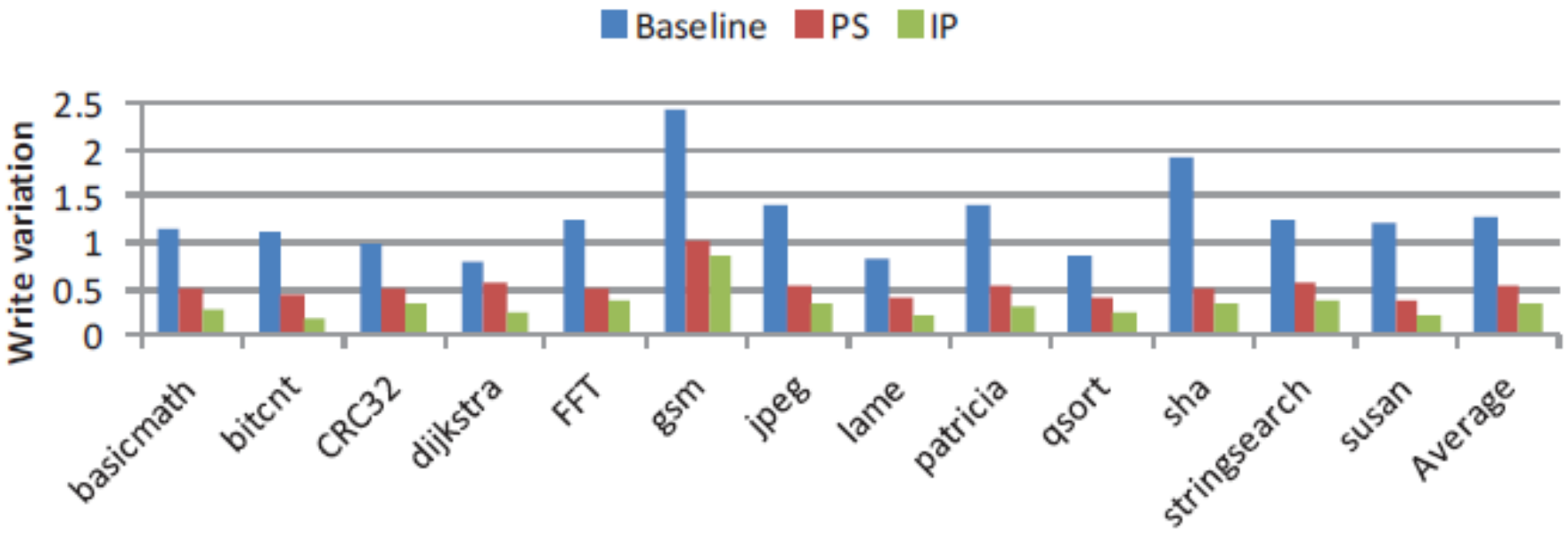


Fig. 7. Improvement of write variations.

The write variation is reduced from 1.28 to 0.52 and 0.32 for PS and IP, respectively.

Pumping Efficiency

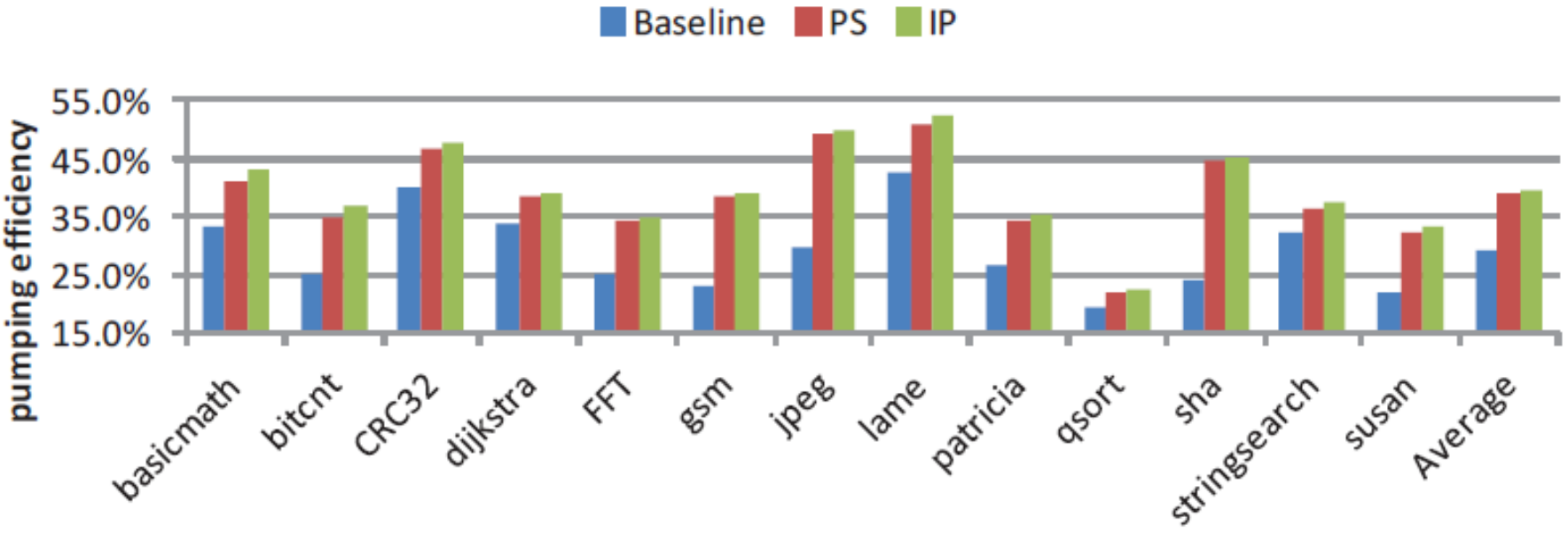


Fig. 8. Improvement of pumping efficiencies under various benchmarks.

PS achieves 38.8% on average while IP achieves 39.8% pumping efficiencies.

Energy Saving

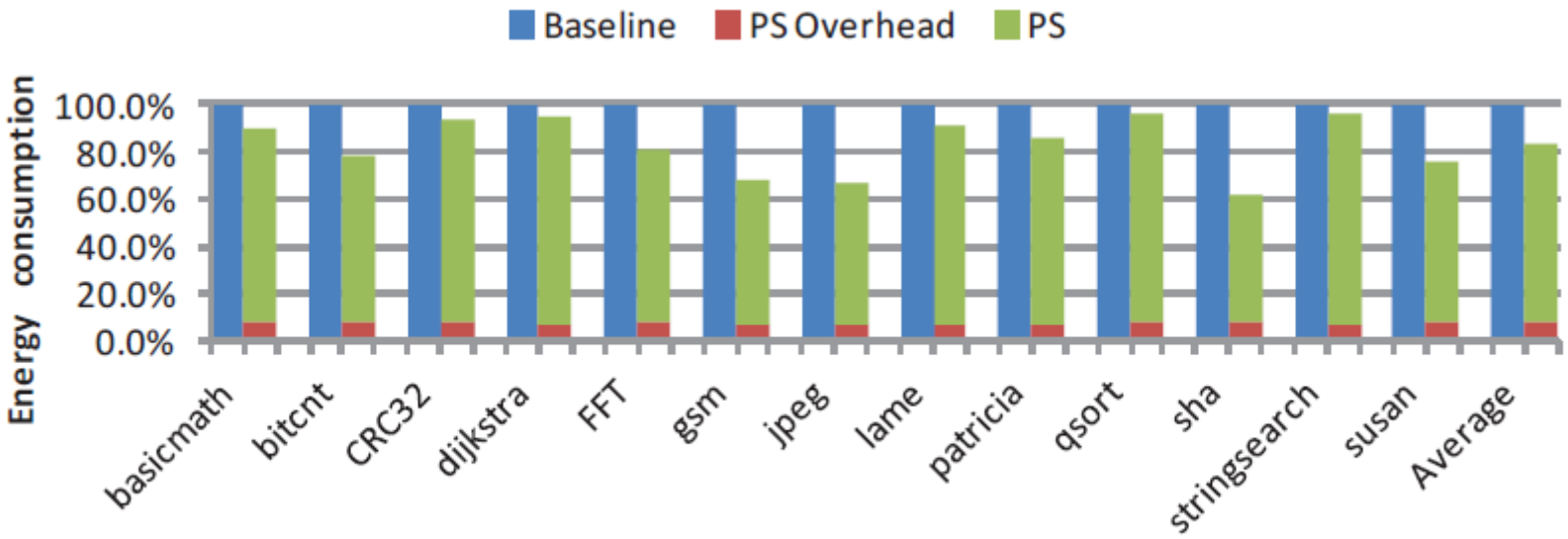


Fig. 9. Energy saving of memory system.

PS achieved 17.0% energy saving compared to the baseline.

Outline

1. Introduction
2. Background
3. Problem Definition
4. Peak To Average Write Scheme
5. Experiment
6. Conclusion

Conclusion

- We have proposed a write scheme, peak-to-average (PTA), to improve the pumping efficiencies by regrou ping the sub-write units.

- An off-line optimal Integer Programming (IP) formula tion and an online strategy, PS, are proposed to mak e the regrouping decision.

- The experimental results show that the pumping effi ciency is improved from 29.8% to 38.8% by the prop osed PS strategy.

Thank you!