Phase-change RAM (PRAM)-based Main Memory

Sungjoo Yoo
April 19, 2011
Embedded System Architecture Lab.
POSTECH

sungjoo.yoo@gmail.com

Agenda

- Introduction
- Current status
- Hybrid PRAM/DRAM main memory
- Next challenge: MLC PRAM
- Summary

PRAM: Key Questions

- PRAM Manufacturers
  - Cost: 2x~4x cheaper than DRAM?
  - Performance: 5x~10x larger latency and power is OK?
- Set makers
  - How much gain in which metrics? Standby power? Cost? Performance? or else?
- Hardware and software designers
  - How to maximize lifetime?
    - Minimizing writes (bit updates) and wear leveling
  - How to mitigate the performance loss due to large latency?
  - How to maximize the benefits of low standby power?

PRAM Benefits for End Products: Power

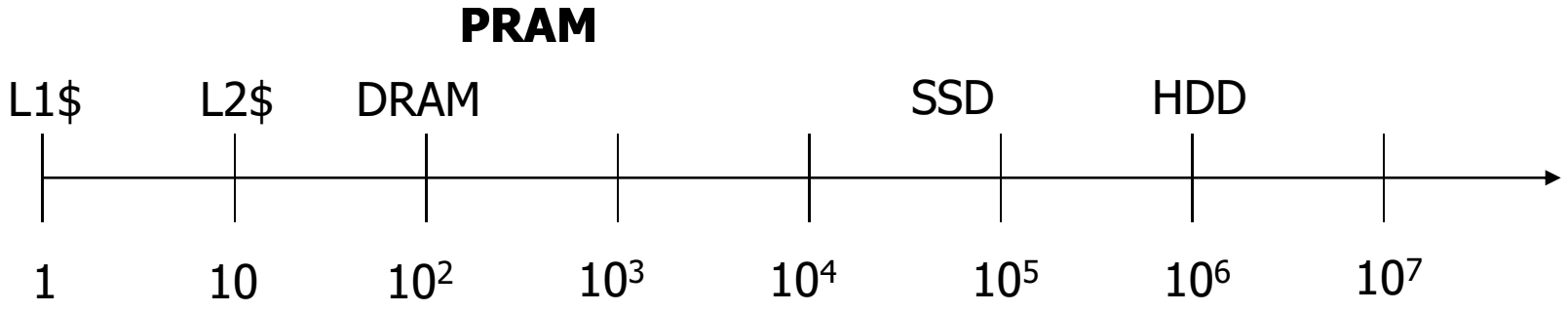
- Smart phone
  - Large DRAM refresh energy for 'Instant on' feature
- Server
  - Main memory (active and refresh) can occupy up to 40% of power consumption in servers
    - E.g., memcached requires very large main memory
- Prospect: DRAM scaling trend
  - DRAM scaling down to 10nm, maybe yes after all, however, with high overhead in refresh energy

PRAM Benefits for End Products: Performance and Lifetime

- Smart phone
  - Fast boot up by pinning OS on PRAM in main memory or storage
- Server
  - Fast checkpointing
    - Checkpoint overhead accounts for over 25% of total application execution time in a petaFLOPS system [Grider07]
- Storage
  - PRAM/NAND in SSD for longer lifetime and write performance
    - PRAM for log buffer, meta data, boot up data, etc.
    - 3x improvement in lifetime and write performance [Xie10]

Where to Put PRAM?

- Main memory scenario
  - DRAM and PRAM can be utilized together to compensate for each other's limitations
- Storage scenario
  - Further improvement in SSD latency and lifetime
  - Simplify power-on reset function
- Combined scenario
  - At both main memory and SSD



*processor clock cycles @ 1GHz

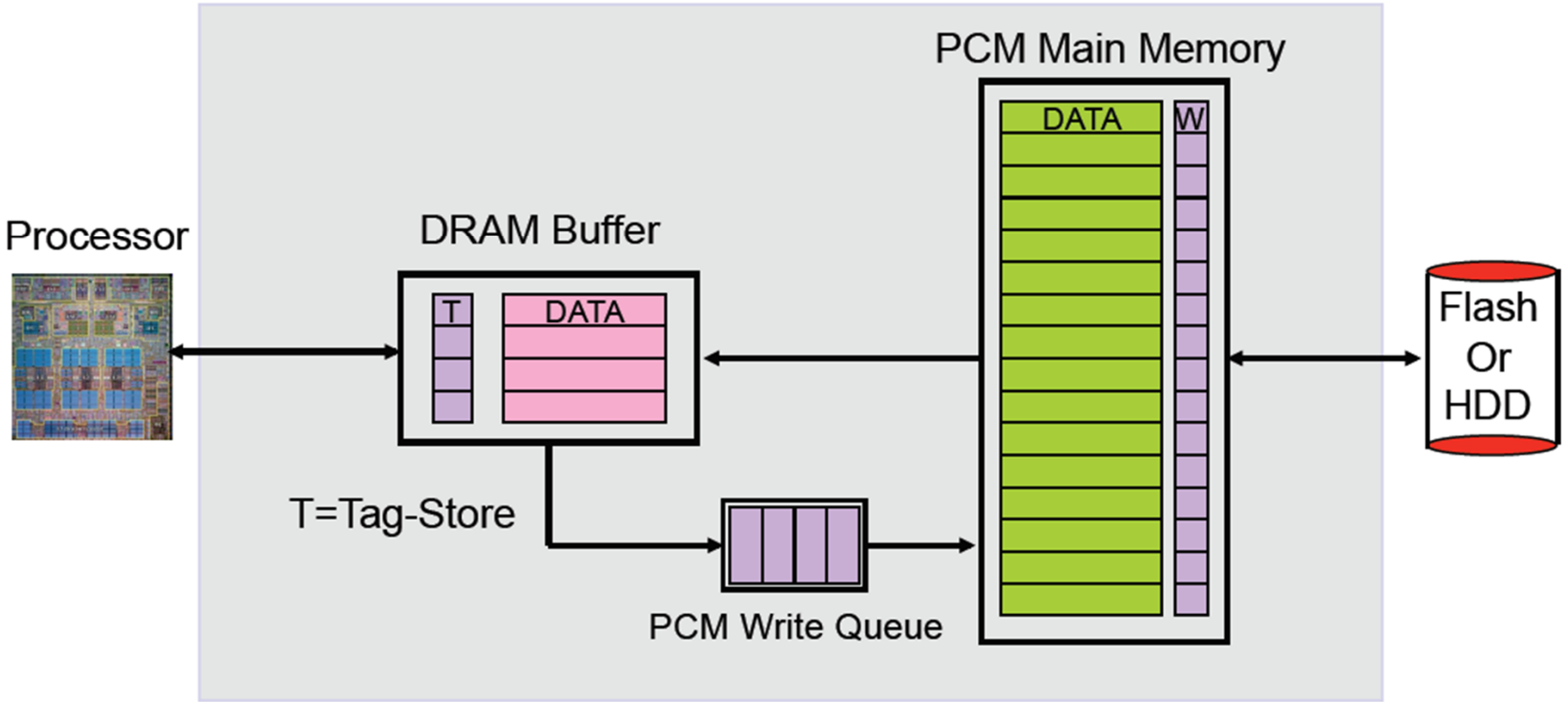
PRAM/DRAM Hybrid Main Memory

- M. Qureshi, et al., “Scalable High Performance Main Memory System Using Phase-Change Memory Technology”, ISCA 2009
- H. Park, S. Yoo, S. Lee, “Dynamic Power Management of PRAM/DRAM Main Memory”

[Qureshi, ISCA 09]

PRAM/DRAM Hybrid Main Memory

Hybrid Memory System



Hybrid Memory System:

1. DRAM as cache to tolerate PCM Rd/Wr latency and Wr bandwidth
2. PCM as main-memory to provide large capacity at good cost/power

[Qureshi, ISCA 09]

PRAM/DRAM Hybrid Main Memory

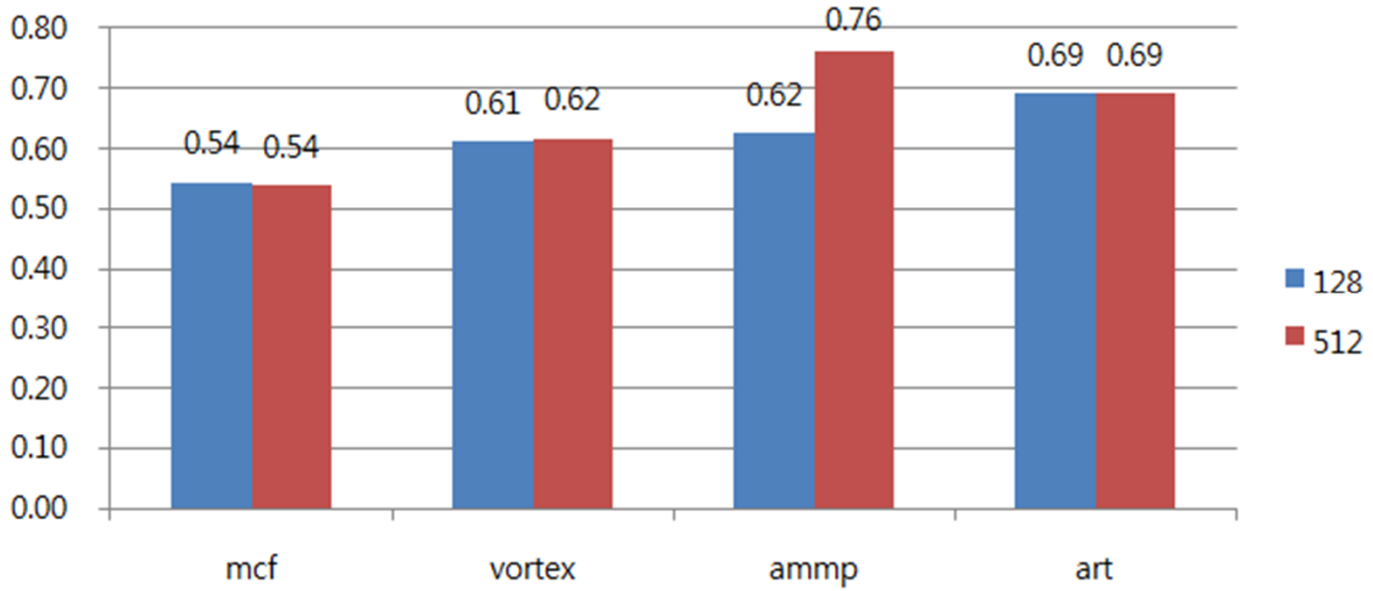
- Taking the benefits of both
  - DRAM as working memory (low latency and write coalescing)
  - PRAM as background memory (large capacity)
- Write reduction
  - Partial write (write dirty cache lines)
- Wear leveling
  - Randomize cache line index in a page

PRAM/DRAM Hybrid Main Memory

- M. Qureshi, et al., “Scalable High Performance Main Memory System Using Phase-Change Memory Technology”, ISCA 2009
- H. Park, S. Yoo, S. Lee, “Dynamic Power Management of PRAM/DRAM Main Memory”

Problem

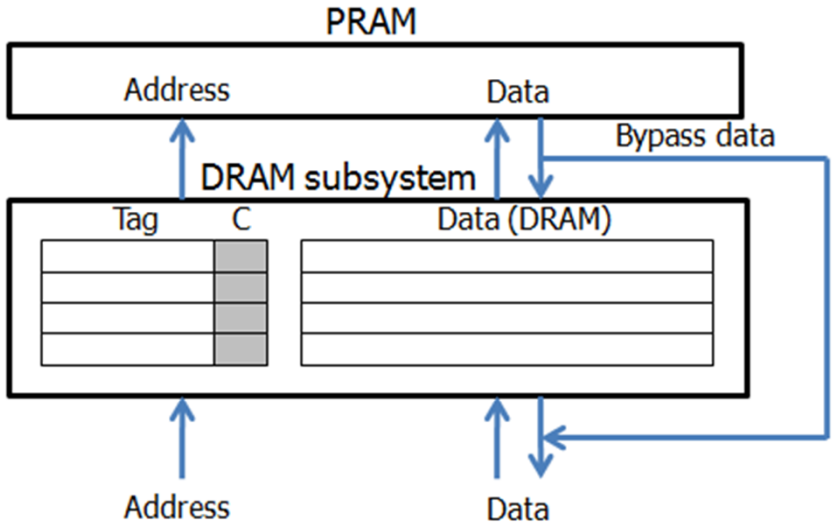
- DRAM dominates total power in hybrid PRAM/DRAM
- DRAM refresh can consume more than 50% of DRAM energy (128MB case study)



Basic Idea

- PRAM: large background main memory
  - Low standby energy & enough memory size
- DRAM: last-level cache with decay
  - Short latency in row access operation
  - DRAM data decay to reduce refresh energy

	PRAM		DRAM	
	Read	Write	Read	Write
Energy per bit	2.47pJ 0.93pJ	16.82pJ 1.02pJ	1.17pJ 0.93pJ	0.39pJ 1.02pJ
Latency at 400MHz	t_{RCD}=22cycles t_{CL}=5cycles t_{WR}=4cycles t_{RP}=60cycles		t_{RCD}=t_{CL}=t_{RP}=5cycles t_{WR}=4cycles	



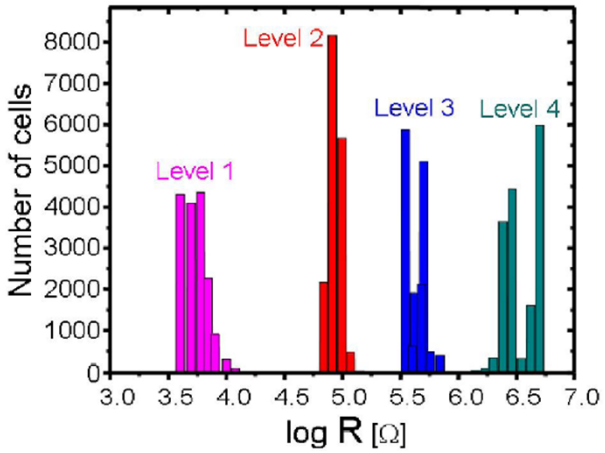
Recent Works on PRAM-based Main Memory Subsystem

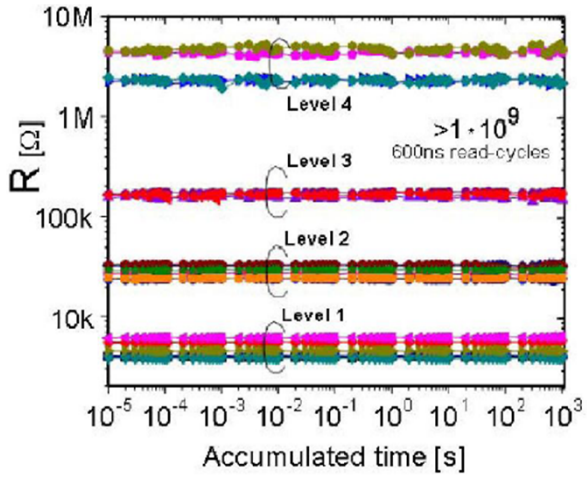
- Reduction in PRAM writes
  - Partial writes [ISCA 09]
  - Differential writes [ISCAS 09]
  - Data encoding including invert coding [MICRO 09]
- Wear leveling
  - Hot/cold swapping [DAC 09]
  - Rotation-based methods: page, cache line, intra cache line, … [ISCA 09]
  - Randomization-based methods: start-gap, security refresh [MICRO 09, ISCA 10]
- Error correction
  - Dynamically replicated memory [ASPLOS 09]
  - Error correction pointer [ISCA 10]
  - SAFER [MICRO 09]
- (Read) performance improvement
  - Write cancellation and pausing [HPCA 10]
  - Morphable memory utilizing both SLC and MLC PRAM selectively [ISCA 10]

[Nirschl, IEDM07]

Next Challenge: MLC PRAM

- Feasible in cell level, but …





- Key issues to be resolved for mass production
  - Trial and error-based programming <-> ISPP in Flash
  - Larger read/write latency than SLC
  - High error rate especially due to R drift

Summary

- PRAM benefits in main memory
  - Low standby power (in near future)
  - Larger capacity
- Problems and solutions
  - PRAM latency/power and write endurance problems → DRAM as working memory
  - Further reduction in standby power → DRAM decay
- Next step
  - Multi-level cell (MLC) PRAM for main memory

References

- [Grider07] G. Grider, et al., "Roadrunner System Management Report," Los Alamos National Laboratory, Tech. Rep. LA-UR-07-7405, 2007.
- [Xie10] G. Sun, et al., "A Hybrid Solid-State Storage Architecture for the Performance, Energy Consumption, and Lifetime Improvement", HPCA, Jan. 2010.
- [Lee09] B. Lee, et al., "Architecting phase change memory as a scalable dram alternative," ISCA, 2009.
- [Qureshi_ISCA09] (IBM ISCA 2009) M. Qureshi, et al., "Scalable high performance main memory system using phase-change memory technology," ISCA, 2009.
- [Qureshi_MICRO09] (IBM MICRO 2009) M. Qureshi, et al., "Enhancing lifetime and security of PCM-based main memory with start-gap wear leveling," MICRO 2009.
- [Seong_ISCA10] N. Seong, et al., "Security Refresh: Prevent Malicious Wear-out and Increase Durability for Phase-Change Memory with Dynamically Randomized Address Mapping," ISCA, 2010.
- [Ipek10] (MS ASPLOS 2010) E. Ipek et al., "Dynamically replicated memory: building reliable systems from nanoscale resistive memories," ASPLOS, 2010.
- [Schechter10] S. Schechter, et al., "Use ECP, not ECC, for Hard Failures in Resistive Memories," Proc. International Symposium on Computer Architecture (ISCA), 2010.
- [Seong_MICRO10] N. Seong, et al., "SAFER: Stuck-At-Fault Error Recovery for Memories," MICRO, 2010.
- [Qureshi_HPCA10] (IBM HPCA 2010) M. Qureshi, et al., "Improving Read Performance of Phase Change Memories via Write Cancellation and Write Pausing," HPCA, 2010.
- [Nirschl_IEDM07] T. Nirschl et al., Write Strategies for 2 and 4-bit Multi-Level Phase-Change Memory, IEDM 2007.