Onyx: A Prototype Phase-Change Memory Storage Array

Ameen Akel*
Adrian Caulfield, Todor Mollov, Rajesh Gupta, Steven Swanson

Non-Volatile Systems Laboratory, Department of Computer Science and Engineering
University of California, San Diego

*Now at Micron Technology
4 KB Operation Request Latencies

- **Disk**
- **Flash**
- **Current PCM**
- **Projected PCM**

![Bar chart showing operation request latencies for write and read operations.](#)
Advantages of Studying PCM SSDs

• Understand current PCM performance
  – With current storage infrastructure
  – Versus other NV tech: e.g. Flash SSDs

• PCM performance may differ from simulation
  – Variance in write latency due to data
  – Wear-out characteristics

• Use real applications to gauge performance

• Understand how software should change for PCM

• Prepare to integrate future-generation PCM
Overview

• Motivation
• PCM Devices
  – Technology Overview
  – Micron P8P Devices
• Onyx Architecture
  – Logical Architecture
  – PCM DIMMs
  – Physical Architecture
• Performance Analysis
• Applications and Conclusions
PCM: The Device Level

- PCM storage medium: Chalcogenide
  - Resistance depends on molecular phase
- Writes
  - Heaters are attached to the chalcogenide
  - Current passed through heaters to change phase
  - Allows bit-alterable writes
- Reads
  - Measure resistance through chalcogenide area
  - Resistance sensed by ability to sink current

M. Breitwisch et al VLSI '07
PCM: The Device Level

- PCM storage medium: Chalcogenide
  - Resistance depends on molecular phase
- Writes
  - Heaters are attached to the chalcogenide
  - Current passed through heaters to change phase
  - Allows bit-alterable writes
- Reads
  - Measure resistance through chalcogenide area
  - Resistance sensed by ability to sink current

M. Wuttig, et. al., FP6 Project CAMELS.
PCM Write Operations in Depth

• Material heated to...
  – > 600°C then cooled quickly ➔ Amorphous
  – ~ 350°C then cooled slowly ➔ Crystalline

• Set and reset
  – Reset – 0 state
  – Set – 1 state
PCM Projections

• Future PCM latency projections*:

<table>
<thead>
<tr>
<th>Operation</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>48 ns</td>
</tr>
<tr>
<td>Set</td>
<td>150 ns</td>
</tr>
<tr>
<td>Reset</td>
<td>40 ns</td>
</tr>
</tbody>
</table>

• Process node progression: 90, 45, 32, 20, 9 nm

P8P PCM

- First-generation NOR-flash replacement
- **Part:** NP8P128A13B1760E (P8P)
- **Process Node:** 90 nm
- **Capacity:** 16 MB
- **Per Device Bandwidth, Latency, Current**
  - Write (64 bytes): 0.5 MB/s, 120 us, 35 mA
  - Read (16 bytes): 48.6 MB/s, 314 ns, 15 mA
- **Lifetime:** One million writes until first bit error
Overview

• Motivation
• PCM Devices
  – Technology Overview
  – Micron P8P Devices
• Onyx Architecture
  – Logical Architecture
  – PCM DIMMs
  – Physical Architecture
• Performance Analysis
• Applications and Conclusions
Moneta: SSD for Emulated Fast NVMs

- DRAM-based NV-SSD emulator
- Learn by building
  - Hardware – Controller & interconnect
  - Software – Driver, file system, apps
- Uses optimized software stack
  - Decreases request latency
  - Improves request concurrency
Onyx: Phase-Change Memory SSD

- Based on Moneta*
  - Shares hardware
  - Shares software stack
- PCM replaces DRAM
  - Uses real PCM
  - Custom PCM controller

*A. M. Caulfield, et. al. Moneta: A high-performance storage array architecture for next-generation, non-volatile memories. MICRO 2010
Moneta/Onyx Architecture

- Host via PIO
- Request Queue
- Scoreboard
- Tag Status Registers
- Ring Control
- Transfer Buffers
- DMA Control
- Host via DMA
- Ring (4 GB/s)
- 2GB PCM
- 2GB PCM
- 2GB PCM
- 2GB PCM
Onyx PCM Controller

• Request Completion
  – Late Completion – On PCM write completion
  – Early Completion – On request reception

• Start-Gap Wear Leveling*
  – Low overhead wear leveling (two registers + logic)
  – Prevents hot spots from wearing out memory
  – Rotates line in memory every gap interval

*M. K. Qureshi, et. al. Enhancing lifetime and security of PCM-based main memory with start-gap wear leveling. MICRO 42.
Closer Look at a PCM DIMM

- 8 Ranks of 5 PCM devices
  - 64 data bits + 16 ECC bits
  - Effectively 16 ranks per memory interface
- Shared control and data lines
- Capacity: 640 MB / DIMM
Prototyping Advanced SSDs

• Built on RAMP’s BEE3 board
  – Four FPGAs connected in a ring
  – Four DIMM slots per FPGA
  – PCIe 1.1 x8 host connection
• System capacity: 10 GB
Overview

• Motivation
• PCM Devices
  – Technology Overview
  – Micron P8P Devices
• Onyx Architecture
  – Logical Architecture
  – PCM DIMMs
  – Physical Architecture
• Performance Analysis
• Applications and Conclusions
Read Performance

- Onyx
- FusionIO
- Moneta

Bandwidth (MB/s) vs Request Size (KB)
Write Performance

- Onyx-Late
- Onyx-Early
- FusionIO
- Moneta

Bandwidth (MB/s) vs Request Size (KB) chart.
BerkeleyDB Performance

![Graph showing performance comparison between Onyx, FusionIO, and Moneta in BDB Benchmark.](image)
Potential PCM Applications

• As a read cache
  – First-gen PCM read speeds compete with flash
  – Next-gen PCM should improve read performance

• Replace DRAM in high-performance apps
  – PCM cost will likely drop below DRAM
  – Will scale aggressively past DRAM

• Outpace flash in high-performance SSDs
  – Reduces complexity of management
  – Provides higher-rated lifetime
  – Saves power, logic, and design time
Conclusions

• Onyx designed to maximize PCM performance
• More improvements possible as PCM scales
  – Onyx architecture will scale with PCM
  – Onyx will benefit from faster reads and writes
• PCM simplifies SSD management relative to flash and improves small access performance
Thank You!

Questions?