Phase Change Memory
An Architecture and Systems Perspective

Benjamin C. Lee
Stanford University
bcclee@stanford.edu
Memory Scaling

- ↑ density, capacity; ↓ cost-capability ratio
- Emerging challenges for prevalent technologies [ITRS07]
Technology Alternatives

● **Charge Memory**
  ○ Write data by capturing charge $Q$
  ○ Read data by detecting voltage $V$
  ○ Examples: Flash, DRAM

● **Resistive Memory**
  ○ Write data by pulsing current $dQ/dt$
  ○ Read data by detecting resistance $R$
  ○ Examples: PCM, STT-MRAM, memristor
Limits of Charge Memory

- Difficult charge placement and control
- Flash: floating gate charge
- DRAM: capacitor charge, transistor leakage
Towards Resistive Memory

- **Scalable**
  - Program cell with scalable mechanisms
  - Map resistance to logical state

- **Non-Volatile**
  - Set atomic structure of cell
  - Incur activation cost to alter properties

- **Competitive**
  - Achieve viable latency, power, endurance
  - Scale to improve performance metrics
Technology

Phase Change Memory

- Store data within phase change material
- Set phase via current pulse
- Detect phase via resistance (amorphous/crystalline)
PCM Scalability

- Program with current pulses, which scale linearly
- PCM roadmap to 30nm [Raoux+08]
- Flash/DRAM roadmap to 40nm [ITRS07]
PCM Non-Volatility

- **Atomic Structure**
  - Program with current pulses
  - Melt material at $650\,^\circ C$
  - Cool material to desired phase

- **Activation Cost**
  - Crystallize with high activation energy
  - Isolate thermal effects to target cell
  - Retain data for >10 years at $85\,^\circ C$
Technology Parameters

- Survey prototypes from 2003-2008 [ISSCC][VLSI][IEDM][ITRS]
- Derive parameters for $F=90\text{nm}$

**Size**
- $9 - 12F^2$ using BJT
- $1.5 \times \text{DRAM}$

**Endurance**
- $1E+08$ writes
- $1E-08 \times \text{DRAM}$

**Latency**
- $50\text{ns Rd}, 150\text{ns Wr}$
- $4 \times, 12 \times \text{DRAM}$

**Energy**
- $40\mu\text{A Rd}, 150\mu\text{A Wr}$
- $2 \times, 43 \times \text{DRAM}$
Technology Parameters

- Survey prototypes from 2003-2008 [ISSCC][VLSI][IEDM][ITRS]
- Derive parameters for $F=90$nm

Size
- 9 - $12F^2$ using BJT
- $1.5 \times$ DRAM

Endurance
- $1E+08$ writes
- $1E-08 \times$ DRAM

Latency
- 50ns Rd, 150ns Wr
- $4 \times$, $12 \times$ DRAM

Energy
- $40 \mu A$ Rd, $150 \mu A$ Wr
- $2 \times$, $43 \times$ DRAM
Technology Parameters

- Survey prototypes from 2003-2008 \([\text{ISSCC}] [\text{VLSI}] [\text{IEDM}] [\text{ITRS}]\)
- Derive parameters for $F=90\text{nm}$

### Size
- $9 - 12F^2$ using BJT
- $1.5 \times$ DRAM

### Endurance
- $1\times 10^8$ writes
- $1\times 10^{-8} \times$ DRAM

### Latency
- $50\text{ns} \text{ Rd}, 150\text{ns} \text{ Wr}$
- $4 \times, 12 \times$ DRAM

### Energy
- $40\mu\text{A} \text{ Rd}, 150\mu\text{A} \text{ Wr}$
- $2 \times, 43 \times$ DRAM
Technology Parameters

- Survey prototypes from 2003-2008 [ISSCC][VLSI][IEDM][ITRS]
- Derive parameters for $F=90\text{nm}$

**Size**
- $9 - 12F^2$ using BJT
- $1.5 \times$ DRAM

**Latency**
- $50\text{ns}$ Rd, $150\text{ns}$ Wr
- $4 \times, 12 \times$ DRAM

**Endurance**
- $1\text{E}+08$ writes
- $1\text{E}-08 \times$ DRAM

**Energy**
- $40\mu\text{A}$ Rd, $150\mu\text{A}$ Wr
- $2 \times, 43 \times$ DRAM
PCM Deployment

- Deploy PCM on memory bus
- Begin by co-locating PCM, DRAM
Price of Scalability

- Replace DRAM with PCM in present architectures
- $1.6 \times$ delay, $2.2 \times$ energy, 500-hour lifetime
Architecture and Scalability

Architecture Objectives

- **DRAM-Competitive**
  - Reorganize row buffer to mitigate delay, energy
  - Implement partial writes to mitigate wear mechanism

- **Area-Efficient**
  - Minimize disruption to density trends
  - Impacts row buffer organization

- **Complexity-Effective**
  - Encourage adoption with modest mechanisms
  - Impacts partial writes
Buffer Organization

- **On-Chip Buffers**
  - Use DRAM-like buffer and interface
  - Evict modified rows into array

- **Narrow Rows**
  - Reduce write energy \( \propto \) buffer width
  - Reduce peripheral circuitry, associated area

- **Multiple Rows**
  - Reduce eviction frequency
  - Improve locality, write coalescing
Buffer Area Strategy

- Narrow rows :: fewer expensive S/A’s (44T)
- Multiple rows :: additional inexpensive latches (8T)
Buffer Design Space

- Derive DRAM, PCM area model
- Explore space of area-neutral buffer designs
Wear Reduction

- **Wear Mechanism**
  - Writes induce phase change at 650 °C
  - Contacts degrade from thermal expansion/contraction
  - Current injection is less reliable after 1E+08 writes

- **Partial Writes**
  - Reduce writes to PCM array
  - Write only stored lines (64B), words (4B)
  - Add cache line state with 0.2%, 3.1% overhead
Partial Writes

- Derive PCM lifetime model
- Quantify eliminated writes during buffer eviction
Scalable Performance

- $1.2 \times$ delay, $1.0 \times$ energy, 5.6-year lifetime
- Scaling improves energy, endurance
Systems and Non-Volatility

Storage Systems

- Persistent data in slow, non-volatile memory
- Buffered data in fast, volatile memory
Storage System Trade-offs

● **Design Objectives**
  ○ Safety :: secure against crashes
  ○ Consistency :: correctness in non-volatile memory
  ○ Performance :: buffering in volatile memory

● **Byte-addressable Persistence (BPRAM)**
  ○ Narrows gap between volatile/non-volatile memory
  ○ Addressable like DRAM
  ○ Persistent like disk, Flash
Byte-addressable Persistent File System (BPFS)

- **Safety**
  - Use PCM as DRAM alternative
  - Reflect writes to PCM in $O(\text{ms})$, not $O(\text{s})$

- **Consistency**
  - Enforce atomicity, ordering in hardware
  - Support shadow paging, copy-on-write

- **Performance**
  - Use short-circuit shadow paging
  - Exploit addressability for small, in-place writes
Tree-Based File System

- root pointer
- indirect blocks
- inodes
- inode file
- file
- directory
- file
Disks & Journaling

- Write to journal before write to file system
- Requires twice the writes
Disks & Journaling

- Write to journal before write to file system
- Requires twice the writes
Disks & Journaling

- Write to journal before write to file system
- Requires twice the writes

![Diagram illustrating file system and journal with nodes A' and B']
Disks & Shadow Paging

- Copy-on-write up to file system root
- Requires recursion up file system tree
Disks & Shadow Paging

- Copy-on-write up to file system root
- Requires recursion up file system tree
Disks & Shadow Paging

- Copy-on-write up to file system root
- Requires recursion up file system tree
PCM & Short-Circuit Shadow Paging

- Update in-place for small writes
- Lowers copying overhead
Hardware Support

- **Atomicity**
  - BPFS assumes atomic 64-bit pointer updates
  - PCM writes atomically into memory array
  - Capacitors guard against power failures

- **Ordering**
  - Caches, memory controller reorder writes
  - Epochs define barrier-delimited BPFS writes
BPFS Evaluation

- Improved safety, consistency
- Improved performance
Conclusions

- **Scaling Challenges**
  - Fundamental limits in charge memory
  - Transition towards resistive memory

- **Architecture and Scalability**
  - Scalable, non-volatile, DRAM-competitive
  - Efficient buffers mitigate latency, energy
  - Partial writes reduce wear

- **Systems and Non-Volatility**
  - BPFS changes storage system trade-offs
  - Short-circuit shadow paging, hardware support
  - Improves durability, performance
Phase Change Memory
An Architecture and Systems Perspective

Benjamin C. Lee
Stanford University
bcclee@stanford.edu