Computer Architecture
Lecture 14b: Emerging Memory Technologies

Prof. Onur Mutlu
ETH Zürich
Fall 2021
12 November 2021
Limits of Charge Memory

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

- Reliable sensing becomes difficult as charge storage unit size reduces
Solution 1: New Memory Architectures

- Overcome memory shortcomings with
  - Memory-centric system design
  - Novel memory architectures, interfaces, functions
  - Better waste management (efficient utilization)

- Key issues to tackle
  - Enable reliability at low cost $\rightarrow$ high capacity
  - Reduce energy
  - Reduce latency
  - Improve bandwidth
  - Reduce waste (capacity, bandwidth, latency)
  - Enable computation close to data
Solution 1: New Memory Architectures

Kim+, "The Experimental Study of Data Retention Behavior in Modern DRAM Devices," ISCA 2022.
Sashabah+, "Resilient Fast and Efficient DRAM Copy and Initialization of Bulk Data," MDSO 2021.
Kim+, "Characterizing Application Memory Error Vulnerability to Optimize Data Center Cost," ISCA 2014.
Wen+, "A Suitable Processing-In-Memory Accelerator for Parallel Graph Processing," EWSN 2019.
Sashabah+, "Gather-Side DRAM Translation to Improve the Spatial Locality of Non-Sequential Access," MWT 2015.
Hashemi+, "Continuous Randomized Transparent Hardware Acceleration for Memory Intensive Workloads," MWD 2018.
Khan+, "A Case for Memory Content-Based Detection and Mitigation of Data Dependent Failures in DRAM," IEEE CCL 2016.
Hashemi+, "The RowHammer Problem and Other Issues We May Face as Memory Becomes Denser," DATE 2017.
Khan+, "Detecting and Mitigating Data-Dependent DRAM Failures by Exploiting Current Memory Content," MWT 2017.
Khan+, "Detecting and Mitigating DRAM Access Latency by Exploiting the Spatial Locality of References in DRAM," ISCA 2019.
Khan+, "Detecting and Mitigating Data-Dependent DRAM Failures by Exploiting Current Memory Content," MWT 2017.
Khan+, "Detecting and Mitigating DRAM Access Latency by Exploiting the Spatial Locality of References in DRAM," ISCA 2019.
Solution 2: Emerging Memory Technologies

- Some emerging **resistive** memory technologies seem more scalable than DRAM (and they are non-volatile)

- Example: Phase Change Memory
  - Data stored by changing phase of material
  - Data read by detecting material’s resistance
  - Expected to scale to 9nm (2022 [ITRS 2009])
  - Prototyped at 20nm (Raoux+, IBM JRD 2008)
  - Expected to be denser than DRAM: can store multiple bits/cell

- But, emerging technologies have (many) shortcomings
  - Can they be enabled to replace/augment/surpass DRAM?
Solution 2: Emerging Memory Technologies

- Zhao+, “FIRM: Fair and High-Performance Memory Control for Persistent Memory Systems,” MICRO 2014.
Intel Optane Persistent Memory (2019)

- Non-volatile main memory
- Based on 3D-XPoint Technology

https://www.storagereview.com/intel_optane_dc_persistent_memory_module_pmm
PCM as Main Memory: Idea in 2009

- Benjamin C. Lee, Engin Ipek, Onur Mutlu, and Doug Burger,
  "Architecting Phase Change Memory as a Scalable DRAM Alternative"
  Proceedings of the 36th International Symposium on Computer
  Architecture (ISCA), pages 2-13, Austin, TX, June 2009. Slides (pdf)
  One of the 13 computer architecture papers of 2009 selected as Top
  Picks by IEEE Micro.
  Selected as a CACM Research Highlight.

Architecting Phase Change Memory as a Scalable DRAM Alternative

Benjamin C. Lee†  Engin Ipek†  Onur Mutlu‡  Doug Burger¶

†Computer Architecture Group
Microsoft Research
Redmond, WA
{blee, ipek, dburger}@microsoft.com

‡Computer Architecture Laboratory
Carnegie Mellon University
Pittsburgh, PA
onur@cmu.edu
PCM as Main Memory: Idea in 2009

- Benjamin C. Lee, Ping Zhou, Jun Yang, Youtao Zhang, Bo Zhao, Engin Ipek, Onur Mutlu, and Doug Burger,
  "Phase Change Technology and the Future of Main Memory"

**Phase-Change Technology and the Future of Main Memory**
Charge vs. Resistive Memories

- **Charge Memory** (e.g., DRAM, Flash)
  - Write data by capturing charge Q
  - Read data by detecting voltage V

- **Resistive Memory** (e.g., PCM, STT-MRAM, memristors)
  - Write data by pulsing current \( \frac{dQ}{dt} \)
  - Read data by detecting resistance R
Promising Resistive Memory Technologies

- **PCM**
  - Inject current to change material phase
  - Resistance determined by phase

- **STT-MRAM**
  - Inject current to change magnet polarity
  - Resistance determined by polarity

- **Memristors/RRAM/ReRAM**
  - Inject current to change atomic structure
  - Resistance determined by atom distance
What is Phase Change Memory?

- Phase change material (chalcogenide glass) exists in two states:
  - Amorphous: Low optical reflexivity and high electrical resistivity
  - Crystalline: High optical reflexivity and low electrical resistivity

PCM is resistive memory: High resistance (0), Low resistance (1)
PCM cell can be switched between states reliably and quickly
How Does PCM Work?

- **Write**: change phase via current injection
  - SET: sustained current to heat cell above $T_{cryst}$
  - RESET: cell heated above $T_{melt}$ and quenched
- **Read**: detect phase via material resistance
  - amorphous/crystalline

![Diagram of PCM operation](image)

**SET** (cryst) Low resistance
- $10^3-10^4 \Omega$

**RESET** (amorph) High resistance
- $10^6-10^7 \Omega$

Photo Courtesy: Bipin Rajendran, IBM  
Slide Courtesy: Moinuddin Qureshi, IBM
Opportunity: PCM Advantages

- **Scales better than DRAM, Flash**
  - Requires current pulses, which scale linearly with feature size
  - Expected to scale to 9nm (2022 [ITRS])
  - Prototyped at 20nm (Raoux+, IBM JRD 2008)

- **Can be denser than DRAM**
  - Can store multiple bits per cell due to large resistance range
  - Prototypes with 2 bits/cell in ISSCC’ 08, 4 bits/cell by 2012

- **Non-volatile**
  - Retain data for >10 years at 85C

- **No refresh needed, low idle power**
 PCM Resistance → Value

Cell value:

1

0

Cell resistance
Multi-Level Cell PCM

- Multi-level cell: more than 1 bit per cell
  - Further increases density by 2 to 4x [Lee+, ISCA'09]

- But MLC-PCM also has drawbacks
  - Higher latency and energy than single-level cell PCM
MLC-PCM Resistance → Value

Cell value:

- Bit 1: 11
- Bit 0: 10, 01, 00

Cell resistance
MLC-PCM Resistance $\rightarrow$ Value

*Less margin between values*

$\rightarrow$ need more precise sensing/modification of cell contents

$\rightarrow$ higher latency/energy ($\sim 2x$ for reads and $4x$ for writes)
Phase Change Memory Properties

- Surveyed prototypes from 2003-2008 (ITRS, IEDM, VLSI, ISSCC)
- Derived PCM parameters for F=90nm

<table>
<thead>
<tr>
<th>Parameter*</th>
<th>Horri\textsuperscript{6}</th>
<th>Ahn\textsuperscript{12}</th>
<th>Bedeschi\textsuperscript{13}</th>
<th>Oh\textsuperscript{14}</th>
<th>Pellizer\textsuperscript{15}</th>
<th>Chen\textsuperscript{5}</th>
<th>Kang\textsuperscript{16}</th>
<th>Bedeschi\textsuperscript{9}</th>
<th>Lee\textsuperscript{10}</th>
<th>Lee\textsuperscript{2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process, F,(nm)</td>
<td>**</td>
<td>120</td>
<td>180</td>
<td>120</td>
<td>90</td>
<td>**</td>
<td>100</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Array size (Mbytes)</td>
<td>**</td>
<td>64</td>
<td>8</td>
<td>64</td>
<td>**</td>
<td>**</td>
<td>256</td>
<td>256</td>
<td>512</td>
<td>**</td>
</tr>
<tr>
<td>Material</td>
<td>GST, N-d</td>
<td>GST, N-d</td>
<td>GST</td>
<td>GST</td>
<td>GST</td>
<td>GST</td>
<td>GST</td>
<td>GST</td>
<td>GST, N-d</td>
<td>GST, N-d</td>
</tr>
<tr>
<td>Cell size ((\mu m^2))</td>
<td>**</td>
<td>0.290</td>
<td>0.290</td>
<td>**</td>
<td>0.097</td>
<td>60 nm(^2)</td>
<td>0.166</td>
<td>0.097</td>
<td>0.047</td>
<td>0.065 to 0.097</td>
</tr>
<tr>
<td>Cell size, F(^2)</td>
<td>**</td>
<td>20.1</td>
<td>9.0</td>
<td>**</td>
<td>12.0</td>
<td>**</td>
<td>16.6</td>
<td>12.0</td>
<td>5.8</td>
<td>9.0 to 12.0</td>
</tr>
<tr>
<td>Access device</td>
<td>**</td>
<td>**</td>
<td>BJT</td>
<td>FET</td>
<td>BJT</td>
<td>**</td>
<td>FET</td>
<td>BJT</td>
<td>Diode</td>
<td>BJT</td>
</tr>
<tr>
<td>Read time (ns)</td>
<td>**</td>
<td>70</td>
<td>48</td>
<td>68</td>
<td>**</td>
<td>62</td>
<td>**</td>
<td>55</td>
<td>48</td>
<td>**</td>
</tr>
<tr>
<td>Read current ((\mu A))</td>
<td>**</td>
<td>40</td>
<td>**</td>
<td>**</td>
<td>40</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Read voltage (V)</td>
<td>**</td>
<td>3.0</td>
<td>1.0</td>
<td>1.8</td>
<td>1.6</td>
<td>**</td>
<td>1.8</td>
<td>**</td>
<td>1.8</td>
<td>1.0</td>
</tr>
<tr>
<td>Read power ((\mu W))</td>
<td>**</td>
<td>40</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Read energy (pJ)</td>
<td>**</td>
<td>2.0</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>2.0</td>
</tr>
<tr>
<td>Set time (ns)</td>
<td>100</td>
<td>150</td>
<td>150</td>
<td>180</td>
<td>**</td>
<td>80</td>
<td>300</td>
<td>**</td>
<td>400</td>
<td>150</td>
</tr>
<tr>
<td>Set current ((\mu A))</td>
<td>200</td>
<td>**</td>
<td>300</td>
<td>200</td>
<td>**</td>
<td>55</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>150</td>
</tr>
<tr>
<td>Set voltage (V)</td>
<td>**</td>
<td>2.0</td>
<td>**</td>
<td>1.25</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>1.2</td>
</tr>
<tr>
<td>Set power ((\mu W))</td>
<td>**</td>
<td>300</td>
<td>**</td>
<td>34.4</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Set energy (pJ)</td>
<td>**</td>
<td>45</td>
<td>**</td>
<td>2.8</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>13.5</td>
</tr>
<tr>
<td>Reset time (ns)</td>
<td>50</td>
<td>10</td>
<td>40</td>
<td>10</td>
<td>**</td>
<td>60</td>
<td>50</td>
<td>**</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>Reset current ((\mu A))</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>400</td>
<td>90</td>
<td>600</td>
<td>**</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Reset voltage (V)</td>
<td>**</td>
<td>2.7</td>
<td>**</td>
<td>1.8</td>
<td>1.6</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>1.6</td>
</tr>
<tr>
<td>Reset power ((\mu W))</td>
<td>**</td>
<td>1620</td>
<td>**</td>
<td>80.4</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>480</td>
</tr>
<tr>
<td>Reset energy (pJ)</td>
<td>**</td>
<td>64.8</td>
<td>**</td>
<td>4.8</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>19.2</td>
</tr>
<tr>
<td>Write endurance</td>
<td>(10^7)</td>
<td>(10^9)</td>
<td>(10^6)</td>
<td>**</td>
<td>(10^8)</td>
<td>**</td>
<td>(10^4)</td>
<td>**</td>
<td>(10^5)</td>
<td>(10^8)</td>
</tr>
</tbody>
</table>

\(^*\) BJT: bipolar junction transistor; FET: field-effect transistor; GST: Ge\(_2\)Sb\(_2\)Te\(_5\); MLC: multilevel cells; N-d: nitrogen doped.

\(^{**}\) This information is not available in the publication cited.
Phase Change Memory Properties: Latency

- Latency comparable to, but slower than DRAM
- Read Latency
  - 50ns: 4x DRAM, 10^{-3}x NAND Flash
- Write Latency
  - 150ns: 12x DRAM
- Write Bandwidth
  - 5-10 MB/s: 0.1x DRAM, 1x NAND Flash

Phase Change Memory Properties

- Dynamic Energy
  - 40 uA Rd, 150 uA Wr
  - 2-43x DRAM, 1x NAND Flash

- Endurance
  - Writes induce phase change at 650°C
  - Contacts degrade from thermal expansion/contraction
  - $10^8$ writes per cell
  - $10^{-8}$x DRAM, $10^3$x NAND Flash

- Cell Size
  - 9-12F^2 using BJT, single-level cells
  - 1.5x DRAM, 2-3x NAND (will scale with feature size, MLC)
Phase Change Memory: Pros and Cons

- **Pros over DRAM**
  - Better technology scaling (capacity and cost)
  - Non volatile → Persistent
  - Low idle power (no refresh)

- **Cons**
  - Higher latencies: ~4-15x DRAM (especially write)
  - Higher active energy: ~2-50x DRAM (especially write)
  - Lower endurance (a cell dies after ~$10^8$ writes)
  - Reliability issues (resistance drift)

- **Challenges in enabling PCM as DRAM replacement/helper:**
  - Mitigate PCM shortcomings
  - Find the right way to place PCM in the system
**PCM-based Main Memory (I)**

- How should PCM-based (main) memory be organized?

- **Hybrid PCM+DRAM** [Qureshi+ ISCA’09, Dhiman+ DAC’09]:
  - How to partition/migrate data between PCM and DRAM
PCM-based Main Memory (II)

- How should PCM-based (main) memory be organized?

- Pure PCM main memory [Lee et al., ISCA’09, Top Picks’10]:
  - How to redesign entire hierarchy (and cores) to overcome PCM shortcomings
An Initial Study: Replace DRAM with PCM

  - Surveyed prototypes from 2003-2008 (e.g. IEDM, VLSI, ISSCC)
  - Derived “average” PCM parameters for F=90nm

### Density
- 9 - 12\(F^2\) using BJT
- 1.5× DRAM

### Latency
- 50ns Rd, 150ns Wr
- 4×, 12× DRAM

### Endurance
- 1E+08 writes
- 1E-08× DRAM

### Energy
- 40\(\mu\)A Rd, 150\(\mu\)A Wr
- 2×, 43× DRAM
Results: Naïve Replacement of DRAM with PCM

- Replace DRAM with PCM in a 4-core, 4MB L2 system
- PCM organized the same as DRAM: row buffers, banks, peripherals
- 1.6x delay, 2.2x energy, 500-hour average lifetime

Architecting PCM to Mitigate Shortcomings

- Idea 1: Use multiple narrow row buffers in each PCM chip
  → Reduces array reads/writes → better endurance, latency, energy

- Idea 2: Write into array at cache block or word granularity
  → Reduces unnecessary wear
Results: Architected PCM as Main Memory

- 1.2x delay, 1.0x energy, 5.6-year average lifetime
- Scaling improves energy, endurance, density

![PCM Performance :: 512Bx4 Buffer](chart1)

- Caveat 1: Worst-case lifetime is much shorter (no guarantees)
- Caveat 2: Intensive applications see large performance and energy hits
- Caveat 3: Optimistic PCM parameters?
PCM As Main Memory

Benjamin C. Lee, Engin Ipek, Onur Mutlu, and Doug Burger,
"Architecting Phase Change Memory as a Scalable DRAM Alternative"
One of the 13 computer architecture papers of 2009 selected as Top Picks by IEEE Micro.
Selected as a CACM Research Highlight.

Architecting Phase Change Memory as a Scalable DRAM Alternative

Benjamin C. Lee† Engin Ipek† Onur Mutlu‡ Doug Burger†

†Computer Architecture Group
Microsoft Research
Redmond, WA
{blee, ipek, dburger}@microsoft.com

‡Computer Architecture Laboratory
Carnegie Mellon University
Pittsburgh, PA
onur@cmu.edu
More on PCM As Main Memory (II)


Phase-Change Technology and the Future of Main Memory
Intel Optane Memory (Idea Realized in 2019)

- Non-volatile main memory
- Based on 3D-XPoint Technology
More on PCM Based Main Memory

HanBin Yoon, Justin Meza, Naveen Muralimanohar, Norman P. Jouppi, and Onur Mutlu, "Efficient Data Mapping and Buffering Techniques for Multi-Level Cell Phase-Change Memories"

*ACM Transactions on Architecture and Code Optimization* (TACO), Vol. 11, No. 4, December 2014. [Slides (ppt) (pdf)]

Presented at the 10th HiPEAC Conference, Amsterdam, Netherlands, January 2015. [Slides (ppt) (pdf)]

*Best (student) presentation award.*

Efficient Data Mapping and Buffering Techniques for Multilevel Cell Phase-Change Memories

HANBIN YOON* and JUSTIN MEZA, Carnegie Mellon University
NAVEEN MURALIMANOHAR, Hewlett-Packard Labs
NORMAN P. JOUPPI**, Google Inc.
ONUR MUTLU, Carnegie Mellon University

SAFARI
Some PCM Bits Take Longer to Read…

(a) Sensing time is longer for higher cell resistances.

(b) One bit is determined before the other.

Fig. 3. MLC PCM cell read operation [Qureshi et al. 2010b].
Observation 1: Read Asymmetry

- The read latency/energy of Bit 1 is lower than that of Bit 0

- This is due to how MLC-PCM cells are read
Observation 1: Read Asymmetry

Simplified example

Capacitor filled with reference voltage
MLC-PCM cell with unknown resistance
Observation 1: Read Asymmetry

Simplified example
Observation 1: Read Asymmetry

Simplified example

Infer data value
Observation 1: Read Asymmetry
Observation 1: Read Asymmetry
Observation 1: Read Asymmetry

Initial voltage (fully charged capacitor)

Voltage

11 10 01 00

Time
Observation 1: Read Asymmetry

PCM cell connected → draining capacitor
Observation 1: Read Asymmetry

Capacitor drained → data value known (01)
Observation 1: Read Asymmetry

- In existing devices
  - Both MLC bits are read at the same time
  - Must wait *maximum time* to read both bits

- However, *we can infer information about Bit 1 before this time*
Observation 1: Read Asymmetry

Voltage

Time

11 10 01 00
Observation 1: Read Asymmetry

Voltage

11 10 01 00

Time
Observation 1: Read Asymmetry

Time to determine Bit 1's value
Observation 1: Read Asymmetry

Time to determine Bit 0's value
Some PCM Bits Take Longer to Write…

Efficient Data Mapping and Buffering Techniques for MLC PCM

(a) All possible cell state transitions.

(b) Cell state transitions when modifying only the MSB or the LSB.

Fig. 4. MLC PCM cell write latencies [Joshi et al. 2011; Nirschl et al. 2007; Happ et al. 2006].
More on PCM Latencies and Exploiting Them

HanBin Yoon, Justin Meza, Naveen Muralimanohar, Norman P. Jouppi, and Onur Mutlu,
"Efficient Data Mapping and Buffering Techniques for Multi-Level Cell Phase-Change Memories"

ACM Transactions on Architecture and Code Optimization (TACO), Vol. 11, No. 4, December 2014. [Slides (ppt) (pdf)]
Presented at the 10th HiPEAC Conference, Amsterdam, Netherlands, January 2015. [Slides (ppt) (pdf)]

Best (student) presentation award.

Efficient Data Mapping and Buffering Techniques for Multilevel Cell Phase-Change Memories

HANBIN YOON* and JUSTIN MEZA, Carnegie Mellon University
NAVEEN MURALIMANOHAR, Hewlett-Packard Labs
NORMAN P. JOUPPI**, Google Inc.
ONUR MUTLU, Carnegie Mellon University
STT-RAM as Main Memory
STT-MRAM as Main Memory

- Magnetic Tunnel Junction (MTJ) device
  - Reference layer: Fixed magnetic orientation
  - Free layer: Parallel or anti-parallel

- Magnetic orientation of the free layer determines logical state of device
  - High vs. low resistance

- Write: Push large current through MTJ to change orientation of free layer
- Read: Sense current flow

STT-MRAM: Pros and Cons

- **Pros over DRAM**
  - Better technology scaling (capacity and cost)
  - Non volatile → Persistent
  - Low idle power (no refresh)

- **Cons**
  - Higher write latency
  - Higher write energy
  - Poor density (currently)
  - Reliability?

- **Another level of freedom**
  - Can trade off non-volatility for lower write latency/energy (by reducing the size of the MTJ)
Architected STT-MRAM as Main Memory

- 4-core, 4GB main memory, multiprogrammed workloads
- ~6% performance loss, ~60% energy savings vs. DRAM

More on STT-MRAM as Main Memory

- Emre Kultursay, Mahmut Kandemir, Anand Sivasubramaniam, and Onur Mutlu,
"Evaluating STT-RAM as an Energy-Efficient Main Memory Alternative"
Proceedings of the 2013 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), Austin, TX, April 2013. Slides (pptx) (pdf)

Evaluating STT-RAM as an Energy-Efficient Main Memory Alternative

Emre Kultursay*, Mahmut Kandemir*, Anand Sivasubramaniam*, and Onur Mutlu†
*The Pennsylvania State University and †Carnegie Mellon University