Lecture 6: Chipkill, PCM

• Topics: error correction, PCM basics, PCM writes and errors
Chipkill

- Chipkill correct systems can withstand failure of an entire DRAM chip

- For chipkill correctness
  - the 72-bit word must be spread across 72 DRAM chips
  - or, a 13-bit word (8-bit data and 5-bit ECC) must be spread across 13 DRAM chips
RAID-like DRAM Designs

• DRAM chips do not have built-in error detection

• Can employ a 9-chip rank with ECC to detect and recover from a single error; in case of a multi-bit error, rely on a second tier of error correction

• Can do parity across DIMMs (needs an extra DIMM); use ECC within a DIMM to recover from 1-bit errors; use parity across DIMMs to recover from multi-bit errors in 1 DIMM

• Reads are cheap (must only access 1 DIMM); writes are expensive (must read and write 2 DIMMs)

Used in some HP servers
• Add a checksum to every row in DRAM; verified at the memory controller

• Adds area overhead, but provides self-contained error detection

• When a chip fails, can re-construct data by examining another parity DRAM chip

• Can control overheads by having checksum for a large row or one parity chip for many data chips

• Writes are again problematic
The cache line is organized into multi-bit symbols

Two symbols are required for error detection and 3/4 symbols are used for error correction (can handle complete failure in one symbol, i.e., each symbol is fetched from a different DRAM chip)

3-symbol codes are not popular because it leads to non-standard DIMMs

4-symbol codes are more popular, but are used as 32+4 so that standard ECC DIMMs can be used (high activation energy and low rank-level parallelism) (16+4 would require a non-standard DIMM)
Virtualized ECC

• Also builds a two-tier error protection scheme, but does the second tier in software

• The second-tier codes are stored in the regular physical address space (not specialized DRAM chips); software has flexibility in terms of the types of codes to use and the types of pages that are protected

• Reads are cheap; writes are expensive as usual; but, the second-tier codes can now be cached; greatly helps reduce the number of DRAM writes

• Requires a 144-bit datapath (increases overfetch)
LoT-ECC

• Use checksums to detect errors and parity codes to fix
• Requires access of only 9 DRAM chips per read, but the storage overhead grows to 26%
Phase Change Memory

- Emerging NVM technology that can replace Flash and DRAM; there are other competing technologies too

- Much higher density; much better scalability; can do multi-level cells

- When materials (GST) are heated (with electrical pulses) and then cooled, they form either crystalline or amorphous materials depending on the intensity and duration of the pulses; crystalline materials have low resistance (1 state) and amorphous materials have high resistance (0 state)

- Non-volatile, fast reads (~50ns), slow and energy-hungry writes; limited lifetime (~$10^8$ writes per cell), no leakage
## PCM as a Main Memory

Lee et al., ISCA 2009

<table>
<thead>
<tr>
<th></th>
<th>PCM</th>
<th>DRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Delay &amp; Timing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tRCD (cy)</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td>tCL (cy)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>tWL (cy)</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>tCCD (cy)</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>tWTR (cy)</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>tWR (cy)</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>tRTP (cy)</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>tRP (cy)</td>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>tRRDact (cy)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>tRRDpre (cy)</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td><strong>Energy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Array read (pJ/bit)</td>
<td>2.47</td>
<td>1.17</td>
</tr>
<tr>
<td>Array write (pJ/bit)</td>
<td>16.82</td>
<td>0.39</td>
</tr>
<tr>
<td>Buffer read (pJ/bit)</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>Buffer write (pJ/bit)</td>
<td>1.02</td>
<td>1.02</td>
</tr>
<tr>
<td>Background power (pJ/bit)</td>
<td>0.08</td>
<td>0.08</td>
</tr>
</tbody>
</table>
PCM as a Main Memory

Lee et al., ISCA 2009

- Two main innovations to overcome these drawbacks:
  - decoupled row buffers and non-destructive PCM reads
  - multiple narrow row buffers (row buffer cache)
Optimizations for Writes (Energy, Lifetime)

- Read a line before writing and only write the modified bits
  Zhou et al., ISCA’09

- Write either the line or its inverted version, whichever causes fewer bit-flips
  Cho and Lee, MICRO’09

- Only write dirty lines in a PCM page (when a page is evicted from a DRAM cache)
  Lee et al., Qureshi et al., ISCA’09

- When a page is brought from disk, place it only in DRAM cache and place in PCM upon eviction
  Qureshi et al., ISCA’09

- Wear-leveling: rotate every new page, shift a row periodically, swap segments
  Zhou et al., Qureshi et al., ISCA’09
Hard Error Tolerance in PCM

• PCM cells will eventually fail; important to cause gradual capacity degradation when this happens

• Pairing: among the pool of faulty pages, pair two pages that have faults in different locations; replicate data across the two pages

• Errors are detected with parity bits; replica reads are issued if the initial read is faulty
Instead of using ECC to handle a few transient faults in DRAM, use error-correcting pointers to handle hard errors in specific locations.

For a 512-bit line with 1 failed bit, maintain a 9-bit field to track the failed location and another bit to store the value in that location.

Can store multiple such pointers and can recover from faults in the pointers too.

ECC has similar storage overhead and can handle soft errors; but ECC has high entropy and can hasten wearout.
• Most PCM hard errors are stuck-at faults (stuck at 0 or stuck at 1)

• Either write the word or its flipped version so that the failed bit is made to store the stuck-at value

• For multi-bit errors, the line can be partitioned such that each partition has a single error

• Errors are detected by verifying a write; recently failed bit locations are cached so multiple writes can be avoided
FREE-p

- When a PCM block (64B) is unusable because the number of hard errors has exceeded the ECC capability, it is remapped to another address; the pointer to this address is stored in the failed block; need another bit per block

- The pointer can be replicated many times in the failed block to tolerate the multiple errors in the failed block

- Requires two accesses when handling failed blocks; this overhead can be reduced by caching the pointer at the memory controller
Title

• Bullet