Lecture 5: Refresh, Chipkill

- Topics: refresh basics and innovations, error correction

Refresh Basics

- A cell is expected to have a retention time of 64ms; every cell must be refreshed within a 64ms window
- The refresh task is broken into 8K refresh operations; a refresh operation is issued every tREFI = 7.8 us
- If you assume that a row of cells on a chip is 8Kb and there are 8 banks, then every refresh operation in a 4Gb chip must handle 8 rows in each bank
- Each refresh operation takes time tRFC = 300ns
- Larger chips have more cells and tRFC will grow

More Refresh Details

- To refresh a row, it needs to be activated and precharged
- Refresh pipeline: the first bank draws the max available current to refresh a row in many subarrays in parallel; each bank is handled sequentially; the process ends with a recovery period to restore charge pumps
- “Row” on the previous slide refers to the size of the available row buffer when you do an Activate; an Activate only deals with some of the subarrays in a bank; refresh performs an activate in all subarrays in a bank, so it can do multiple rows in a bank in parallel

Fine Granularity Refresh

- Will be used in DDR4
- Breaks refresh into small tasks; helps reduce read queuing delays (see example)
- In a future 32Gb chip, tRFC = 640ns, tRFC_2x = 480ns, tRFC_4x = 350ns – note the high overhead from the recovery period

What Makes Refresh Worse

- Refresh operations are issued per rank; LPDDR does allow per bank refresh
- Can refresh all ranks simultaneously – this reduces memory unavailable time, but increases memory peak power
- Can refresh ranks in staggered manner – increases memory unavailable time, but reduces memory peak power
- High temperatures will increase the leakage rate and require faster refresh rates (> 85 degrees C ➔ 3.9us tREFI)

Refresh Innovations

- Smart refresh (Ghosh et al.): do not refresh rows that have been accessed recently
- Elastic refresh (Stuecheli et al.): perform refresh during periods of inactivity
- Flikker (Liu et al.): lower refresh rate for non-critical pages
- Refresh pausing (Nair et al.): interrupt the refresh process when demand requests arrive
- Preemptive command drain (Mukundan et al.): prioritize requests to ranks that will soon be refreshed

Refresh Innovations II

- Concurrent refresh and accesses (Chang et al. and Zhang et al.): makes refresh less efficient, but helps reduce queuing delays for pending reads
- RAIDR (Liu et al.): profile at run-time to identify weak cells; track such weak-cell rows in retention time bins with Bloom Filters; refreshes are skipped based on membership in a retention time bin
- Empirical study (Liu et al.): use a custom memory controller on an FPGA to measure retention time in many DIMMs; shows that retention time varies with time and is a function of data in neighboring cells

Basic Reliability

- Every 64-bit data transfer is accompanied by an 8-bit (Hamming) code – typically stored in a x8 DRAM chip
- Guaranteed to detect any 2 errors and recover from any single-bit error (SEC-DED)
- Such DIMMs are commodities and are sufficient for most applications
- 12.5% overhead in storage and energy
- For a BCH code, to correct t errors in k-bit data, need an r-bit code, $r = t * \text{ceil}(\log_2 k) + 1$

Terminology

- Hard errors: caused by permanent device-level faults
- Soft errors: caused by particle strikes, noise, etc.
- SDC: silent data corruption (error was never detected)
- DUE: detected uncorrectable error
- DUE in memory caused by a hard error will typically lead to DIMM replacement
- Scrubbing: a background scan of memory (1GB every 45 mins) to detect and correct 1-bit errors

Field Studies

Schroeder et al., SIGMETRICS 2009

- Memory errors are the top causes for hw failures in servers and DIMMs are the top component replacements in servers
- Study examined Google servers in 2006-2008, using DDR1, DDR2, and FBDIMM

Field Studies

Schroeder et al., SIGMETRICS 2009

Table 1: Memory errors per year:

Platf.	Tech.	Per machine				
		CE Incid. (%)	CE Rate Mean	CE Rate C.V.	CE Median Affct.	UE Incid. (%)
A	DDR1	45.4	19,509	3.5	611	0.17
B	DDR1	46.2	23,243	3.4	366	–
C	DDR1	22.3	27,500	17.7	100	2.15
D	DDR2	12.3	20,501	19.0	63	1.21
E	FBD	–	–	–	–	0.27
F	DDR2	26.9	48,621	16.1	25	4.15
Overall	–	32.2	22,696	14.0	277	1.29

Platf.	Tech.	Per DIMM				
		CE Incid. (%)	CE Rate Mean	CE Rate C.V.	CE Median Affct.	UE Incid. (%)
A	DDR1	21.2	4530	6.7	167	0.05
B	DDR1	19.6	4086	7.4	76	–
C	DDR1	3.7	3351	46.5	59	0.28
D	DDR2	2.8	3918	42.4	45	0.25
E	FBD	–	–	–	–	0.08
F	DDR2	2.9	3408	51.9	15	0.39
Overall	–	8.2	3751	36.3	64	0.22

Field Studies

Schroeder et al., SIGMETRICS 2009

- A machine with past errors is more likely to have future errors
- 20% of DIMMs account for 94% of errors
- DIMMs in platforms C and D see higher UE rates because they do not have chipkill
- 65-80% of uncorrectable errors are preceded by a correctable error in the same month – but predicting a UE is very difficult
- Chip/DIMM capacity does not strongly influence error rates
- Higher temperature by itself does not cause more errors, but higher system utilization does
- Error rates do increase with age; the increase is steep in the 10-18 month range and then flattens out

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

RAID-like DRAM

Udipi et al., ISCA'10

- 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

SSC-DSD

- 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

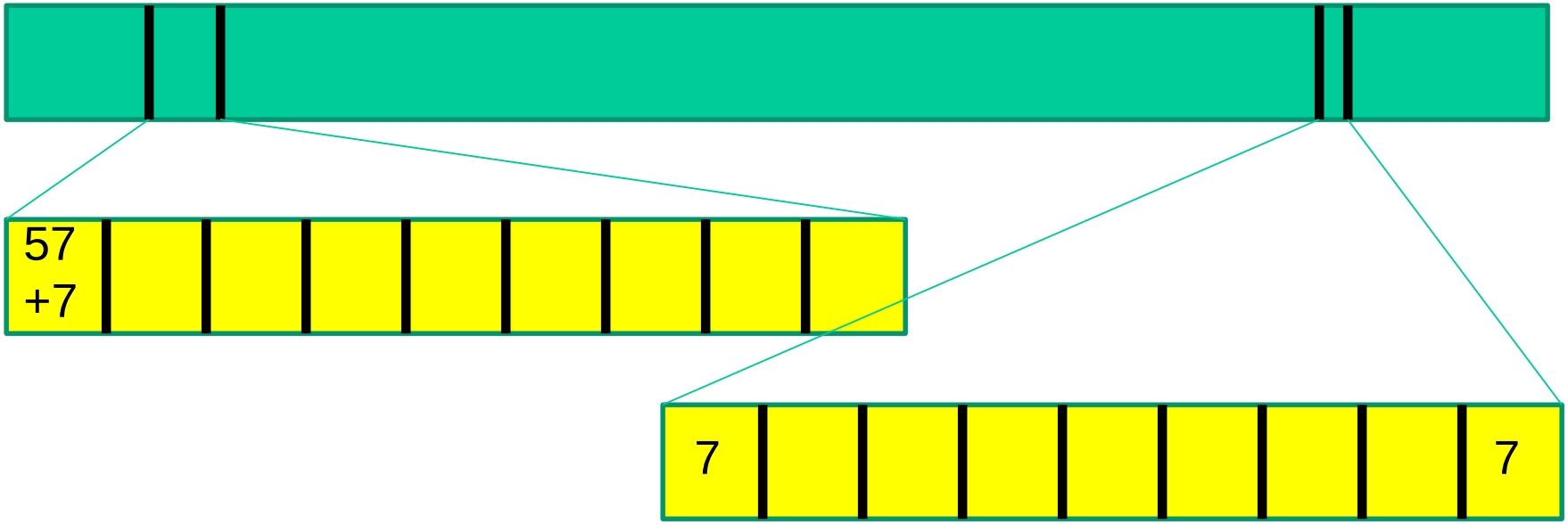
Yoon and Erez, ASPLOS'10

- 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

Udipi et al., ISCA 2012

- 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%



Title

- Bullet