Last time…

- Dynamic RAM (DRAM) is the main form of main memory storage in use today
  - Holds values on small capacitors, need refreshing (hence dynamic)
  - Slow multi-step access: precharge, read row, read column
- Static RAM (SRAM) is faster but more expensive
  - Used to build on-chip memory for caches
- Cache holds a small set of values in fast memory (SRAM) close to the processor
  - Need to develop a search scheme to find the values in cache and a replacement policy to make space for newly accessed locations
- Caches exploit two forms of predictability in memory reference streams
  - Temporal locality, same location likely to be accessed again soon
  - Spatial locality, neighboring location likely to be accessed soon

Some Basics (Again)

- Block: the unit of access/storage in cache
- Word: the unit of access by CPU
- A block contains multiple words.
  - Why multiple words?
- On cache miss,
  - Memory access
  - Cache block replacement
  - Why keep it?
- Five things to decide
  - After fetching a block from the memory, where do we place it inside the cache?
  - If the line is taken or the cache is full already, which block to evict?
  - How many words per block?
  - How big?
  - What happens on write?

Placement Policy

- Direct-Mapped Cache
  - Tag
  - Index
  - Block Offset
  - Block: \( V, t, k, b \)
  - Data Word or Byte
  - \( V \) = \( k \) \( \mod 2^n \)

- 2-Way Set-Associative Cache
  - Tag
  - Index
  - Block Offset
  - Block: \( V, t, k, b \)
  - Data Word or Byte
  - \( V \) = \( t \)
Fully Associative Cache

Block Size and Spatial Locality

CPU-Cache Interaction

Improving Cache Performance

Replacement Policy

Serial-versus-Parallel Cache and Memory access
## CSE 490/590 Administrivia

- **Feedback on lectures**
  - If you have any feedback/concern, please send it along to me
  - Thanks to those who already did
  - Please ask questions if things are not clear
  - Or you can simply scream, “TOO FAST!”
  - Please utilize my office hours (I will change to sometime in the afternoon)
- **Very important to attend**
  - Recitations this week & next week
- **Quiz 1**
  - Fri, 2/11
  - Closed book, in-class
  - Includes lectures until last Monday (1/31)
  - Review: Wed (2/9)

## Acknowledgements

- These slides heavily contain material developed and copyright by
  - Krste Asanovic (MIT/UCB)
  - David Patterson (UCB)
- And also by:
  - Arvind (MIT)
  - Joel Emer (Intel/MIT)
  - James Hoe (CMU)
  - John Kubiatowicz (UCB)
- MIT material derived from course 6.823
- UCB material derived from course CS252