Today’s Outline

- Admin:
  - HW #5 due Thursday, May 20 at 11:45pm
  - Midterm #2, Wed May 19th.
- Memory Hierarchy and Locality
- B-Trees

Why do we need to know about the memory hierarchy/locality?

- One of the assumptions that Big-Oh makes is that all operations take the same amount of time.
- Is that really true?

Definitions

- **Cycle** – (for our purposes) the time it takes to execute a single simple instruction. (ex. Add 2 registers together)
- **Memory Latency** – time it takes to access memory

Moore’s Law
Processor-Memory Performance Gap

- x86 CPU speed (100x over 10 years)

What can be done?
- **Goal**: Attempt to reduce the number of accesses to the slower levels.
- **How**?

Locality

**Temporal Locality** (locality in time) – If an item is referenced, it will tend to be referenced again soon.

**Spatial Locality** (locality in space) – If an item is referenced, items whose addresses are close by will tend to be referenced soon.

Caches

- Each level is a sub-set of the level below.

- **Cache Hit** – address requested is in cache
- **Cache Miss** – address requested is NOT in cache
- **Cache line size** (chunk size) – the number of contiguous bytes that are moved into the cache at one time

Examples

\[
\begin{align*}
\text{x} &= a + 6; & \text{x} &= a[0] + 6; \\
\text{y} &= a + 5; & \text{y} &= a[1] + 5; \\
\text{z} &= 8 \times a; & \text{z} &= 8 \times a[2]; 
\end{align*}
\]

Locality and Data Structures

- Which has (at least the potential for) better spatial locality, arrays or linked lists?