Memory Hierarchy

CSE 373
Data Structures & Algorithms
Ruth Anderson

Today's Outline

- **Admin**:
  - HW #5 due Tuesday, Feb 24th at 11:45pm
  - If you are working with a partner, must email Sean by Friday, Feb 20th at 11:45pm
  - Midterm #2 next Friday, Feb 27th
- **Hashing**
- **Memory Hierarchy and Locality**

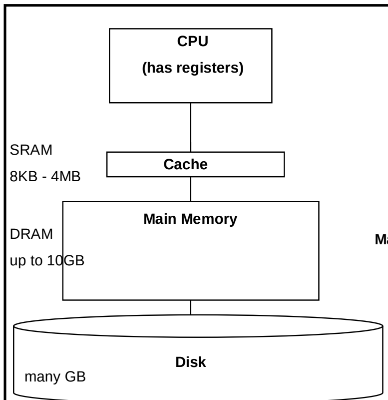
Why do we need to know about the memory hierarchy/locality?

- 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



Time to access:

1 ns per instruction

Cache

2-10 ns

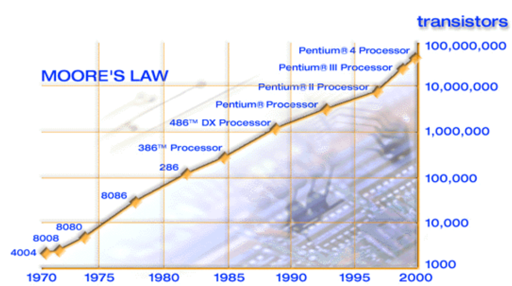
Main Memory

40-100 ns

Disk a few *milli*seconds

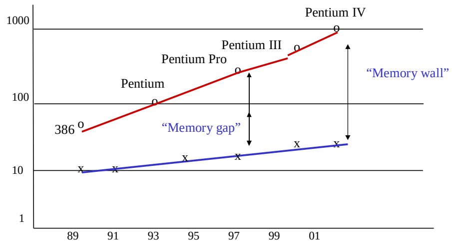
(5-10 *Million* ns)

Moore's Law



Processor-Memory Performance Gap

- x86 CPU speed (100x over 10 years)



2/18/2009 7

What can be done?

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

2/18/2009 8

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.

2/18/2009 9

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

2/18/2009 10

Examples

```
x = a + 6;        x = a[0] + 6;
y = a + 5;        y = a[1] + 5;
z = 8 * a;        z = 8 * a[2];
```

2/18/2009 11

Locality and Data Structures

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

2/18/2009 12