

## Topics for today:

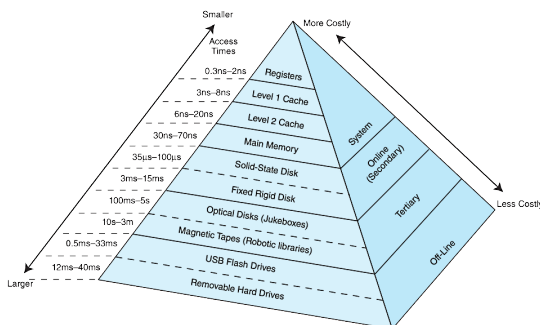
### Memory management

- Cache memory

## Memory vs. Hard disks

- What's the difference?
  - Memory is smaller (less capacity in bytes)
  - Memory is faster (costs fewer CPU cycles to access it)
  - Memory is more expensive per byte

## The Memory Hierarchy



## Some terminology

- If it's there and you can see it, it's *real*.
- If it's there and you can't see it, it's *transparent*.
- If it's not there and you can see it, it's *virtual*.
- If it's not there and you can't see it, you erased it.

### Cache memory

Using a *cache* is a way to improve speed.

Accessing main memory is slow compared to the CPU (or, at least, fast memory is expensive). Cache memory is high-speed memory—smaller, better technology, and physically closer to the CPU than main memory.

### Locality principle

Use of a cache is effective because of the *locality principle*:

Usually, when the CPU references one address in memory, the next reference is a nearby address.

*(Examples?)*

### Cache Hits vs. Misses

Cache hit – when you need a location and it turns out it's in the cache already

Cache miss – when you need a location, and it's not in the cache

Goal: make hit rate high

Cost of cache hit: lowish

Cost of cache miss: high