Processor Architecture Memory System Basics

Last modified 4/4/24

Memory System Basics

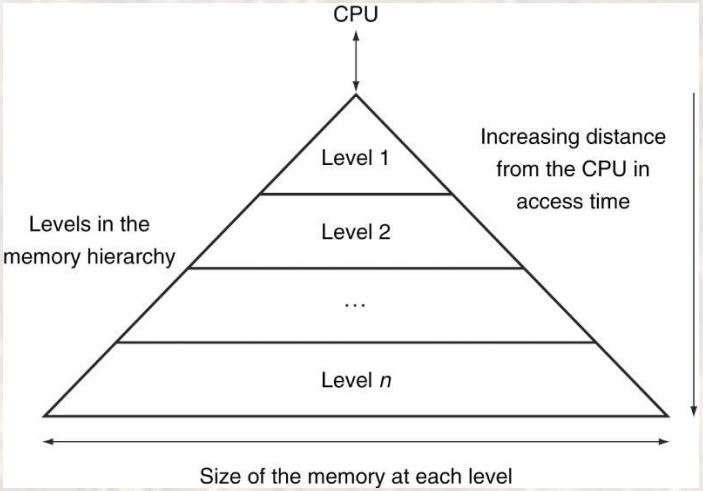
- Memory Hierarchy
  - Typical System

    Registers

    Cache (SRAM)

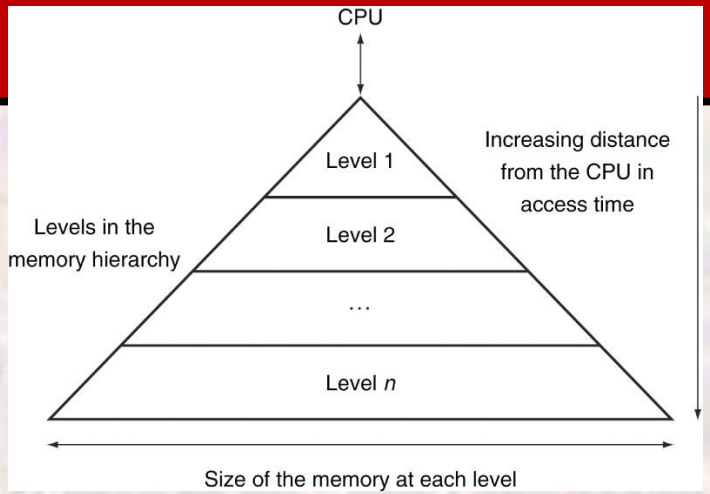
    Main Memory (DRAM)

    Storage (HDD, SSD or Flash)



  - Advanced systems may have 2,3,4 levels of cache
    - Each is progressively slower and larger
    - Size is targeted at holding entire applications

Memory System Basics



- Memory Hierarchy
  - Typical System - 2GHz

    Registers - 1 Clk access - 32

    Cache (SRAM) – 1 – 20 Clk access – 32KB – 16MB

    Main Memory (DRAM) – 30 - 50 Clk access – 4GB – 64GB

    Storage (HDD) – 1M - 10M Clk access – 128GB – 2TB
    Storage (SSD) – 1K - 10K Clk access – 16GB – 256GB

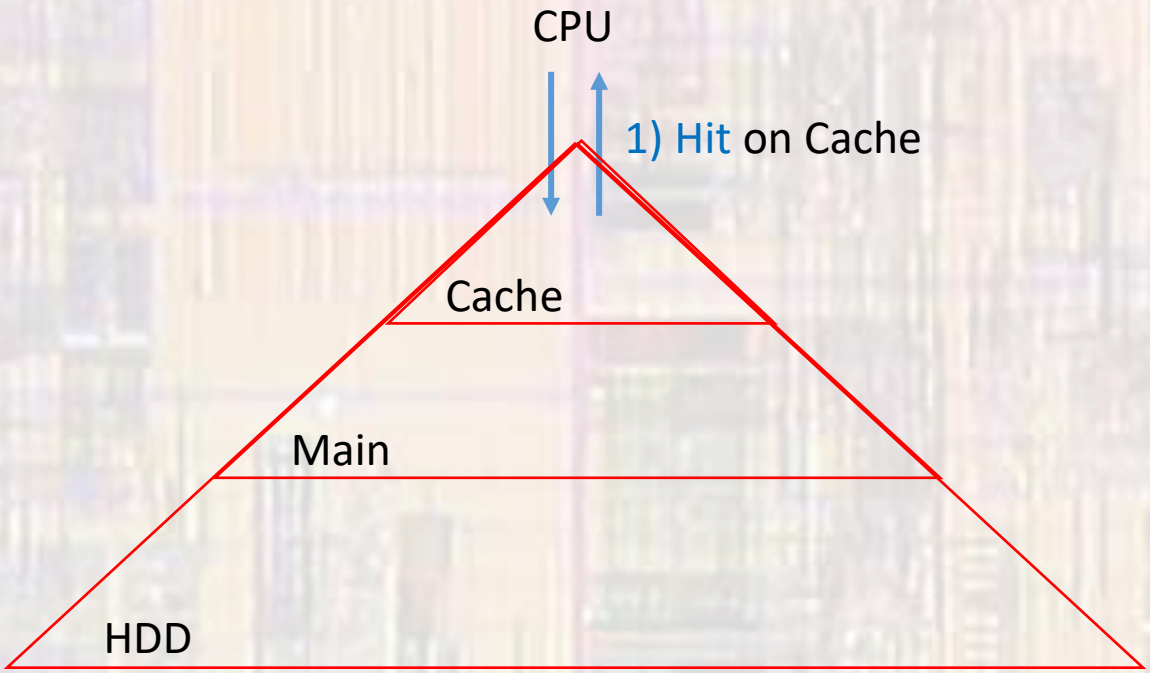
  - Advanced systems may have 2,3,4 levels of cache
    - Each is progressively slower and larger
    - Size is targeted at holding entire applications

Memory System Basics

- Memory Hierarchy

CPU accesses something already in the cache

2 clks

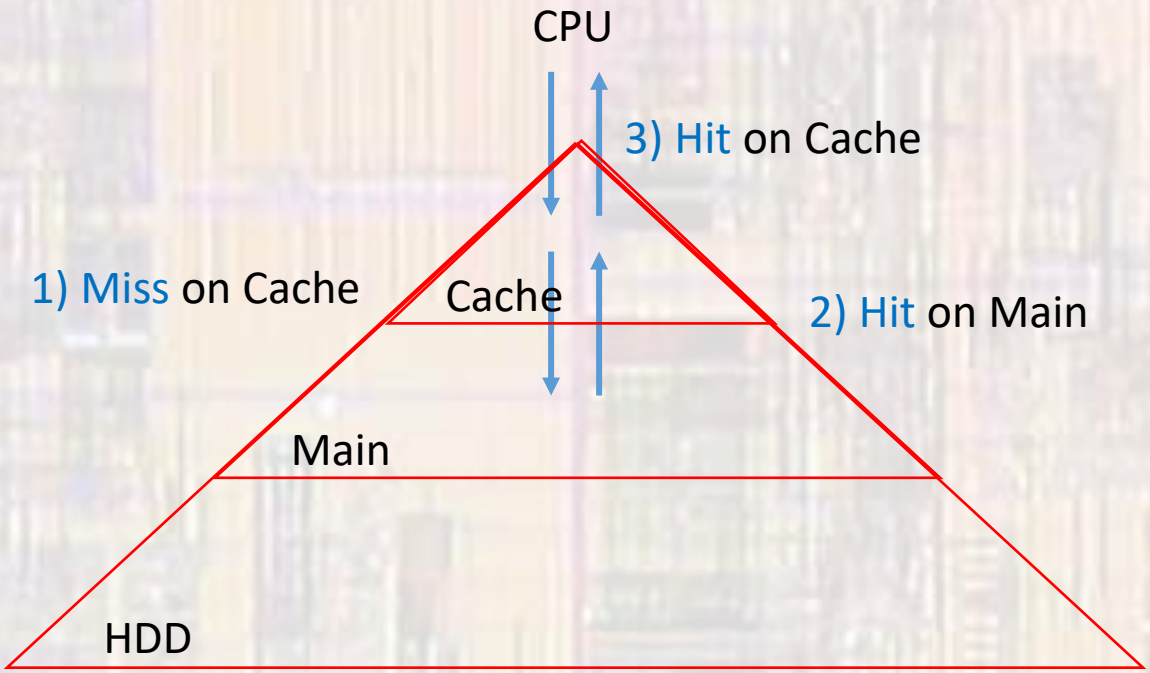


Memory System Basics

- Memory Hierarchy

CPU accesses something not in cache, but in main memory

2 + 40 + 2 clks

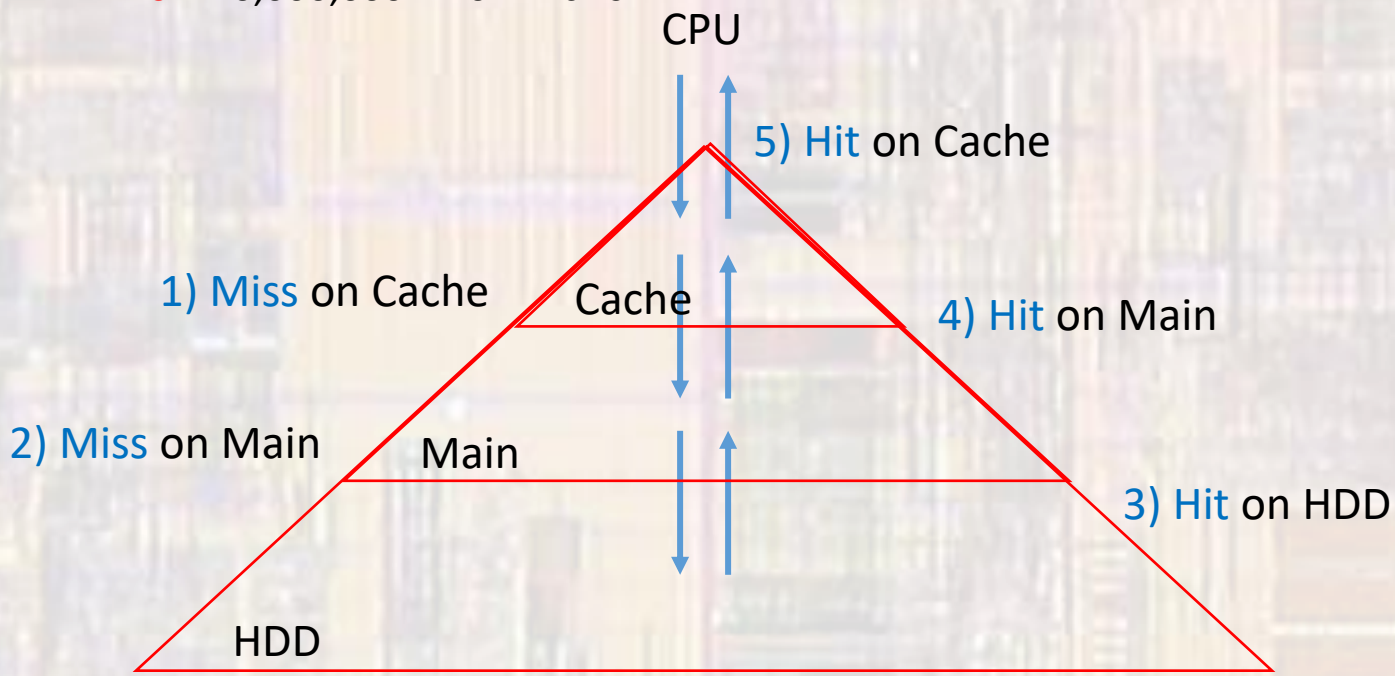


Memory System Basics

- Memory Hierarchy

CPU accesses something
not in cache or main memory,
but on the HDD

2 + 40 + 10,000,000 + 40 + 2 clks



Memory System Basics

- At time of system configuration
  - BIOS
    - Separate chip or portion of a chip with non-volatile memory
    - Stores specifics of the hardware and software to perform startup
    - Pointed to by the reset/power-up vector in the processor
  - Swap Partition
    - Section of the HDD or SSD set aside for use by the processor to optimize performance
    - Designed to allow fastest possible access
      - Contiguous memory (large burst transactions)
      - In the case of HDD – fastest access portion of the physical disk

Memory System Basics

- Startup
  - Early startup
    - BIOS is read
    - Patches or fixes are applied
    - Limited hardware is powered up and configured
  - Middle startup
    - Firmware is copied to the Main Memory (DRAM)
    - Firmware starts executing
    - Additional hardware is powered up and configured
  - Late startup
    - OS is loaded from the HDD or SSD

Memory System Basics

- Application Loading (including the OS)
  - Applications are stored on the HDD or SSD
  - When requested
    - The application is copied from HDD or SSD to the Main Memory (SDRAM)
    - If it does not fit, a portion of the application is copied
  - During operation, if additional portions of the application are required – or additional applications are requested
    1. Currently unused portions are written back to the HDD/SSD to the Swap partition (not their original location) to free up space on the SDRAM
    2. The newly required portion is loaded into Main Memory
  - This continues as new applications are started

Memory System Basics

- Application Shutdown (including power off)
  - Any portions of Cache marked as changed are written back to the main memory
  - Any portions of main memory marked as changed are written back to the swap partition
  - Any portions of memory in the Swap partition that have been marked as changed are written back to the primary portion of the HDD/SSD