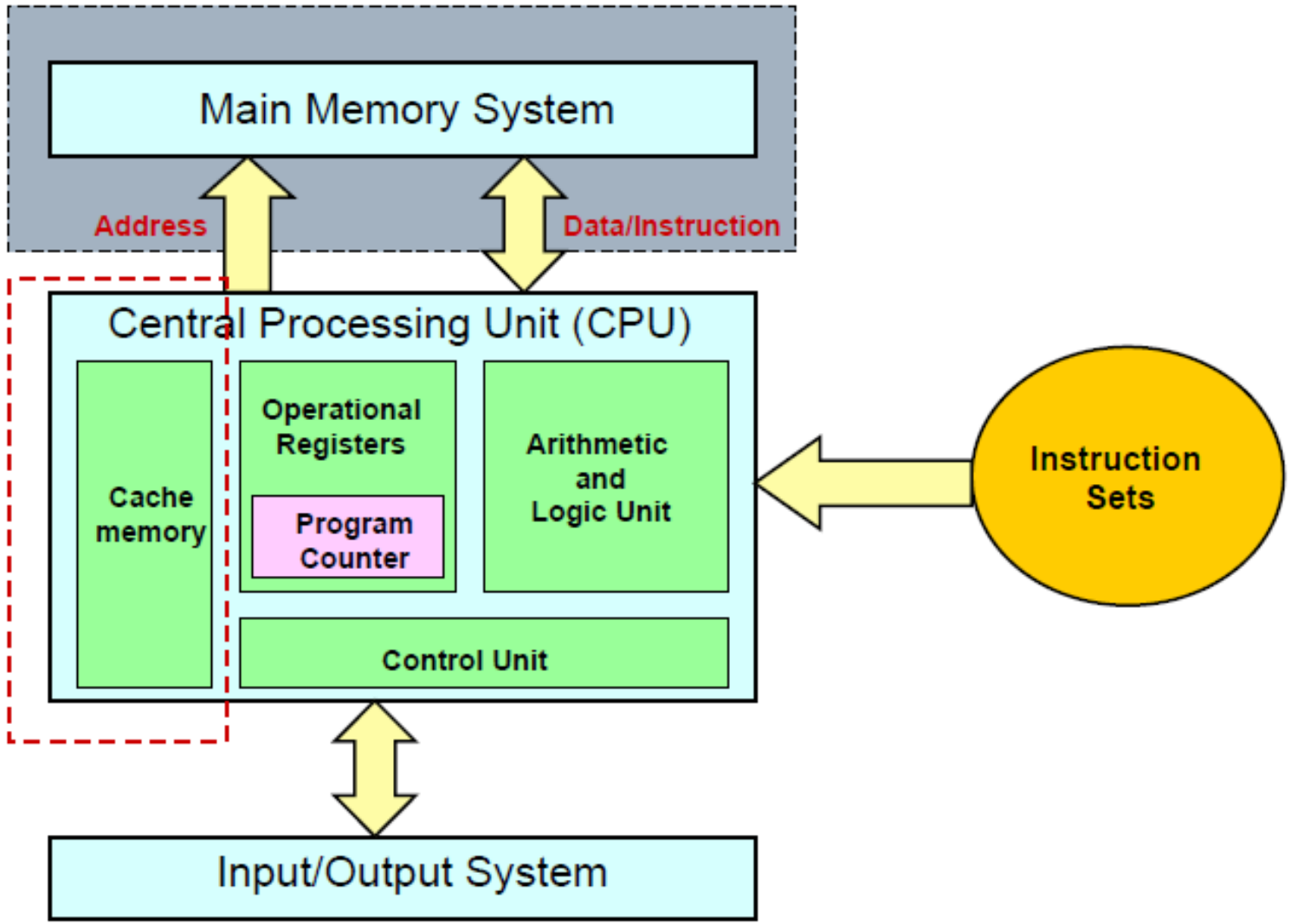
Jiwaji University
SOS in Computer Science and Applications
B.C.A. (IV Semester)
Paper -402 Advance Computer Architecture

Unit 2

Topic:

Computer Memory System
their Characteristic's

By Arun Singh


Main Memory System
Address
Data/Instruction
Central Processing Unit (CPU)
Cache memory
Operational Registers
Program Counter
Arithmetic and Logic Unit
Control Unit
Instruction Sets
Input/Output System

Basic Concepts

- The maximum size of the memory that can be used in any computer is determined by the addressing scheme
  - For example, a 16-bit computer that generates 16-bit addresses is capable of addressing up to 216=64K memory locations.
  - Similarly, machines whose instructions generate 32-bit addresses can utilize a memory that contains up to 232=4G memory locations
- Most modern computers are byte addressable Form the system standpoint, we can view the memory unit as a block box
  - Data transfer between the memory and processor takes place through the use of two processor registers, MAR and MDR
- A useful measure of the speed of memory units is the time that elapses between the initiation of an operation and the completion of that operation. This is referred to as the memory access time.
- Another important measure is the memory cycle time, which is the minimum time delay required between the initiation of two successive memory operations
- A memory unit is called random-access memory (RAM) if any location can be accessed for a Read or Write peration in some fixed amount of time that is independent of the location's address
- The memory cycle time is the bottleneck in the systemOne way to reduce the memory access time is to use a cache memory
- Cache memory is a small, fast memory that is inserted between the larger, smaller main memory and the processor.
- Virtual memory is used to increase the apparent size of the physical memory. Data are addressed in a virtual address space that can be as large as the addressing capability of the processor. But at any given time, only the active portion of this space is mapped onto locations in the physical memory. The remaining virtual addresses are mapped onto the bulk storage devices used, such as magnetic disks

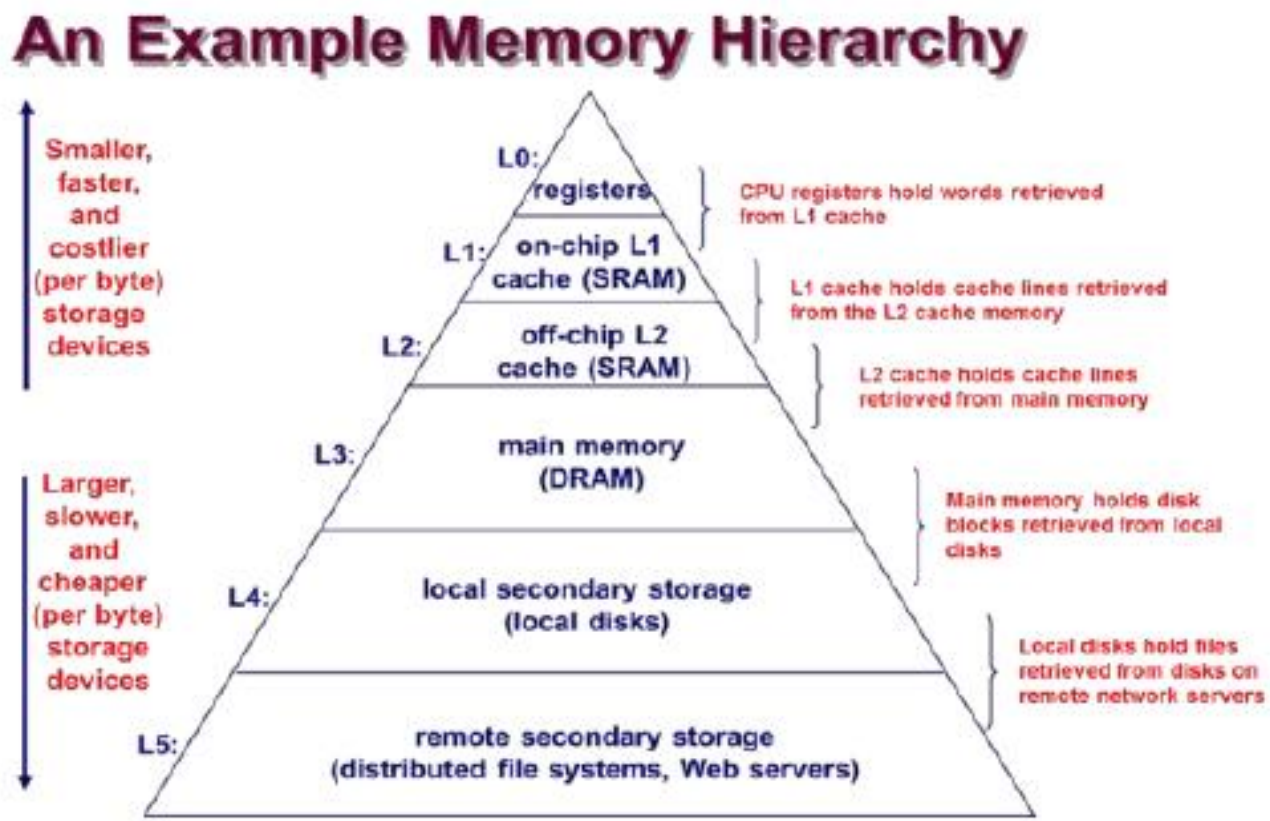
Microcomputer Memory

• Memory is an essential component of the microcomputer system.
• It stores binary instructions and datum for the microcomputer.
• The memory is the place where the computer holds current programs and data that are in use.
• None technology is optimal in satisfying the memory requirements for a computer system.
• Computer memory exhibits perhaps the widest range of type, technology, organization, performance and cost of any feature of a computer system.
• The memory unit that communicates directly with the CPU is called main memory.
• Devices that provide backup storage are called auxiliary memory or secondary memory

Memory Hierarchies

- Some fundamental and enduring properties of hardware and software:
- Fast storage technologies cost more per byte and have less capacity
- Gap between CPU and main memory speed is widening
- Well-written programs tend to exhibit good locality
- These fundamental properties complement each other beautifully
- They suggest an approach for organizing memory and storage systems known as a memory hierarchy

An Example Memory Hierarchy



Characteristics of memory systems

• The memory system can be characterized with their Location, Capacity, Unit of transfer, Access method, Performance, Physical type, Physical characteristics, Organization.

Location

• Processor memory: The memory like registers is included within the processor and termed as processor memory.
• • Internal memory: It is often termed as main memory and resides within the CPU.
• • External memory: It consists of peripheral storage devices such as disk and magnetic tape that are accessible to processor via i/o controllers

Capacity

• Word size: Capacity is expressed in terms of words or bytes.
— The natural unit of organization
• Number of words: Common word lengths are 8, 16, 32 bits etc.
— or Byte

Unit of Transfer

• Internal: For internal memory, the unit of transfer is equal to the number of data lines into and out of the memory module.
• External: For external memory, they are transferred in block which is larger than a word.
• Addressable unit
— Smallest location which can be uniquely addressed
— Word internally
— Cluster on Magnetic disks

Access Method

• Sequential access: In this access, it must start with beginning and read through a specific linear sequence. This means access time of data unit depends on position of records (unit of data) and previous
location. — e.g. tape

• Direct Access: Individual blocks of records have unique address based on location. Access is accomplished by jumping (direct access) to general vicinity plus a sequential search to reach the final location. — e.g. disk

• Random access: The time to access a given location is independent of the sequence of prior accesses and is constant. Thus any location can be selected out randomly and directly addressed and accessed. — e.g. RAM

• Associative access: This is random access type of memory that enables one to make a comparison of desired bit locations within a word for a specified match, and to do this for all words simultaneously.

• e.g. cache

Performance

• Access time: For random access memory, access time is the time it takes to perform a read or write operation i.e. time taken to address a memory plus to read / write from addressed memory location. Whereas for non-random access, it is the time needed to position read / write mechanism at desired location.

— Time between presenting the address and getting the valid data

• Memory Cycle time: It is the total time that is required to store next memory access operation from the previous memory access operation.

Memory cycle time = access time plus transient time (any additional time required before a second access can commence).

— Time may be required for the memory to “recover” before next access

— Cycle time is access + recovery

• Transfer Rate: This is the rate at which data can be transferred in and out of a memory unit.
— Rate at which data can be moved
— For random access, R = 1 / cycle time
— For non-random access, Tn = Ta + N / R; where Tn – average time to read or write N bits, Ta – average access time, N – number of bits, R – Transfer rate in bits per second (bps).

Physical Characteristics

• Decay: Information decays mean data loss.
• Volatility: Information decays when electrical power is switched off.
• Erasable: Erasable means permission to erase.
• Power consumption: how much power consumes?

Organization

• Physical arrangement of bits into words
• Not always obvious
- e.g. interleaved