Computer Organization and Embedded System

Computer Arithmetic and Memory System (ACtE0402)

Ishwar Kumar Singh
Computer Engineer
Government of Nepal

4.2 Computer Arithmetic and Memory System (ACtE0402):

- Arithmetic and Logical operation,
- The Memory Hierarchy,
- Internal and External memory,
- Memory Write Ability,
- Storage Permanence,
- Composing Memory
- Cache memory principles,
- Elements of Cache design –
  - Cache size,
  - Mapping function,
  - Replacement algorithm,
  - Write policy,
  - Number of caches,

Memory

- Memory is an essential component of the microcomputer system which stores binary instructions and data for the microcomputer.
- The memory is the place where the computer holds current programs and data that are in use.
- 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.

Characteristics of Memory System

The memory system can be characterized with their

- Location
- Capacity
- Unit of transfer
- Access method
- Performance
- Physical type
- Physical characteristics

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

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

Unit of Transfer

- For **Internal Memory**, the unit of transfer is equal to the number of data lines into and out of the memory module.
- For **External Memory**, they are transferred in block which is larger than a word.

Access Method

- **Sequential Access:**
  - In this access, it 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.
  - Example: Tape

- **Direct Access:**
  - In this 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.
  - Example: Disk

Access Method

- **Random Access:**
  - The time to access a given location is independent of the sequence of prior accesses and is constant.
  - Any location can be selected out randomly and directly addressed and accessed.
  - Example: RAM

- **Associative Access:**
  - Associative Access is a special type of random access method which enables comparison of desired bit locations within a word for a specific match and to do this for all words simultaneously.
  - Thus based on portion of word's content, word is retrieved rather than its address.
  - Example: Cache

Performance

- **Memory access time:** The amount of time taken by the processor to read data, instructions, and information from memory.
- **Memory Cycle time:** It is the total time that is required to store next memory access operation from the previous memory access operation.
  - Any additional time required before a second access can commence.
  - Time may be required for the memory to "recover" before next access i.e. Cycle time = access time + recovery time
- **Transfer Rate:** This is the rate at which data can be transferred in and out of a memory unit.

Physical Types

- Semiconductor: RAM
- Magnetic: Disk & Tape
- Optical: CD & DVD
- Others

Physical Characteristics

- Decay: Data can be lost
- Volatile: Data will be lost if power is switched off.
- Erasable: Permission to Erase
- Power Consumption

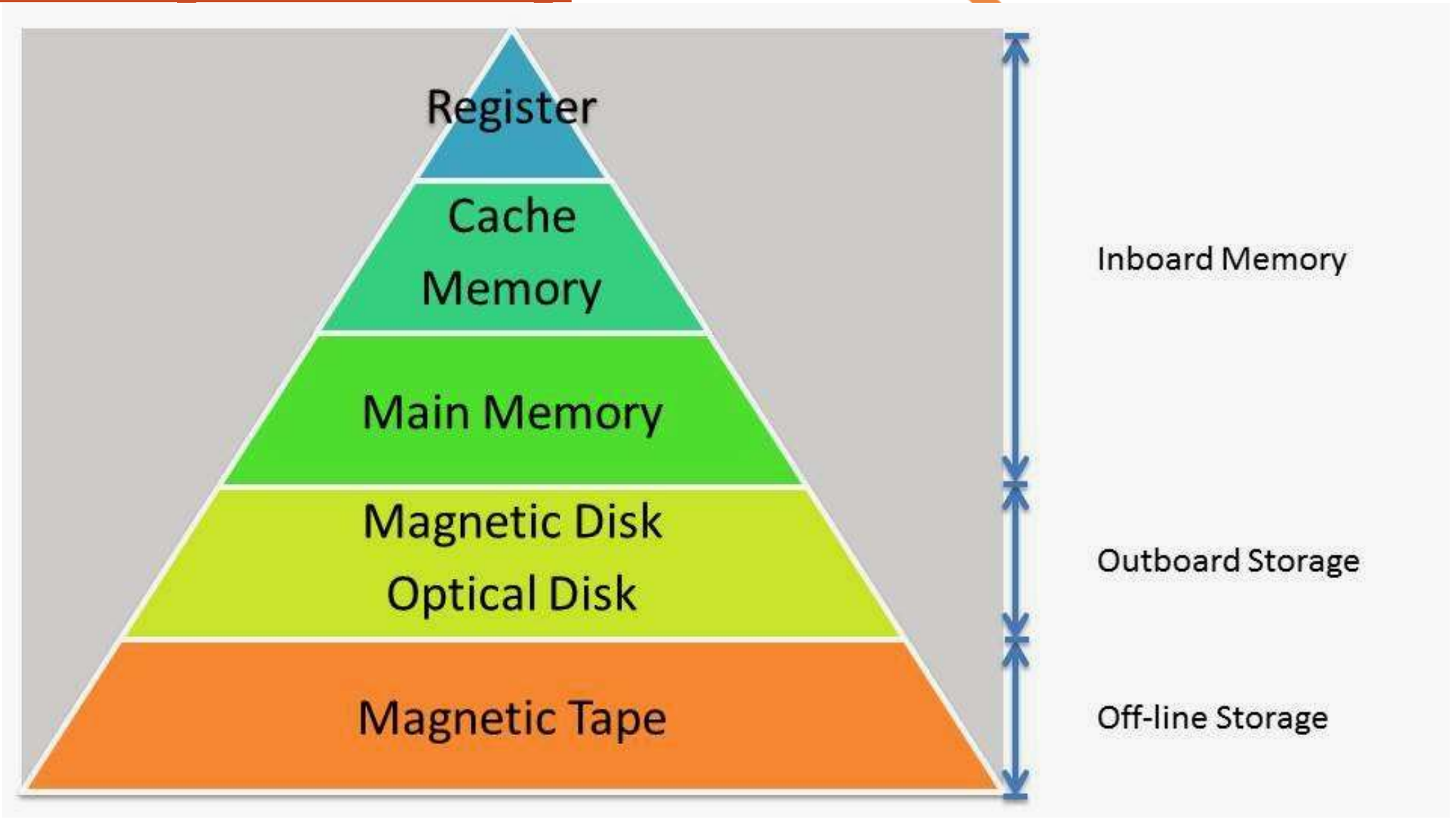
Memory Hierarchy

- Capacity, cost and speed of different types of memory play a vital role while designing a memory system for computers
- There is a tradeoff between these three characteristics cost, capacity and access time. One cannot achieve all these quantities in same memory module because
  - If capacity increases, access time increases (slower) and due to which cost per bit decreases.
  - If access time decreases (faster), capacity decreases and due to which cost per bit increases.
- The designer tries to increase capacity because cost per bit decreases and the more application program can be accommodated.
  - But at the same time, access time increases and hence decreases the performance.

Memory Hierarchy

- Memory Hierarchy is to obtain the highest possible access speed while minimizing the total cost of the memory system.
- The memory hierarchy system consists of all storage devices employed in a computer system from the slow but high-capacity auxiliary memory to a relatively faster main memory, to an even smaller and faster cache memory
- The memory unit that directly communicate with CPU is called the main memory and devices that provide backup storage are called auxiliary memory
- A special very-high-speed memory called cache is used to increase the speed of processing by making current programs and data available to the CPU at a rapid rate.

Memory Hierarchy



Memory Write Ability

- Write ability is the manner and speed that a particular memory can be written. The ranges of write ability are:

  1. High End (Processor writes to memory quickly – RAM, Register)
  2. Middle Range (Processor writes to memory slowly – FLASH, EEPROM)
  3. Lower Range (Special equipment used to write to memory – EPROM , PROM)
  4. Low End (bits stored only during fabrication – MROM)

- A basic writing process involves providing address values to the address lines and data to the data lines and selecting write function.

Memory Write Ability

High End:

- The memory with high end write ability is the easiest one to write data into.
- These are **flip-flop based memory** that a processor can write directly into.
- Examples are **Registers and RAM**, which are also the memory closest to the processor and used by the processor during its computations.

Memory Write Ability

Middle Range:

- The memories in this range are a little bit difficult to write and are slower than high end ones.
- The processor can still write in them, but with a slower speed.
- These memories are not accessed frequently and used for storing data for a longer period of time.
- Examples are **FLASH, EEPROM**.
- The middle range memories can be used in design and testing phase of an embedded system.

Memory Write Ability

Lower Range:

- A **Special Programmer Device** is required to write data into this type of memory.
- It is slower memory.
- Examples:
  - **EPROM** in which data is written by using voltages higher than that of normal operation,
  - **PROM** in which data is written by blowing connections which represent bit values and can be programmed only once.

Memory Write Ability

Low End:

- In the Low End memory devices, data is written only during manufacturing.
- The data writing process starts with the design of the chip, locations and the data to be held by the location, and completes with actual manufacturing of the chip.
- Once manufactured, the data cannot be re-written.
- Example: **MROM**
- In embedded systems, such memories can be used to hold program or some data that are used very frequently, but the values to be stored must be final.

Storage Permanence

- Storage permanence is the ability of memory to hold its stored bits after they have been written. The ranges are :-

  1. High End (Never loss bits – MROM, PROM)

  2.Middle Range (holds bits for days or years after power cut off – NVRAM)

  3. Lower Range (holds bits as long as power is supplied – SRAM)

  4. Lower End (lose bits immediately after written – DRAM)

- The memory based upon the **combinational logic circuit** stores data almost permanently while **RAM** cannot hold it even for a short period of time without power.

Storage Permanence

High End:

- Once the data are written, the high end memory devices can hold data almost permanently.
- Examples are **MROM, PROM**.
- The program of an embedded system can be put into this type of memory.

Storage Permanence

Middle Range :

- The memory with middle range storage permanence can store for certain period of time.
- The examples are **NVRAM**, a battery-backed RAM, or Flash Memory.

Storage Permanence

Lower Range :

- Lower range memory devices can hold data as long as power is available.
- The data are lost once the power supply is plugged off.
- These are used by the processor during its operations, the data are not written for storage purpose.
- An example is **SRAM** in which bits are stored in transistors/flip flops which can hold data as long as power is available.

Storage Permanence

Low End:

- From data retention point of view, the low end memory devices are the worst ones.
- They lose data even when power is available.
- The stored data are represented by charge stored in capacitor which, practically, loses its charge if not refreshed periodically.
- Example: **DRAM**.
- They are used because the packing density is high and occupies less space.

Thank You.