RAM

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

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 Bytes

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 Types

- Semiconductor

— RAM

- Magnetic — Disk & Tape
- Optical — CD & DVD
- Others

— Bubble — Hologram

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

The Memory Hierarchy

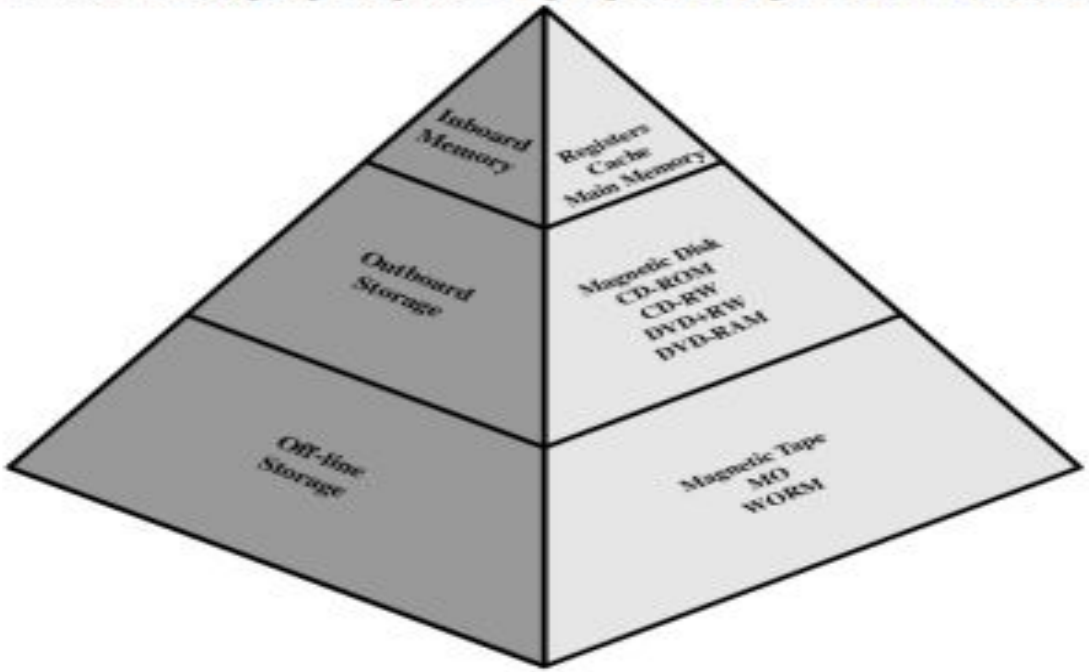
- Capacity, cost and speed of different types of memory play a vital role while designing a memory system for computers.
- If the memory has larger capacity, more application will get space to run smoothly.
- It's better to have fastest memory as far as possible to achieve a greater performance. Moreover for the practical system, the cost should be reasonable.
- 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.

So the best idea will be to use memory hierarchy

- Memory Hierarchy is to obtain the highest possible access speed while minimizing the total cost of the memory system.
- Not all accumulated information is needed by the CPU at the same time.
- Therefore, it is more economical to use low-cost storage devices to serve as a backup for storing the information that is not currently used by CPU
- The memory unit that directly communicate with CPU is called the main memory
- Devices that provide backup storage are called auxiliary memory
- The memory hierarchy system consists of all storage devices employed in a computer system from the slow by high-capacity auxiliary memory to a relatively faster main memory, to an even smaller and faster cache memory

CONT…

- The main memory occupies a central position by being able to communicate directly with the CPU and with auxiliary memory devices through an I/O processor
- 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
- CPU logic is usually faster than main memory access time, with the result that processing speed is limited primarily by the speed of main memory
- The cache is used for storing segments of programs currently being executed in the CPU and temporary data frequently needed in the present calculations
- The memory hierarchy system consists of all storage devices employed in a computer system from slow but high capacity auxiliary memory to a relatively faster cache memory accessible to high speed processing logic. The figure below illustrates memory hierarchy.



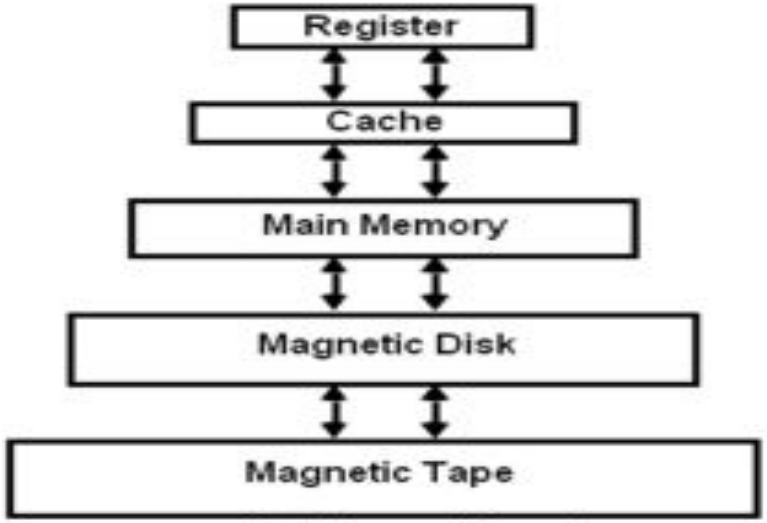


Fig: Memory Hierarchy

- **As we go down in the hierarchy**

- Cost per bit decreases
- Capacity of memory increases
- Access time increases
- Frequency of access of memory by processor also decreases.

- **Hierarchy List**

- Registers
- L1 Cache
- L2 Cache
- L3 Cache
- Main memory
- Disk cache
- Disk
- Optical
- Tape

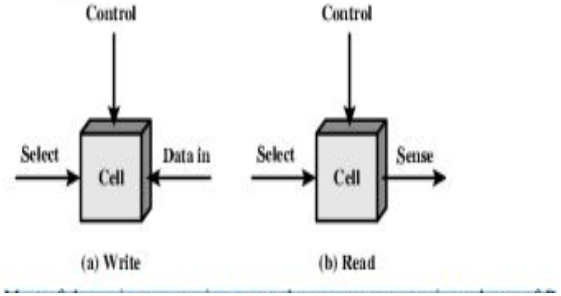
Internal and External memory

- Internal or Main Memory

The main memory is the central unit of the computer system. It is relatively large and fast memory to store programs and data during the computer operation. These memories employ semiconductor integrated circuits. The basic element of the semiconductor memory is the memory cell.

The memory cell has three functional terminals which carries the electrical signal. o The select terminal: It selects the cell. o The data in terminal: It is used to input data as 0 or 1 and data out or sense terminal is used for the output of the cell's state. o The control terminal: It controls the function i.e. it indicates read and write

- Most of the main memory in a general purpose computer is made up of RAM integrated circuits chips, but a portion of the memory may be constructed with ROM chips



RAM– Random Access memory :
Memory cells can be accessed for information transfer from any desired random location. The process of locating a word in memory is the same and requires of locating a word in memory is the same and requires an equal amount of time no matter where the cells are located physically in memory thus named 'Random access'. Integrated RAM are available in two possible operating modes, Static and Dynamic

Static RAM (SRAM)

- The static RAM consists of flip flop that stores binary information and this stored information remains valid as long as power is applied to the unit. Fig: SRAM structure
- Four transistors T1, T2, T3 and t4 are cross connected in an arrangement that produces a stable logical state. In logic state 1, point C1 is high and point C2 is low. In this state, T1 & T4 are off and T2 & T3 are on.
- In logic state 0, point C1 is low and C2 is high. In this state, T1 & T4 are on and T2 & T3 are off. The address line controls the two transistors T5 & T6. When a signal is applied to this line, the two transistors are switched on allowing for read and write operation. For a write operation, the desired bit value is applied to line B while it's complement is applied to line B complement. This forces the four transistors T1, T2, T3 & T4 into a proper state.
- For the read operation, the bit value is read from line B.

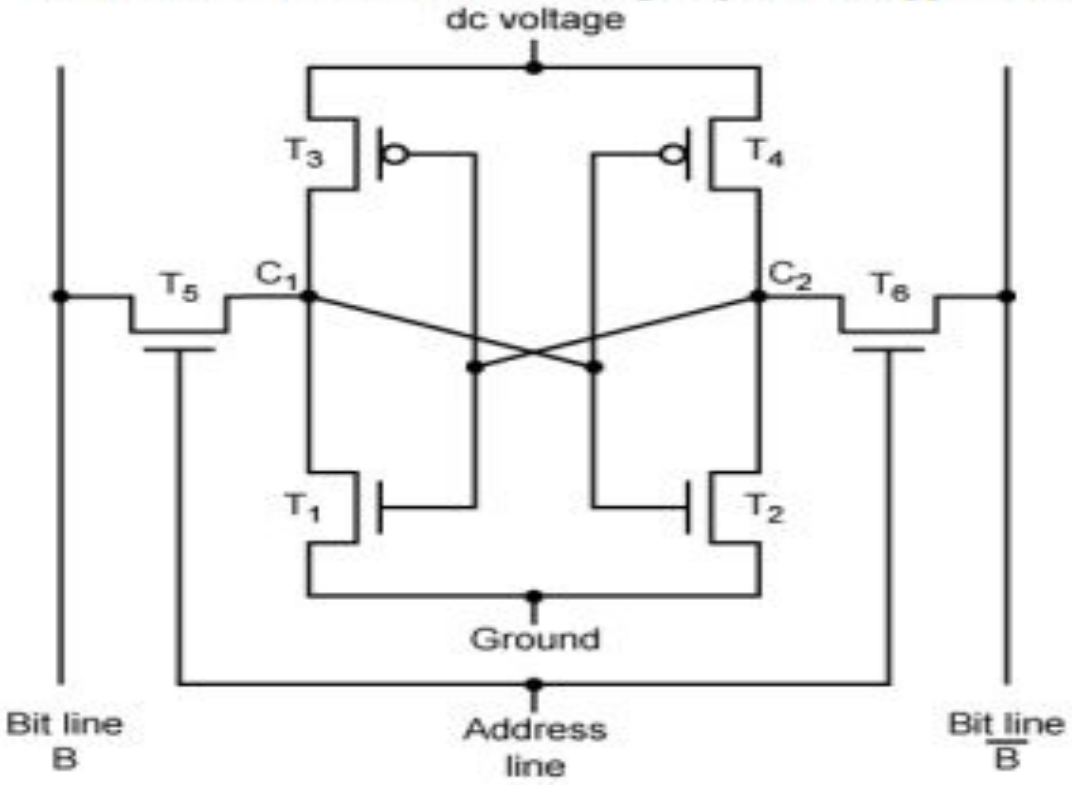
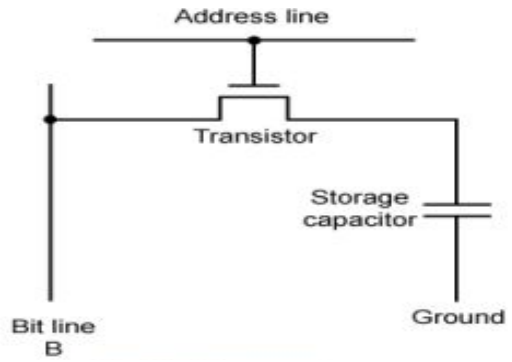


Fig: SRAM structure

Dynamic RAM (DRAM)

- The dynamic RAM stores the binary information in the form of electrical charges and capacitor is used for this purpose. Since charge stored in capacitor discharges with time, capacitor must be periodically recharged and which is also called refreshing memory. Fig: DRAM structure The address line is activated when the bit value from this cell is to be read or written. The transistor acts as switch that is closed i.e. allowed current to flow, if voltage is applied to the address line; and opened i.e. no current to flow, if no voltage is present in the address line.



SRAM versus DRAM

- Both volatile
  - Power needed to preserve data

Static RAM

- Uses flip flop to store information
- Needs more space
- Faster, digital device
- Expensive, big in size
- Don't require refreshing circuit
- Used in cache memory

Dynamic RAM

- Uses capacitor to store information
- More dense i.e. more cells can be accommodated per unit area Slower, analog device
- Less expensive, small in size
- Needs refreshing circuit
- Used in main memory, larger memory units