Part (2)
Memory system

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.

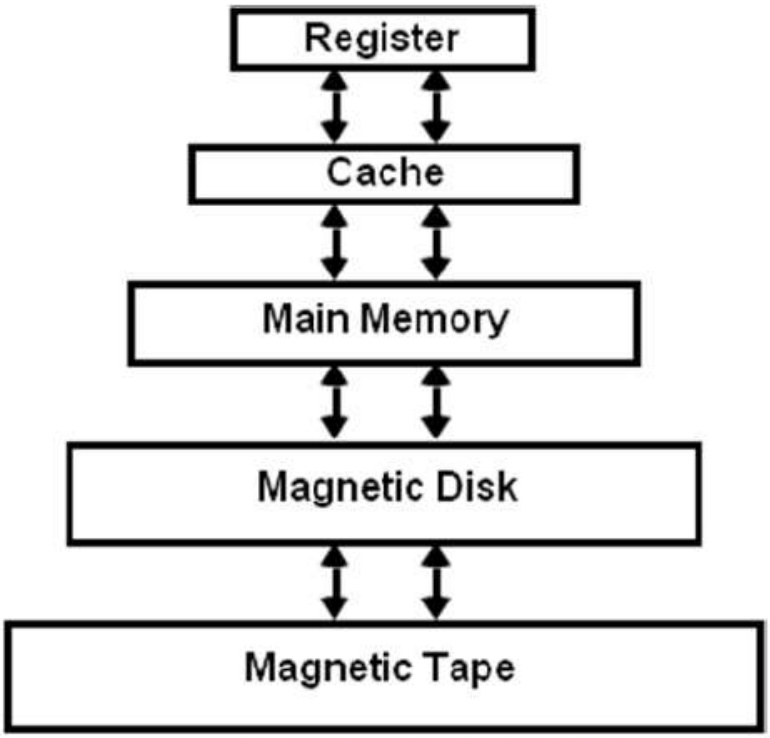


Fig: Memory Hierarchy

Note that the registers in the processing unit are temporary storage devices. They are the fastest components of the computer system memory.

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.

Thus, in a general purpose computer system, the highest speed memory is closest to the processing unit and is most expensive. The least expensive and slowest memory devices are farthest from the processing unit.

MEMORY SYSTEM characteristics:

The most important characteristics of any memory system are its ***capacity***, ***data access time***, ***the data transfer rate***, ***the cycle time***, and ***cost.***

The capacity of the storage system: Is the maximum number of units (bits, bytes, or words) of data it can store:
RAM capacity= no. of words X word size

The access time: Is the time taken by the memory module to access the data after an address is provided to the module. The data appear in the MDR at the end of this time in a RAM.

The data transfer rate: Is the number of bits per second at which the data can be read out of the memory. This rate is the product of the reciprocal of access time and the number of bits in the unit of data (data word) being read(Transfer rate=(1/ t).B)

The cycle time :Is a measure of how often the memory can be accessed.

The cost: Is the product of capacity and the price of memory device per bit. RAMs are usually more costly than other memory devices.

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

MAIN MEMORY

As the name implies, the main memory provides the main storage for a computer.
The figure below shows **a *typical interface*** between the main memory and the CPU:

*****Two CPU registers are used to interface the CPU to the main memory. These are the memory address register (**MAR**) and the memory data register (**MDR**).

The MDR is used to hold the data to be stored and/or retrieved in/from the memory location whose address is held in the MAR.

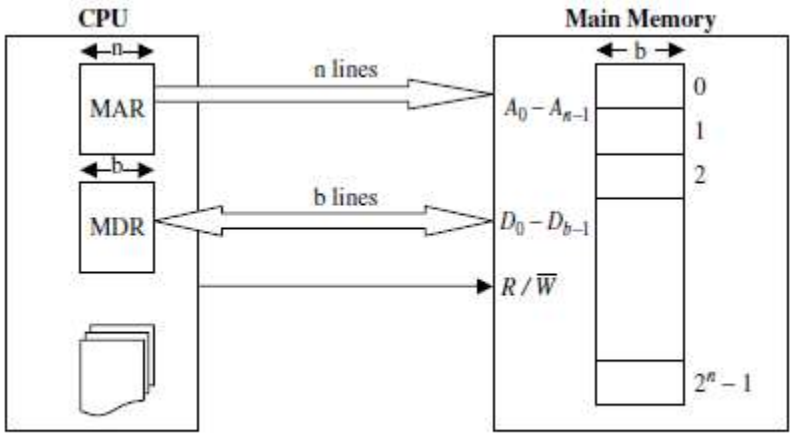


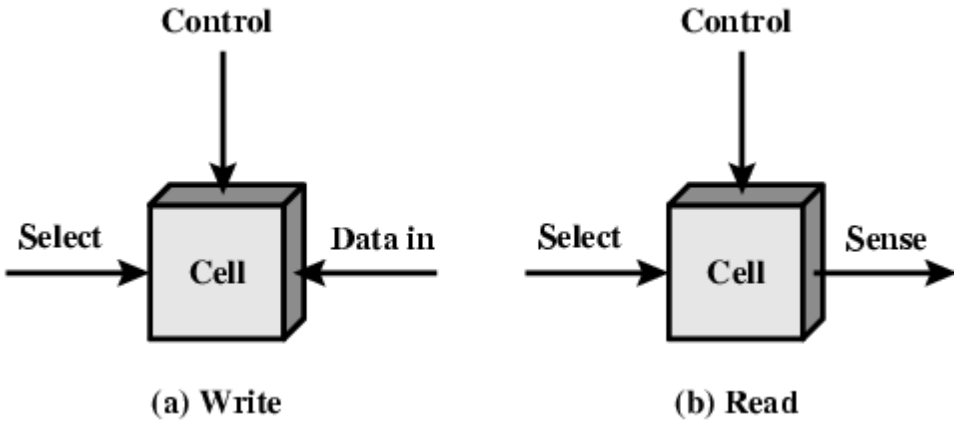
Figure 7.1 A typical CPU and main memory interface

☐ 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**.

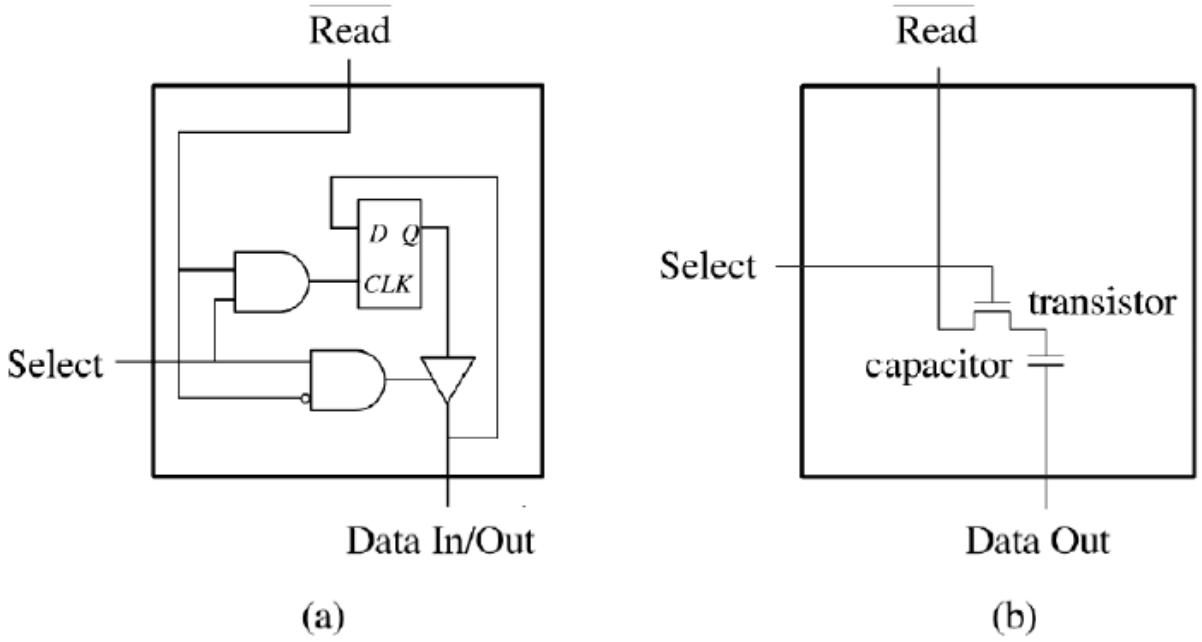


(a) Write (b) Read

☐ 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 chip

Integrated RAM are available in two possible operating modes, S*tatic and Dynamic*

Functional Behavior of a RAM Cell



Static RAM cell (a) and dynamic RAM cell (b).

SRAM versus DRAM
☐ Both volatile
o Power needed to preserve data

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

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

MEMORY SYSTEM DESIGN USING ICs

The major steps in such memory designs are the following:

1. Based on speed and cost parameters, determining the type of memory ICs (static or dynamic) to be used in the design.

2. It is generally better to select an IC with the largest capacity in order to reduce the number of ICs in the system.

3. Determining the number of ICs needed
N = (total memory capacity)// chip capacity).

4. Arranging the above N ICs in a P x Q matrix, where
Q = (number of bits per word in memory system)/(number of bits per word in the IC) and P = N/Q.

5. Designing the decoding circuitry to select a unique word corresponding to each address.

The following examples illustrates the design:

EX: construct (64K x 16) memory using (16K x 1) memory chip:

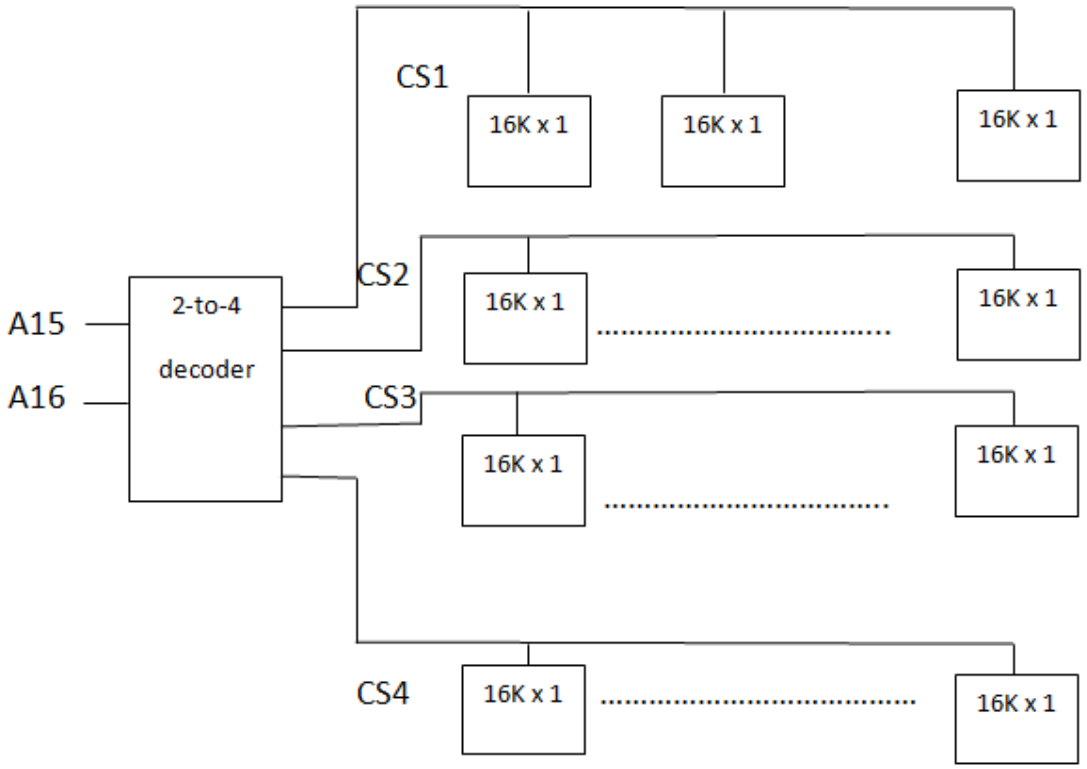
64K = 2^{16} 16 address line (A1……A16)

16K= 2^{14} 14 address line (A1……A14)

We need 64 K/16 K = 4 row =P

We need 16/1= 16 column=Q

16-14=2 (A15 ,A16) we use them in a 2- to-4 decoder to select rows:



EX: construct (4K x 4) memory using (1K x 4) memory chip.

4K/1K = 4 row =P , 4/4=1 column =Q

$4K = 2^{12}$

$1K = 2^{10}$ ⟹ 12-10=2 then we use 2-to-4 decoder