K.T.S.P.Mandal's

Hutatma Rajguru Mahavidyalaya, Rajgurunagar

Tal-Khed, Dist.-Pune 410505.

F.Y.BSc (Computer Science)

Semester-II

Electronics Paper-II

Subject- ELC-122 – Basics of Computer Organization

According to new CBCS syllabus w.e.f.2019-2020

Prof.Aparna P.Kulkarni

Department of Computer Science

Hutatma Rajguru Mahavidyalaya, Rajgurunagar.

Unit-IV: Memory Organization

- Memory Architecture, Memory hierarchy, Types of Memories
- Data Read/Write process
- Vertical and Horizontal Memory Expansion,
- Role of Cache memory
- Virtual Memory

Memory Architecture:

A memory typically consists of large number of cells organized as a matrix of mx n. Hence 'm' represents the number of memory words and 'n' represents the size of each memory word. For example, a 4k x 8 memory consists of 4k, i.e. 4096 memory locations with each location storing 8 bits or 1 byte of data. So a 4k x 8 memory is also referred to as 4kB memory.
The number of address bits is given by

$2^{(\text{No. of address bits})}$ = No. of words

Here the number of words is 4k

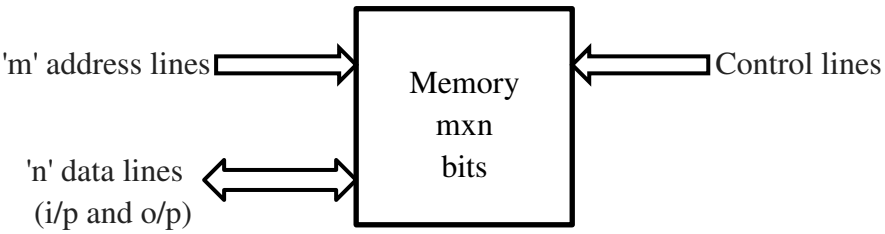
$2^{(\text{No. of address bits})}$ = 4k = 4.1k

$2^2 . 2^{10}$

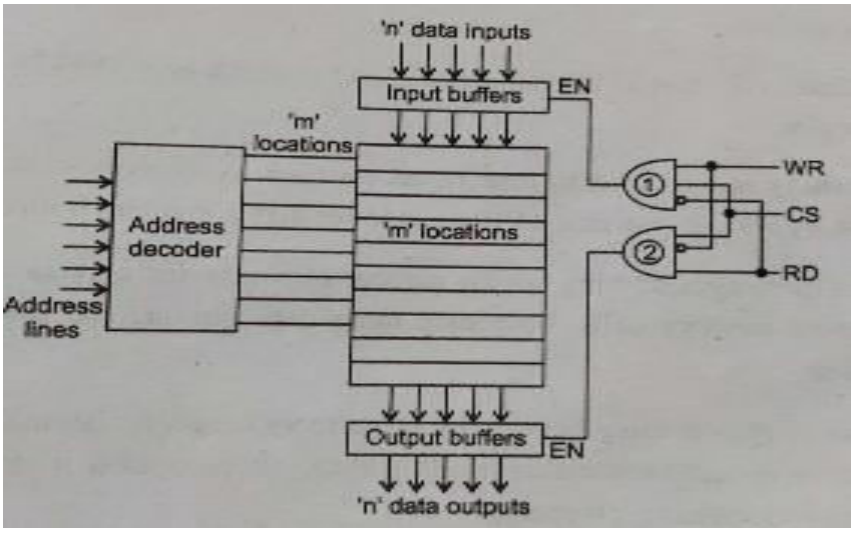
$2^{(\text{No. of address bits})} = 2^{12}$

Hence, the number of address bits required to address any one of the 4k locations is 12.

In general, the structure of a typical memory device is as shown:



Internal architecture of a typical memory:



Here when CS is 1 and WR = 1, RD= 0 output of gate 1 = 1 and input buffers are enabled. So input data is stored/written into the addressed location. When CS = 1, WR = 0 and RD = 1, output of gate 2 is 1. Now the data is read out from the selected location and comes on the output lines. Both read and write cannot be active at the same time.

Memory Parameters:

1. Memory access time T_A

The memory access time T_A is defined as the maximum time from the start of the valid address present and the time when valid data is put on the data bus or is made available on the data bus. Access time varies from memory to memory.

It can be 100s of us for slower memory to few us or ns for faster memories.

2. Data hold time:

It is defined as the minimum time for which data must be present after a valid write pulse ends.

3. Data setup time:

It is defined as the minimum time for which data must be present, before a valid write pulse begins.

4. Speed of memory

Speed of memory is measured in time of access time. A memory with smaller access time is faster. A memory with access time of 10 μs is faster than a memory with access time of 250 μs.

5. Capacity

Capacity of a memory is nothing but the number of words that a memory can store. If there are more number of memory cells, obviously more data bits can be stored in memory. So its capacity is more.

6. Cost:

The cost of memory is nothing but cost/bit of memory usage. It also includes the manufacturing cost of the memory and freight charges. A memory is cheaper if it provides more capacity at lesser cost, or good speed at lower cost.

Memory Hierarchy

Memory hierarchy is the arrangement of memory based on cost, capacity and speed.

Two Level Hierarchy:

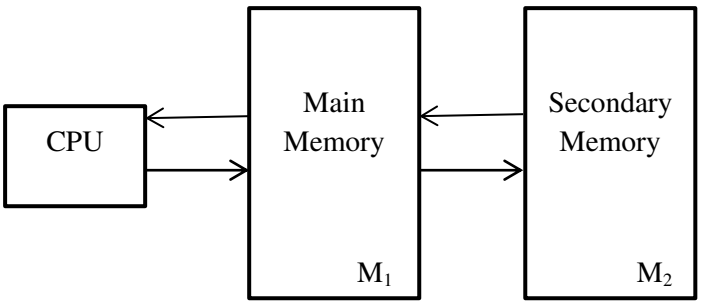


Fig. shows two level of memories used in the hierarchy. The main memory occupies a central position by being able to communicate directly with the CPU. The secondary memory cannot communicate directly with the CPU. It is only through the I/O processor that the data from secondary memory is brought into main memory. Programs not currently needed in main memory are transferred into secondary memory to provide space for currently used programs and data.

Main memory and secondary memory differ from each other in cost, access time and capacity.

Main memory is high speed, costly but limited in capacity as compared to secondary memory. Secondary memory is slower, less costly but has very large storage capacity.

Three Level Memory Hierarchy:

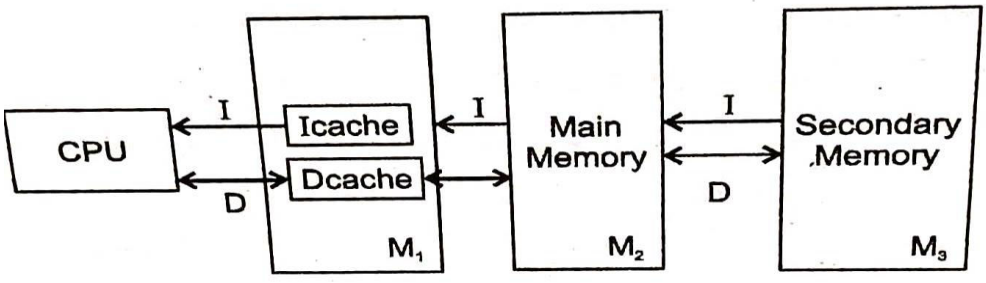


Figure 4.4: Three level hierarchy

As shown in fig., a cache memory is added. Split cache is added since it has separate areas for storing instructions (the I-cache) and data (the D-cache).

The CPU and other processors can communicate directly with M_1; only M_1 can communicate with M_2 and so on.

The following relation normally holds between adjacent memory levels M_i and M_{i+1}.

Cost per bit = $C_i > C_{i+1}$

Access time t_{Ai}, < t_{Ai+1}

Storage capacity $S_i < S_{i+1}$

Four Level Memory Hierarchy:

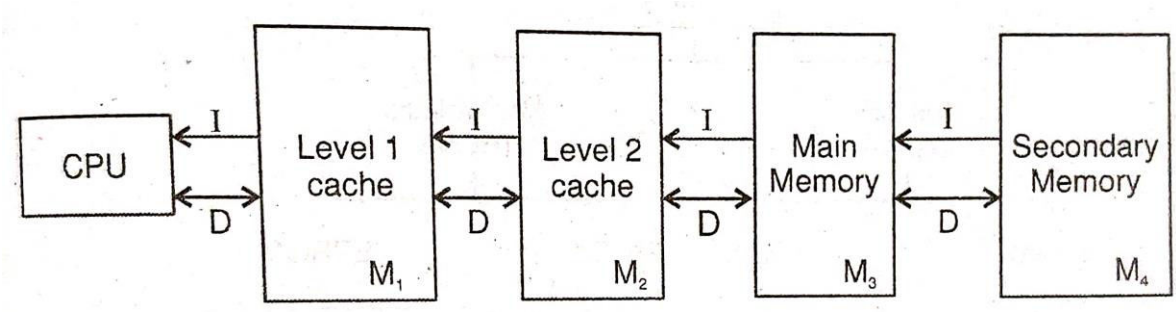


Figure 4.5: Four level Memory hierarchy

In four level hierarchy two non-split cache levels are added, followed by main memory and secondary memory. Slower devices are put at lower level and faster devices are at higher level.

The CPU and other processors can communicate directly with M_1; only M_1 can communicate with M_2 and so on.