UNIT-V

The Memory Systems

Basic memory circuits, Memory System consideration,

Read only Memory :- ROM, Programmable ROM (PROM), EPROM, EEPROM, Flash Memory

Cache Memory — Mapping functions, Inter-leaving

Secondary Storage — Magnetic Hard disk, optical disks.

→ Memory is used to store data temporarily & to execute programs.

→ What ever we write it first stores in main memory

→ Memory systems must be fast, inexpensive.

→ when MAR connects to memory (Read to memory) & that data is written in MDR (write from memory) the time required by this is less then we can guess the memory speed

→ virtual memory is a technique which is used to organise memory efficiently. It is not a physical memory.

→ Memory address is called physical address.

→ The address which exists in processor is logical address another name for logical address is virtual address

Basic concepts:-

→ MAR stores 32 bits so it need 32 address lines

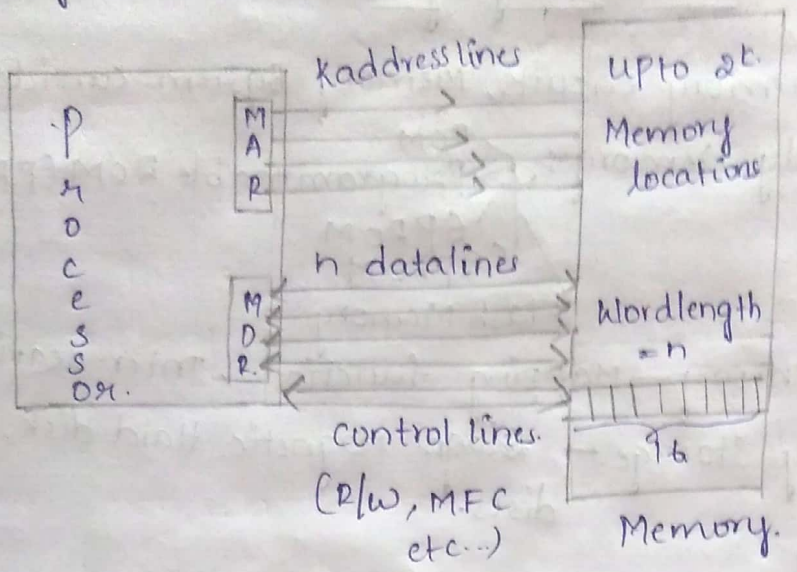
→ if we have k address lines then we have 2^k words

→ The count of data lines are lesser than the count of address lines

→ If we have 16 cells then will have 8 data lines

→ Here wordlength is 8

→ In single unit of clk cycle we can transfer only one word (i.e, datalines)



→ one of the control line must acknowledgement line.

→ whenever data is ready to write in to datalines then memory sends Memory function completed signal to processor using one of the control line. Then processor enables to read data using data lines

→ MFC means Memory function control line.

→ Another control line which contains information the total ~~cont~~ no of bits transferred from memory to processor & processor to memory.

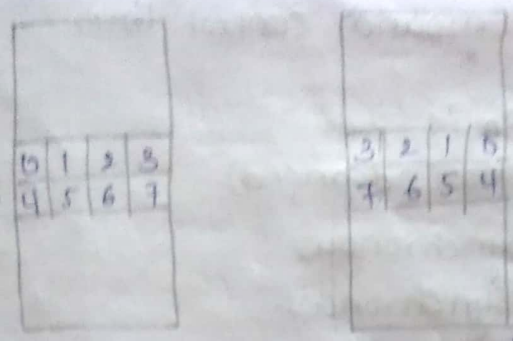
→ There is another control line which is responsible for count of data transfer between memory & processor.

→ Modern computers are byte addressable. (i.e, 16 bit computer) 16 bit $= 2^{16} = 64$ kilo byte.

32 bit $= 2^{32} = 4$ giga byte.

40 bit $= 2^{40} = 1$ Tera byte memory locations

→ There are two ways of ordering & storing data inside memory i.e; Big Endian Byte order & little Endian byte order.



Big Endian — little Endian.

Memory Acess time:-

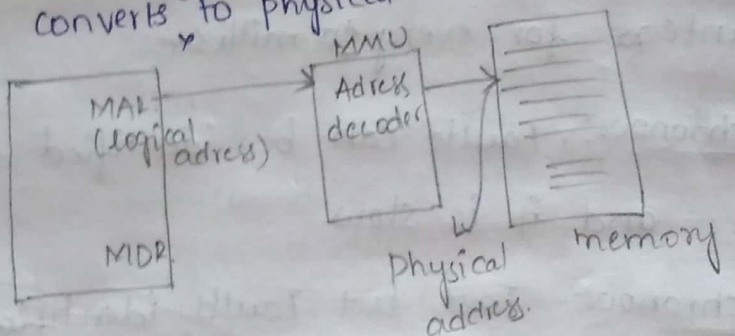
→ The time between MAR & MDR is the MAT (Memory Acess time)

→ The time required for Memory Read & MFC signal called is, MAT.

→ MAT is the parameter to compare the speed of memories.

→ Memory cycle time.

2/3/19 → MAR is not directly connected to memory. There is a interface called Memory mapped unit. The logical address in MAR is then mapped to memory and it converts (address decoder decodes) to physical address



Semiconductor Circuits:

→ Before 1960's magnetic pore droms were used to Store 0's & 1's. drawback is requires more space.

→ After 1960's Semi conductors were introduced. adva -ntage less space. disadvantage more expensive.

→ with resonable cost wit VLSIT technology Incorporated. to brogh this Semi conducting devices

→ Ram chip stores 0 & 1's information and this chips

always covered withe ground control lines

types:- Static Ram :- Basic
CMOS (Advanced)

Dynamic Ram - Synchronous
Asynchronous.

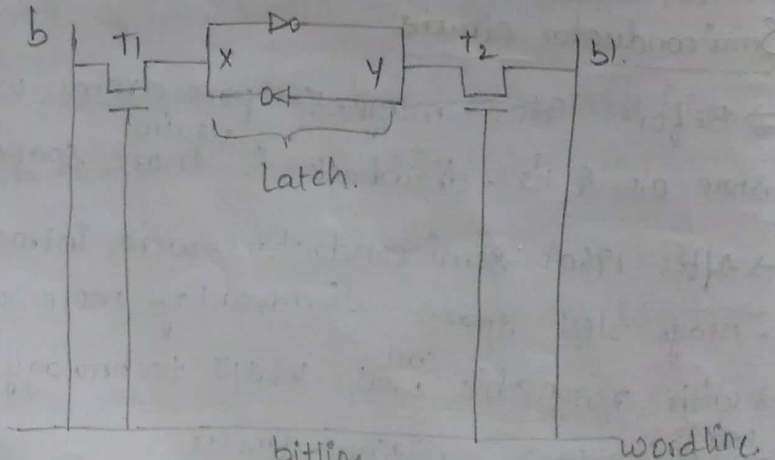
Basic:- This is non-volatile.
→ cheaper than Semi conductor devices
→ 2 not gates & 2 transistors.

CMOS:- Non-volatile, In expensive.
→ Total 6 transistors
No of gates
→ draw back complexity is high.

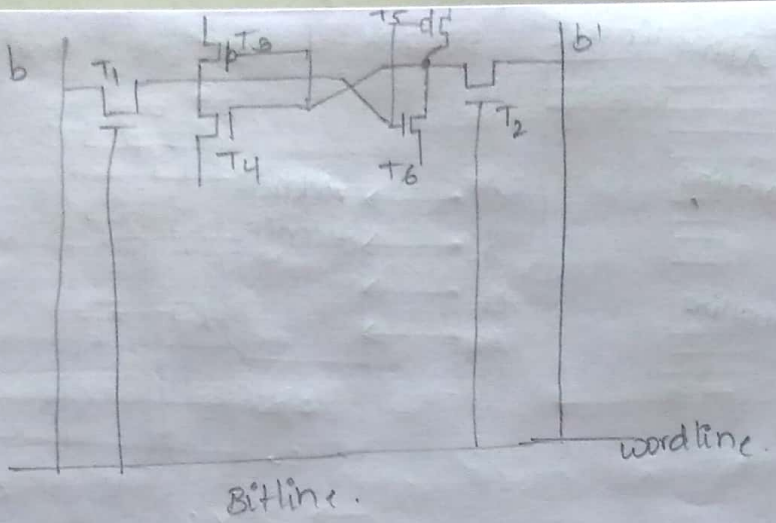
Dynamic:- Only one transistor, one capacitor.
it is very cheap. Draw back is volatile.
this can be overcome by giving power refersh counters. for every 10 millisec.

Synchronous:- Faults can be identified easily and it is slow.

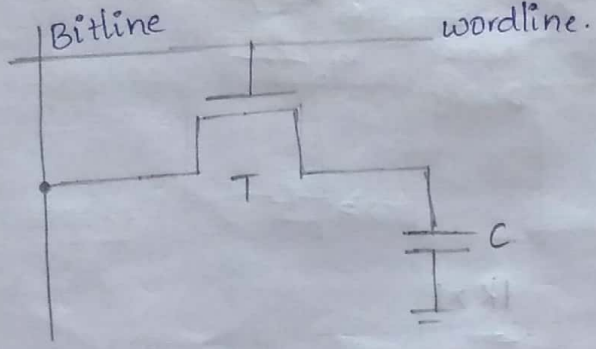
Asynchronous:- Fast but Faults identification is difficult.



Static RAM.



CMOS



DRAM Cell.

→ we have 16×8 RAM How many address lines, data lines required.

$= 2^4 \times 8$ = 4 address lines = 8 data lines.

we also require 4 control lines Read/write, chip sele ct (cs), Power, ground for memory.

→ For DRAM we have asychornous chip select & synchro -nous chip select along with the above lines.

→ The total lines required by this 16, 4 addres, 8 data, 4 control lines

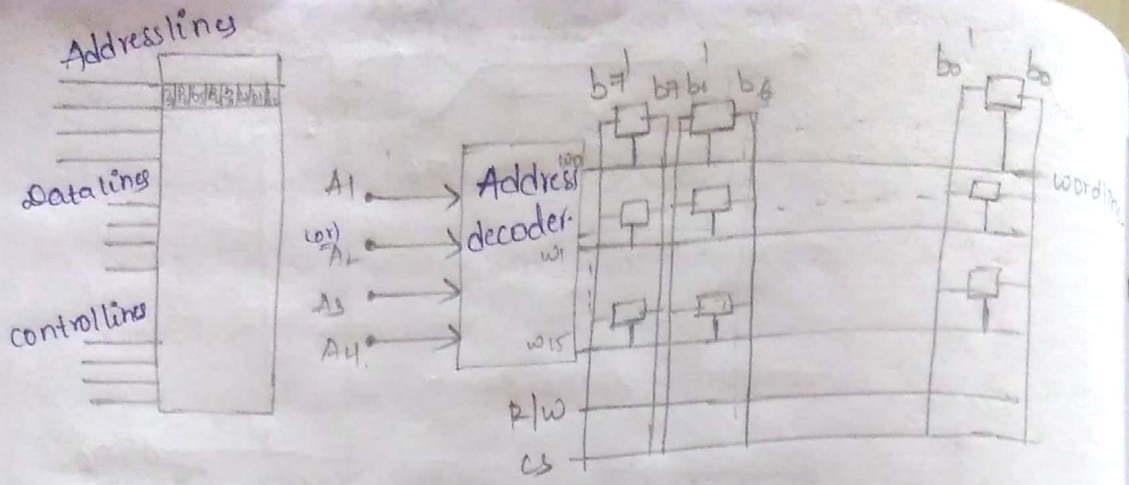
→ Suppose we need 1K×1 RAM it require 1 data lines

1024 × 1

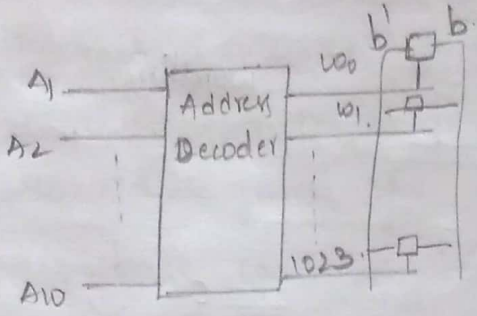
$= 2^{10} \times 1$ = 10 address lines

= 4 control lines

total = 10+4+1 = 15 lines required



→ $2^4 \times 8$



1k x 1

6/3/19

2M x 32 Here 2Mega = 2048k.

2048k x 32

4 (512k) x 4x8

4 (512 x 2^{10} x 8)

Four 512K x 8 RAMs.

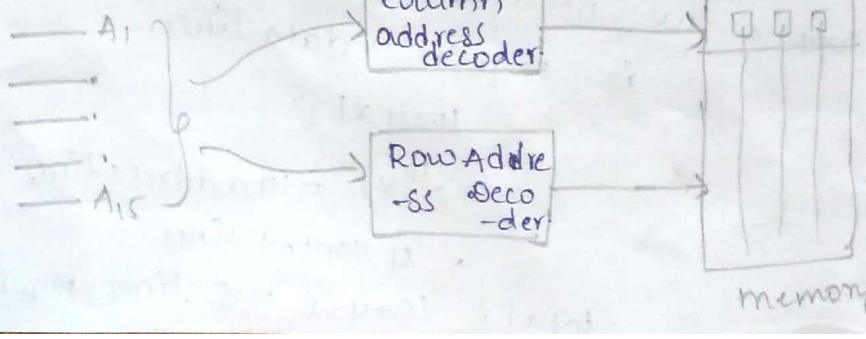
512 x 2^{10} x 8

2^9 x 2^{10} x 8.

→ we require 19 Address lines, 8 data lines and 4 other line Total = 31 lines

Refresh counter

→ In memory cells are organised 2 dimensionally.



Address lines are mapped to column address decoder as well as row address decoder.

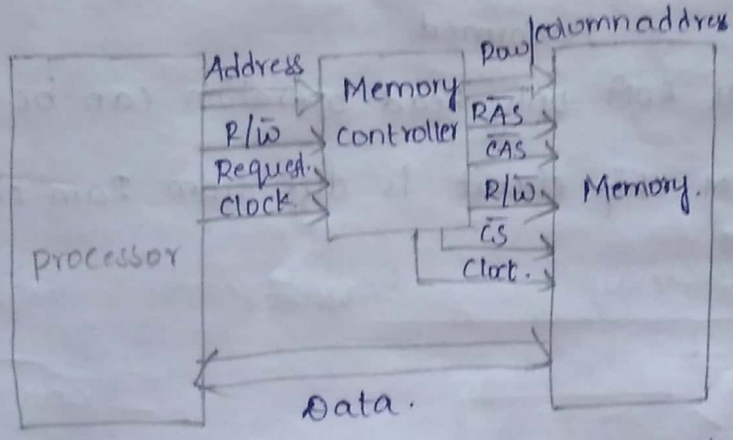
Suppose I want to design 128MB RAM using dynamic RAM. the drawback is it needs more circuit connections to solve this there are advanced DRAM's they are SIMM (Single Inline Memory Module), DIMM (Dual Inline Memory Module)

Memory System considerations:-

→ The modern SIMM, DIMM are advanced DRAM technologies.

→ DRAM's are little slow compared to advanced technologies but gives better for performance in case of larger memories

→ Depends on our requirement it will select corresponding chip technology.



RAS - Row Address selector CS - chip selector.

CAS - column " " SDRAM - Synchronous Dynamic RAM

ASDRAM - Asynchronous " "

→ Refresh overhead:-

→ All dynamic memories have to be refreshed.

→ In older DRAMs a typical period for refreshing all rows was 16milli sec.

→ In typic SDRAM atypical period is 64millisec.

consider an SDRAM whose cells are arranged in 8k = 8192 rows.

Suppose that it takes 4 clock cycles to access each row then how much time it takes to refresh all rows.

→ then it takes 32768 clock cycles to refresh all rows

At a clock rate of 133 Megahertz. i.e; $\frac{32768}{133 \times 10^6}$

is the ti 246×10^{-3} sec is the time required to refresh all rows.

Thus the refreshing processor occupies 0.64 milliseconds in each time interval

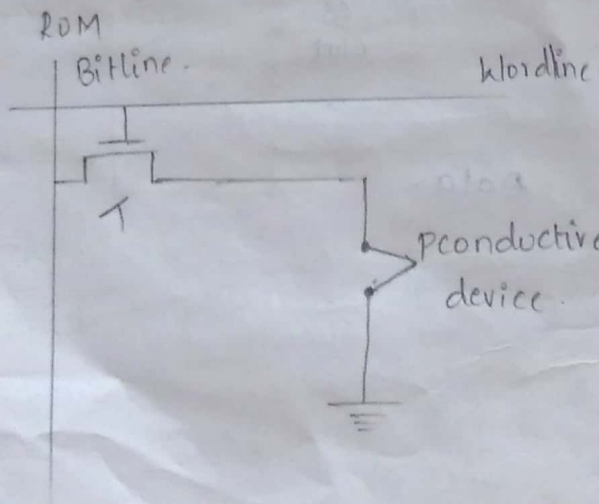
→ The refresh overhead is $\frac{0.246}{64} = 0.38$

19 ROM:-

→ For RAM Read & Write operations can be performed

For ROM only Read operation can be performed.

→ ROM chip circuit is diff from Ram chip circuit.



→ any hardware components connected to it it only reads data.

→ If it is connected it is zero, if it is not connected it is 1 (ie, able to store)

→ It can only perform read operation since P conductive device is grounded.

→ Bit line is connected to power supply.

Prom:- Programmable readonly memory.

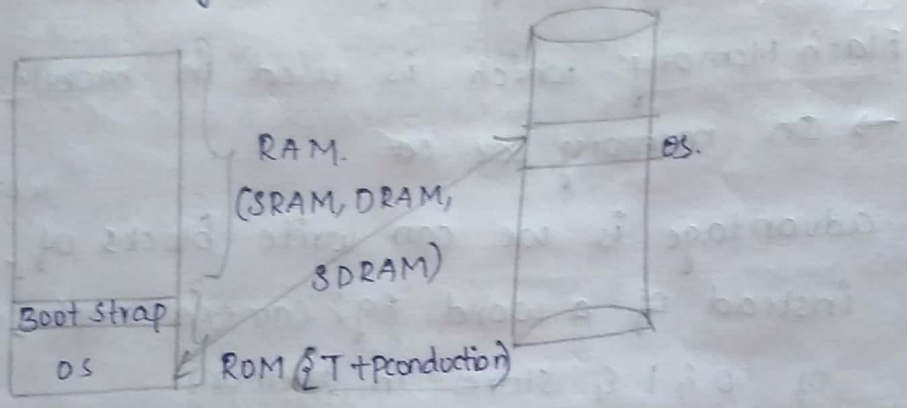
→ SROM, DROMs are voltaile when power is lost data is lost.

Why ROM?

→ ROM is having instructions Read eg:s as Files, Bootstrap loader files (or) Instructions

→ suppose if we write any bit in above instr -uctions ~~os will not~~ os, Bootstrap loader problem occurs (i.e, it does not work).

→ That's why these instructions.



PROM

→ user can asess os when they are running and can change some os functionalities using PROM.

Eg:- Task bar can be placed at bottom side (or) top

→ PROMs provide convience & flexibility.

→ programming ability is achieved by inserting a fuse at point p

→ Advantage of fuse is masking (i.e, 0 can be masked to 1 in ROMcell)

EPROM:- Erasable & Reprogramable ROM

→ It remove some set of instructions & replace some set of instructions.

→ If we replace Boot strap loader & write program we

→ Eg:- we can write our own program to bring os to main memory by changing boot code. A special transistor is used which can use erase & reprogramble.

→ In EPROM to erase the code uv light will be used., Its costly.

EEROM:- Electrical Erasable ROM

→ By using electrical singal we can erase the code.

→ Different voltage signals are needed. (disadvantage)

Flash Memory:- which is used in mobiles

→ ~~In memory we~~

advantage is we can write blocks of word of data/code instead of a word. Eg:- camera; captures blocks of 0 & 1 & stores in memory.

→ previous ROMs requires continous power supply where as ~~previous ROMs~~ Flash. require less power.

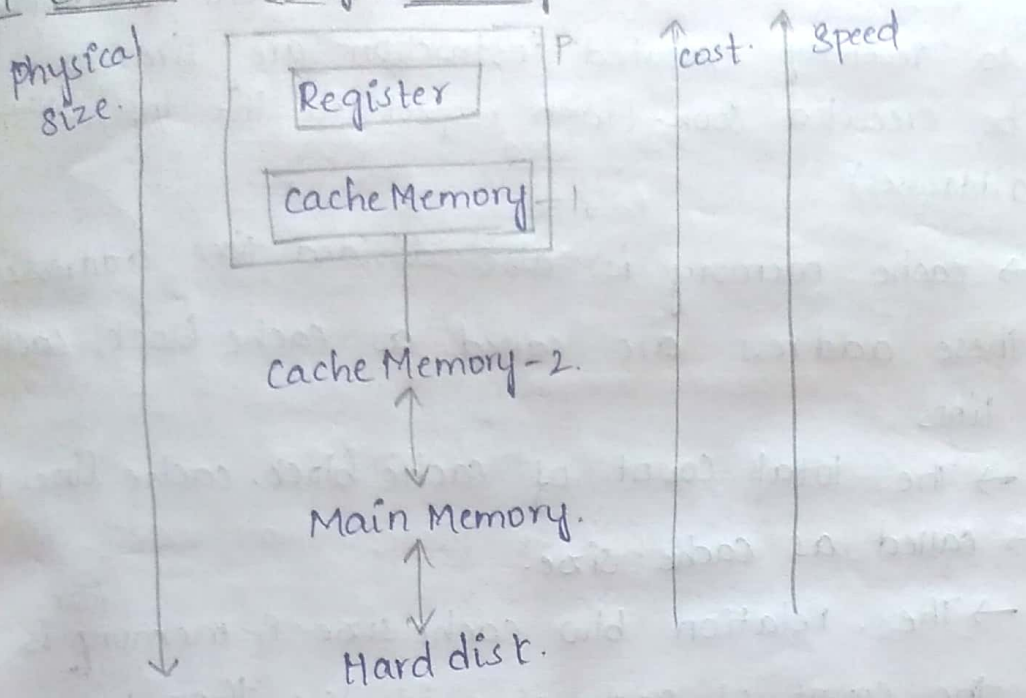
→ while listening MP3 it represents sound on screens

→ ~~Smart~~ Flash cards (or) SD cards that size 8, 16, 32, 64, Flash drive permenant storage for Flash memory

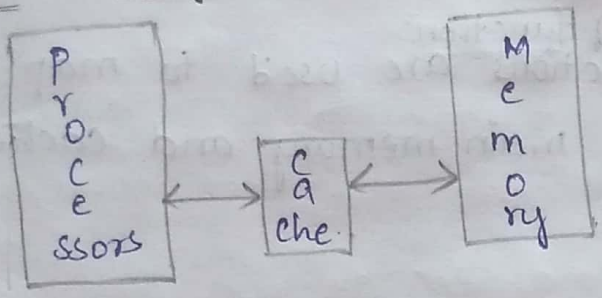
→ Flash drags ~~provides~~ responsible for insensitive vibrations

→ The draw back of flash memory is it supports with less Floppy drives i.e, Flash drive storage capacity is very less.

8/3/19 Cache Memory Hierarchy:-



cache memory:-



→ The speed of maing memory is very low compa -red to the modern processors.
→ static Rams are used to design cache memory
→ It is used to increase the speed.
→ cache memory works on the principle of locality of repference.

There are 2 types of locality of reference:-

(i) Temporal.
(ii) spatial

(i) Temporal: Recently executed instruction is likely to be executed again very soon eg:- For loop.

```
for(i=0; i<10; i++)
  {
    i= i+10;
  }
```

(ii) Spatial:- This aspect means in close proximity. to recently executed instruction are like to be executed soon. (with respect to the instruction addresses)

→ cache memory is also divided like addresses so these address are called as cache block, cache lines

→ The total count of cache block, cache lines is called as cache size.

→ The relation blw cache size & memory size. is the count of cache size is less than memory size count.

Mapping function:-

→ Mapping functions are used to map blocks of data between main memory and cache.

readhit
readmiss
writehit
writemiss

Replacement algorithm

Least recently used

Right through protocal is done by memory (or) write back as well as copy back, is done by processor.

→ In cache memory we have dirty bit (or) modify bit it says whether it is updated or not. if it gets update, flag bit get update.

→ There are three types:-

(i) Direct mapping
(ii) Associative mapping
(iii) Set associative mapping.

→ cache has 120 blocks, main memory 4096 blocks

1 block = 16 words

Sir → Main memory data has to map to cache memory.

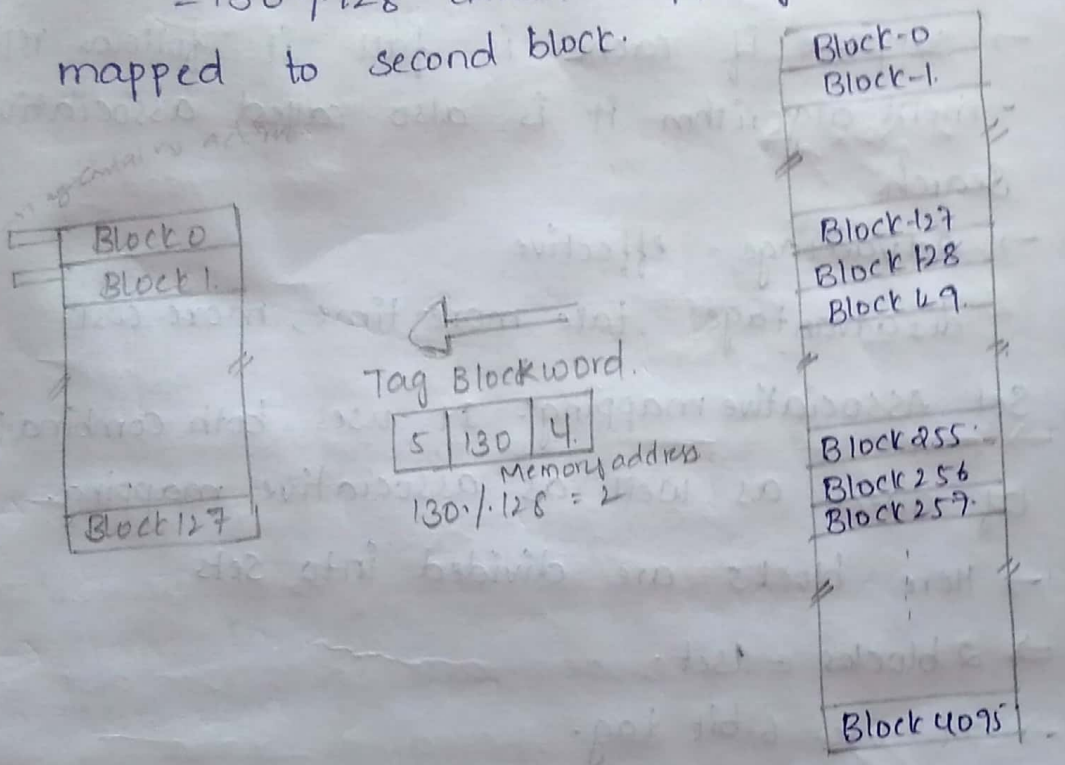
→ consider cache consists of 128 blocks
main memory is of 4096 block.

→ in both memory each block is of 16 words

→ we need to map a ~~memory~~ word from memory to cache. which block of cache has to choose? & corresponding word has to choose.

Eg:- Two 130 block data has to map to cache main memory memory.

= 130 % 128 direct mapping. so it is mapped to second block.



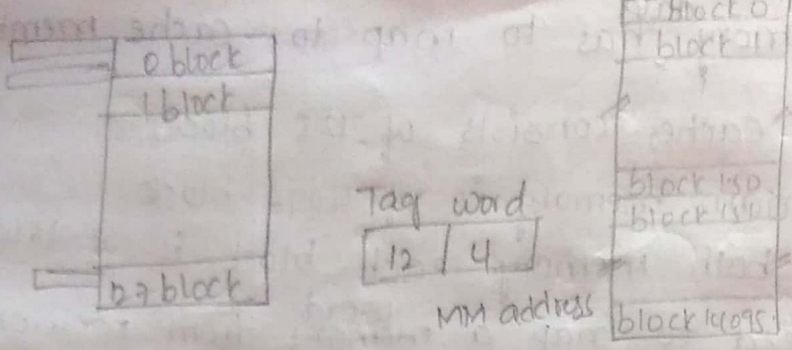
→ It will not search in all blocks wheter it is already present (or) not (duplicates)

→ In chache only 2 block is filled remaining all blocks are vacant.

9/5/19 → In Associative mapping if all blocks are filled then it use replacement algorithm.

→ Direct mapping is very flexible, ~~Tats~~ Tag is useful it show the location where that ~~data~~ data is present.

Associative mapping:-



→ In associative mapping it searches whethers corresponding word exits or not.

→ If not matches it Places that word in any block.

→ Suppose if cache is full it follows replace-ment algorithm. it is also called associative Search.

→ Advantage - effective

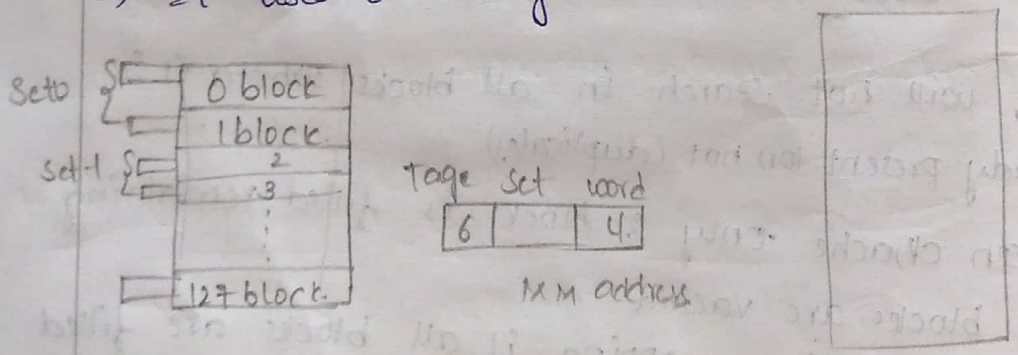
disadvantage - take more time, more cost.

Set Associative mapping:- It uses both combinations of direct as well as associative mapping.

→ Here blocks are divided into sets

→ 2 blocks = 1set.

→ It uses 6 bit tag.



Here Eg:- 0, 128, 256, 512

⇒ 0 % 128, 128 % 128, 256 % 256, 512 % 512

Directly all are mapped to zero.

→ Associative search in all blocks of seto

→ If Search unsucess replace in vacant block.

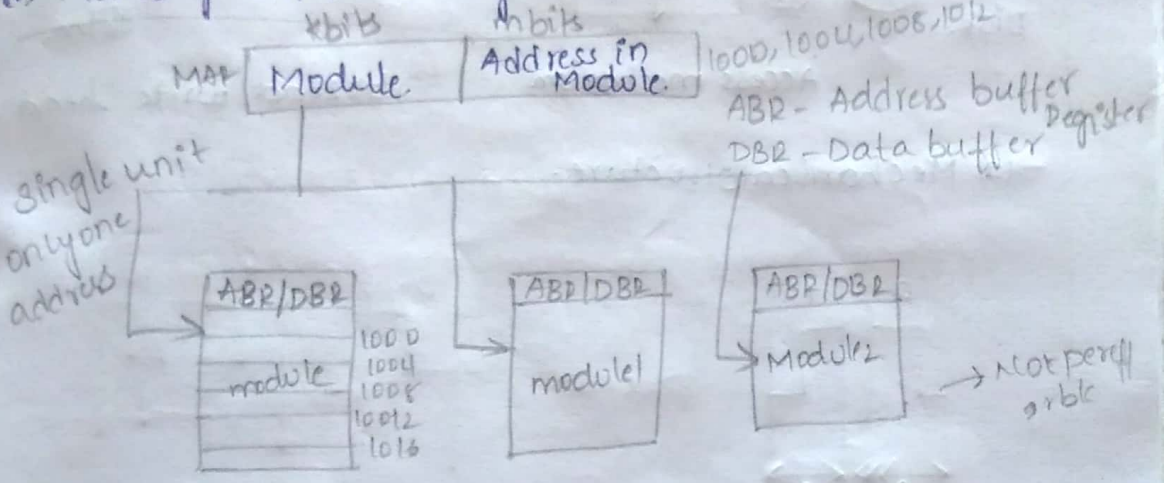
word.

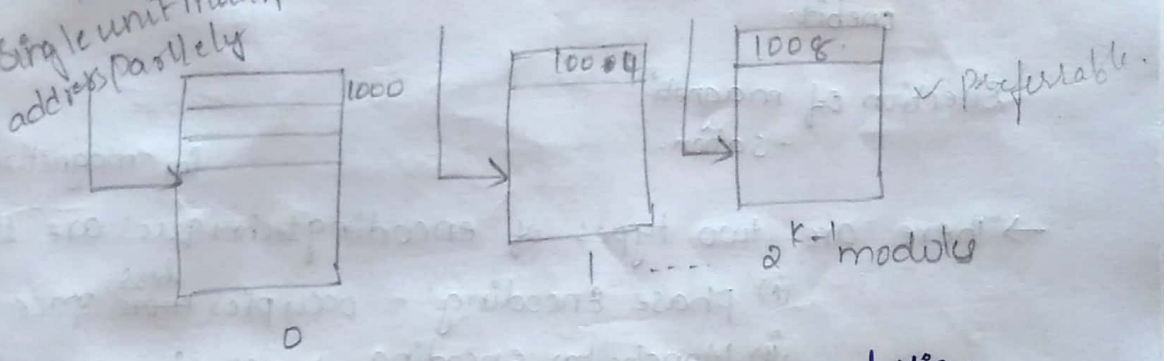
speed b/w memory & processor

System Performance measurements:

(i) Hit/Fail Ration:- Read hits Read miss, write hit, write miss

(ii) Memory Inter-leaving





11/3/19 → It maintain secondary storage devices.

→ Need for secondary storage device is, if we use dynamic RAM's every time for storing entire data we need more number of DRAM's.

→ disadvantage:- cost is high

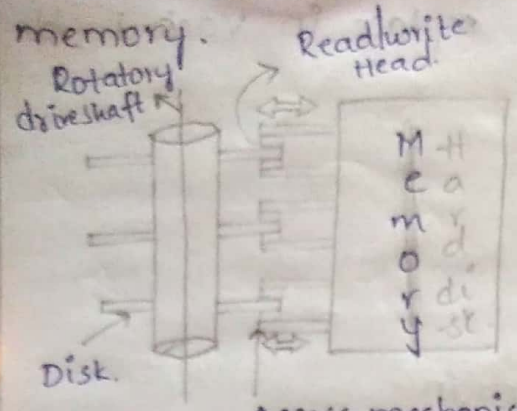
→ Solution is use secondary storage device concept with less cost.

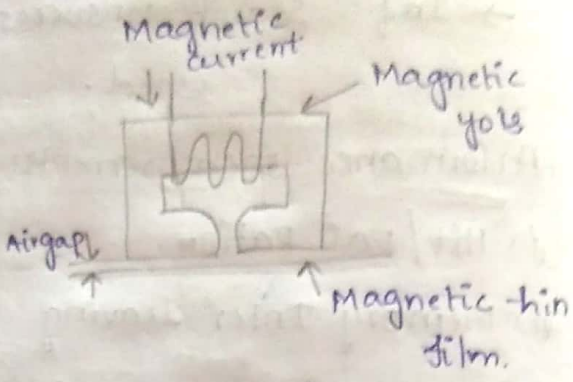
Examples of Secondary storage devices :-

(i) Magnetic Hard disk.

(ii) Optical disk.

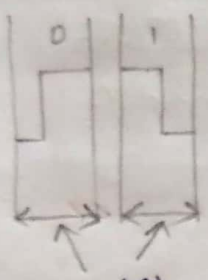
→ Magnetic Hard disks:- There should be read/write head from hard dist which moves data to main



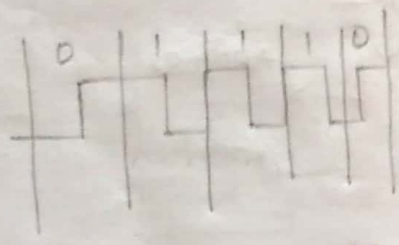


(b) Read/write head detail

(a) Mechanical structure.



Direction of magnetisation.



Phase encoding.

for magnetized data

→ There are two types of encoding techniques are there,

(i) phase encoding - occupies ~~more~~ less space

(ii) Monchestor Encoding → occupies more space.

→ In the most modern disk units the disk & ~~to~~ read/write are placed in sealed Air filled enconder this approach is known as Winchester technology.

15/3/19 Organisation & accessing of Data on a Disk:-

→ Disk is collection of surfaces. (logically called cylinder)

each surface is having 1000's of tracks and each track is having sectors

→ Sector contains our data.

→ If we want to access data from hard disk.

Surface number. tracknumber. sector number.

there are also concepts like sector header, error correction code.

→ There will be gap between sectors that is called as intersector gap.

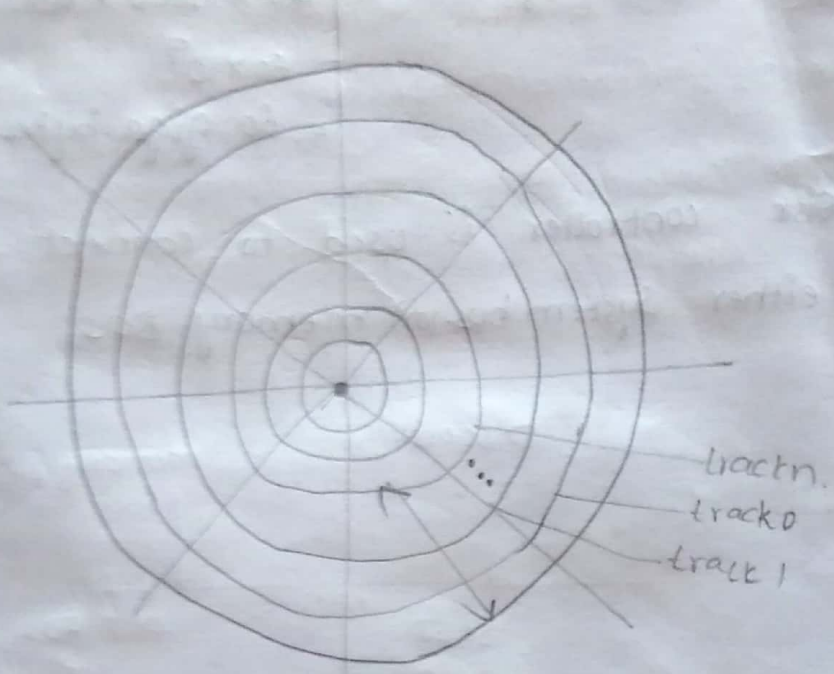


Fig:- Structure of a one surface.

Seek Time:- when you place a header, the time required to place head from surface to tract.

General seek time range is 5-8 millisec.

Rotational delay (or) latency:-
The time required from track to the starting address to the address of the sector.

Acess time:- It is the sum of seek time + laten -cy time.

Parameters:- Total no of surfaces,
Tracks per each surface,
Sectors in each track.

Eg:
→ Total no of surfaces = 20

Tracks per each surface is 15000

Sectors for each track is 400

→ each sector can hold 512 bytes of data.

Sol: Size of disk = $20 \times 15000 \times 400 \times 512$

$= 60 \times 10^9$

= 60 giga bytes.

→ Disk controller is used to connect Disk to either System bus (or) memory Bus.

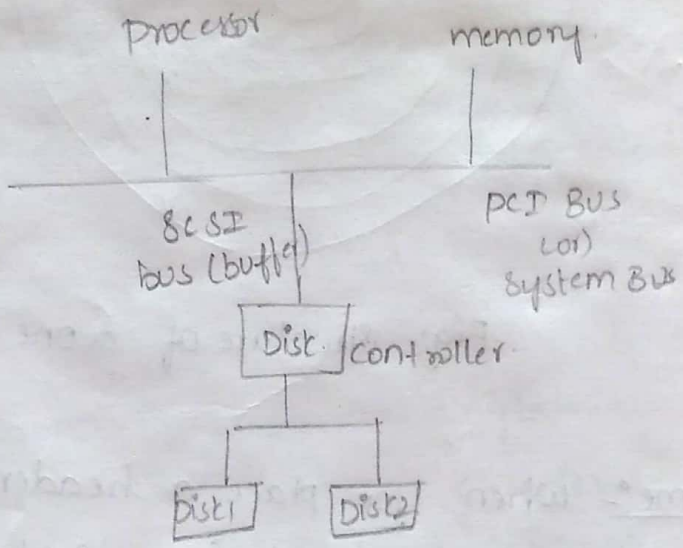


Fig:- Disk connected to system bus.

→ Hard disk is cheaper compared to ~~Ram~~ RAM chips.

→ Harddisk can store booting files & os files.

→ even though it can able to store small progra-ms i.e, booting programs permenantly. i.e Hardisk is non-voltaile.

→ RAM, ROM, is voltaile as these are made of semiconductor devices. it can store data only to milli sec. if it need to store more time we has to provide continous power supply. Power supply is given by refresh counters

→ Memory mapping unit is a part of memory controller.

→ In disk non need continous power supply. because sectors contain power.

Floppy disk:-

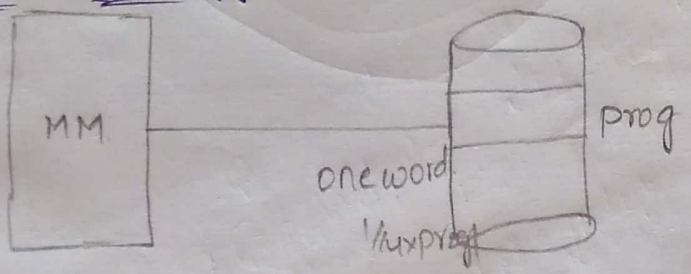
Advantage is very simple, flexible, movable.

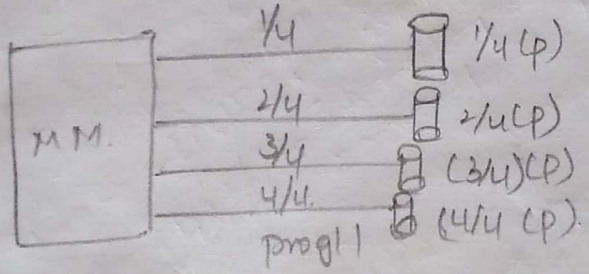
→ It also uses magnetic plates, phase (or) monchestor encoding.

→ It can store 2MB, but magnetic disk can store 60 giga bytes.

→ It has two types: single density, double density.

RAID Technology:-





RAID = Redundant Array Inexpensive data.

→ The above is RAID 0 technology.

→ The challenging task of hard disk is to reduce access time.

→ RAID 0 technology increases speed.

→ RAID 1 divides the hard disk. in to two one contains ~~one~~ original data and other is duplicate. when one trashes other remains.

→ RAID 2, RAID 3, RAID 4, technologies follows parity checking scheme

→ RAID 2, RAID 3, RAID 4. no need full duplication of dist. for realiability it uses parity checking scheme.

→ RAID 5 technology. It uses parity based error Recovery scheme.

→ The combination of these technologies is used.

RAID 10 (combination of RAID 1 & RAID 0)

→ It contains some other technologies such as ATA/EIDE Disks, SCSI Disks, RAID Disks.