# The Memory System

# The Memory System [Unit-III]

Basic Concepts, Semiconductor RAM Memories, Read-Only Memories, Speed, Size and Cost, Cache Memories, Performance Considerations, Virtual Memories, Memory Management Requirements, Secondary Storage.

## Some Basic Concepts :-

The maximum size of the memory that can be used in any Computer is determined by the addressing scheme.

For eg, a 16-bit Computer that generates 16-bit addresses is capable of addressing upto $2^{16} = 64K$ memory locations. Similarly machines whose instructions generate 32-bit addresses can utilize a memory that contains $2^{32} = 4G$ (giga) memory locations. Similarly 40-bit $\Rightarrow 2^{40} = 1T$ (tetra) locations.

Most modern Computers are byte addressable (show in fig) shows the possible address assignment for a byte-addressable 32-bit Computer.

The memory is usually designed to store and retrive data in word-length quantities. In fact, number of bits actually stored or retrived in one memory access is the most common definition of the word

length of a Computer.

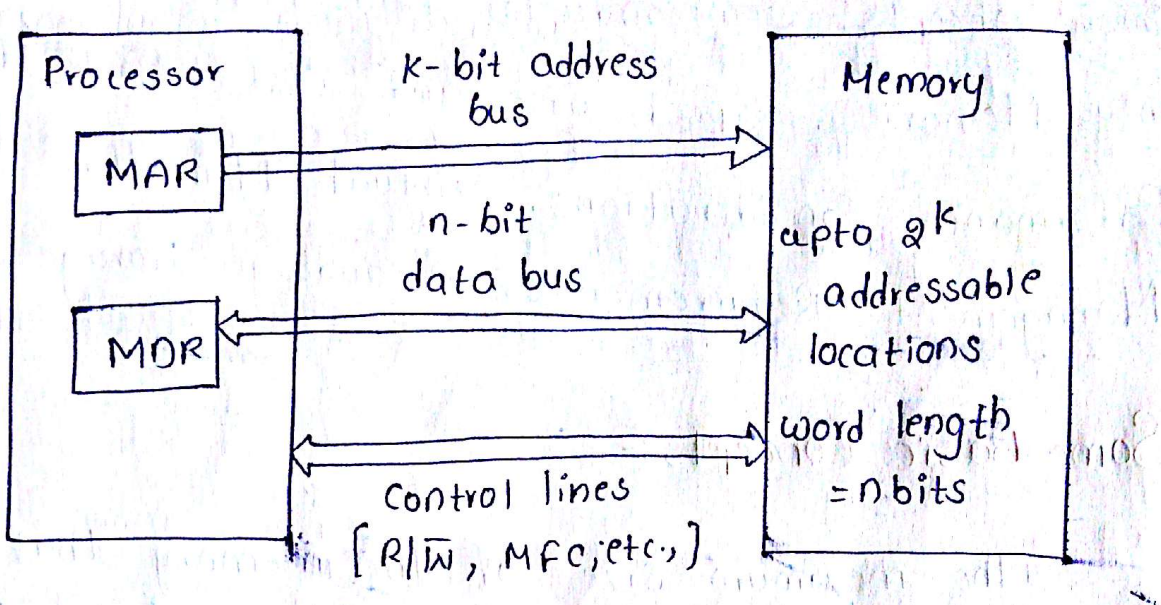



fig.: Connection of memory to the processor

From the above fig., it shows that the memory and processer are represented in two different blocks. Data transfer between the memory and the processor takes place through the use of two processor registers, usually called MAR (Memory Address Register) and MDR (Memory Data Register).

If MAR is k-bits long then MDR is n-bits long. then the memory unit may contain up to $2^k$ addressable locations. During the memory cycle, n bits of data are transferred between the memory and the processor.

This transfer takes place over the processor bus, which has k-address lines and n-data lines. The bus also includes R/$\bar{w}$ and memory function

Completed (MFC) for co-ordinating data transferes.

If read or write operations involve consecutive address location in the main memory, then a "block transfer" operation can be performed. In which the only address sent to the memory is the one that identifies the first location.

A usefull measure of the speed of the memory units is the time that elapses between the initiation of an operation and the completion of that operation. This can be referred to as "memory access time".

Another important measure is the "memory cycle time", in which is the minimum time delay required between the initiation of two successive memory operation.

One way to reduce the memory access time is to use a "cache memory". This is a small, fast memory that is inserted between the larger, slower main memory and the processor. It holds the currently active segment of a program and their data.

The address generated by the processor is referred to as "virtual" or "logic address." The maping function is implemented by a special memory control circuit,

often called the "memory management Unit".

## Semiconductor RAM memories :-

The semiconductor memory includes a memory cell array including a no.of memory cells arranged in a rows and colums, a no.of word lines each connecting all memories cells in a row and a no.of bit cells each one connecting all memory cells of a column.

Semiconductor memories are available in wide range of speeds. Their cycles times range from 100ns to less than 10ns. When first introduced in the late 1960's they were much more expensive than the magnetic core memories they replaced. Becaus of rapid advances in VLSII [very Large Scale Integration] technology, the cost of semiconductor memories has dropped dramatically.

## Internal Organization of Memory Chips :-

Memory cells are usually Organized in the form of array, in which each cell is capable of storing one bit of information.

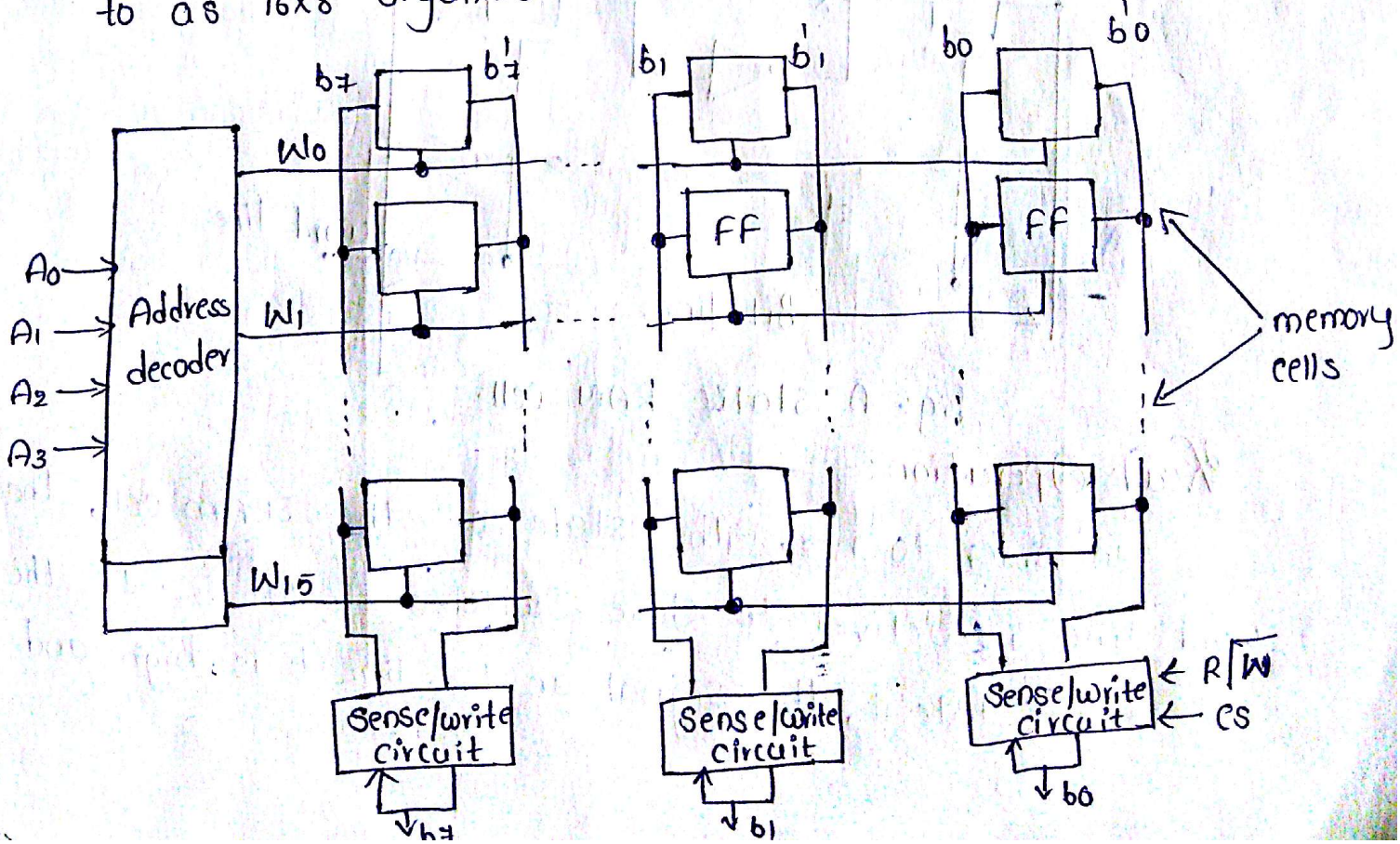
Each row of cells constitute of a memory word, and all cells of a row are connected to a common line reffered to as the "word line" which is driven

by the address decoder on the chip. The Sense/Write circuits are connected to the data input/output lines of the chip. The cells in each column are connected to a Sense/Write circuit by two "bit lines".

During the read operation, these circuit sense, or read, the information stored in the cells selected by a word line and transmits the information to the output data lines.

During write Operation the sense/Write circuit receive input information and store in the cells of the selected word.

The fig is an eg of very small memory chip consisting of 16 words of 8 bits each. This is reffered to as 16x8 organization.



## Static Memories:-

Memories that consists of circuits capable of retaining their state as long as power is applied are known as "static Memories".

The fig illustrated how static Ram or SRAM cell may be implemented. Two inverters are cross-connected to form a latch. The latch is connected to two bit lines by transistors $T_1$ and $T_2$. These transistors act as switches that can be opened or closed under control of the word line. When the word line is at the ground level, the transistor are turned off and the latch retains its state.

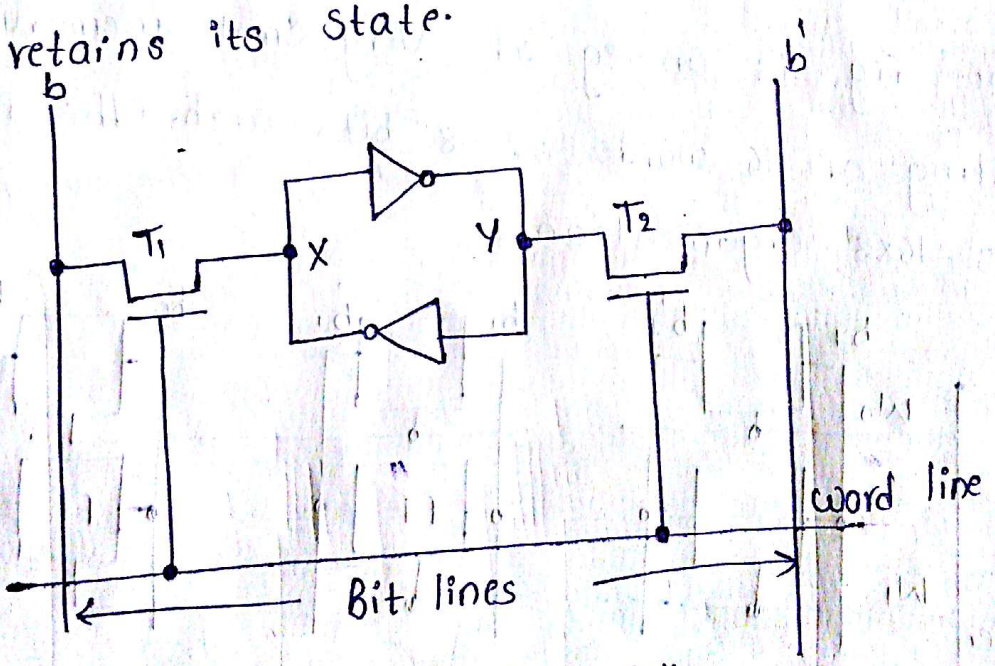



fig: A static RAM cell.

## Read Operation:-

In order to read the state of the SRAM cell, the word line is actived to close Switches $T_1$ and $T_2$. If the cell is in state 1, the signal on bit line b is high and

and the signal on bit line b' is low. The opposite is true if the cell is in state 0. Thus b and b' are complements of each other.

## Write Operation:-

The state of the cell is set by placing the appropriate value on the bit line b and its complement on b', and then activating the word line. This forces the cell into corresponding state. The required signals on the bit lines are generated by the Sense/Write Circuit.

## CMOS Cell:-

The CMOS realization of the memory cell is given below the fig. Transistor pairs ($T_3$, $T_5$) and ($T_4$, $T_6$) from the inverters in the latch. The state of the cell is read or written as just explained.

For eg. in state 1, the voltage at point x is maintained by high, vcc by having transistors $T_3$ & $T_6$ on, while $T_4$ & $T_5$ are off. Thus if $T_1$ & $T_2$ are turned on[closed], bit lines b & b' will have high and low signals respectively.

The power supply voltage $V_{supply}$ is 5V in older CMOS SRAMs or 3.3 V is new low-voltage

versions. If the power is interrupted the cell's content will be lost. When power is restored, the latch will settle into stable state, but it will not necessarily be the same state the cell was in before the interruption.

Hence, SRAM's are said to be "Volatile memories" because their contents are lost when power is interrupted.

A major advantage of CMOS, SRAMs is their very low power consumption because current flows in the cell only when the cell is being accessed.

SRAMs can be accessed very quickly. Access time of just a few nanoseconds are found Commercially available chips.

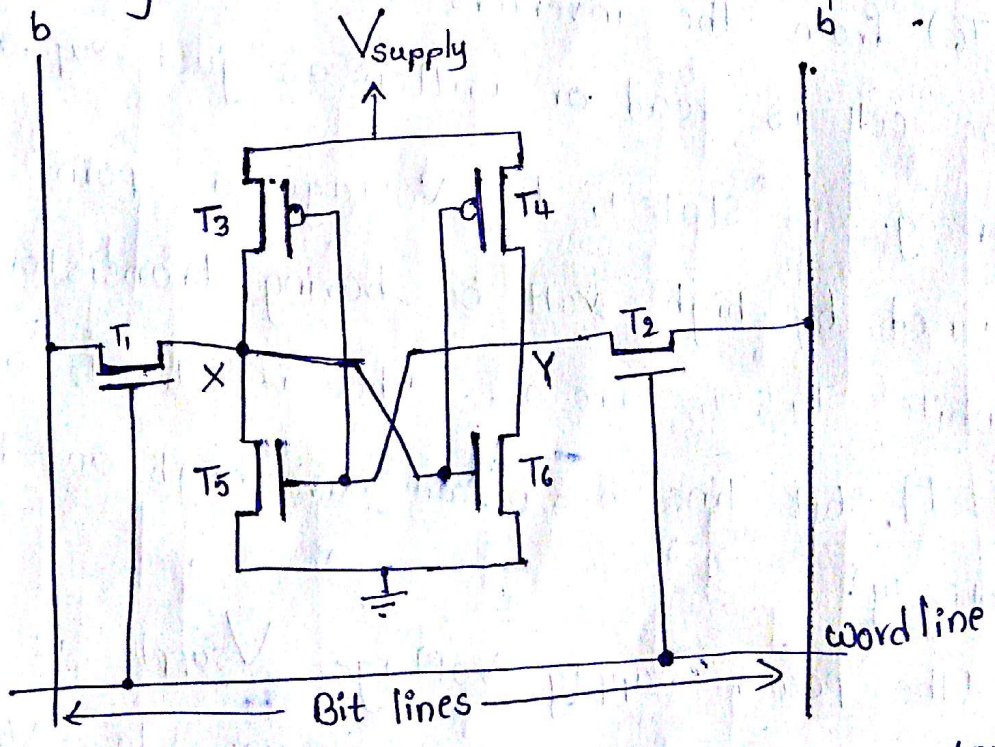



fig: CMOS cell. (Complementary metaloxide semi-conductor)

## Asynchronous DRAMS:-

Static RAM's are fast, but they come at high cost because their cells require several transistors. Less expensive RAM's can be implemented if similar Simpler cells are used. How ever, such cells do not retain their state indefinitely; hence they are called "dynamic RAM's" (DRAM's)

Information stored in DRAM memory cell in the form of a charge on a capacitor, and this Charge can be imaintained for only ten of milliseconds. Since, the cell is required to store information for much more longer time.

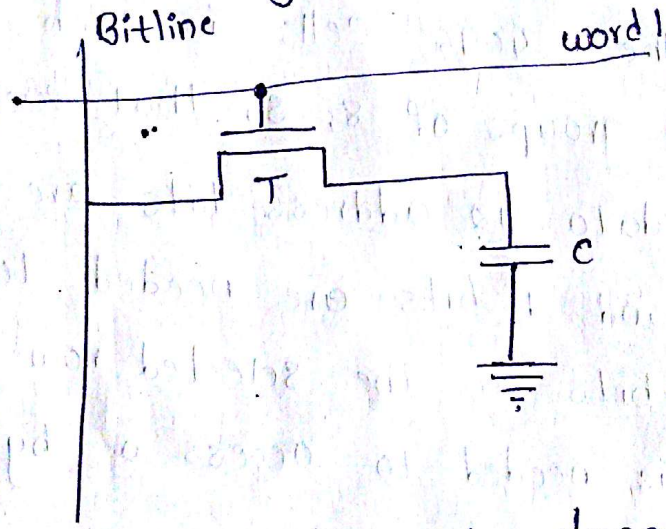
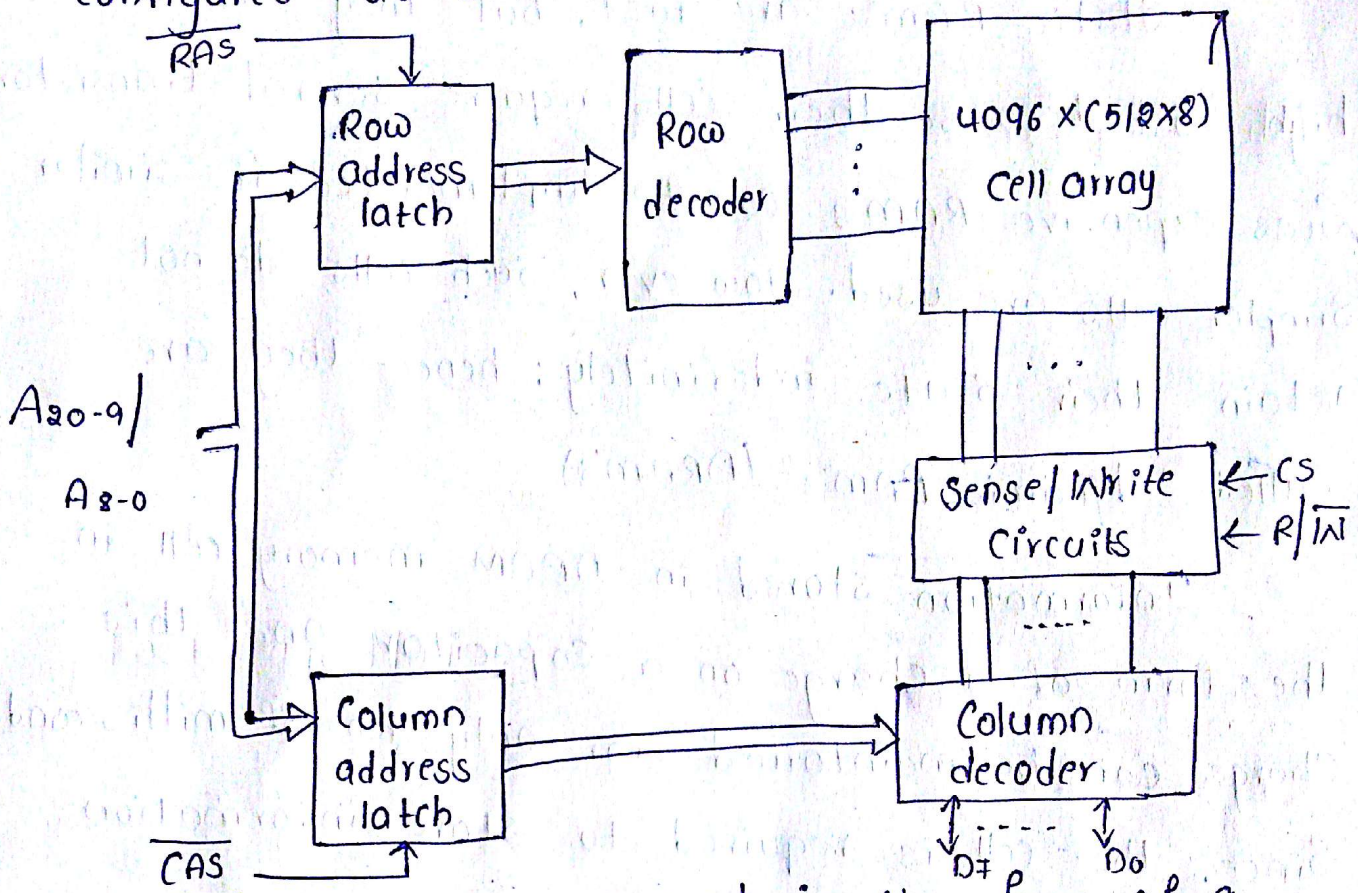



fig: A Single transistor dynamic memory cell.

The above fig shows A Single transistor dynamic memory cell. T is turned on and an appropriate voltage is applied to the bit line. This causes a known amount of charge to be Stored in Capacitor.

The following fig shows the 16-Mega bit DRAM chip, Configured as 2MX8,




The cells are organized in the form of a 4K x 4K Array. The 4096 cells in each row are divided into 512 groups of 8, so that each row can store 512B of data. 12 address bits are needed to select a row and 9-bits are needed to specify a group of 8-bits in the selected row. Thus, a 21-bit address is needed to access a byte in this memory.

The higher order 12-bits and low-order of 9-bits of the address constitute the row & column address of a byte. During a read/write operation, the row address is applied first.

It is loaded into the row address latch is reffered to a signal pulse on the "Row Address Strobe (RAS)" input of the chip. Then a Read Operation is initiated, in which all cells on the selected row are read & refreshed.

After the row address is loaded, the column is applied to the address pins are loaded into the column address latch under the control of "Column Address Strobe (CAS)" signal.

The information in the latch are decoded and the appropriate group of 8 sense/write are selected. If the control signal indicates the Read Operation the o/p values of the selected circuits ared transferred to the datalines $D_{7-0}$. For write Operation the information on the $D_{7-0}$ lines is transferred to the selected circuits. This information is then used to Overwrite the contents of the selected cells in the corresponding 8-Columns.

The timing of the memory device is controlled asynchronously. A specalized memory controller circuit provides the necessary control signals. RAS & CAS signals. The processor must take into account the delay in the response of the memory. Such memories are reffered as "ASYNCHRONOUS DRAM's".

## Advantages :-

=> It has high density & low cost, these are widely used.

=> Available chips are of size 1M to 256M bits & larger chips are developed.

=> A DRAM chip is organized to read/write a number of bits in parallel.

=> It provides flexibility in designing memory systems.

## Fast Page Mode :-

When the DRAM is accessed, the contents of all 4096 cells in the selected row are sensed, but only 8-bits are placed on the data lines $D_{7-0}$. This byte is selected by the column address bits $A_{8-0}$. A simple modification can make it possible that is a latch can be added to the o/p of the sense amplifier in each column.

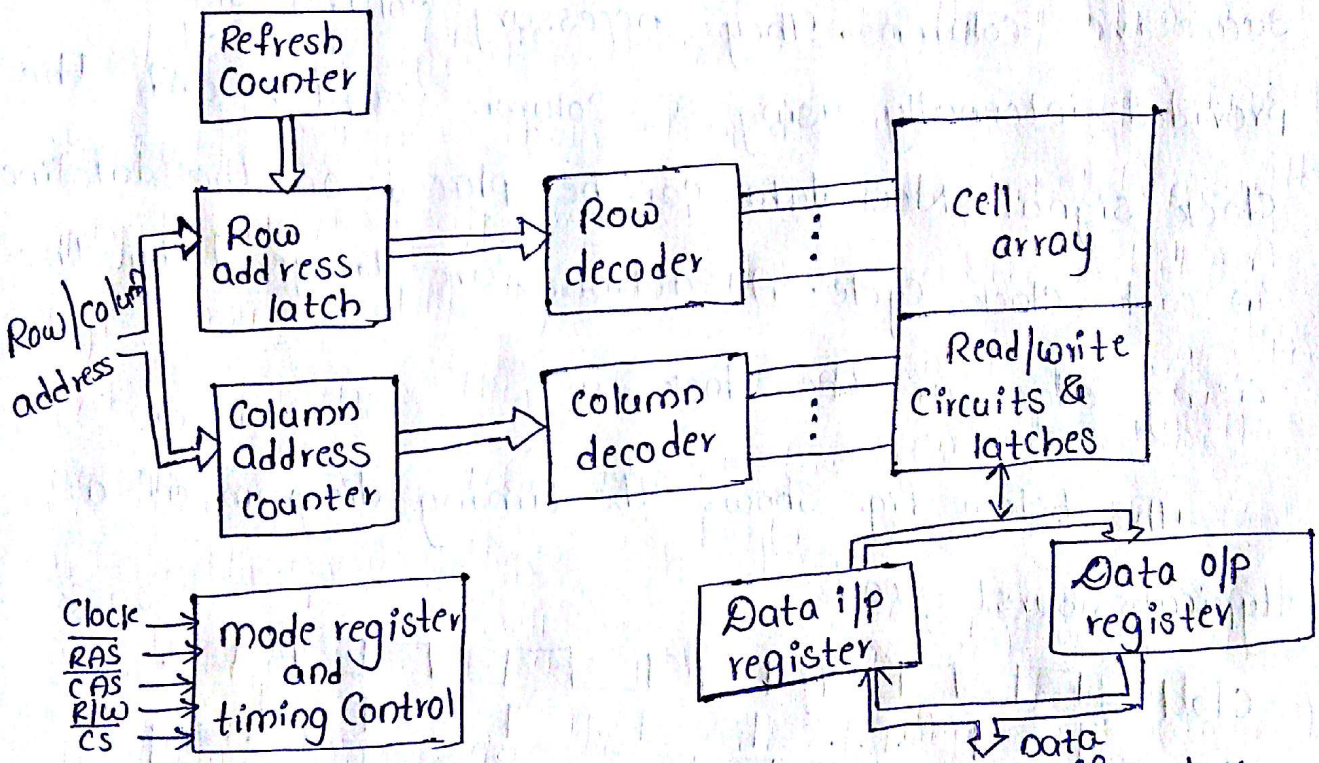
The most usefull arrangement is to transfer the bytes in sequential order, which is achieved by applying a consecutive sequence column addresses under the contorl of sucessive CAS signals. This scheme allows transferring a block of data at a much faster rather than can be acheived for transfers involving random addresses.

The block transfer capability is referred to as "fast Page Mode" features.

## Synchronous DRAMs:-

More recently developments in memory technology have resulted in DRAM's whose operation is directly Synchronized with a clock signal. Such memories are known as Synchronous DRAM's (SDRAM's).

The below fig shows the Structure of SDRAM:-



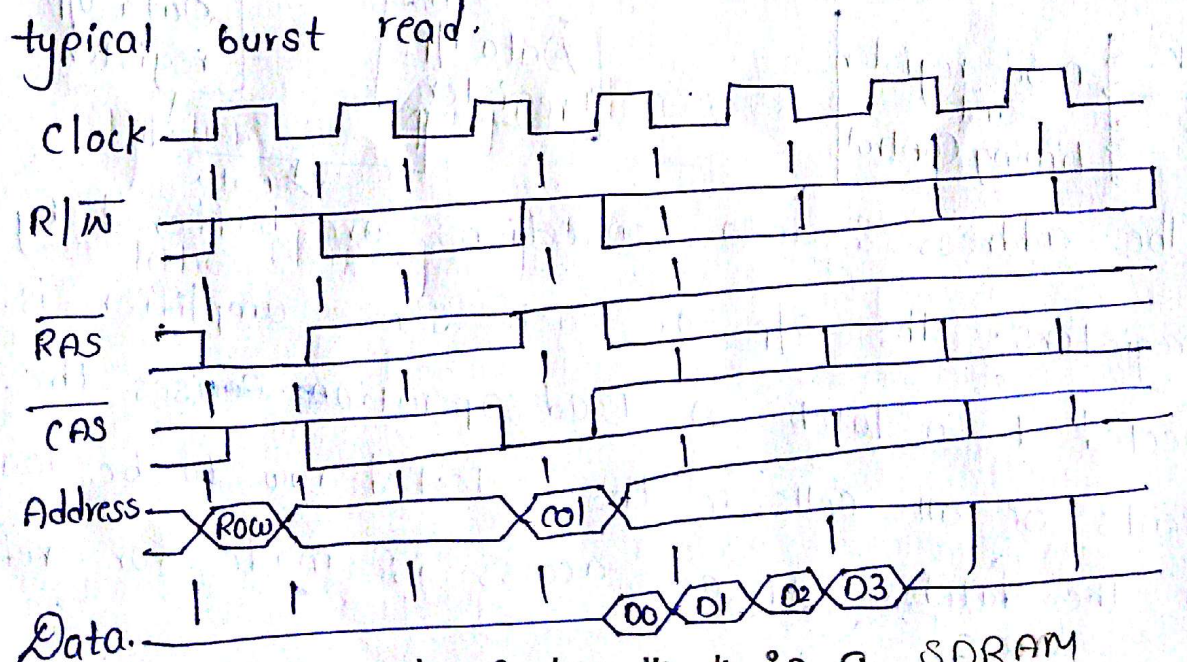

The address & data Connections are buffered by means of registers. The o/p of each Sense amplifier is Connected to a latch. A Read operation Causes the Contents of all cells in the selected row to be loaded into the latches. If an access is made for refreshing

purposes but it will not change the contents of these latches and it will refresh the contents of the cell. Data held in the latches and it will that corresponds to the selected columns are transferred into the data o/p register.

SRAM's have several different modes of operation, which can be selected by using control information into "mode" register.

IR In SDRAM's it is not necessary to provide externally generated pulses on the CAS line to select successive columns. They necessary control signals are provided internally using a column counter and the clock signal. New data can be placed on the datalines in each clock cycle. All actions are triggered by the rising edge of the clock.

The below fig. shows the timing diagram of a typical burst read.




Burst read of length 4 in a SDRAM

First, the row address is latched under control of the RAS signal. The memory typically takes 2 or 3 clock cycles.

## Latency & Band Width:-

Transfer between the memory and the processor involve single words for data or small blocks of words. Large blocks, constituting a page of data are transferred between the memory and the disks. The speed and efficiency of these transfers have large impact on the performance of the computer system. A good indication of the performance is given by two parameters: "Latency & Bandwidth".

The term "memory Latency" is used to refer the amount of time it takes to transfer a word of data to or from the memory. In this case of reading or writing a single word of data, the latency provides a complete indication of memory performance.

When transferring blocks of data, it is of intrest to know how much time is needed to transfer entire block. Since blocks can be variable in size. it is useful to define a performance measure in

terms of number of bits or bytes that can be transferred in one second. This measures is often reffered as "Band Width". The bandwidth of a memory unit depends on the speed of access to the stored data on the number of bits that can be accessed in parallel.

## Double Data Rate SDRAM:-

The standard SDRAM perform all actions on the rising edge of the clock signal. A similar memory device is available, which accesses the cell array in the same way, but transfers data on both edges of the clock. The latency of these devices is the same for standard SDRAM's. But, since they transfer data on both edges of the clock, their bandwidth is essentially doubled for long burst transfers. Such devices are known as "Double Data Rate SDRAM's) or (DDR SDRAM's)

## Structure of Larger Memories:-

### Static Memory System:-

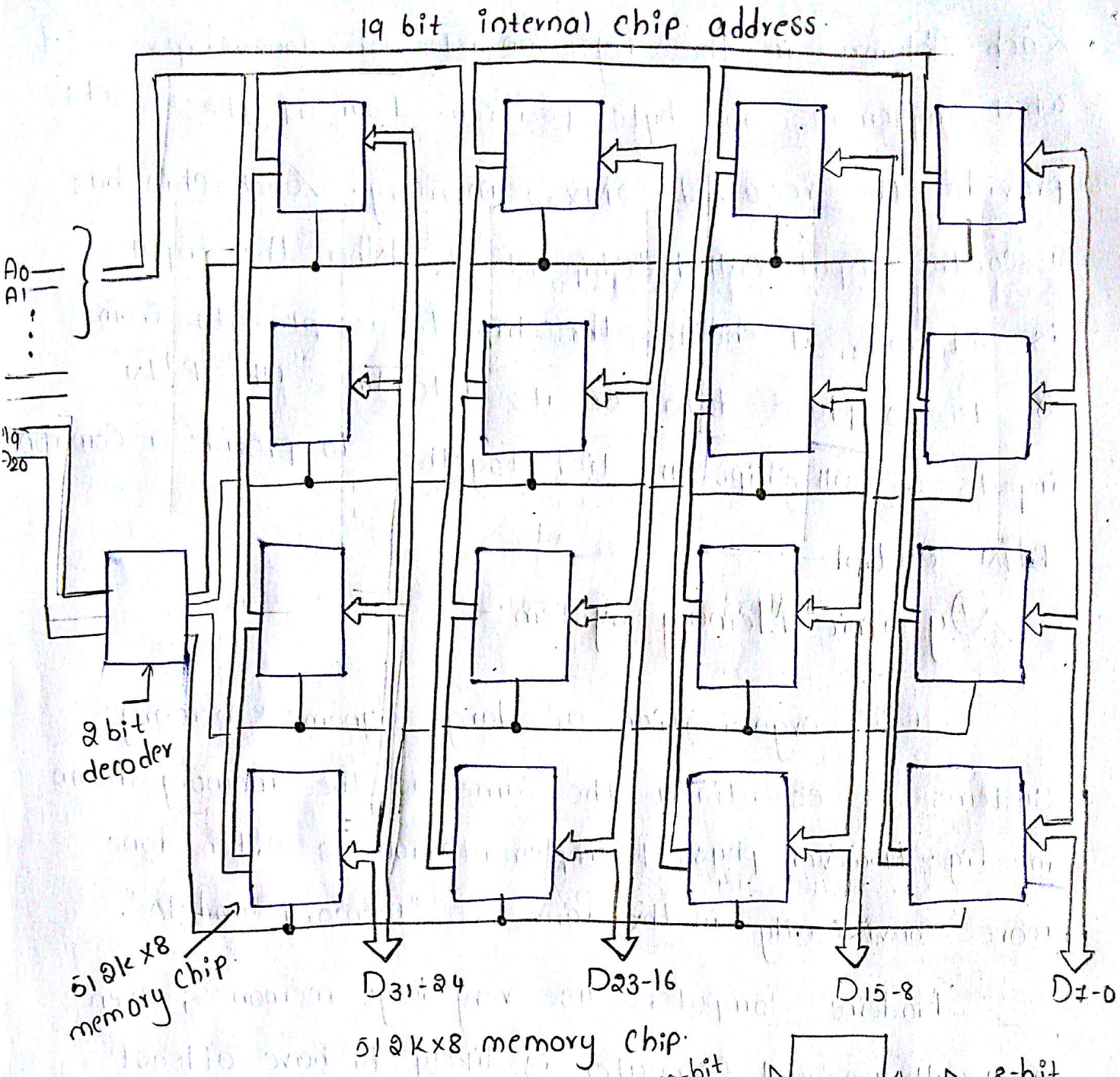
Consider a memory consisting of 2M (2,097,152) words of 32 bit each. The fig. shows how we can implement this memory using 512K×8 static memory chips.

Each Column in the fig Consists of four chips. which implement one byte position. Four of these sets provide the required 2MX32 memory. Each chip has a control input called "Chip Select". When the input is set to 1, it enables the chip to accept data from or to to place data on its data lines. The R/$\overline{W}$ inputs of all chips are tied together to provide a Common R/$\overline{W}$ control.

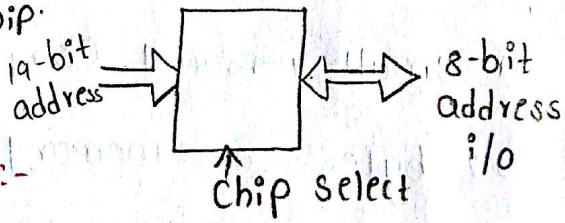
## Dynamic Memory System:-

The organization of large dynamic memory Systems is essentially the same as the memory shown in fig. However physical implementation is often done more Conviniently in the form of "memory modules".

Moderns Computers use very large memories; even a small personal Computer is likely to have atleast 32 bytes of memory. However, if a large memory is built by placing DRAM chips directly on the main memory system printed-circuit board that Contains the processor, often referred to as "mother Board", it will occupy an unacceptably large amount of space on the board.




512K x8 memory chip.




## Memory System Consideration :-

The choice of a RAM chip for a given application depends on several factors. Foremost among these factors are the cost, power dissipation, and size of the chip.

SRAM's are generally used only when very fast Operation is the primary requirement. DRAM's are

the predominant choice for implementing computer main memories. The high densities acheivable in these chips make large memories economically feasible.

## Memory Controller :-

To reduce the no. of pins, the dynamic memory chips use multiplexed address inputs. The address is divided into two parts. The higher-order address bits, which select a row in the cell array, are provided first and latched into the memory chip under control of the RAS signal. Then the lower order address bits, which select a column, are provided on the same address pins and latched together CAS signal.

A typical processor issues all bits of an address at the same time. The required multiplexing of address bits is usually performed by a memory Controller circuit, which is interposed between the processor and the dynamic memory as shown in fig. The Controller accepts a complete address and the R/$\overline{W}$ signal from the processor under control of a Request signal which indicates that a memory access operation is needed. The Controller provides the RAS-CAS timing, in addition

to its address multiplexing function. It also sends the $R/\overline{W}$ and CS signals to the memory.

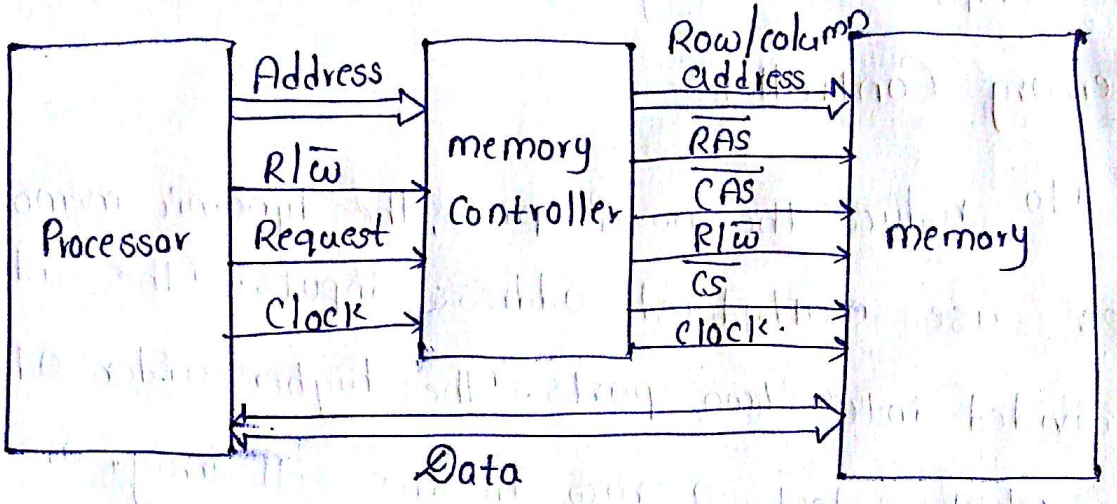



fig use of a memory Controller

The CS signal is usually active low, hence it is shown as $\overline{CS}$ in the fig. Data lines are connected directly between the processor and the memory.

## Refresh Overhead : Rambus Memory:-

The performance of a dynamic memory is characterized by its latency and bandwidth. Since all dynamic memory chips are similar organizations for their cell arrays, their latencies are tend to be similar if the chips are produced using the same manufacturing process.

DDR, SDRAM's and standrand SDRAM's are connected to the processor bus. Thus, the speed of transfers is not just a function of the speed

of the memory device — it also depends on the speed of the bus.

A very wide bus is expensive and requires a lot of space on the motherboard. An alternative approach is to implement a narrow bus that is much faster. This approach was very used by Rambus to develop a proprietary design known as "Rambus." The key feature of Rambus technology is a fast signalling method used to transfer information between chips.

The reference voltage is about 2v & the two logic values are represented by 0.3 V swings above and below Vref. This type of signalling is generally known as "differential Signalling".

Differential signaling and high transmission rates require special techniques for the design of wire connections that serve as communication links. These requirements make it difficult to make the bus wide. It is also necessary to design special circuit interfaces to deal with the differential signals. Rambus provides a complete specification for the design of such communication links, called the "Rambus Channel".

Circutary needed to interface to the Rambus channel is included on the chip. Such chips are known as "Rambus DRAM's (RORAM's)".

The original specification of Rambus provided for a channel consisting of 9 data lines and no. of control and power supply. 8 of data lines are intended for transferring a byte of data. The 9th data line can be used for purposes such as parity checking. A 2-channel Rambus is also known as "Direct RORAM, has 18 lines intended. (transfers 2 bytes of data).

Communication b/w the processor, or some other device that can serve as a master, and RDRAM modules, which serve as slaves, is carried out by means of packets transmitted on the data lines. There are 3 types of packets; request, aknowledge, and data.

Rambus technology competes directly with the DDR and SDRAM techology. Each has certain advantages & disadvantages. Finally ~~as~~ in the memory market, assuming that the performance is adequate, the decisive factor is often the price of components.

## Read Only Memories:-

Both SRAM and DRAM chips are volatile, which means that they lose the stored information if power is turned off. There are different types of ROM's they are:- PROM, EPROM, EEPROM.

## ROM:-

The below fig shows the Configuration for a ROM cell. A logic value "0" is stored in the cell if the transistor is connected to ground P, otherwise a "1" is stored. The bit line is connected through a resistor to the power supply. To read the state of the cell, the word line is activated. Thus the transistor switch is closed and the voltage on the bit line drops to near zero if there is a connection b/w the transistor and ground. A sense circuit at the end of the bit line generates the proper output value. Data are written into a ROM when it is manufactured.

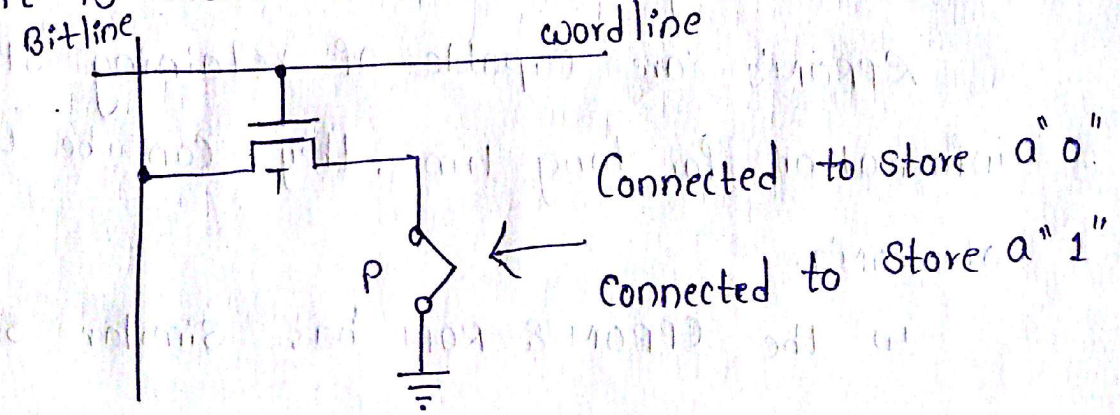



fig: ROM Cell.

## PROM:-

For small quantities, we are going to use a type of ROM called "Programmable Read Only Memory".

In this, the programability is achieved by inserting a fuse at point P (as shown in ROM fig). Before it is programmed, the memory contains all zero's (0's).

The user can insert 1's at the required locations by burning out the fuses at the locations using high-current pulses.

* PROM's provide flexibility & convinience.
* PROM's provide faster & less expensive because they can be programmed directly by the users. The ROM and EPROM are irreversible.

## EPROM:-

This is an type of ROM, which allows the stored data to be erased and new data to be loaded. Such an erasable, Reg Reprogrammable Rom is usually called as "EPROM".

EPROM's are capable of retaining stored information for long time, they can be used inplace of ROM's.

In the EPROM & ROM has similar structure. In

EPROM cell, the connection to ground is always made at point "P" and a special transistor is used, which has the ability to function as either as a normal transistor or as disabled transistor that is always turned off i.e, the transistor can be programmed to behave as a permanently open switch.

The important advantage is that their contents can be closed and reprogrammed.

The disadvantage is that a chip is removed physically from the circuit for reprogramming and the entire contents is erased by exposure to u.v (ultra violet) rays. light.

EEPROM :-

The another version of erasable PROM's that can be both programmed and erased electrically. Such chips are called EEPROM's. It is possible to erase the contents selectively.

The disadvantage is that different voltages are needed for erasing, writing & reading the stored data.

## Flash Memory:-

Flash memory is intermediate b/w EPROM & EEPROM. Like EEPROM, flash memory uses an electrical erasing technology.

An entire flash memory can be erased in one or a few sensor's which is must faster than EPROM.

Flash memory uses only one transistor per bit and so achieves the high density.

Single flash chips don't provide sufficient storage capacity for applications mentioned. Larger memory modules consisting of a no.of chips are needed. There are two popular choices for implementation of such modules: They are.

(i) Flash Cards
(ii) Flash Drives

## Speed Size & Cosst:-

It is very clear that very fast memory can be implemented if SRAM chips are used. But these chips are expensive because their basic cells have "6" transistors, which precludes packing a very large no.of cells into one single chip.

The alternative use is to use "Dynamic Ram Chips" which have much simplier basic cells and thus are much less expensive. But such memories are significantly slower.

Although dynamic memory units in the range of hundred's of megabytes can be implemented at reasonable cost, the affrodable size is still small compared to the demands of large programs with voluminous data. A solution is to provide by using secondary storage mainly, magnetic disks to implement large memory spaces. Very large disks are avaiable at reasonable price, and they are used extensively in memory Computer systems.

The entire Computer memory can be viewed as the hierarch depicted in (fig). The processor registers are at the top in terms of the speed of access. The next level of the hierarchy is relatively small amount of memory that can be implemented directly on the processor chip. This memory is called a "Processor Cache". Another next level of hierarchy is called the "main memory".

This rather large memory is implemented using dynamic memory components, typically in the form of SIMM's, DIMM's or RIMM's.

Disk devices provide a huge amount of inexpensive storage. They are very slow compared to the semiconductor devices used to implement in the main memory.

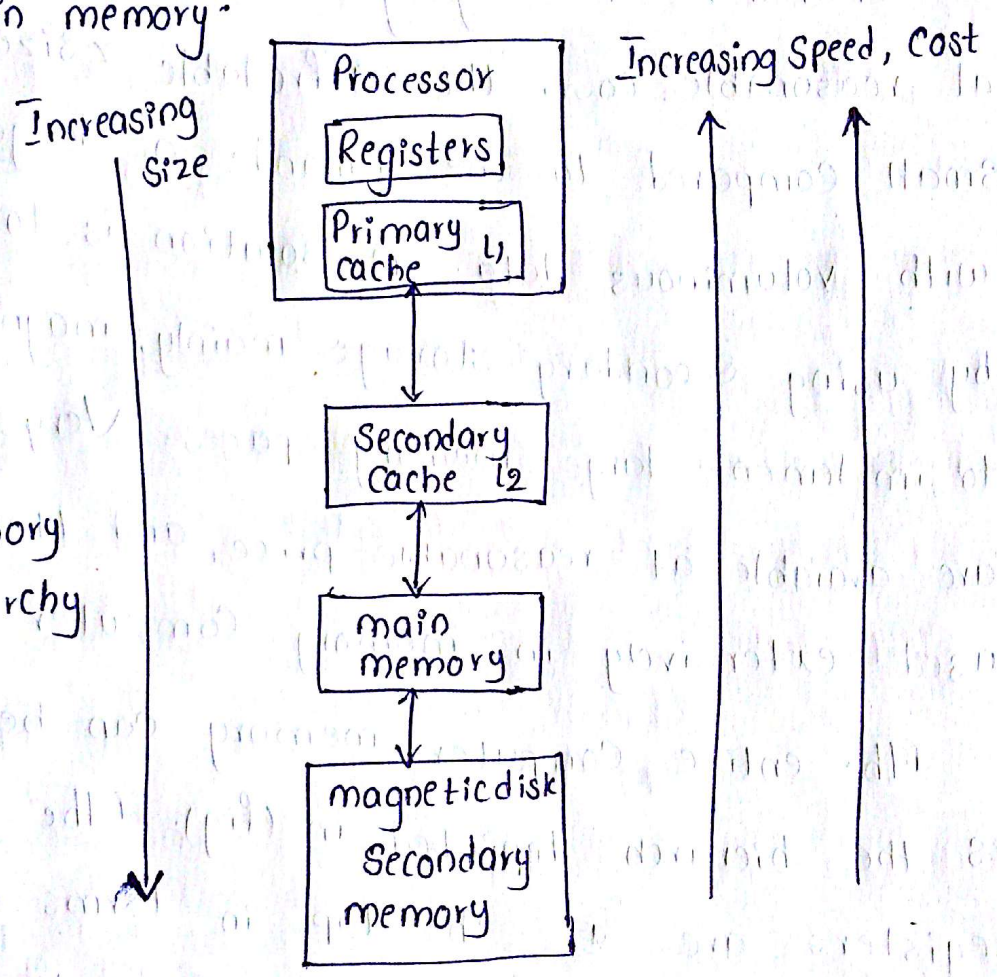



fig: memory hierarchy

## Cache Memories:-

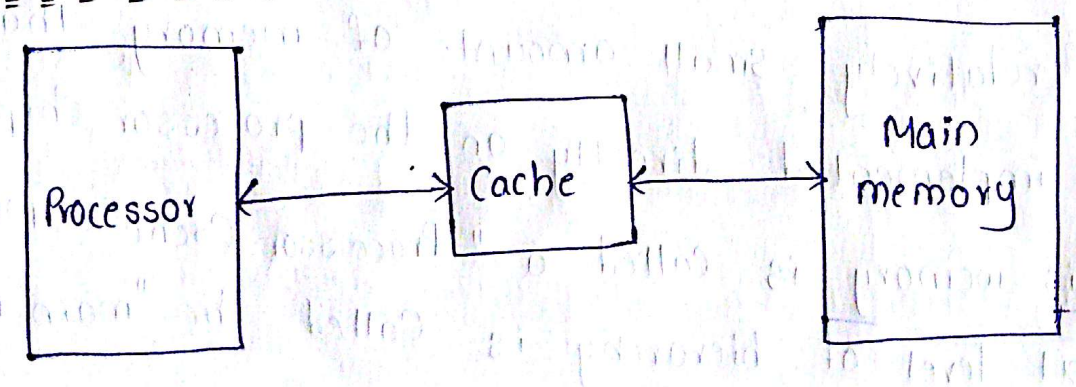



fig: Cache memory

The speed of the memory is very low, because for good performance, the processor cannot spend much of its time waiting to access instructions and data in main memory.

An efficient of cache mechanism is based on a property of Computer programs called "locality of reference'. Analysis of programs shown the most of their excecution time is spent on routines in which many instructions are excecuted repeatedly.

The instructions is localized areas of the program are excecuted repeatedly during some time period and the remainder of the program is accessed relatively infrequently. This is reffered to as "locality of reference". This is divided into two types (i) Temporal (ii) Spatial

The "temporal" means that a recently excecuted instructions is likely to be excecuted again.

The "Spatial" means that instructions in close proximity to a recently excecuted instructions (Based on instruction address).

If the active segments of a program can be placed in a fast cache memory, then the total

execution time is reduced.

The temporal aspect of the locality of reference that whenever an information item is first needed, this item should be brought into cache where it it is needed.

The spatial aspect suggests that instead of fetching one item from the main memory to the cache, it is useful to fetch several items that reside at adjacent addresses.

The term "block" to refer to a set of contiguous address locations of some size. Another item is refer to a cache block is "cache line".

The cache memory can store a reasonable no.of blocks at any given time, but this number is small compared to the total no.of blocks in the main memory.

The correspondence b/w the main memory blocks and those in the cache is specified by a "mapping function".

The cache control must decide which block should be removed to create space for new block that contains the referenced word. The

collection of rules for making this decision is "replacement Algorithm".

The R/W operation is performed on the appropriate cache location. In read operation, the main memory is not involved.

The write operation may be proceed into two ways. (i) Write-through protocal
(ii) Write-back or Copy back
(back) (back)

In write-through protocol, both cache memory and the main memory locations are updated simultaneously.

In write back or copy back, the cache memory is only updated and the updated part is marked with its associated flag bit. The bit is known as "dirty or modified bit". The main memory location is updated later, when the block Containing this marked word is to be removed them from the cache to make room for a new block.

When the addressed word in a Read operation is not in the Cache, a "read miss" occurs.

The later approach, which is called load-through or early restart. In this, it reduces the processor's waiting period. but it is more complex circuitary.

During write operation, if the addressed word is not in cache, a "write miss" occurs. Then, the write-through is used, the information is written directly into the main memory.

In this case of Write back protocol, the block containing the addressed word is first bought into the cache, and then the desired word in the cache is overwritten with the new information.

## Performance Considerations :-

Two key factors of a Computer are performance & cost performance depends on how fast instructions can be brought into the processor for execution & how fast they can be executed.

An effective way to introduce parallelism is to use an interleaved organization.

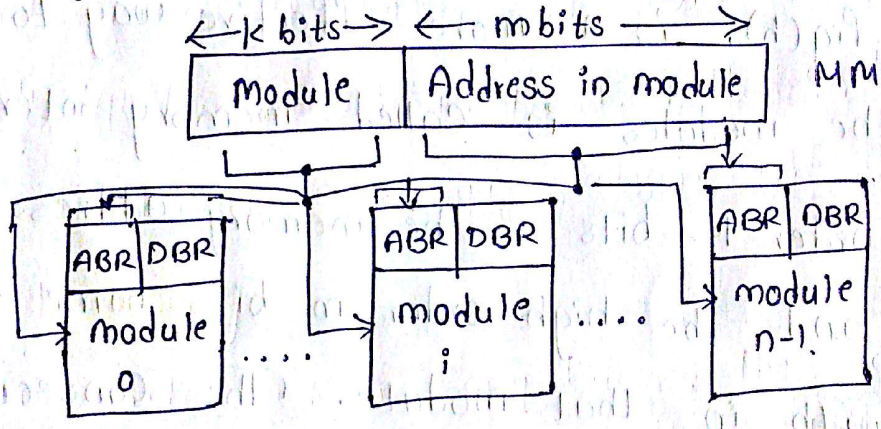
## Interleaving :-

If the main memory of a Computer is

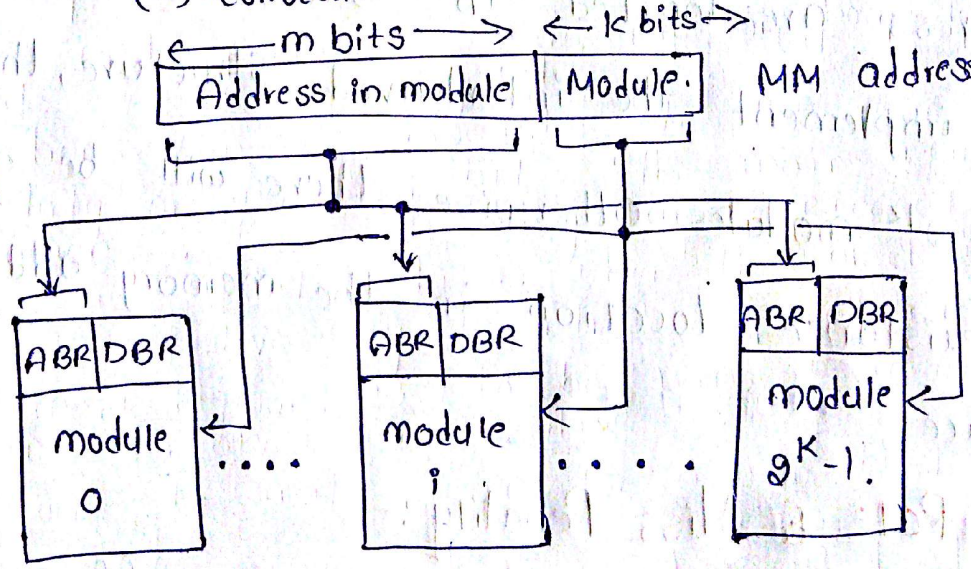
structured as collection of physically seperate modules, each with its own address bytes are address Buffer Registers (ABR). and data Buffer Register (DBR).

The memory access operations may proceed in more than one module at the same time. Thus, the aggregate rate of transmission of words to and from the main memory system can be increased.

Two methods of address layout as shown in fig.




(a) Consecutive words in module.




(b) Consecutive words in Consecutive modules

In fig(a), the memory address generalized by the processor is decoded. The high order k bits name one of 'n' modules and the low-order 'm' bits name a particular word in that module.

When consecutive locations are accessed, when a block of data is transferred to a cache, only a one module is involved. At the same time, however, devices with DMA ability may be accessing information in other memory modules.

The fig(b), is a more effective way to address the modules is called memory intervaling. The low order k bits of the memory address select a module, and the high-order m bits name or location with in that module. The Consecutive address are located in successive modules. To implement the interleaved structure, there must be $2^k$ modules otherwise, there will be gaps of nonexistance location in the memory address space.

## Hit Rate & Miss Penality :-

An excellent indicator of the effectiveness of a particular implementation of the memory

hierarchy is the sucess rate in accessing information at various levels of the hierarchy. A sucessfull access to data in a cache is called a "hit". The no.of hits stated as a fraction of all attempted accesses is called the "hit rate", and the "miss rate" is the no.of misses stated as a fraction of attempted accesses.

The entire memory hierarchy would appear to the processor as a single memory unit that has to access time of a cache on the processor chip and the size of a magnetic disk.

Performance is adversely affected by the actions that must be taken after a miss. The extra time needed to bring the desired information into the cache is called the "miss penality". In general, the miss penality is the time needed to bring a block of data from a slower unit in the memory hierarchy to a faster unit.

The rough estimate of the cache can be obtained as follows: $\frac{\text{Time without Cache}}{\text{Time with Cache}}$.

## Caches on the Processor Chip :-

When information is transferred between different chips, considerably delays are introduced in driver and receiver gates on the chips. Unfortunately, space on the processor chip is needed for many other functions, this limits the size of the cache that can be accommodated.

The average access time experienced by the processor in a system with two levels of cache is

$$t_{ave} = h_1 C_1 + [1-h_1] h_2 C_2 + [1-h_1][1-h_2] M$$

where

$h_1$ = hit rate in the $L_1$ Cache

$h_2$ = hit rate in the $L_2$ Cache

$C_1$ = time to access information in the $L_1$ Cache

$C_2$ = time to access information in the $L_2$ Cache

$M$ = time to access information in main memory.

$(1-h_1)(1-h_2)$ is Should be low. If both $h_1, h_2$ are in the 90 % range, then the no. of misses will be less.

## Other Enhancements :-

### Write Buffer :-

When the write through protocol is used, each write Operation results in writing a new Value

into the main memory. The processor must wait for the memory function to be completed. The processor typically doesnot immediately depend on the result of a write Operation, So it is not necessary for the processor to wait for write requested to be completed. To improve performance, a write buffer can be included for temporary storage of write requests. The write buffer may hold a no. of write requests. Afterward, the Contents of the buffer are written into the main memory. Thus, the write buffer also works well for the "write -back protocol".

## Prefetching:-

The new data are brought into the cache when they are first needed. The processor has to pause untill the new data arrive, which is the effect of miss penalty.

To avoid stalling the processor, it is possible to prefetch the data into the cache before they are needed. The simplest way to do this is through software.

A special prefetch instruction may be provided in the instructions set of the processor. A prefetch instruction is inserted in a program to cause the data to be loaded in the cache by the time they are needed in the program. Prefetch instructions can be inserted into a program either by the programmer or by the compiler. It is obviously preferable to have the compiler insert these instructions, which can be done with good success for many applications.

Prefetching can also be done through hardware. This involves adding circuitry that attempts to discover a pattern in memory references and then prefetches data according to this pattern.

## Lock up Free Cache :-

The software prefetching scheme does not work well if it interferes significantly with the normal execution of instructions. This is the case if the action of prefetching stops other accesses to the cache until the prefetch is completed. A cache of this type is said to be locked while it services a miss. A cache that can support multiple outstanding misses is called "lockup free". Since it

can service only one miss at a time it must circuitry that keeps track of all outstanding misses. Lock up free caches were first used in the early 1980's in the cyber series of computers manufactured by "Control Data Company".

## Virtual Memories:-

In most modern computer systems, the physical main memory is not as large as the address space spanned by an address is issued by the processor. The size of the main memory in a typical computer ranges from a few hundred megabytes to 1G bytes. When a program doesnot completely fit into the main memory, the parts of it not currently being excecuted are stored on secondary storge devices, such as "magnetic disks".

Techniques that automatically move program and data blocks into the physical main memory when they are required for excecution is called "Virtual techniques".

The binary address that the processor issues for either instructions or data are called "virtual" or Logical Addresses.

These addresses are translated into physical addresses by a combination of h/w & s/w Components.

If a virtual address refers to a part of the program or data space that is currently in the physical memory, then the contents of the appropriate location in the main memory are accessed immediately. If the referenced address is not in the main memory, its contents must be brought into a suitable location in the memory.

The below fig, shows an organization that implements virtual memory.

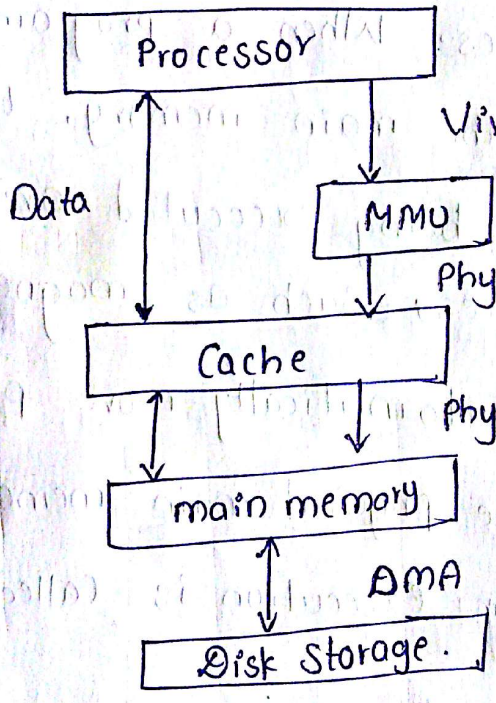



fig: Virtual memory Organization

A special hardware unit, called the "Memory Management Unit" [MMU], translates Virtual

address to physical address. When the desired data [or instructions] are in the main memory the data is fetched as and transfers to cache memory.

If the data are not in the main memory, the MMU causes the OS to bring the data into the memory from the disk. Transfer of data b/w the disk & main memory is performed using DMA.

## Address Translation:-

A simple method for translating virtual address into physical address is to assume that all programs and data are composed of fixed-length unit called "pages", each of which consisting of block of words that occupy contiguous locations in the main memory. Pages commonly range from 2k to 16 k bytes in length. Pages should not be too small, because the access time of magnetic disk is much longer than the access time of the main memory. If pages is too large it is possible that a substantial portion may not be used, yet this unnecessary data will occupy valuble space in the memory.

A virtual memory address translation method

based on the concept of fixed length pages (is shown in fig). Each virtual address generated by the processor, whether it is for an instruction fetch or an operand fetch/store operation is interpreted as "Virtual Page Number" (higher order bits) followed by an offset (low-order bits) that specifies the location of a particular byte (or word) within a page.

This information includes the main memory address where the page is stored and the current status of the page. An area in the main memory that can hold one page is called "Page frame". The starting address of the page table is kept in a "page table base register".

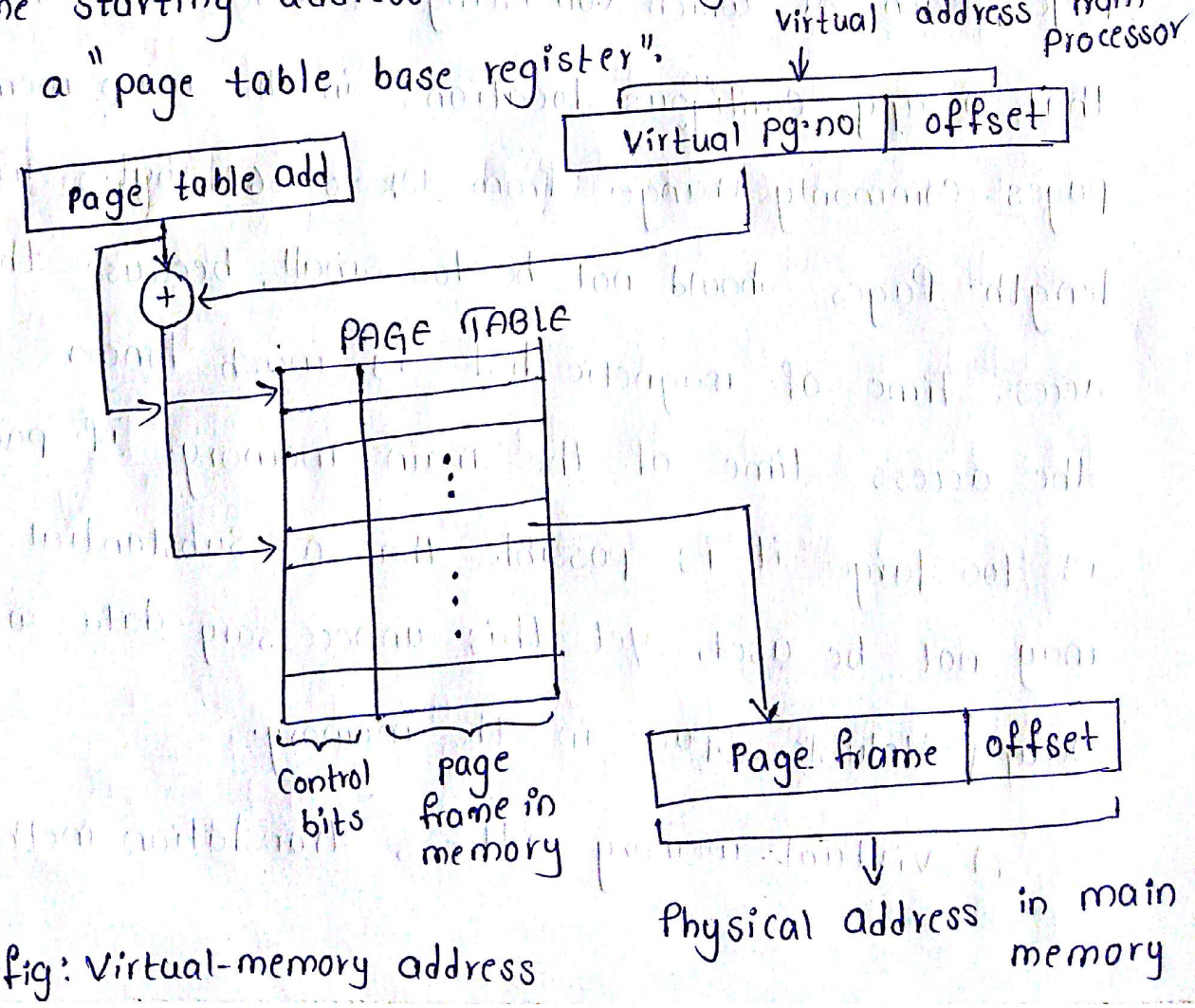



fig: Virtual-memory address

The page table information is used by MMU for every read & write access, so ideally, the page table should be situated within the MMU. Unfortunately the page table may be rather large, and since the MMU is normally implemented as part of the processor chip, it is impossible to include a complete page table on this chip. A small cache usually called the "Translation Lookaside Buffer" (TLB) is incorporated into the MMU for this purpose. The operation of the TLB with respect to the page table in the main memory is essentially the same as the operation we have in conjunction with the cache memory.

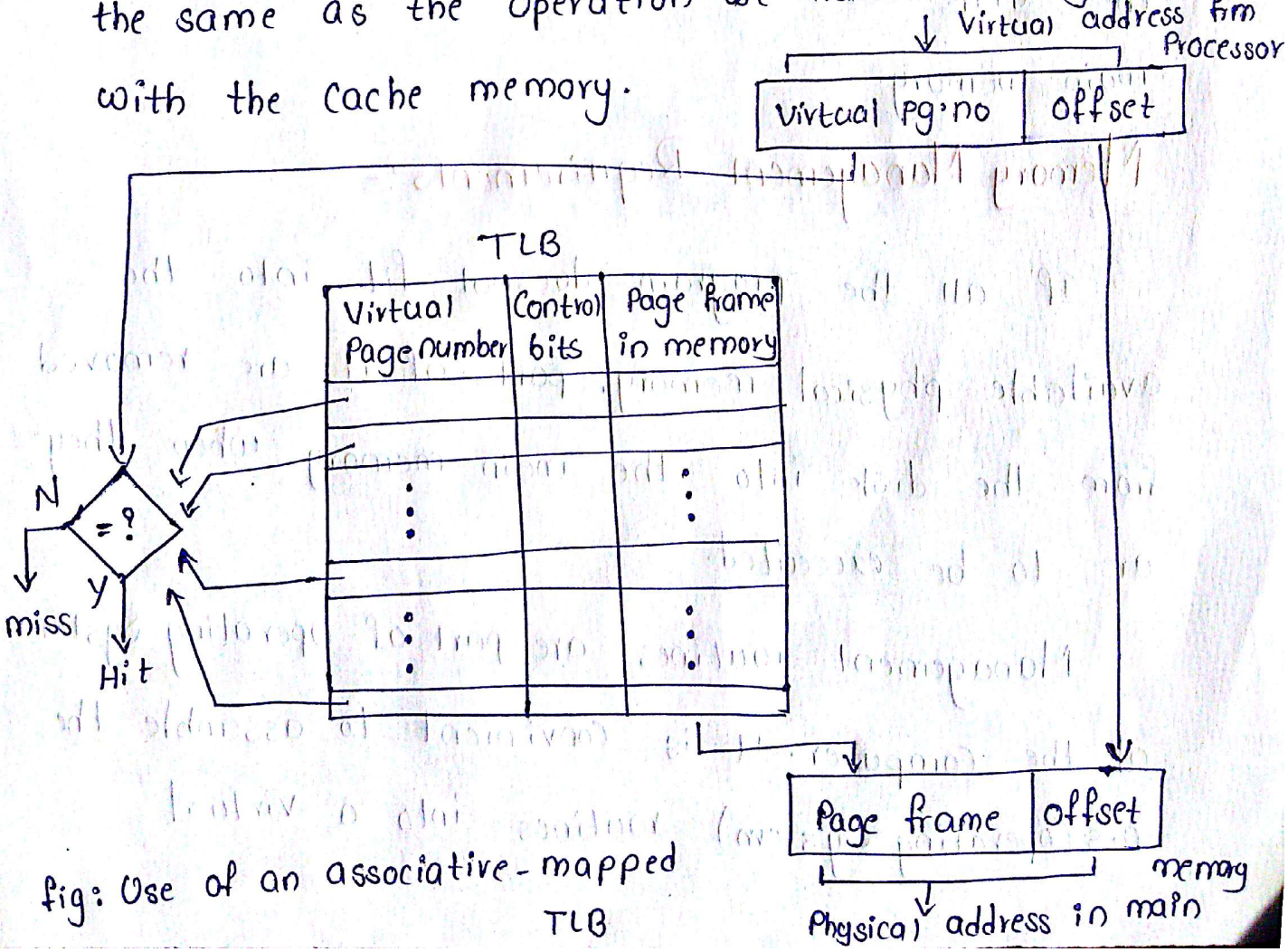



fig: Use of an associative-mapped TLB

When a program generates an access request to a page that is not in the main memory, a "page fault" have occured. The whole page must be brought from the disk into the memory before access can proceed. When it detects the page fault the MMU asks the operating System to intervene by raising an exception (interrupt).

It is essential to ensure that the interupted task can continue correctly when it resumes execution.

A modified page has to be written back to the disk before it is removed from the main memory.

## Memory Management Requirements :-

If all the program doesnot fit into the available physical memory, parts of it are removed from the disk into the main memory when they are to be excecuted.

Management routines are part of Operating System of the Computer. It is convinient to assemble the o.s (operating System) routines into a virtual

address space, called the system space, that is seperate from the virtual space in which user application programs reside. The later space is called the "user space". The MMU uses a page table base register to determine the address of the table used in the translation process.

In any computer system in which independent user programs coexist in the main memory, the notion of protection must be addressed. No program should be allowed to destroy either the data or instructions of other programs in the memory. The processor has two states of protection, the "supervisor state" and the "user state". As the name suggets, the processor is usually placed in the supervisor state when Operating System routines are being excecuted and in user state to excecute user programs. These privileged instructions, which include such operations as modifying the page table base register can only be excecuted while the processor is in the Supervisor state.

It is sometimes desirable for one application program to have access to certain pages belonging to another program. The operating system can arrange this by causing these pages to appear in both spaces.

## Secondary Storage :-

Large storage requirements of most computer systems are economically realized in the form of magnetic disks, optical and magnetic tapes which are usually reffered to as secondary storage devices.

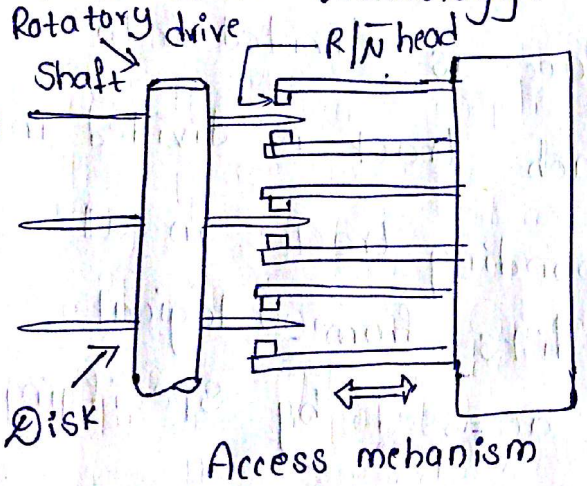
## Magnetic Hard Disk :-

As the name implies, the storage medium in a magnetic disk system consists of one or more disk mounted on a common spindle. A thin magnetic film is deposited on each disk, usually on both sides.

Digital information can be stored on the magnetic field film by applying current pulses of suitable polarity to the magnetized coil. Using the clock signal as reference, the data stored on other tracks can be read correctly.

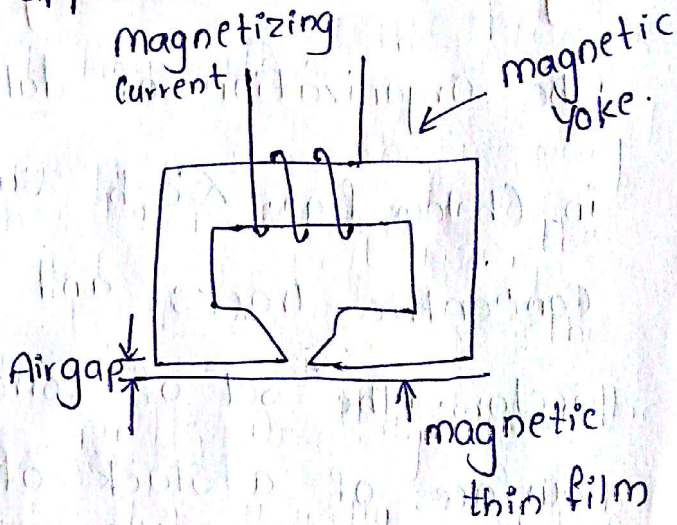
The modern approach is to combine the clocking information with the data. Several different techniques have been developed for such encoding. One simple scheme depicted in (fig.c) is known as "phase encoding or Manchester encoding". The drawback of Manchester encoding is its poor bit storage density.

In most modern disk units, the disks and the read/write heads are placed in a sealed, air-filtered enclosure. This approach is known as "Winchester technology."

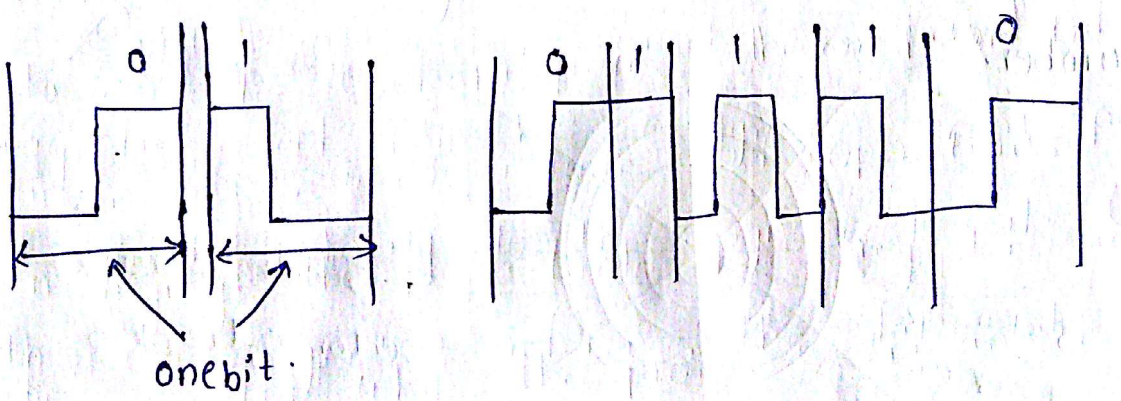



(a) Mechanical Structure

Access mechanism




(b) R/W head detail.




(c) Bit representation by phase encoding

The disk system consists of three key parts. One part is the assembly of disk platters, which is usually reffered as the disk. The second part Comprises the electromechanical mechanism that spins the disk & moves the read/write heads; it is called "disk drive". The 3rd part is a electronic Circuitary that controls the operation of system, which is called "disk Controller".

## Organization & Accessing Data on a Disk:-

The Organization of data on a disk is illustrated in (below fig). Each surface is divided into Concentric tracks, and each track is divided into sectors. The set of Corresponding tracks on all surfaces of a stack of disks form a logical cylinder. The data are accessed by specifying the surface number, track number and the sector number.

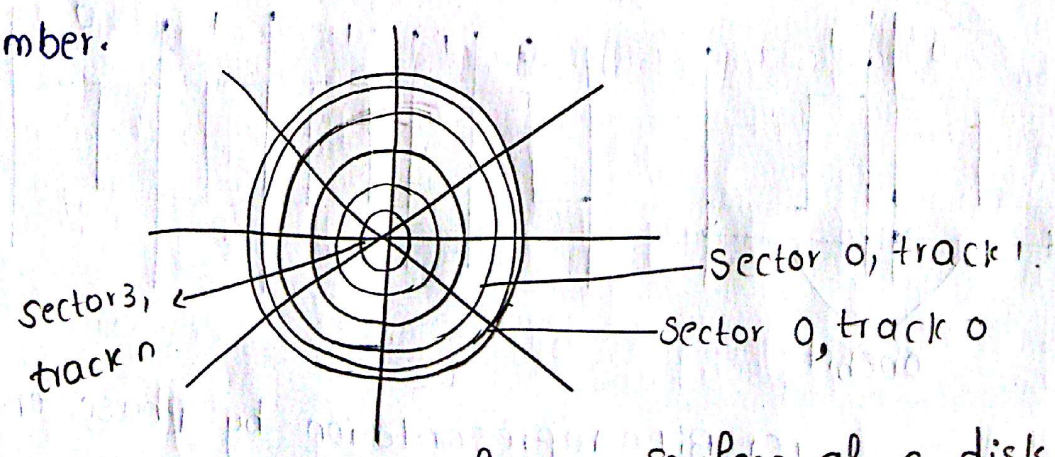



fig: Organization of one surface of a disk.

Data bits are stored serially on each track. Each sector usually contains 512 bytes of data, but other sizes may be used. The data are preceded by a sector header that contains identification information used to find the desired sector and on the selected track. There are additional bits that constitute an "Error Correcting Code" [ECC]. The ECC bits are used to detect & correct errors that may have occured in writing or reading of the 512 data bytes. To easily distinguish b/w two consecutive sectors, there is small intersector gap. In a typical computer, the disk is subsequently divided into logical partitions.

## Access Time :-

There are 2 components involved in the time delay b/w receiving an address and the begining of the actual data transfer. The first, called the "seek time", is the time required to move the R/W head to the proper track.

The second component called "rotational delay," also called "latency time". This is the amount of time that elapses after the head is positioned over the

correct track until the starting position of the addressed sector passes under the R/W head. The sum of these two delays is called "disk access time".

## Disk Controller :-

Operation of a disk drive is controlled by a disk controller circuit. Which also provides an interface b/w the disk drive and the bus that connects it to the rest of Computer system. The disk controller may be used to control more than one drive.

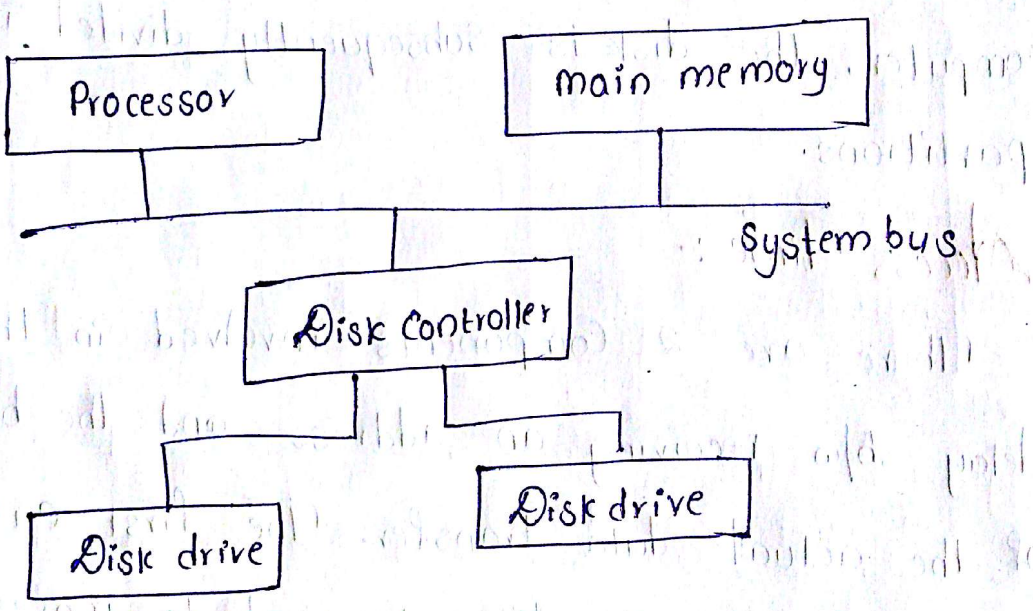



fig: Disk Connected to the system bus.

The disk controller uses the DMA scheme to transfer b/w the disk and the main memory. Actually, these transfers are from/to the data buffer, which is implemented as a part of disk

Controller module.

Main memory Address:- The address of the first main memory location of the block of words involved in the transfer.

Disk Address:- The location of the sector containing the beginning of the desired block of words.

Word Count:- The no. of words in the block to be transferred.

On the disk drive side, the Controller's major functions are:-

Seek:- Causes the disk drive to move the read/write head from its current position to the desired track.

Read:- Initiates a Read operation starting at the address specified in the disk address register. Data read serially from the disk are assembled into words and placed into the data buffer for transfer to the main memory.

Write:- Transfers data to the disk, using a control method similar to that for Read Operations.

Error Checking:- Computes the ECC value for the data read from a given sector & compares it

with the corresponding ECC value read from the disk.

## Software & Operating System Implications:-

All data transfers activities involving disks are initiated by the Operating System. The disk is a nonvolatile storage medium, so the OS itself is stored on a disk.

When power is turned off, the contents of the main memory are lost. When the power is turned on again, the OS has to be loaded into the main memory. Which takes place a part of process known as "booting". ROM stores a small monitor program that can read and write main memory locations as well as read one block of data stored on the disk at address 0. This block reffered to as "boot block", Contains a loaded program. After the boot block is loaded into the memory by the ROM monitor program, it loads the main parts of the OS into the main memory.

In a Computer system that has multiple disks, the OS may require transfer from several disks.

## Floppy Disk :-

The devices previously discussed are known as hard or rigid disk units. "Floppy disks" are smaller, simpler and cheaper disk units that consist of flexible, removable, plastic diskette coated with magnetic material. The diskette is enclosed in a plastic jacket which has opening where the R/W head makes contact with the diskette.

One of the simplest schemes used in the first floppy disks for recording data is phase or Manchester encoding. Disk encoded in this way are said to have "single density". A more complicated varient of this scheme, called "double density".

The main feature of floppy disks is their low cost & shipping convenience. How ever, they have much smaller storage capacities, longer access times, and higher failure rates than hard disks. Current standard floppy disks are 3.25 inches in diameter and store 1.44 or 2M bytes of data.

## Raid Disk Arrays :-

Semiconductor memory speeds have

improved more modestly. The smallest relative improvement in terms of speed has been in disk storage devices, for which access times are still on the order of milliseconds.

High performance devices tend to be expensive. Sometimes it is possible to achieve very high performance at a reasonable cost by using a no.of low cost devices operating in parallel.

In 1988, researchers at the university of California - Berkeley proposed a storage system based on multiple disks. They called it "RAID", for "Redundant Array of Inexpensive Disk". Using multiple disks also makes it possible to improve the reliability of Overall system. 6 different configurations were proposed. They are known as RAID levels even though there is no hierarchy involved.

RAID 0 is the basic configuration intended to enhance performance. A single large file is stored in several seperate disk units by breaking the file up into a no.of smaller pieces and storing these pieces on different disks. This is called "data striping". When the file is accesed for a read, all disks can deliver their data in parallel.

RAID 1 is intended to provide better reliability by storing identical copies of data on two disks rather than just one. The two disks are said to be mirrors of each other. Then, if one disk drive fails, all read & write operations are directed to its mirror drive.

RAID 2, RAID 3, RAID 4 levels achieve increased reliability through various parity checking schemes without requiring a full duplication of disks. All of parity information is kept in one disk.

RAID 5 also makes use of a parity-based error -recovery scheme. However, the parity information is distributed among all disks, rather than being stored on one disk.

## Commodity Disk Considerations:-

Most disk units are designed to connected to standard busses. The performance of a disk unit depends on its internal structure and the interface used to connect it to the rest of the system. The cost depends largerly on the storage capacity, but it is also affected greatly by the sales volume of a particular product.

The different types of disks are (i) ATA/EIDE Disks
(ii) SCSI Disks
(iii) RAID Disks.

## Optical Disks :-

Large storage devices can also be implemented using optical means. The familiar Compact Disk (CD), used for audio system, was the first practical application of this technology. Soon after, the optical technology was adapted to the Computer environment to provide high-capacity read-only storage referred to as "CD-ROM".

The first generation of CD's was developed in mid 1980's by the sony & Philips Companies, which also published a Complete Specification for these devices.

## CD Technology :-

The optical technology that is used for CD system is based on laser light source. A laser beam is directed onto the surface of the spinning disk. Physical indentations in the surface are arranged along the tracks of the disk.

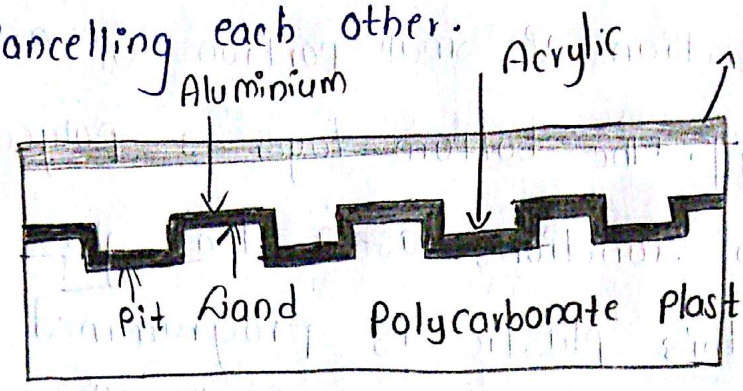
They reflect the focused beam toward a photo detector, which detects the stored binary pattern.

A cross-section of small portion of CD is [shown in fig]. The bottom layer is polycarbonate plastic, which functions as a clear glass base. The surface of this plastic is programmed to store data by indenting it with "pits". The unindented parts are called "lands". A thin layer of reflecting aluminium material is placed on top of a programmed disk.

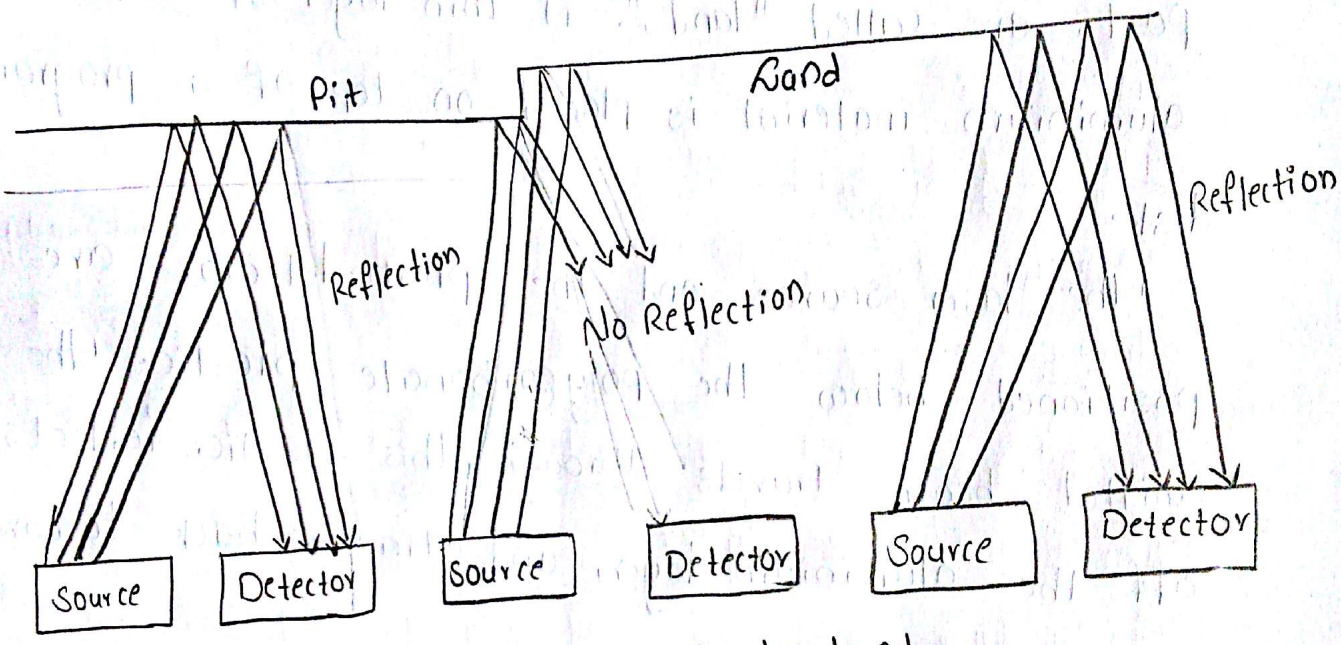
The laser source and the photodetector are positioned below the polycarbonate plastic. The emitted beam travels through this plastic, reflects off the aluminium layer, and travels back toward the photodetector.

When the light reflects from the solely from the pit, or solely from the land, the detector will see the reflected beams as a bright spot. But, a different situations arises when the beam moves through the edge where the pit changes to the land, and vice versa. The pit will is recessed on quarter of the wavelength of the light. Thus

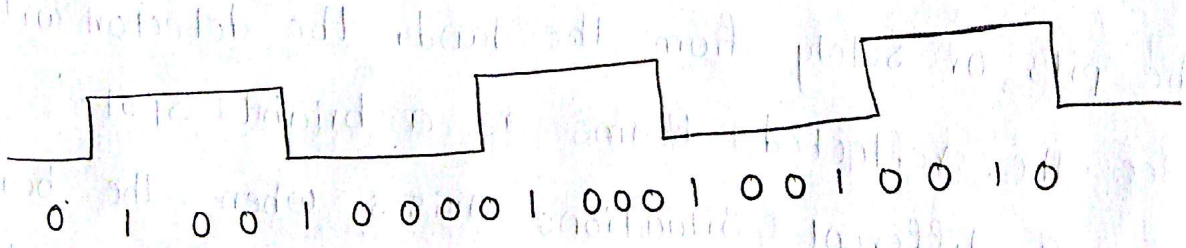
the reflected wave from the pit will be 180° out of phase with the wave reflected from the land, Cancelling each other.




(a) Cross Section




(b) Transition from pit to land.




(c) Stored binary pattern

fig: Optical Disk.

At the pit-land & land-pit transistions the detector will not see a reflected beam & will detect a dark spot. The CD is 120 mm in diameter. There is a 15-mm hole in centre. Data are stored on tracks that Cover pits area from 25-mm radius to 58-mm radius. The Space b/w the tracks is 1.6 microns. pits are 0.5 microns wide and 0.8 to 3 microns long. There are more than 15,000 tracks on a disk. If the entire track Spiral were unraveled, it would be over 5mm km long!

## CD: Rom:-

Since information is stored in binary form in CD's they are suitable for use as a storage medium in Computer system. The biggest problem is to ensure the integrity of stored data. Because the pits are very small, it is difficult to implement all of the pits perfectly. It is necessary to use additional bits to provide error checking and Correcting Capability. CD's are used in Computer applications have such capability. They are called CD-Rom's, because after manufacture their Contents can only be read, as with semiconductor ROM chips.

Stored data are organized on CD ROM tracks in

the form of blocks that are called sectors. There are several different formats for a sector. One format is known as, Mode 1, uses 2352-byte sectors. There is a 16-byte header that contains a synchroni-zation field used to detect the begining of the sector and addressing information used to identify the sector.

Error detection and correction is done at more than one level. As in CD's each byte of stored information is encoded using a 14-bit code that has some error correcting capability. This code can correct single bit errors.

CD-ROM drives operates at a no.of different rotational speeds. The basic speed, known as 1x, is 75 sectors per second. The importance of CD-ROM's for computer systems stems for their large storage capacity and fast access times compared to other inexpensive portable media, such as floppy disks & magnetic tapes. They are widely used for the distribution of software, databases, application programs and Vedio games.

CD-Recordables:-

A new type of CD was developed in the

late 1990's on which data can be easily recorded by a computer user. It is known as CD-Recordable [CD-R]. A spiral track is implemented on a disk during manufacturing process. A laser in a CD-R drive is used to burn pits into a organic dye on the track. When a burned spot is heated beyond a critical temperature, it becomes opaque. The written data stored permanently. Unused portions of disk can be used to store additional data on later time.



## CD-Rewritables:-

The most flexible CD's are those that can be written multiple times by user. They are known as CD-RW's [CD-ReWritables].

The basic structure of CD-RW's is similar to the structure of CD-Rs. Instead of using organic dye in the recording layer, an alloy of silver, indium, antimony and tellurium is used. This alloy has intresting and useful behaviour when it is heated and cooled.

The CD-RW drive uses three different laser powers. The highest power is used to record the pits. The middle power is used to put the alloy

into the crystalline state; it is refferred to as "erase power". The lowest power is used to read the stored information. There is a limit on how many times a CD-RW disk can be rewritten. Presently this can be done upto 1000 times.

CD-RW's provide a low-cost storage medium. They are suitable for archival storage of information that may range from databases to photographic images. The CD-RW drives are used for low-volume distribution of information just like CD-R's.

## DVD Technology :-

The sucess of CD technology an the continuing quest for greater storage capability has led to the development of "Digital Versatile Disk" [DVD]. The first DVD standard was defined in 1996 by consortium companies. The objective is to be able to store a full length movie on one side of DVD disk.

The physical size of DVD disk is the same as for CD's. The disk is 1.2 mm thick and it is 120mm in diameter. Its storage capacity is made larger than that of CD's by several design changes:

⇒ A red light laser with a wavelength of 635 nm is used instead of the infrared light laser used in CD's which has a wavelength of 780m. The shorter wavelength makes it possible to focus the light to a smaller spot.

⇒ Pits are smaller, having a minimum length of 0.4 micron.

⇒ Tracks are placed close together, the distance between tracks is 0.74 micron.

Using these improvements leads to DVD capacity of 4.7 Gbytes.

Access time for DVD drives are similar to CD drives. However, when DVD disk rotates at the same speed, the data transfer rates are much higher because of the higher density of pits.

## DVD-RAM:-

A rewritable version of DVD devices known ans DVD-RAM, has also been developed. It provides a large storage capacity. Its only disadvantages are higher price and relatively slow writing speed. To ensure that the data have been recorded correctly on the disk, a process known as write verification is performed.

## Magnetic Tape Systems:-

Magnetic tapes are suited for off-line storage of large amount of data. They are typically used for hard disk backup purpose and for archival storage. Magnetic tape recording uses the same principles are used in magnetic-disk recording. The main difference is that the magnetic film is deposited on a very thin 0.5 or 0.25 inch wide plastic tape. A seperate R/W head is provided for each bit position on the tape, so that all bits of a character can be done read or written in parallel. One of the character bits is used as parity bit.

Data on the tape are organized in the form of records seperated by gaps [as shown in fig]. The record gaps are long enough to allow the tape to attain its normal speed before the begining of the next record is reached. To help users to organize large amount of data, a group of related books records is called "file". The begining of file is identified by "file mark".

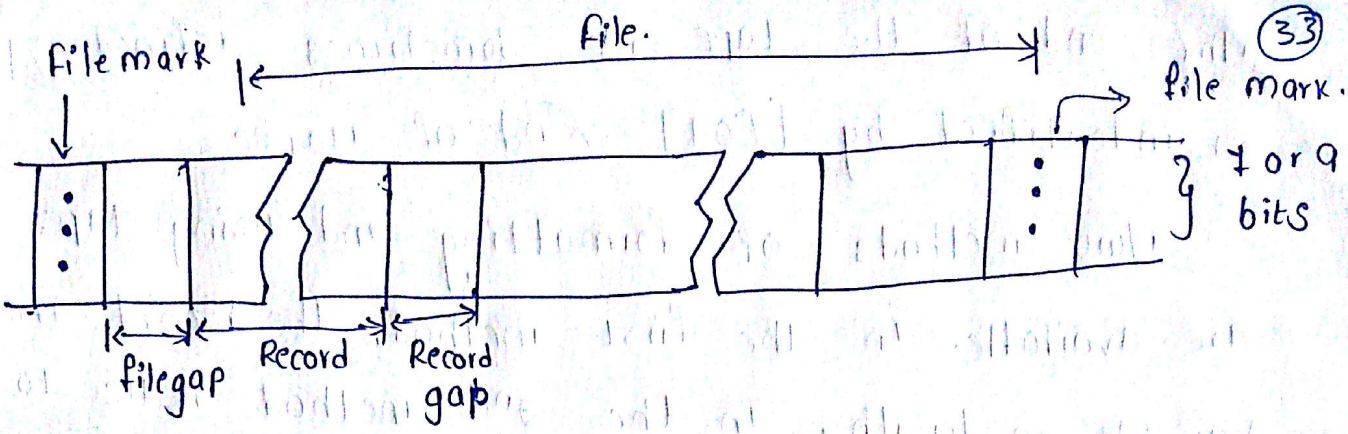



fig: Organization of data on magnetic tape.

The first record following a file mark can be used as "header" or "identifier" for this file.

The Controller of magnetic tape drive enables the execution of a no. of control commands in addition to R/W Commands. Control commands include the following operations:

=> Rewind tape
=> Rewind and upload tape
=> Erase tape
=> Write tape mark
=> Forward space one record
=> Backspace one record
=> Forward space one file
=> Backspace one file.

The tape mark reffered to in operation 'write tape mark" is similar to the file mark except that is used for identifying the begining of the tape.

The end of the tape is sometimes defined by or indentified by [EOT] End of Tape.

Two methods of formatting and using tapes are available. In the first method, the records are variable in length. In the 2nd method is use to fixed-length records, In this case it is possible to update records in place.

## Cartridge Tape System:-

Tape systems have been developed for backup of on-line disk storage. One such system use an 8-mm Vedio format tape housed in Cassette. These units are called "Cartridge tapes". They have Capacities in range of 2 to 5 giga bytes, and handle data transfers at the rate of few hundred kilobytes per second. Multiple Cartridge Systems are available that automate the loading and unloading of Cassettes so that 10's of gigabytes of on-line storage can be backed up unattended.