

Welcome to ACE Engineering Academy - online live class

Subject: **Computer Organization and Architecture**

Faculty: **Y.V. Ramaiah**

9866339106

Subject

Computer organization & Architecture

Chapters (Topics)

- I. Computer Arithmetic ✓
- II. Memory Organization
- III. Secondary Memories
- IV. Basic processor organization and Design
- V. Pipeline organization
- VI. Control unit Design
- VII. IO Organization

Chapter 2 Memory Organization

- Introduction ✓
- Memory Basics ✓
- Memory Classification ✓
- Memory Size Expansion ✓
- Primary Memory
- Secondary Memory ✓
- ROM and its design ✓
- **RAM and its design** ✓
- Memory Hierarchy ✓
- Cache Memory
- Mapping Techniques
- Different misses occurred in cache
- Different block replacement techniques
- Tag directory design
- Associative Memory

Q. If the associativity of a processor cache is doubled while keeping the capacity and block size unchanged, which one of the following is guaranteed to be NOT affected?

$$N = 2 \checkmark$$

$$N = 4 \checkmark$$

- (a) Width of tag comparator ✗
- (b) Width of set index decoder ✗
- (c) Width of way selection multiplexer ✗
- (d) Width of processor to main memory data bus ✓

Q. A CPU has 32-bit memory address and a 256 KB Cache memory. The Cache is organized as a 4-way set associative Cache with Cache block size of 16 bytes.

From the above question, What is the number and size of the comparator required for tag matching?

- (a) four 4-bit comparators
- (b) sixteen 16-bit comparators
- (c) four 16-bit comparators.
- (d) one 4-bit comparator

$$\begin{aligned}
 PA &= 32 \\
 CMW &= 2^8 \\
 CA &= 18 \\
 N &= 4 \\
 \frac{\text{TAG size}}{16 \text{ bit}} & \\
 \text{Comparator size} & \\
 \text{No. of Comp} &= 4
 \end{aligned}$$

Q. An 8KB direct-mapped write-back cache is organized as multiple blocks, each of size 32-bytes. The processor generates 32-bit addresses. The cache controller maintains the tag information for each cache block comprising of the following

- 1 Valid bit
- 1 Modified bit

As many bits as the minimum needed to identify the memory block mapped in the cache.

What is the total size of memory needed at the cache controller to store meta-data (tags) for the cache?

- (a) 4864 bits
- (b) 6144 bits
- (c) 6656 bits
- (d) 5376 bits

$$\begin{aligned}
 W\beta & \\
 CMW &= 2^{13} \\
 BW &= 2^5 \\
 PA &= 32 \\
 CL &= \frac{2^3}{2^5} = 2^8 \\
 &= 256 \\
 \text{Tag size} &= 19 \text{ bits} \\
 (19 + V + M) * 256 &= 19 \text{ bits} \\
 21 * 256 &= 5376 \text{ bits}
 \end{aligned}$$

Q. A computer has a 256 K Byte, 4-way set associative, write back data cache with block size of 32 Bytes. The processor sends 32 bit addresses to the cache controller. Each cache tag directory entry contains, in addition to address tag, 2 valid bits, 1 modified bit and 1 replacement bit.

The size of the cache tag directory is

- (a) 160 K bits
- (b) 136 K bits
- (c) 40 K bits
- (d) 32 K bits

$$(16 + 4) * 8 \text{ K bits} = 160 \text{ K bits}$$

$$\begin{aligned} C_L &= \frac{18}{2^5} \\ N &= 4 \\ B_W &= 2^5 \\ P_A &= 32 \end{aligned}$$

$$\begin{aligned} \text{Tag size} &= (32 - 18) \\ &= 2^{13} = 8 \text{ K} \\ &= 16 \text{ bits} \end{aligned}$$

Q. Consider two cache organizations: The first one is 32KB 2-way set associate with 32-byte block size. The second one is of the same size but direct mapped. The size of an address is 32 bits in both cases. A 2-to-1 multiplexer has latency of 0.6 ns while a k-bit comparator has a latency of $k/10$ ns. The hit latency of the set associative organization is h_1 while that of the direct mapped one is h_2 .

The value of h_1 is

- (a) 2.4 ns
- (b) 2.3 ns
- (c) 1.8 ns
- (d) 1.7 ns

$$T_{\text{comp}} = 1.8 \text{ ns}$$

$$T_{\text{mux}} = 0.6 \text{ ns}$$

$$T_{\text{Total}} = 2.4 \text{ ns}$$

Q. Consider two cache organizations: The first one is 32KB 2-way set associate with 32-byte block size. The second one is of the same size but direct mapped. The size of an address is 32 bits in both cases. A 2-to-1 multiplexer has latency of 0.6 ns while a k-bit comparator has a latency of $k/10$ ns. The hit latency of the set associative organization is h_1 while that of the direct mapped one is h_2 .

The value of h_2 is

(a) 2.4 ns

(c) 1.8 ns

(d) 1.7 ns

$$\left. \begin{array}{l} \text{I} \\ \text{II} \end{array} \right| \quad \left. \begin{array}{l} N = 2^5 \\ Bw = 2^5 \\ C_{imw} = 2^{15} \\ C_A = 15 \end{array} \right| \quad \left. \begin{array}{l} N = 1 \\ Bw = 2^5 \\ C_{imw} = 2^{15} \\ C_A = 15 \end{array} \right| \quad T_{MUX} = 0.6 \text{ ns} \quad PA = 32 \quad T_{cmp} = k/10 \text{ ns}$$

Hit latency
= Amount of time needed to declare hit

TAG size h_2 (direct map)

$$PA - CA = (32 - 15) = 17 \text{ bit}$$

$$Cmp \text{ size} = 17 \text{ bit}$$

$$T_{cmp} = \frac{17}{10} \text{ ns} = \underline{\underline{1.7 \text{ ns}}}$$





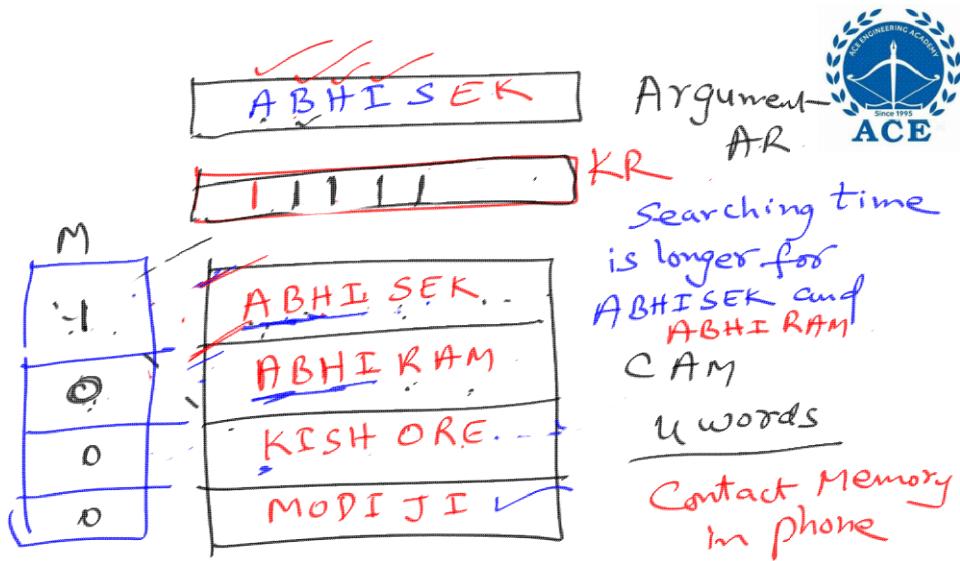
Associative Memory

- It is also known as Content Addressable Memory (CAM) because in this, words are accessed by their Content rather than their Address.
- It is fastest memory and faster than Cache Memory but it is costliest Memory because it requires Complex H/W Logic circuit.

Application:- C.A.M. is used in the design of Tag directory in different mapping techniques.

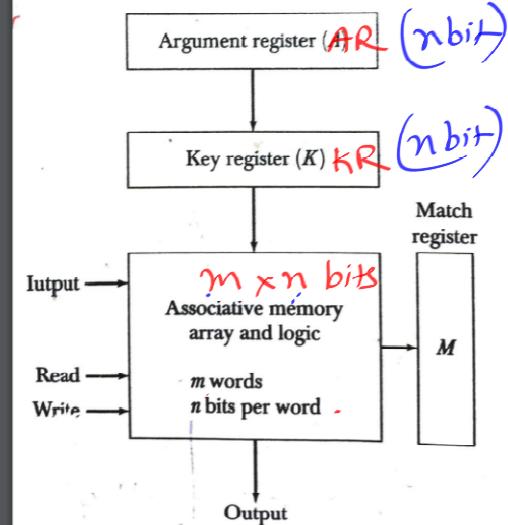
This principle is also used in Contact Search in mobile phones and Google Search.

Our Example



AR is used to hold the word to be searched
→ KR is used for masking application

Figure 12-6 Block diagram of associative memory.



KR is used to limit the searching time
→ size of A-R and K-R = size of word



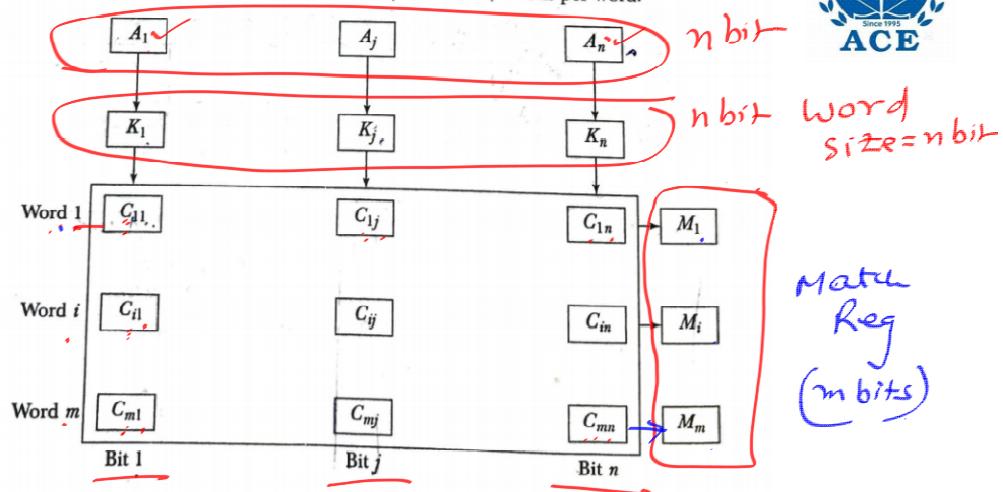


Match Register:- Its size is m bits

because one bit is needed for one word

- During word search if match occurs for certain word; the respective bit is '1' otherwise it is Reset.
- After setting only one bit in Match Register, the respective word will be sent to the output

Figure 12-7 Associative memory of m words, n cells per word.



H/W logic diagram for declaring Hit/Miss in N-way Block Set Associative Mapped Cache



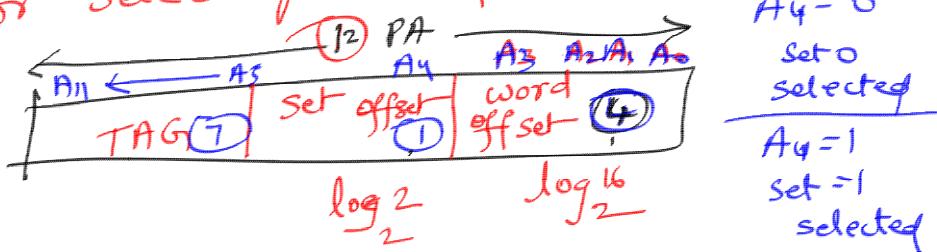
- No. of Comparators required = N because each set has 'N' blocks and all blocks Tag information (in the Selected set) will be compared with physical address TAG information in parallel.
- $N \times 1$ MUX is needed for selecting the Target data (from the selected line) on Data Bus.

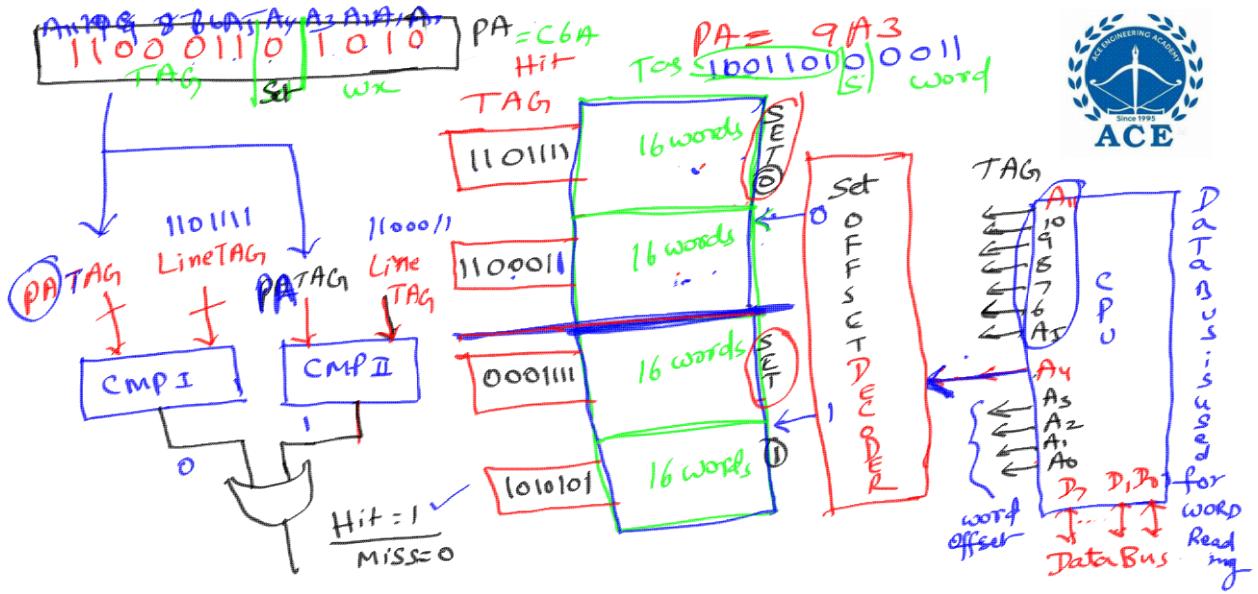
$$\begin{aligned} \text{let } N &= 2, CL = 4, BW = 16, MB = 256 = 2^8 \\ CMW &= 2^2 \times 2^4 = 2^6 = 64 \quad (2^2) \quad = 2^6 \\ S &= 2 \quad S_0 \text{ and } S_1 \quad MMW = 2^8 \times 2^4 \\ &\quad = 2^{12} \quad = 2^12, PA = 12 \end{aligned}$$



It requires one set off set Decoder

for selecting one of the sets.





→ Set offset Decoder is used to select one of the sets, in the selected set two no. of cache lines are available

→ Let Target word (M_{mw}) Address is $(C6A)_{16}$

$1100011|0|1010$ word

TAG Hit

→ Let Target word (M_{mw}) Address is $(9A3)_{16} = 10011010 0011$

MISS occurs

- Each Cache line is designed with one word offset Decoder
- To select one of the Cache lines in the Selected Set, the 2×1 mux is used, after selecting the target line; the target word will be transferred to CPU through Data Bus.

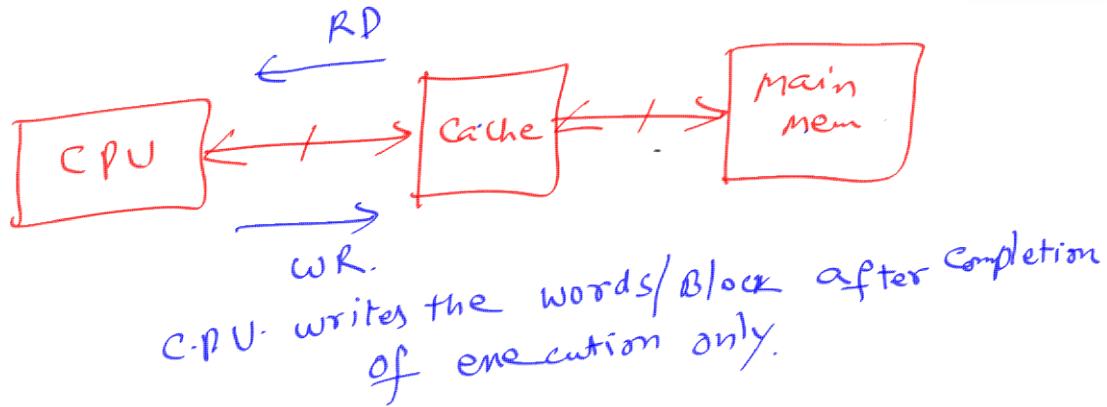


Different writing techniques in Cache and Main Memory



Read miss: while reading, let word is not available in Cache, it known as Read Miss.

Write Hit/miss:— while performing writing operation; let the Cache memory is having same word/block (without modification) is available; it is known as write hit



Different writing technique used in
Cache mapping techniques



-
- (i) write through
 - (ii) write Back



Write through:- CPU writes the target information word/Block in both Cache and Main Memory in parallel.

In this, the Recently executed block is available to Secondary Memories

(through DMA controller)

Demerit:- Valuable memory space is wasted

Merit:- It does not require extra bits (like Dirty, clean, Replacement) cost is cheaper

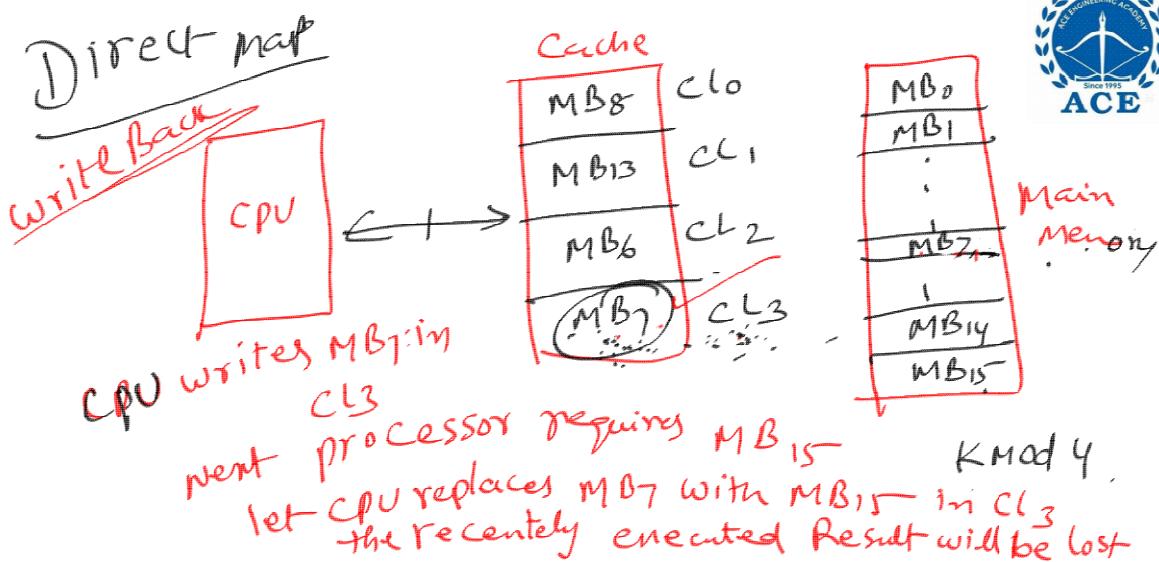
Write Back Cache:- In this, CPU first writes the information only in Cache Memory. Whenever CPU wants to evict that block then CPU first sends that block to Main Memory and it will be evicted in Cache Memory.

Merit: Valuable memory space is not wasted.



In write back Cache, each cache line design requires additional bits (like Dirty bit, clean bit, modified bit and Replacement bit etc) along with its Tag field.

(These bits are not fixed for all systems) Designer defines the no. of bits





Dirty bit (modified bit) :-

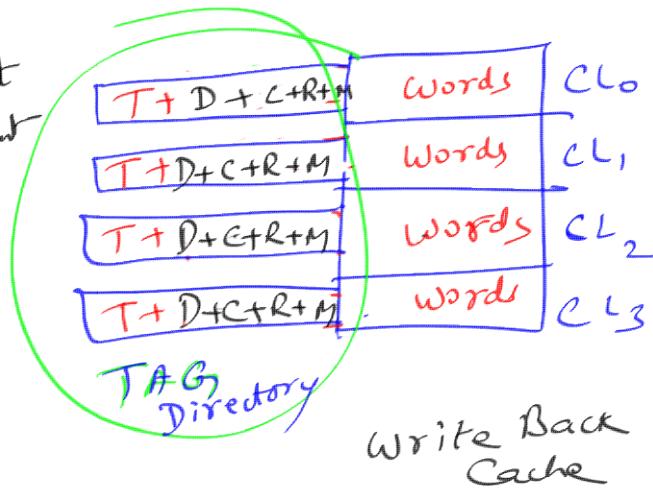
It is set ($D=1$), when CPU writes the (modified) information in Cache Mem otherwise means, no modification has been done in Cache block then $D=0$ (Reset) and clean bit = 1 set

Replacement bit - It is set when the old block is replaced with new block only

Tag directory size in write Back Cache

$$(T + \text{extra bits}) * CL .$$

D = Dirty bit
 C = Clean bit
 M = Modified bit
 R = Replace bit



These additional bits are used for protecting the data in Cache lines

MSQ

- Q. Let WB and WT be two set associative cache organizations that use LRU algorithm for cache block replacement. WB is a write back cache and WT is a write through cache. Which of the following statements is/are FALSE?
- (A) Each cache block in WB and WT has a dirty bit. False
 - (B) Every write hit in WB leads to a data transfer from cache to main memory. False
 - (C) Eviction of a block from WT will not lead to data transfer from cache to main memory. TRUE
 - (D) A read miss in WB will never lead to eviction of a dirty block from WB. False



a
b
and
d



a) In Content Addressable Memory

- ~~a)~~ Size of A.R = size of a word MSQ
- ~~b)~~ Size of K.R = No. of words in CAM
- ~~c)~~ Size of M.R = No. of words in CAM
- ~~d)~~ Size of KR = size of a word
a, c and d

Chapter 3 Secondary Memories



This topic is very essential
for understanding Memory
management unit topic
in O.S. subject

Chapter-3 Secondary memories



- (i) Role of secondary memories in a system ✓
- (ii) Different secondary memories used in a system ✓
- (iii) Magnetic surface memory ✓ VV
Imap
- (iv) Magnetic Hard disk memory ✓
- (v) Magnetic tape memory
- (vi) CD and DVD

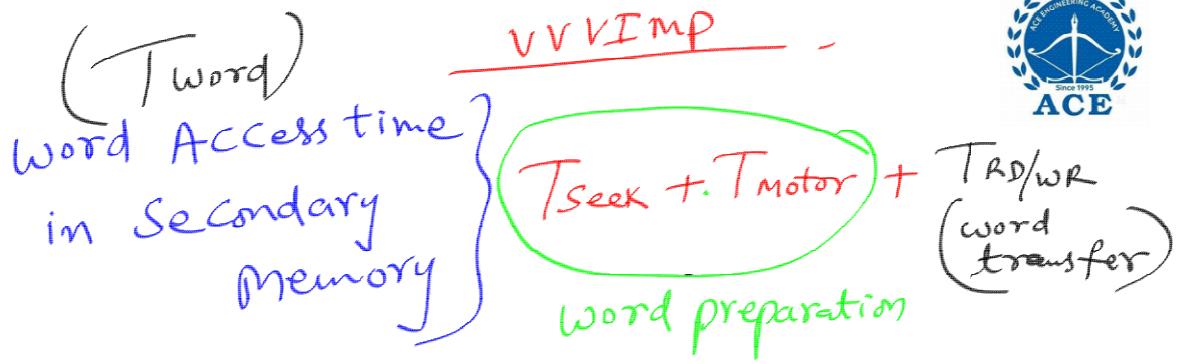
→ Secondary memory is operated with electro mechanical devices HEAD and MOTOR

→ HEAD is used to perform Read and write operation and Motor is used to Rotate the Data Recording surface in clockwise direction

→ Head and motor are operated in milliseconds time only



- The amount of time required to position the R/W Head from the current track to the desired track is known as Seek time (T_{seek})
- After positioning the Head in the target track; the motor is needed to rotate for bringing the target sector under the Head, this time is known as Rotational delay.



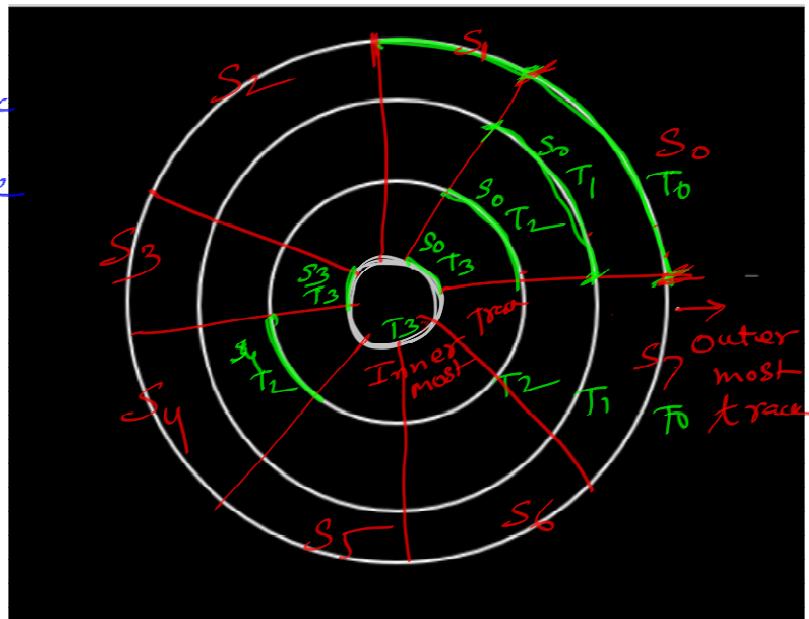
In electronic memory } - Word Access (T_{word}) = TRD/WR
(primary memory) time

single magnetic Memory Surface



The process of writing the data in magnetic and optical memory is known as Data Recording

Single magnetic surface



$$T = 4$$
$$T_0 \text{ to } T_3$$

$$S = \text{no of sectors/track} = 8$$
$$S_0 \text{ to } S_7$$



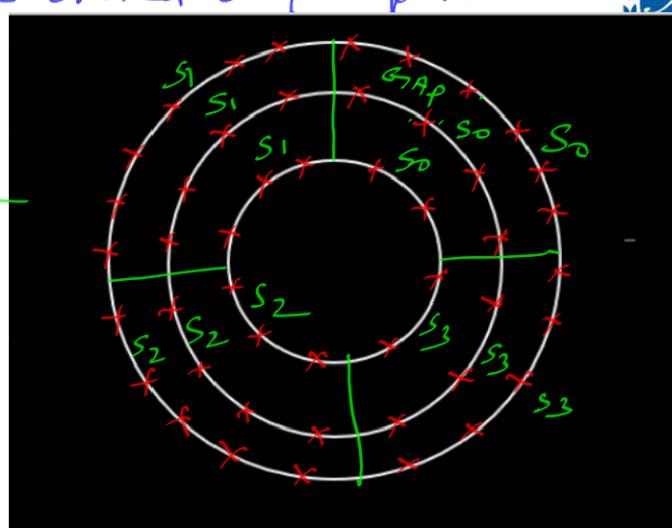
- magnetic surface is divided into the no. of concentric circles known as **Tracks**
- Track is divided into no. of mangable units, known as **Sectors**.
- while performing RD/WR operation, the system selects only one sector size data. Hence sector is known as **mangable unit**.



- Each track has 8 sectors.
- Surface has 4 tracks.
- Total no. of sectors in this Surface
 $= T \times S = 4 \times 8 = 32$

→ Data is recorded only on perimeter boundaries.

$$x = \text{bit } \frac{0}{1}$$



Data can be Recorded in a magnetic surface in Two techniques.



- (i) Constant bit Recording Density.
- (ii) Variable bit Recording Density