CS & IT ENGINEERING

COMPUTER ORGANIZATION AND ARCHITECTURE

Cache Organization

Lecture No.- 01

By- Vishvadeep Gothi sir

Recap of Previous Lecture

- Topic: RAM Chip
- Topic: ROM Chip
- Topic: Multiple Chips in Single Memory System
- Topic: DRAM Refresh

Topics to be Covered

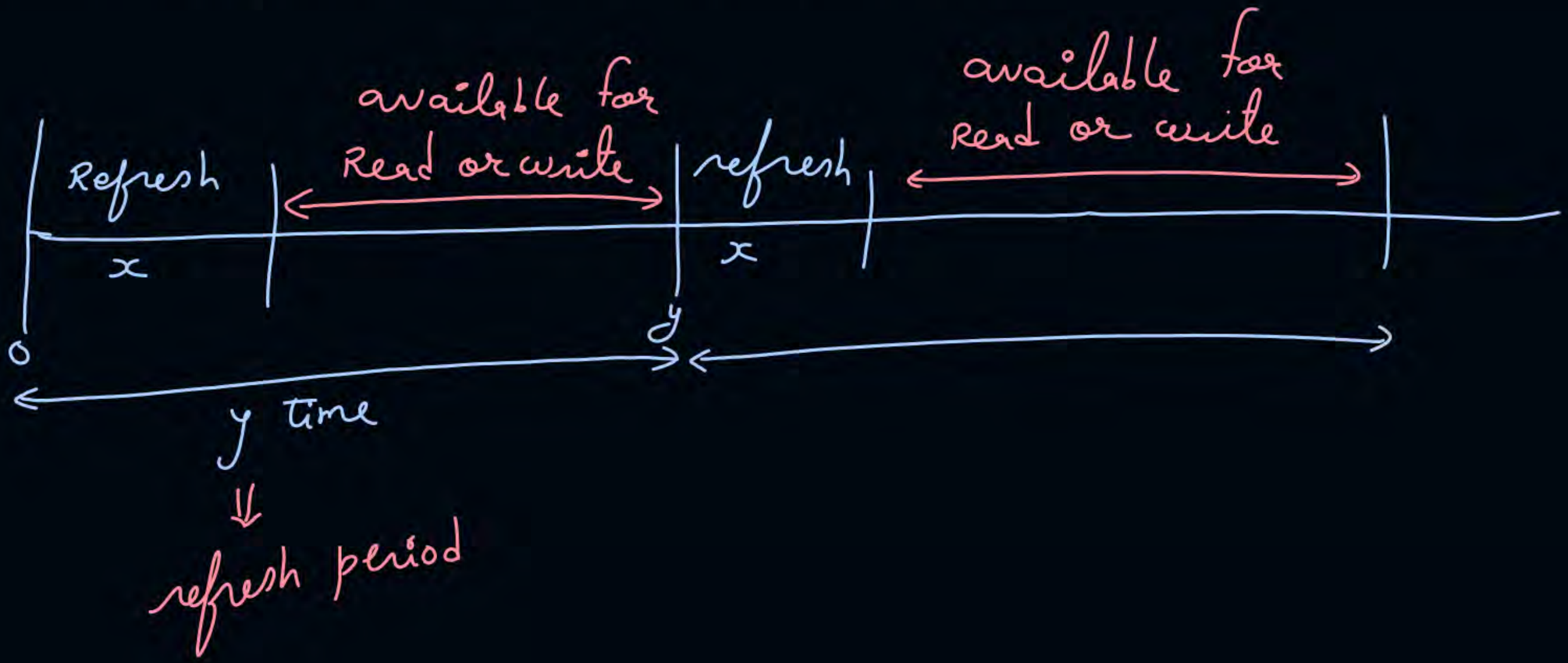
- Topic: Associative Memory
- Topic: Locality of Reference
- Topic: Cache Memory

Topic : DRAM Refresh

H.W. Questⁿ $= \dfrac{4 \times 8 \text{ bits}}{1 \times 1} = 32$ chips

arrangement

DRAM refresh is done periodically.



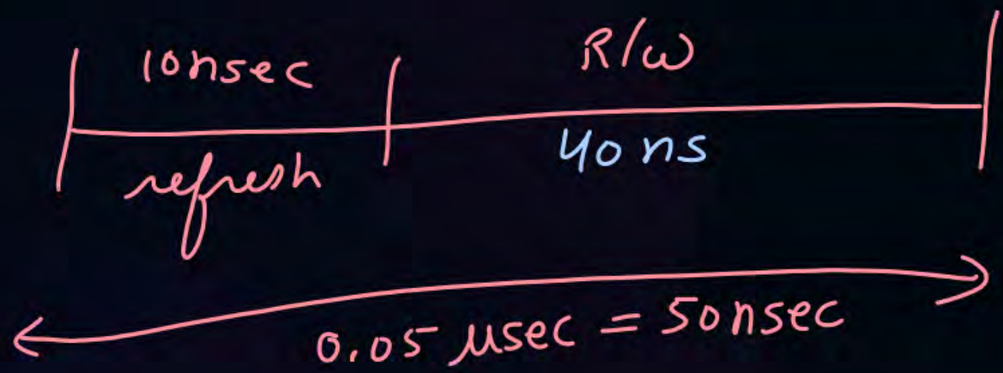
⇓

refresh period

[NAT]

#Q. Consider a DRAM which can be refreshed in 10ns. The refresh period is 0.05 microseconds.

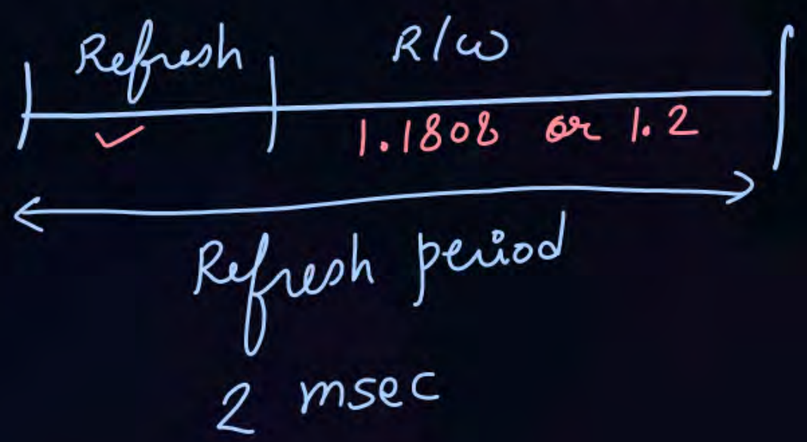
1. % of time taken in refresh? $\frac{10}{50} * 100\% = 20\%$
2. % of time remaining for read write is? $= \frac{40}{50} * 100\% = 80\%$



[NAT]

GATE-2018

#Q. A 32–bit wide main memory unit with a capacity of 1 GB is built using 256M X 4-bit DRAM chips. The number of rows of memory cells in the DRAM chip is 2^{14}. The time taken to perform one refresh operation is 50 nanoseconds. The refresh period is 2 milliseconds. The percentage (rounded to the closet integer) of the time available for performing the memory read/write operations in the main memory unit is _____?



Total refresh time $= 2^{14} * 50$ nsec

$= 2^{10} * 2^{4} * 50$ ns

$= 800 * 2^{10}$ ns

2^{10} taken as 1000

$= 800$ usec

$= 0.8$ msec

Time remaining for R/W $= 2 - 0.8 = 1.2$

% of time for R/W $= \frac{1.2}{2} * 100\%$

$= 60\%$

2^{10} taken as 1024

$= 819200$ nsec

$= 0.8192$ msec

$= 2 - 0.8192 = 1.1808$

$= \frac{1.1808}{2} * 100\%$

$= 59.04\% = 59\%$

[NAT]

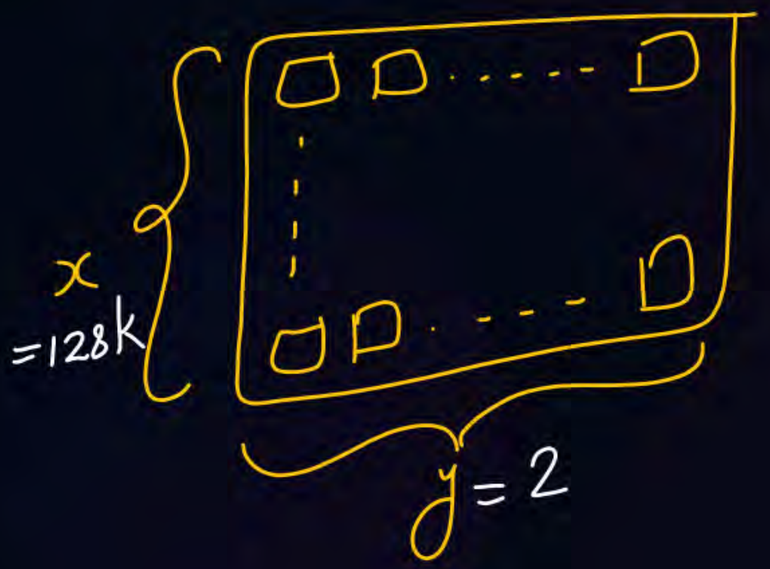
no. of cells

$\{ x = 128k, \; y = 2 \}$

#Q. A DRAM chip of 256K×8 bits has x rows of cells with y cells in each row? If DRAM takes 20ns for 1 refresh and 2.56 milliseconds for entire chip refresh then the value of x , y are ____?

no. of cells = 256k = 2^{18}

$x * y = 256k$



chip refresh time = no. of rows of cells * 1 refresh time

$2.56\ msec = x * 20\ nsec$

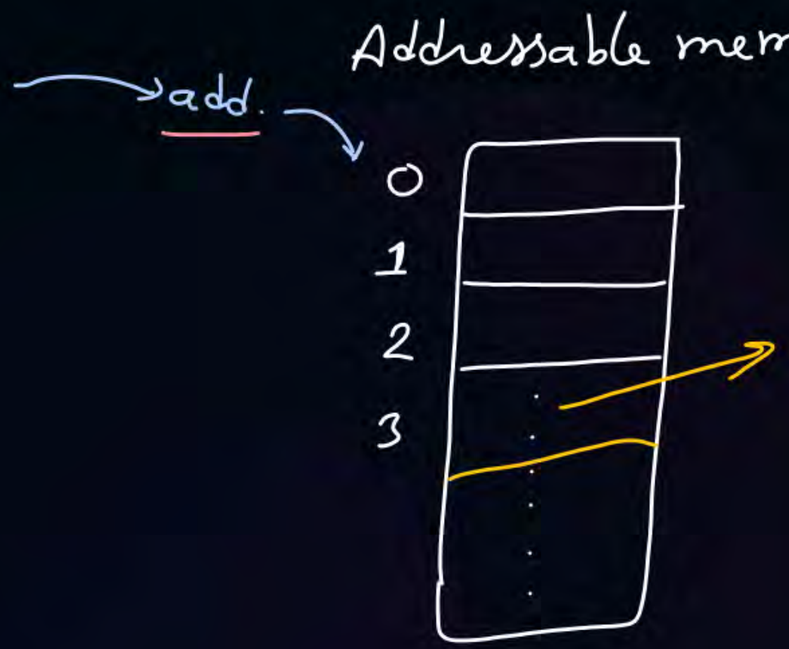
$x = \frac{2.56\ ms}{20\ ns}$

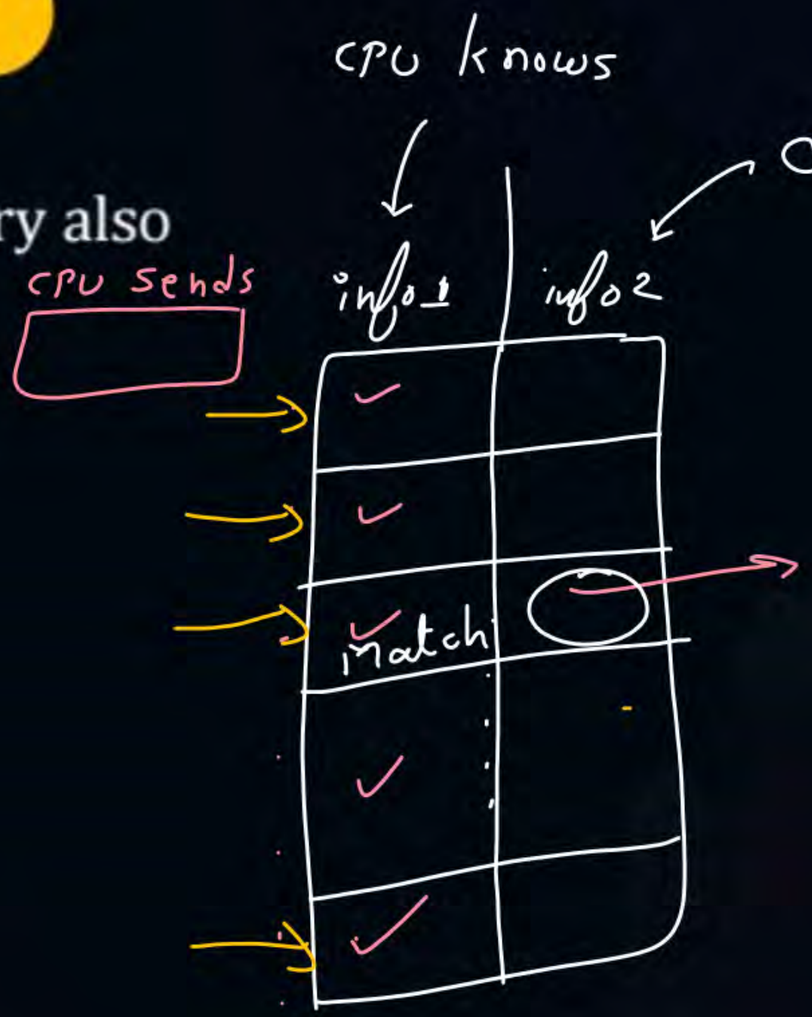
$x = \frac{0.128\ ms}{ns} = \frac{128 * 2^{-10} * 2^{-10}}{2^{-30}}$

$= 128k$

Topic : Associative Memory

Known as content addressable memory also





→ Cells do not have addresses

→ Each cell contains 2 infons
  - ① CPU knows
  - ② CPU wants

→ CPU's generated value is compared with first infon of each cell.

→ whichever cell's content is matching, it's associated infon 2 is sent to CPU for access.

→ Comparison in each cell is done in parallel; hence this memory is very-very fast.

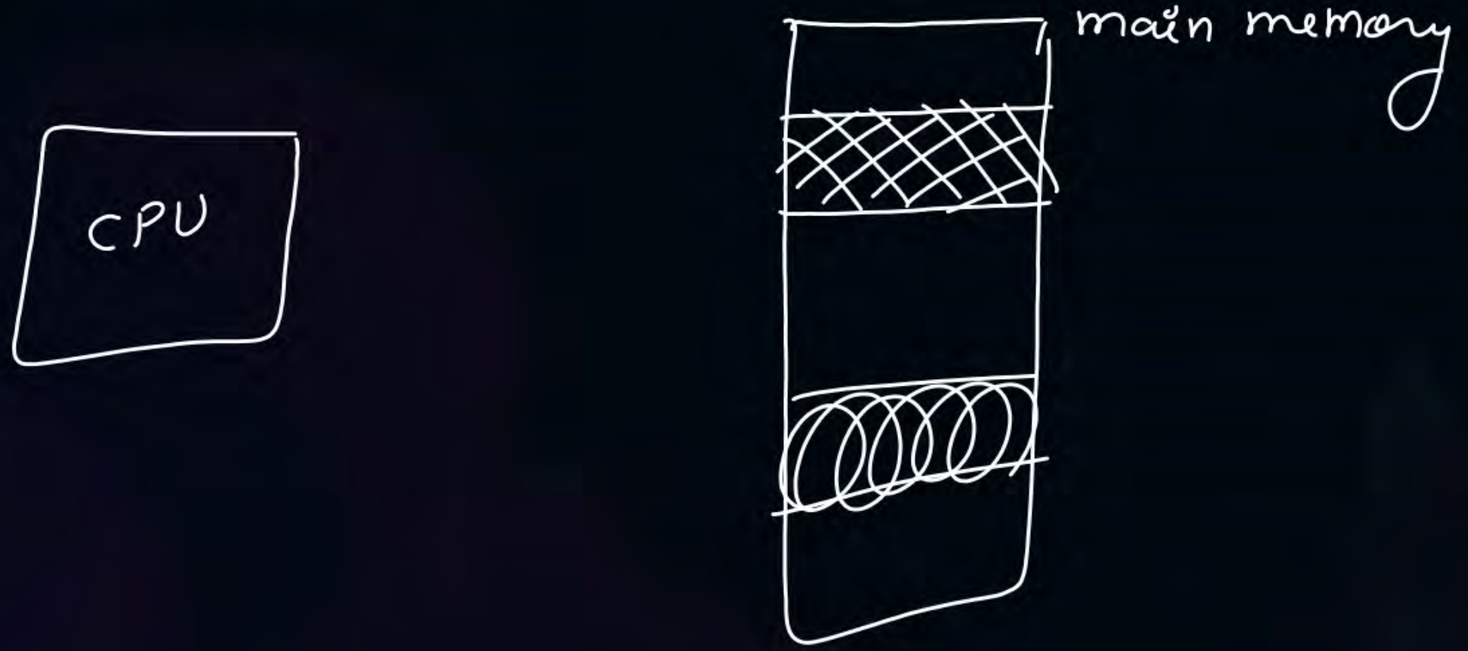
→ very-very expensive.

Associative mem. used for

1. Cache
2. TLB (Translation lookaside Buffer)

Topic : Locality of Reference

If CPU has requested one address for memory access, then that particular address or near by addresses will be accessed soon.

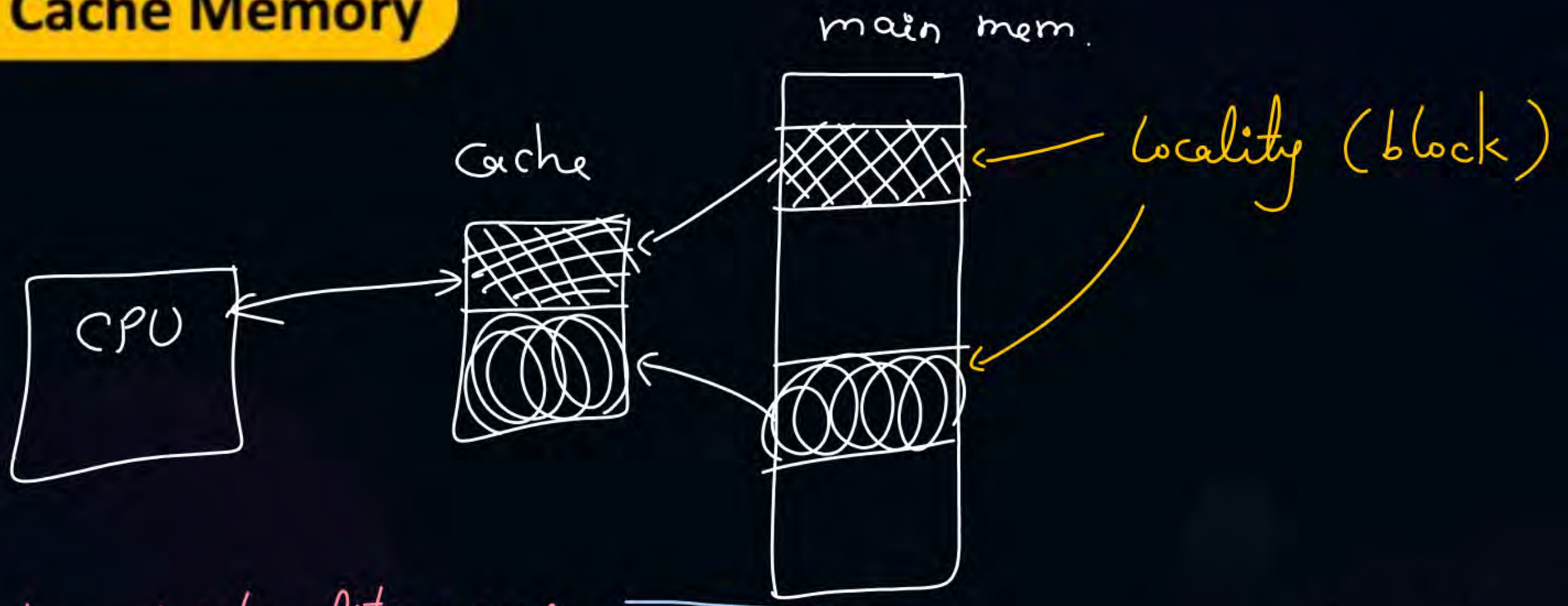


Topic : Locality of Reference

Types:

1. Spatial (according to space) ⇒ if nearby addresses accessed soon
2. Temporal (acc. to time) ⇒ if same address accessed soon

Topic : Cache Memory

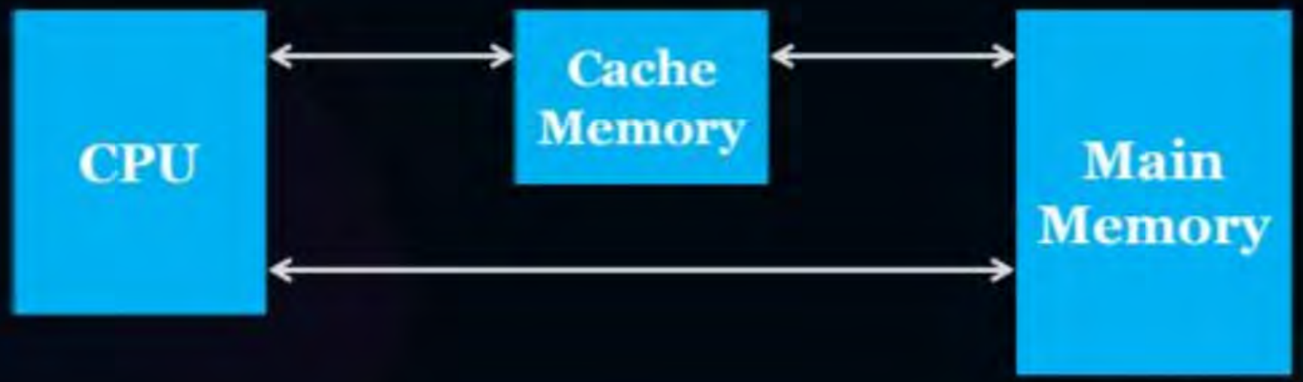


currently demanded localities are kept into a smaller and faster memory, called as Cache.

→ on basis of locality of reference.

Topic : Cache Memory

Use of cache reduces avg. mem. access time.



Topic : Working of Cache Memory

1. Cache Hit :- when demanded content of CPU is present in cache
2. Cache Miss :- —"— —"— —"— is not —"— —"—
3. Hit Ratio :- fraction of time CPU experiences hit in cache.
   (h or H)

$$H = \frac{\text{no. of hits}}{\text{Total mem. references}}$$

miss ratio = $1 - H$

When there is cache miss, then the demanded content is sent from mm to CPU.

And along with it, the block (which contains missed content) is copied from mm to cache for future references.

→ only for read operations

2 mins Summary

Topic Associative Memory

Topic Locality of Reference

Topic Cache Memory

Happy Learning

THANK - YOU