ALL Mid Solution

Subject.....

Date:.....Time:.....

Computer Architecture

Mid: 20-21
17-18

(Book page=60, problem 2.9)

~~Solution is (a)~~

1) Assuming that ENIAC had the capability to have multiple vacuum ~~tubb~~ tubes in the ON and OFF state simultaneously, why is his representation "wasteful" and what range of integer values could we represent using the 10/20/30 vacuum tubes [10]

Solution: This representation is wasteful because to represent a single decimal digit from 0 to 9 we need to have ten tubes.
If we could have an arbitary number of these tubes ON at the same time, then those same tubes could be treated as binary bits.

With 10/20/30 vacuum tubes or bits, we can represent 2^n patterns ($2^{10}, 2^{20}, 2^{30}$) or 1048576 (for example 2^{10}).
So the range of integer values could we represent using 10/20/30 bit is 0 to $(2^n - 1)$. or 1048575

02

Subject.....
Date:.....Time:.....

Mid : 20-21, 17-18, 19-20

(1) Suppose three interrupt handlers A, B, and C (having priority level A > B > C) with Interrupt Service Routine (ISR) 10, 30 and 20 respectively. Graphically show the transfer of control for interrupt sequence of C, A and B at time t = 10, 25 and 45 respectively. [10]

Solution:

Given Data,

interrupts = A, B, and C

priority = A > B > C

ISR execution time;

A → 10 units, B → 30 units, C → 20 units.

Interrupt sequence (Time);

C occurs at t = 10

A occurs at t = 25

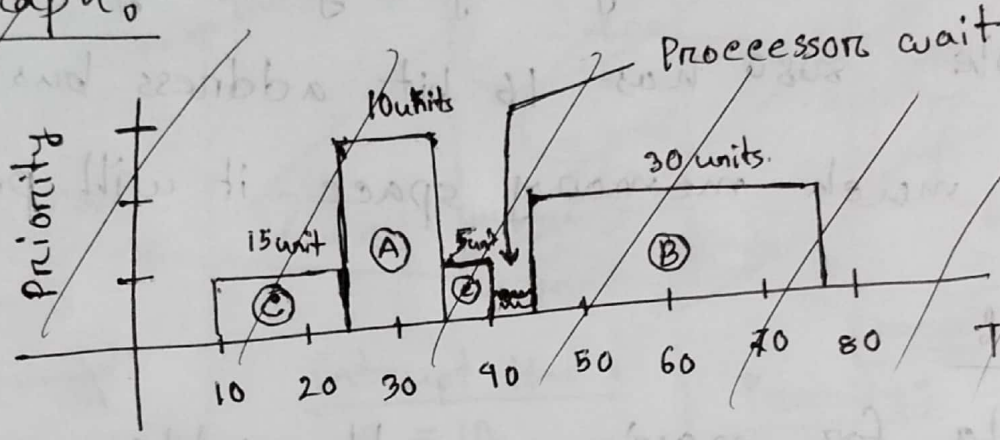
B occurs at t = 45

Subject.....

Date:.....Time:.....

Now, ~~ther~~ Here is the graph and explainations

Graph:



Explaination:

When t.=10 C starts execution. C needs 20 times units (scheduled t.=10 to t=30).

But at t=25 A interrupt because A has higher priority than C, so C is preempted and A starts execution for 10 times units (till t=25 to t=35)

At t=35, A ends execution and again C starts and execute the remaining part untill t=40.

Then t=40 to t=45 processor remain Idle and at ~~45~~ t=45 B start execution. B execute ~~xxx~~ 30 times unit. ~~is~~ till t.=75.

04

Subject.....
Date:.....Time:.....

(3) Regarding Address Bus, Bus width determines maximum memory capacity of system. For example 8080 has 16 bit address bus. Then how much memory space it will provide? [5]

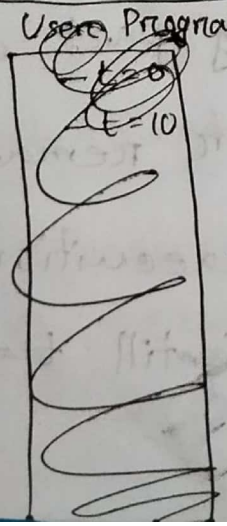
Solution:

Formula for maximum addressable memory,

$$\text{memory capacity} = 2^{\text{Address Bus width}} \times \text{word size}$$

since each address typically refers to one byte (word size = 1 byte). So the total memory capacity is $2^{16} = 65536$ byte = 64 KB.

[1] Graph picture of Q: no: 2



$$2l+1 = \frac{2^m+1}{3}$$

$$\Rightarrow 6l+3 = 2^m+1$$

$$\Rightarrow 5l+3 = 2^m$$

$$\Rightarrow l = \frac{1}{3}(2^m-3) \quad \text{P.T.O}$$

05

Subject.....

Date:.....Time:.....

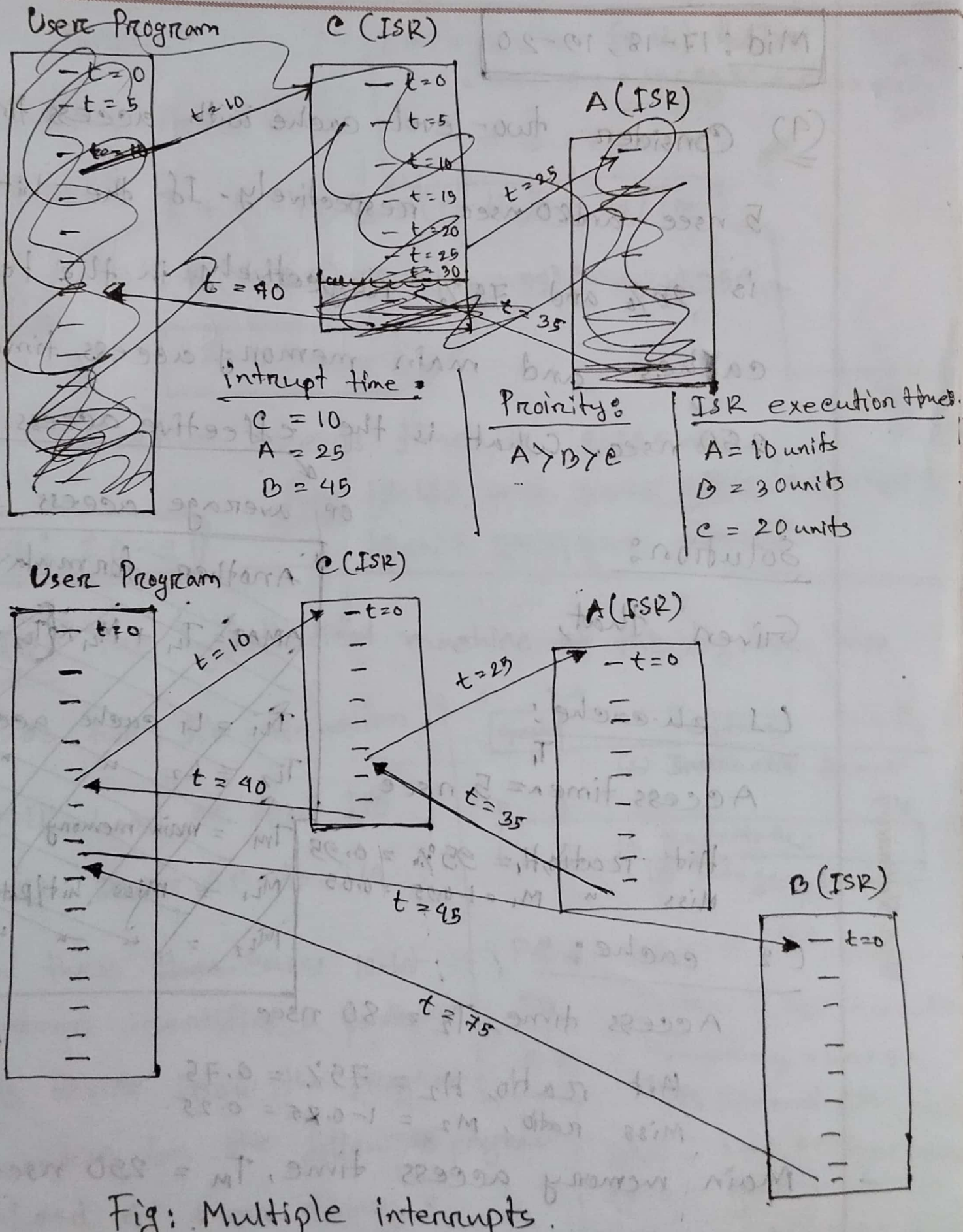


Fig: Multiple interrupts.

Subject.....
Date:.....Time:.....

Mid: 17-18, 19-20

(4) Consider two-level cache with access time 5 nsec and 80 nsec respectively. If the hit ratio is 95% and 75% respectively in the two caches and main memory access time is 250 nsec. What is the effective access time? [10]

or, average access time.

Another formula is

$AMAT = T_{L_1} + M_{L_1} \times \{T_{L_2} + (M_{L_2} \times T_m)\}$

T_{L_1} = L_1 cache access time

T_{L_2} = L_2 " " "

T_m = main memory " "

M_{L_1} = Miss hit/Rate of L_1 cache

M_{L_2} = " " " L_2 "

Solution:

Given that,

L1 ch-cache:

T_1

Access time = 5 nsec

Hit ratio, H_1 = 95% = 0.95

Miss " M_1 = 1 - 0.95 = 0.05

L2 cache:

Access time, T_2 = 80 nsec

Hit ratio, H_2 = 75% = 0.75

Miss ratio, M_2 = 1 - 0.25 = 0.25

Main memory access time, T_m = 250 nsec

$EAT = T_{L_1} \times H_{L_1} + M_{L_1} \times T_{L_2} \times H_{L_2} + M_{L_1} M_{L_2} T_M$ ⟶ Mahmud

$EAT = T_{L_1} + M_{L_1} T_2 + M_{L_1} M_{L_2} T_M$

$EAT = T_1 \times H_1 + (1-H_1) \times H_2 \times T_2 + (1-H_1)(1-H_2) \times T_M$

$= 5 \times 0.95 + 0.05 \times 0.95 \times 80 + 0.05 \times 0.25 \times 250$ (0.75)

$= 10.875$ nsec.

We know,

Effective Access time,

$EAT = T_1 + (1-H_1)T_2 + (1-H_1)(1-H_2)T_M$

$= 5 + (1-0.95) \times 80 + (1-0.95)(1-0.75).250$

$=$ **12.125 nsec.**

∴ Effective access time is 12.125 nsec.

Ans 12.125 and 10.875 দুটোই উত্তর হবে
12.125 তুলনামূলক বেশি সঠিক.

Mid: 20-21

(5) The hypothetical machine of the figure also has two I/O instruction.

0011 Load AC from I/O

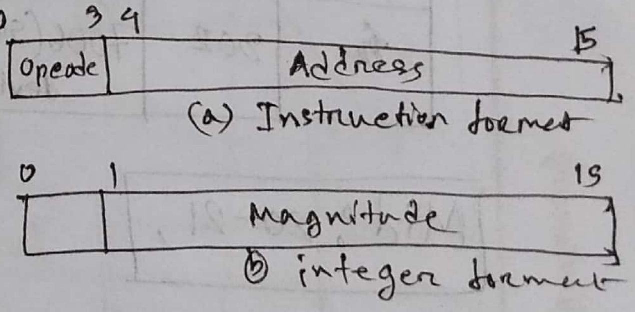
0011 store AC to I/O

In these cases 12-bit address identifies a particular I/O device. Show the program Execution for the following program

(1) Load AC from device 5

(2) Add contents of memory location 940

(3) Store AC to device 6.



PC = Address of instru...
IR = Instru... begin execute
AC = Temporary storage
(c) Internal CPU regi...

0001 – Load AC from memory
0010 – Store " to "
0101 – Add to AC from ...

Subject.....
Date:.....Time:.....

Assume that the next value retrieved from device 5 is 3 and that location 940 contains a value of 2. [10]

~~Solo~~ Solution: This solution is not good, Better soln at page 19

Given Program Execution Steps:

Steps	PC	Instruction (IR)	Action	AC
1	300	3005 (Load AC from device 5)	Load from device 5 → AC = 3	3
2	301	5940 (Add memory 940 to AC)	Memory location contains 2, so AC = 3 + 2 = 5	5
3	302	7006 (Store AC to device 6)	Store AC (5) to device 6	5

Mid: 20-21,

(6) What are the four main components of any general-purpose component? [5]

The four main components are given below,

(1) Main Memory: It stores both data and instructions.

Bari Stationary

Subject.....

Date:.....Time:.....

② Arithmetic and logic unit (ALU): Perform arithmetic and logical operations on data.

③ Control Unit: Directs and manages execution of instructions.

④ Input and Output (I/O) Devices: Enable communication between the computer and external environment.

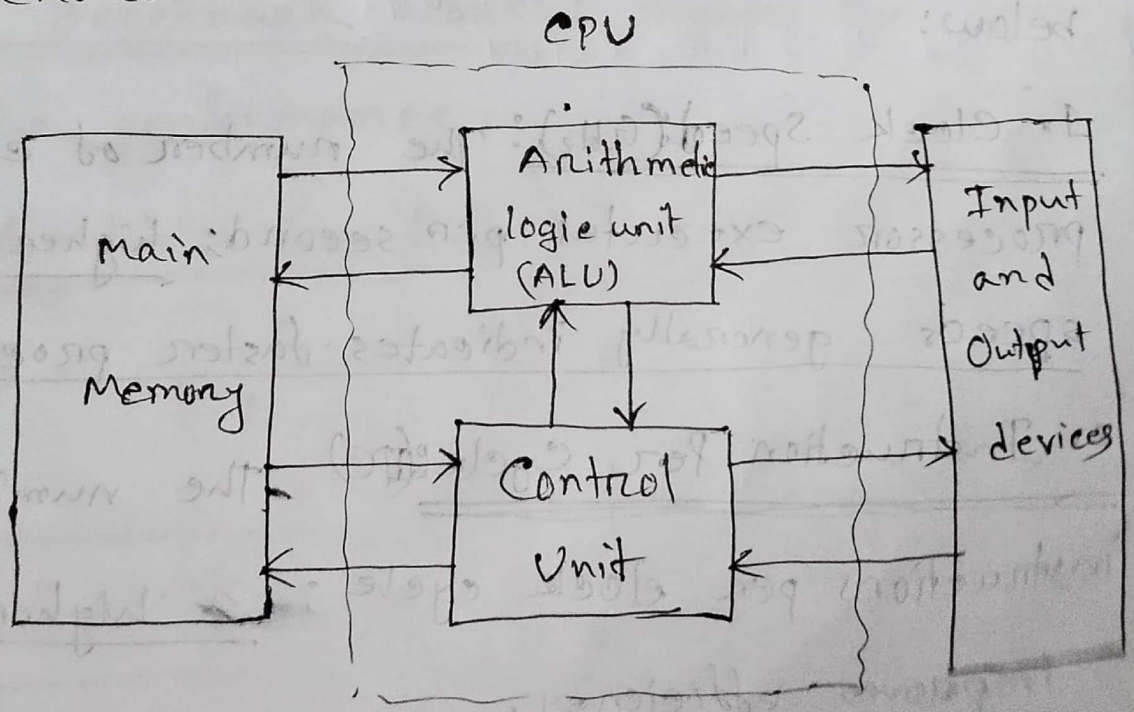


Fig: Four main components of computer.

Subject.....
Date:.....Time:.....

Mid: 17-18

(7) How to determine the performance of a processor? Explain. [5]

Underline কর বড় বড় অংশ লিখলে হবে।

Solution:

The performance of a processor is determined in several ways. Some ways are given below:

1. Clock Speed (GHz): The number of cycles a processor executes per second; higher clock speeds generally indicates faster processing.

2. Instruction Per Cycle (IPC): The number of instructions per clock cycle; a higher IPC improves effeciency.

3. CPI (Cycles Per Instruction): Lower CPI means faster execution of instructions.

Formula: $T = \text{Instruction count } (I_c) \times CPI \times \text{cycle time } (T)$

Subject.....
Date:.....Time:.....

Another formulas are,

$$\text{MIPS rate} = \frac{f}{CPI \times 10^6}.$$

$$\text{MFLOPS rate} = \frac{\text{Number of executed floating point}}{\text{Execution Time} \times 10^6}$$

Both of the are common performance measurement formula of a processor.

4: Benchmark scores: Higher the scores, higher the performance.

5: Amdahl's law (Parallel processing speed up):

$$S = \frac{1}{(1-P) + \frac{P}{N}}.$$

Mid : 17-18, 19-20

(8) Describe the bottleneck for cache Design. [10]

P.T.O.

Subject.....
Date:.....Time:.....

A bottleneck refers to a limitation or constraints that ~~slow~~ slow down the overall performance of a system. It is the weakest point that restricts efficiency of a system and prevents the system to achive its full potential.

There are several bottlenecks ~~that impacts of system~~ we faces to design a cache which impacts on system performance. such as,

1. Cache Size vs Access Time:

Larger caches store more data but increase access time.

2. Cache Hit vs Miss Rate: A high miss rate forces the processor to access slower main memory, increasing latency.

3. Memory Bandwidth: Cache misses increase memory traffic, causing delays in data retrieval.

Subject.....
Date:.....Time:.....

4: Cache Coherency: In multi-core system, maintaining ~~cons~~ consistency between caches leads to additional overhead.

5. Power Consumption: Larger and faster caches consume more power.

6. Associativity Trade-offs: It reduces conflicts but increase lookup time and hardware complexity.

Mid: 20-21

(9) We would like to design a cache of 64 kByte. Main Memory is 16 Mbytes. Cache Block is 4 bytes (2 bit word identifier) and rest of bits is block identifier. Use the concept of direct mapping address. [10]

Solution:

Main memory = 16 Mbytes

$= (2^4 \times 2^{20})$ bytes

$= 2^{24}$ bytes.

Subject.....
Date:.....Time:.....

Cache memory = 64 k bytes

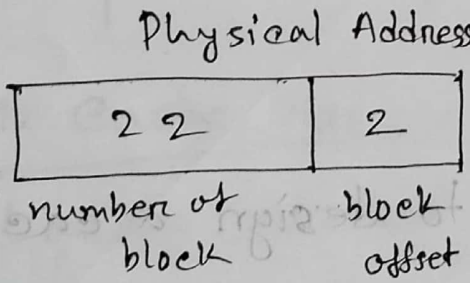
$= (2^6 \times 2^{10})$

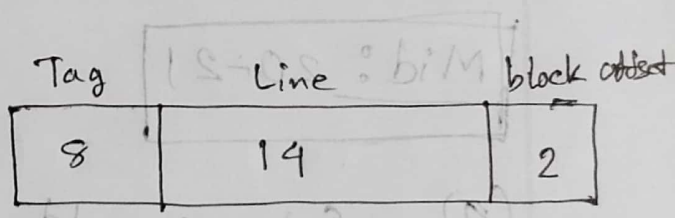
$= 2^{16}$ bytes

Total number of blocks $= \frac{2^{24}}{2^2} = 2^{22}$

Total number of lines $= \frac{2^{16}}{2^2} = 2^{14}$

Line size = block size = $4 = 2^2$





So the design is,

Fig: Cache Memory design with Direct mapping.

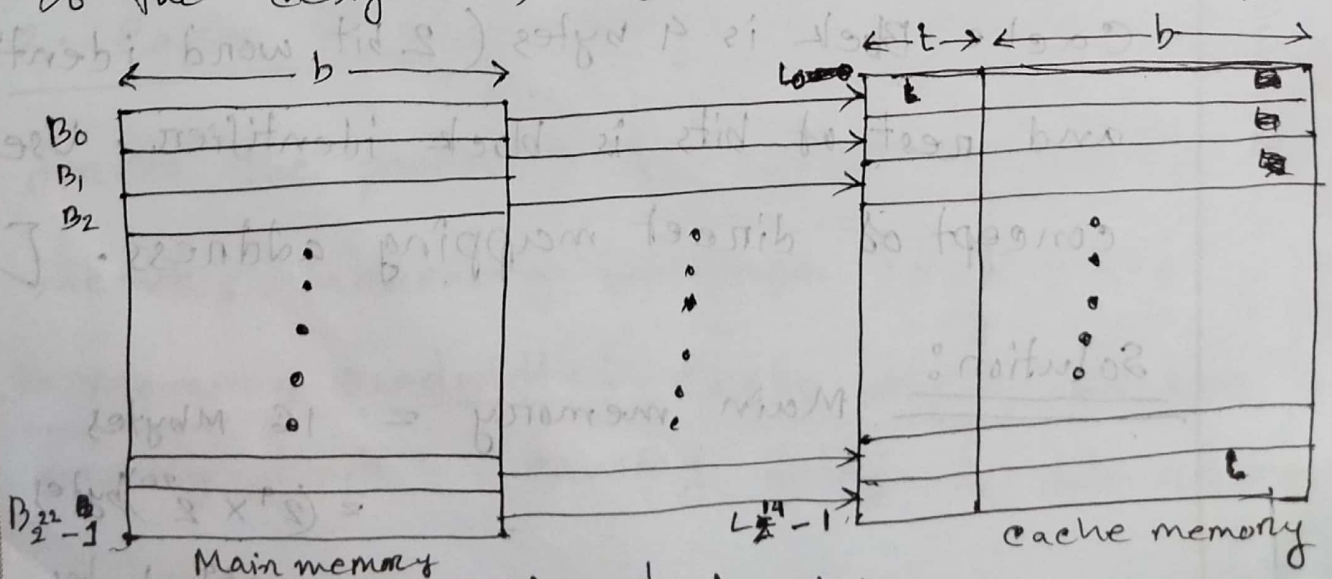


Fig: Direct Mapping.

Subject.....
Date:.....Time:.....

Mid : 20-21

(10) When a block that is resident in the cache is to be replaced, there are two cases to consider. Describe the write through and write back. [10]

Solution:

When a block in the cache is updated or replaced, there are two common write policies:

1. Write-Through:

- It updates both cache and main memory simultaneously.
- Multiple CPUs can monitor main memory traffic to keep local CPU up to date.
- Lots of traffic
- Slows down writes and bogus write through caches.

Subject.....
Date:.....Time:.....

2. Write Back:

- Updates initially made in cache only.
- Main memory updated later, when the cache block is replaced.
- I/O must access main memory through cache.
- When an update occurs, a dirty bit associated with the line is set.

Mid: 20-21

(11) It is possible to build a computer which uses only static RAM. This would be very fast and would need no cache. Analyze its cost. [10]

Solution: Using only static RAM (SRAM) in a computer would be extremly fast but highly expensive due to its 6-transistor per bit design compare to DRAM's 1-transistor, 1-capacitor

Subject.....
Date:.....Time:.....

design. SRAM ~~consumes more power~~, takes more (less density to store data in a unit, complex design and harder to manufacture) space, and ~~limits storage capacity~~, making it impractical for the large memory system.

Instead ~~more~~ modern computers use a memory hierarchy (SRAM cache + DRAM main memory) to balance speed and cost efficiently.

Overall SRAM computer would be fast than other but ~~it~~ it is ~~impra~~ impractical for large scale implementation for its high cost and complexity.

—x—

~~EAT = (Hit Ratio L1 x Access time L1) + (Miss Ratio L1 x Hit Ratio L2 x Access time L2) + (Miss Ratio L1 x Miss Ratio L2 x main memory)~~

$= (H_{L_1} \times T_{L_1}) + \{(1 - H_{L_1}) \times H_{L_2} \times T_{L_2}\} + \{(1-H_{L_1})(1-H_{L_2}) T_m\}$

$= H_{L_1} T_{L_1} + H_{L_2} T_{L_2} - H_{L_1} H_{L_2} T_{L_2} + (1 - H_{L_2} - H_{L_1} + H_{L_1} H_{L_2}) T_m$

Better Solution of Q: 5

19

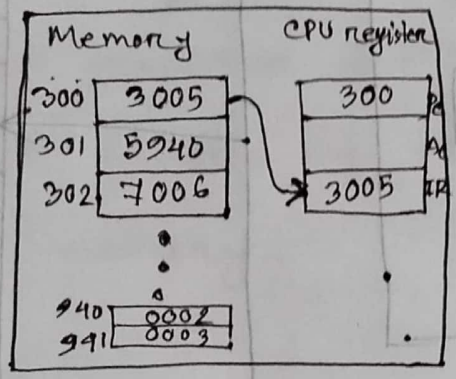
Subject: ////

Date: Time:

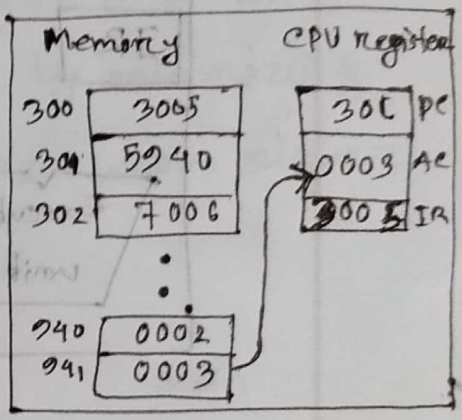
Here's the program execution step by step. 3 is given below.

(a) Load AC from device 5

step: 1

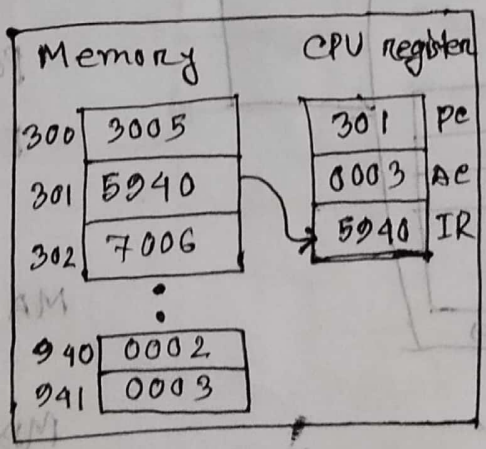


step 2:

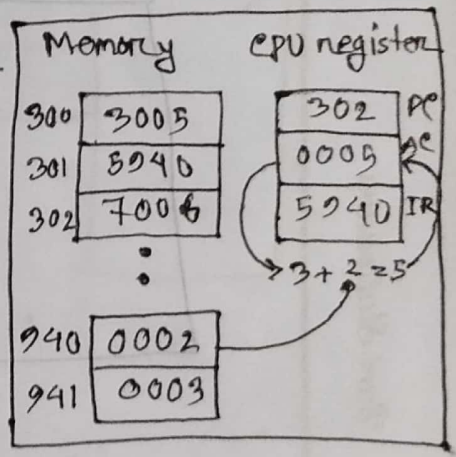


(b) Add contents of memory location 940

step: 3

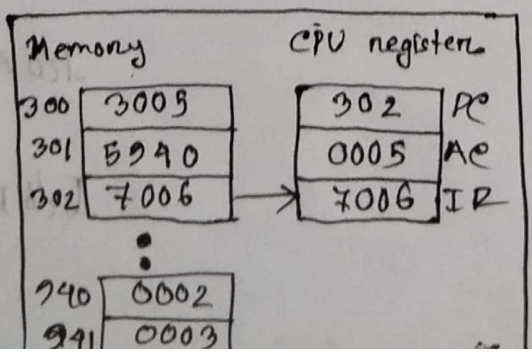


step: 4

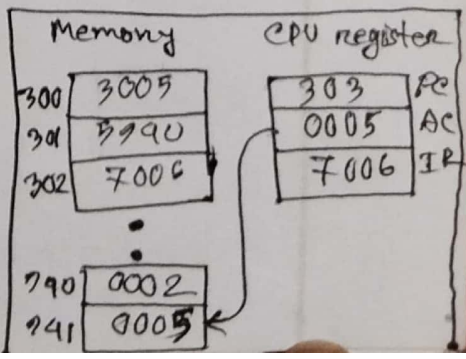


~~step: 5~~ (c) store AC to device 6

step: 5

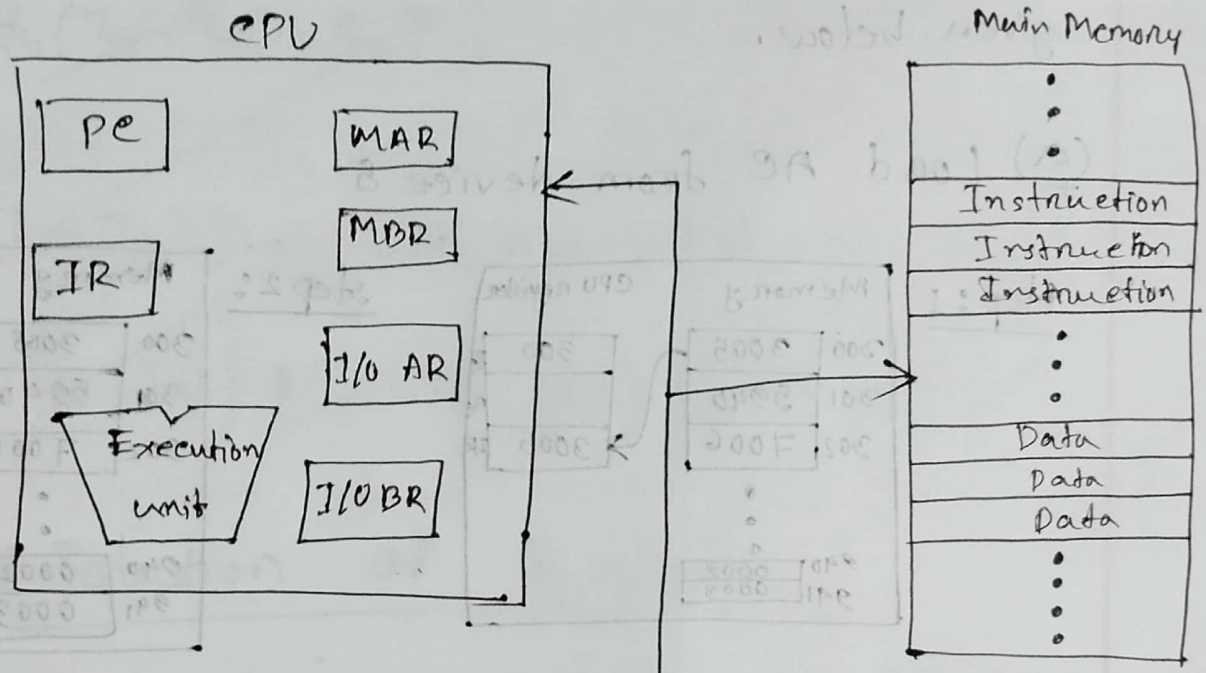


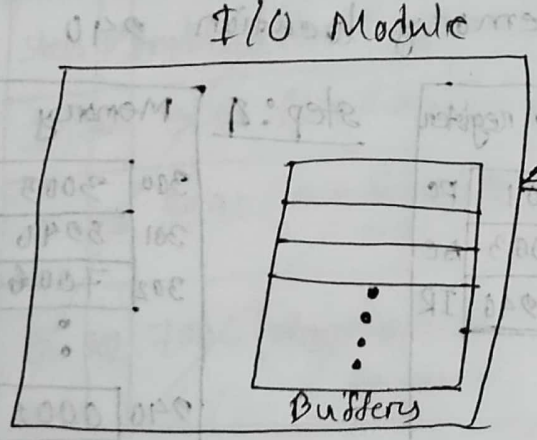
step: 6



Subject.....
Date:.....Time:.....

Computer Components top-level view.





PC = Program Counter

IR = Instruction Register

MAR = Memory Address Register

MBR = Memory buffer Register

I/O AR = I/O Address Register

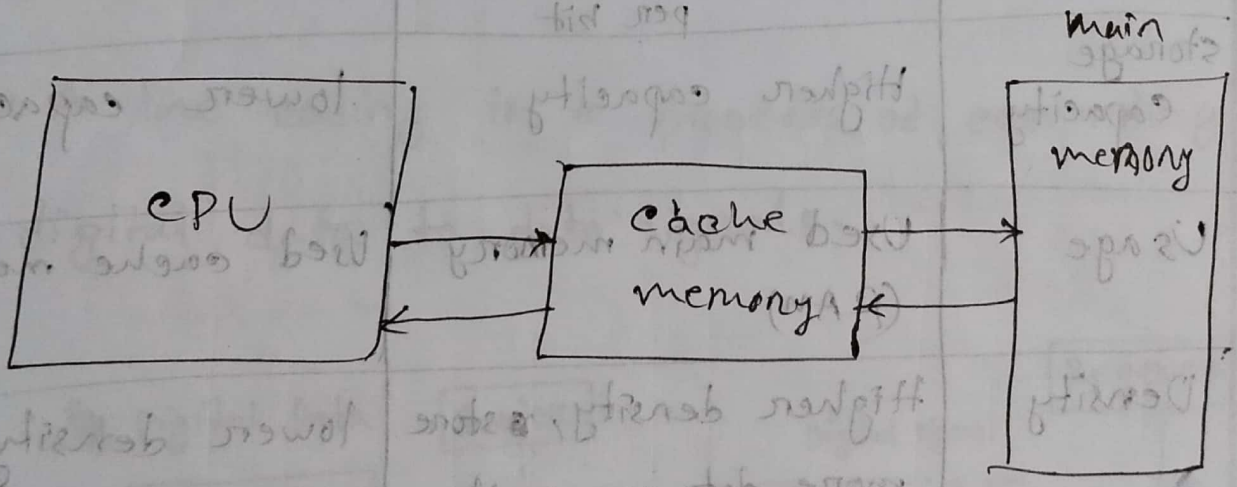
I/O BR = I/O Buffer Register

Subject.....

Date:.....Time:.....

What is cache?

=> A cache is a small, high-speed memory located between the CPU and main memory (RAM) that stores frequently memory access time and improve processor performance by minimizing delays caused by fetching data from slower main memory.



22

Subject.....
Date:.....Time:.....

DRAM vs SRAM?

Aspect	DRAM	SRAM
speed	slower due to refresh cycle	Faster and no refresh is needed.
Power Consumption	consume more power	consume less power
cost	cheaper (1 transistor + 1 capacitor) per bit	Expensive (6 transistor) per bit
storage capacity	Higher capacity	lower capacity
Usage	Used main memory (RAM)	Used cache memory
Density	Higher density, store more data per unit	lower density, store less data per unit
complexity	simpler design	complex design and harder to manufacture

Hivemind, 5
Hallowery. — 5
Matador — 5
Heckler — 10