

Mumbai University

Terna Engineering College, Nerul, Navi Mumbai

**Department of Computer Engineering**

<b>Course Code</b>	CSC403	<b>Program</b>	B.E. (CMPN)
<b>Semester</b>	IV	<b>Year</b>	II
<b>Name of the Faculty</b>	Rohini Palve	<b>Class / Div</b>	A & B
<b>Course Title</b>	Computer Organization and Architecture	<b>Academic year</b>	2019-20

**Amey Thakur B-50****ASSIGNMENT 5**

- Q1) Explain requirement of I/O module  
 Q2) Explain Interrupt driven method of data transfer with flow chart and comment how it is superior to Programmed I/O and explain Programmed I/O  
 Q3) State various types of data transfer techniques and explain DMA in detail alongwith various modes.

{Explain different data transfer techniques} {10M} (No need to write answer in assignment but explain all three types in brief if asked in Exam)

- Q4) Differentiate between Memory Mapped I/O and I/O Mapped I/O  
 Q5) What is bus arbitration? Explain any two techniques for the same.  
 Q6) Explain bus contention and different methods to resolve it  
 Q7) Explain Super Scaler Architecture  
 Q8) Explain Flynn's classification  
 Q9) explain multicore architecture in detail  
 Q10) Draw and explain VLIW processor architecture  
 Q11) What is dynamic instruction scheduling  
 Q12) Explain in brief pipelining in superscalar processor with degree m=2

CO4: To describe superscalar architectures, multi-core architecture and their advantages (Q7 to Q12)

CO6: To Identify various types of buses, interrupts and I/O operations in a computer system (Q1 to Q6)

# Terna Engineering College, Navi Mumbai

Q.1.

Ans:

## ① Processor communication:

This involves the following task

- Exchange of data between processor and I/O module.
- Command decoding - I/O module accepts commands sent from the processor.

Eg. the I/O module for a disk drive may accept the following commands from the processor: READ SECTOR, WRITE SECTOR, SEEK track, etc.

- Status reporting - The device must be able to report its status to the processor, e.g., disk drive busy, ready etc. Status reporting may also involve reporting various errors.
- Address recognition - Each I/O device has a unique address and the I/O module must recognize this address.

## ② Device Communication:

- The I/O module must be able to perform device communication such as status reporting.

## ③ Control & timing:

- The I/O module must be able to co-ordinate the flow of data between the internal resources (such as processor, memory) and external devices

# Terna Engineering College, Navi Mumbai

## ④ Data buffering :

- This is necessary as there is a speed mismatch between speed of data transfer between processor and memory and external devices. Data coming from the main memory are sent to an I/O module in a rapid burst. The data is buffered in the I/O module and then sent to the peripheral device at its rate.

## ⑤ Error detection :

- The I/O module must also be able to detect errors and report them to the processor. These errors may be mechanical errors (such as paper jam in a printer) or changes in the bit pattern of transmitted data. A common way of detecting such errors is by using parity bits.

# Terna Engineering College, Navi Mumbai

Q2

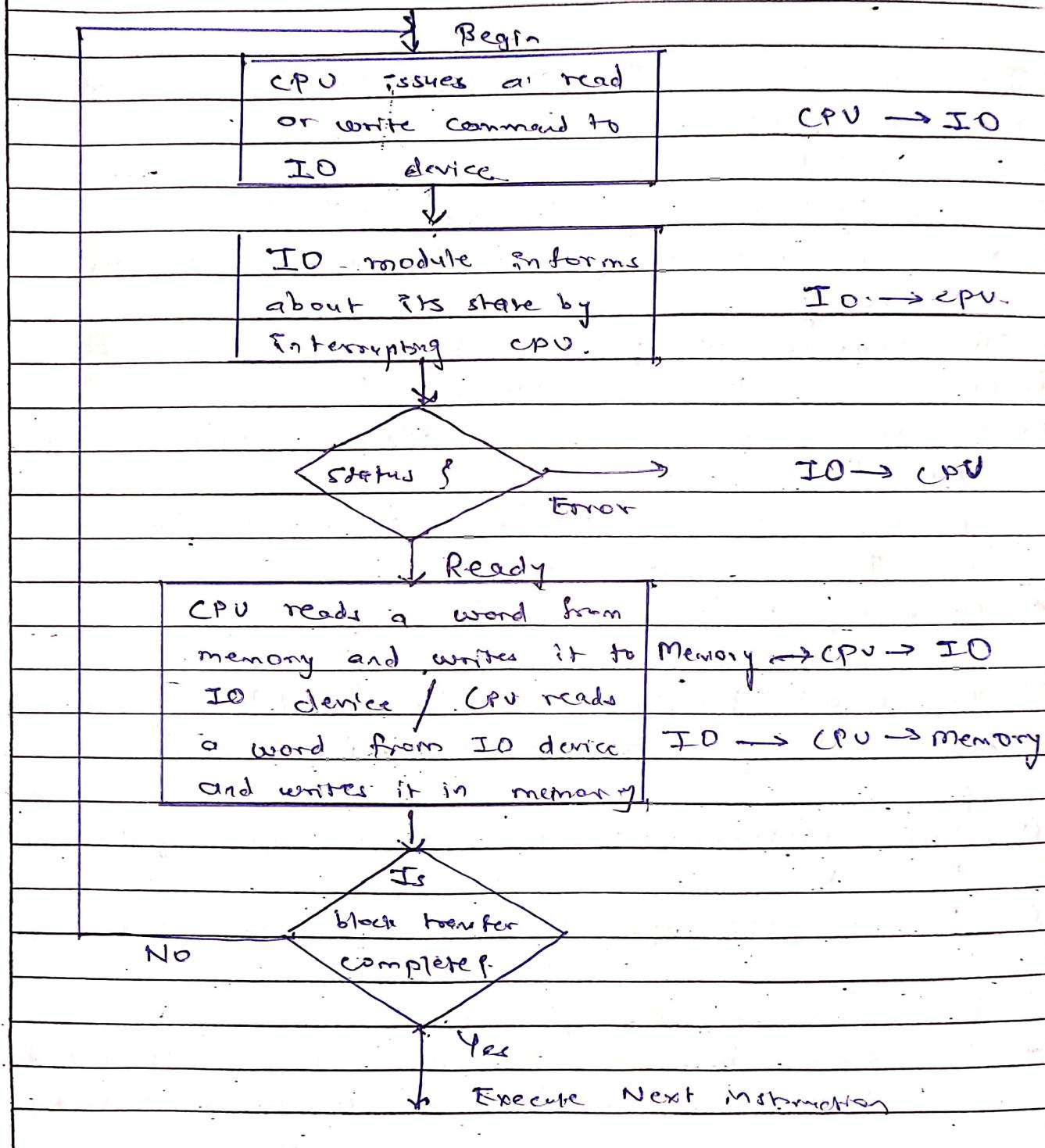
Ans:

## Interrupt Driven I/O

- Interrupt Driven I/O overcomes the disadvantage of programmed I/O, i.e., CPU waiting for I/O device.
- This disadvantage is overcome by CPU not repeatedly checking for the device being ready or not instead the I/O module interrupt when ready.
- The sequence of operation for interrupt driven I/O.
  - ① CPU issues the read command to I/O device.
  - ② I/O module gets data from peripheral while CPU does other work.
  - ③ Once the I/O module completes the data transfer from I/O device, it interrupts CPU.
  - ④ On getting the interrupt, CPU requests data from I/O module.
  - ⑤ I/O module transfers the data to CPU.
- When CPU gets an interrupt, it performs the operations →
  - ① Save context, i.e., the contents of the register on the stack.
  - ② Processes interrupt by executing the corresponding ISR.
  - ③ Restore the register context from the stack.
- IC 8259 has 8 interrupt lines and is used as I/O module when interrupt driven I/O is used.

# Terna Engineering College, Navi Mumbai

- Transferring a block of data using interrupt driven IO



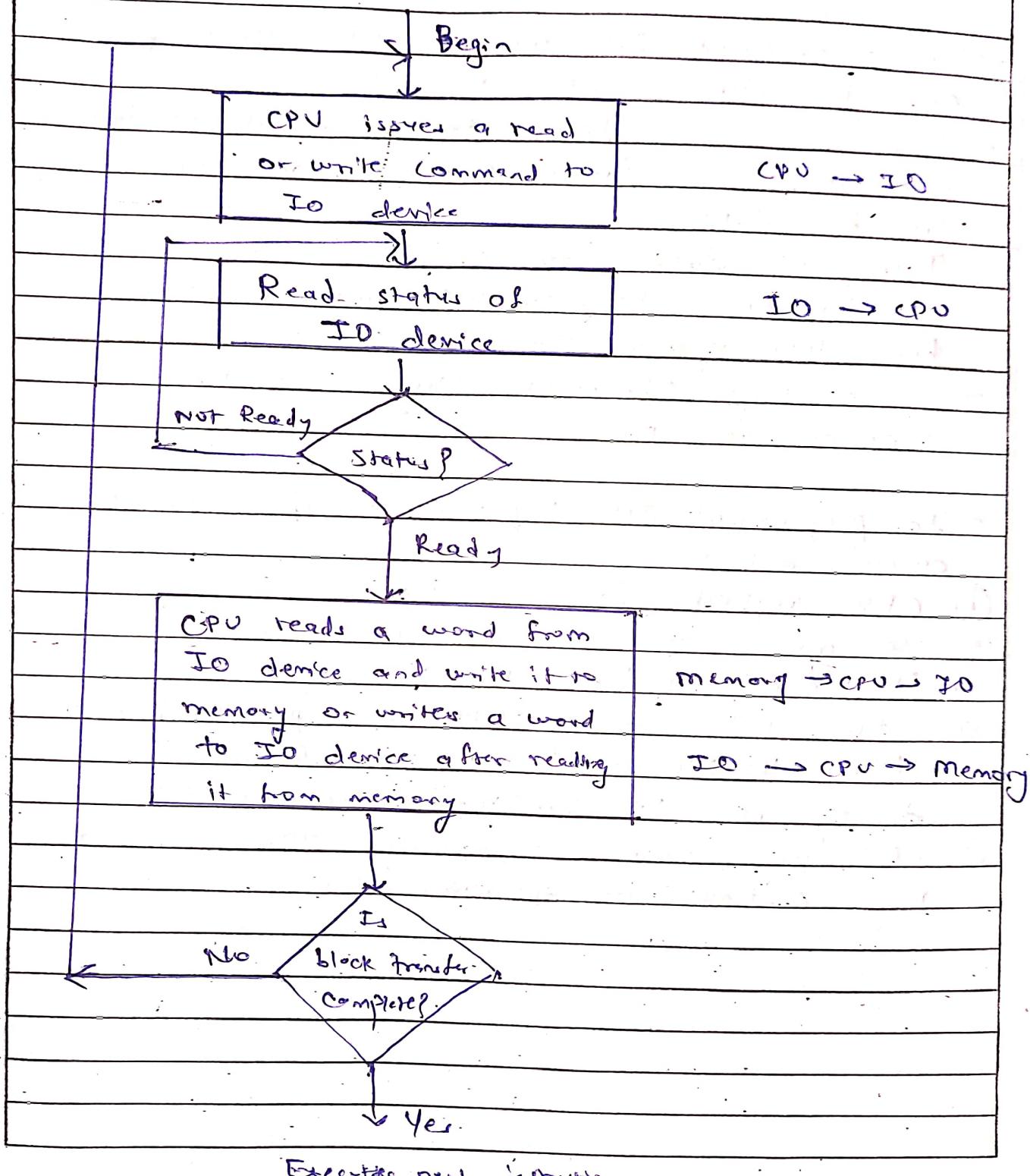
# Terna Engineering College, Navi Mumbai

## Programmed I/O

- In programmed I/O method of interfacing, CPU has direct control over I/O
- The processor checks the status of the devices and issues read / write command and then transfers data  
During the data transfer, CPU waits for I/O module to complete operation and hence this system waits the CPU time
- That's why Interrupt Driven I/O is Superior than programmed I/O.
- In programmed I/O, sequence of operations to be carried out →
  - ① CPU requests for I/O operation
  - ② I/O module performs the said operation
  - ③ I/O module updates the status bits
  - ④ CPU checks these status bits periodically.  
Neither the I/O module cannot inform CPU directly nor can I/O module interrupt CPU
  - ⑤ CPU may wait for operation to complete or may continue the operation later

# Terna Engineering College, Navi Mumbai

- Transferring a block of data using programmed I/O



Execution next instruction

# Terna Engineering College, Navi Mumbai

Q.3.

Ans:

There are 3 modes of transfer for data, commands and status as follows:

- ① Programmed IO
- ② Interrupt driven IO
- ③ Direct memory Access

## Direct Memory Access

- DMA is a form of IO in which a special module called a DMA module, controls the exchange of data between main memory and an IO module. The CPU sends a request for the transfer of a block of data to the DMA module and is interrupted only after the entire block has been transferred.
- In DMA based data transfer, the transfer operation is carried out by the DMA controller which acts as a master in microprocessor based system. The data is transferred directly between I/O and memory and data transfer is controlled by either I/O device or DMA controller. Microprocessor does not participate in this data transfer method.

# Terna Engineering College, Navi Mumbai

- The peripheral device (such as the disk controller) will request the service of DMA by pulling DREQ (DMA request) high.
- The DMA will put a high on its HRQ (hold request) signaling the CPU, through its HOLD pin that it needs to use the buses.
- The CPU will finish the present bus cycle (not necessarily the present instruction) and respond to the DMA request by putting high on its HDKA (Hold acknowledge), thus telling the 8237 DMA that it can go ahead and use the buses to perform its task. HOLD must remain active high as long as DMA is performing its task.
- DMA will activate DACK, which tells the peripheral device that it will start to transfer the data.
- DMA starts to transfer the data from memory to peripheral by putting the address of the first byte of the block on the address bus and activating MEMR, thereby reading the byte from memory into the data bus; it then activates low to write it to the peripheral. Then DMA decrements the counter and increments the address pointer and repeats this process until the count reaches zero and the task is finished.
- After the DMA has finished its job it will deactivate HRQ, signaling the CPU that it can regain control over its buses.

# Terna Engineering College, Navi Mumbai

- The DMA controller has following modes of data transfer.

They are :-

## ① Cycle Steal :

- A read or write signal is generated by the DMAc, and IO devices either generates or latches the data. The DMAc effectively steals cycles from the processor in order to transfer the byte, so single byte transfer is also known as cycle stealing.

② Requests by DMA devices for using the bus are always given higher priority than processor requests.

③ Among different DMA devices, top priority is given to high speed peripherals such as disks, high speed network interface and graphics display device.

④ Since the processor initiates most memory access cycles, it is often stated that DMA steals memory cycles from the processor for its purpose.

⑤ If DMA controller is given exclusive access to the main memory to transfer a block of data without interruption, this is called block or burst mode.

# Terna Engineering College, Navi Mumbai

## ② Burst Transfer

- To achieve block transfers, some DMAC's incorporate an automatic sequencing of the value presented on the address bus. A register is used as byte count, being decremented for each byte transfer and upon the byte count reaching zero, the DMAC will release the bus. When the DMAC operates in burst mode, the CPU is halted for the duration of the data transfer.
- In burst mode, the processor is stopped completely until the DMA transfer is completed. Although the processor has no control over its system during such a delay, this mode appears to be more appropriate when predictability is the main goal.

The main disadvantage being that the CPU is halted for the time when the DMA is in control of the bus.

## ③ Hidden mode/ Transparent mode

- It is possible to perform hidden DMA, which is transparent to the normal operation of the CPU. In other words, the bus is grabbed by the DMAC when processor is not using it. The DMAC monitors the execution of the processor, and when it recognizes the processor executing an instruction which has sufficient empty clock cycles to perform a byte transfer, it waits till the processor is decoding the op code and then grabs the bus during this time.
- The processor is not slowed down but continues processing normally. Naturally the data transfer by the DMAC must be completed before the processor starts.

# Terna Engineering College, Navi Mumbai

Q.4.

Ans:

## IO Mapped IO

## Memory Mapped IO

① IO device is treated as an IO device and hence given an IO address	① IO device is treated like a memory device and hence given a memory address
② IO device has an 8/16 bit IO address	② IO device has a 20 bit Memory address
③ IO device is given IOR# and IOW# control signals	③ IO device is given MEMR# and MEMW# control signals
④ Decoding is easier due to lesser address lines	④ Decoding is more complex due to more address lines
⑤ Decoding is cheaper.	⑤ Decoding is more expensive
⑥ Works faster due to less delays	⑥ More gates add more delays hence slower.
⑦ Allows max $2^{16} = 65536$ IO devices	⑦ Allows many more IO devices as IO addresses are now 20 bits
⑧ IO devices can only be accessed by IN and OUT instructions	⑧ IO devices can now be accessed using any memory instruction.
⑨ ONLY AL/AH/AX registers can be used to transfer data with the IO device.	⑨ Any registers can be used to transfer data with the IO device
⑩ Popular technique in microprocessors	⑩ Popular technique in microcontrollers.

# Terna Engineering College, Navi Mumbai

Q5.

Ans

## Bus Arbitration :

- The process of determining which competing bus master will be allowed access to the bus is called Bus Arbitration.
- Multiple devices may need to use the bus at the same time so must have a way to arbitrate multiple requests.
- Arbitration allows more than one module to control the bus at one particular time.
- Arbitration may be centralized or distributed.

### ① Centralized Arbitration :

- Single hardware device controls bus access: Bus controller / Arbiter.
- It is usually a part of processor.

### ② Distributed Arbitration :

- Each module may claim the bus
  - Access control logic is on all modules
  - Modules work together to control bus
- Bus arbitration schemes try to balance

### ① Bus priority :

- The highest priority device should be serviced first

### ② Fairness :

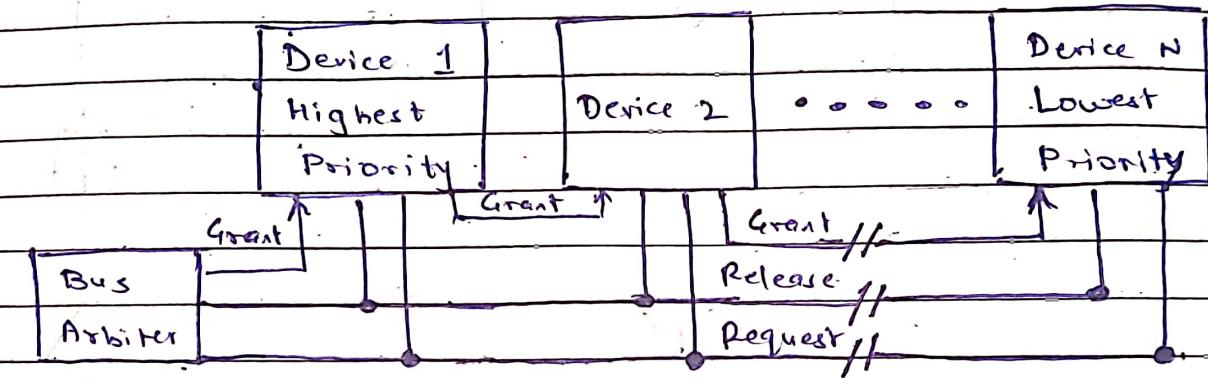
- Even the lowest priority device should never be completely locked out from the bus.

# Terna Engineering College, Navi Mumbai

- Bus arbitration schemes can be divided into 4 as -

- ① Daisy chain arbitration
- ② centralized, parallel arbitration
- ③ Distributed arbitration by self-selection
- ④ Distributed arbitration by collision detection

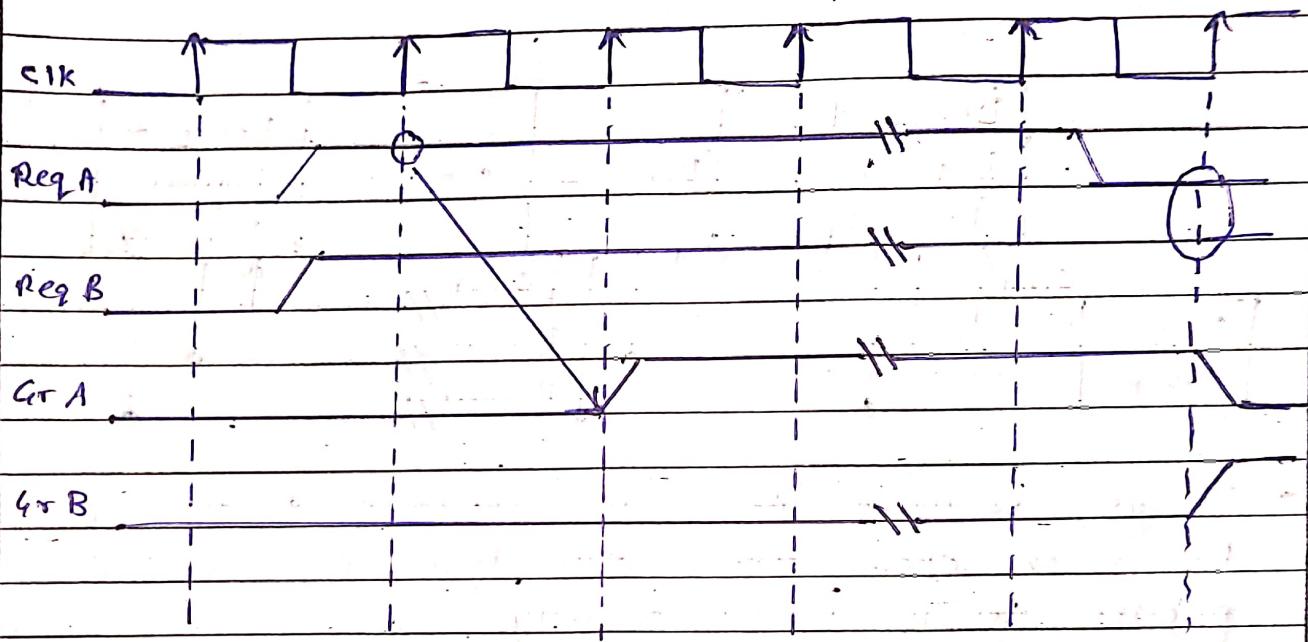
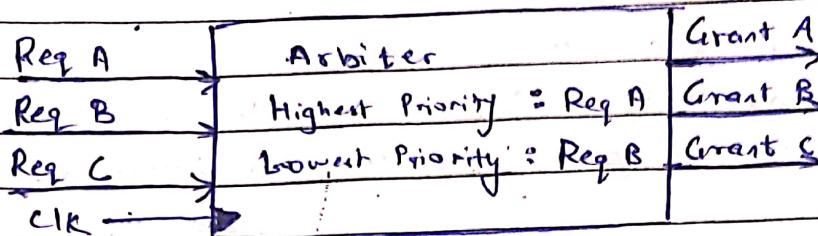
## Daisy Chain Arbitration



- It got its name from the structure for the grant line which chains through each device from the highest priority to the lowest priority.  
The higher priority device will pass the grant line to the lower priority device only if it does not want it so priority is built into the scheme.
- The advantage of this scheme is that it is simple.
- The disadvantages are:
  - It cannot ensure fairness. A low priority device may be locked out indefinitely.
  - Also, the daisy chain grant line will limit the bus speed.

# Terna Engineering College, Navi Mumbai

## Centralized Arbitration



- In the centralized parallel arbitration scheme, the device independently request the bus by using multiple request lines.
- A centralized arbiter chooses from among the devices requesting bus access and notifies the selected device that it is now the bus master via one of the grant line.

# Terna Engineering College, Navi Mumbai

- Consider an example, where A has the highest priority and C the lowest. Figure shows how bus is granted to different bus masters for the example. Since A has the highest priority, Grant A will be asserted even though both requests A and B are asserted.
- Device A will keep request A asserted until it no longer needs the bus so when Request A goes low, the arbiter will disable Grant A.
- Since request B remains asserted at this time, the arbiter will then assert Grant B to grant Device B access to the bus.
- Similarly, Device B will not disable Request B until it is done with the bus.

# Terna Engineering College, Navi Mumbai

Q.6.

Ans:

- Bus contention is when more than one device tries to drive data line at the same time.
- If one device is trying to drive it high and other tries to drive low then, we call it bus contention.
- Methods to avoid Bus contention
  - ① By having good arbitration.
  - ② The master device should always control who has access to the bus at any time.
  - ③ All slave devices need to be able to tri-state their output to bus.

# Terna Engineering College, Navi Mumbai

Q.7.

Ans:

- Superscalar architecture is a type of microprocessor design and construction that makes it possible for a processor to work on multiple sets of instruction at the same time. By sending them through separate execution units. Each unit can still only handle one set of instructions in order at a time, however it is possible to have multiple units run concurrently.
- Superscalar architecture requires the use of a built in scheduler that looks through the instruction queue and identifies groups and sets of instructions that don't conflict with one another.

# Terna Engineering College, Navi Mumbai

Q.8.

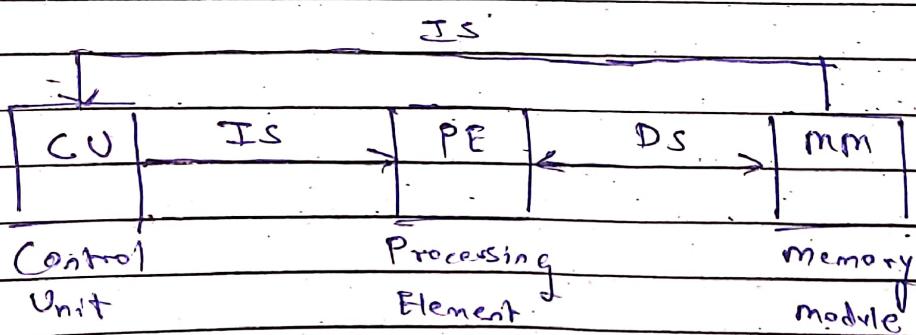
Ans:

Flynn's Classification.

- A method introduced by Flynn, for classification of parallel processors, is most common.
- This classification is based on the no. of Instruction streams (IS) and Data streams (DS) in the system.
- Single Instruction single Data (SISD)
- Single Instruction multiple Data (SIMD)
- Multiple Instruction Single Data (MISD)
- Multiple Instruction multiple Data (MIMD)

## ① Single Instruction single Data (SISD)

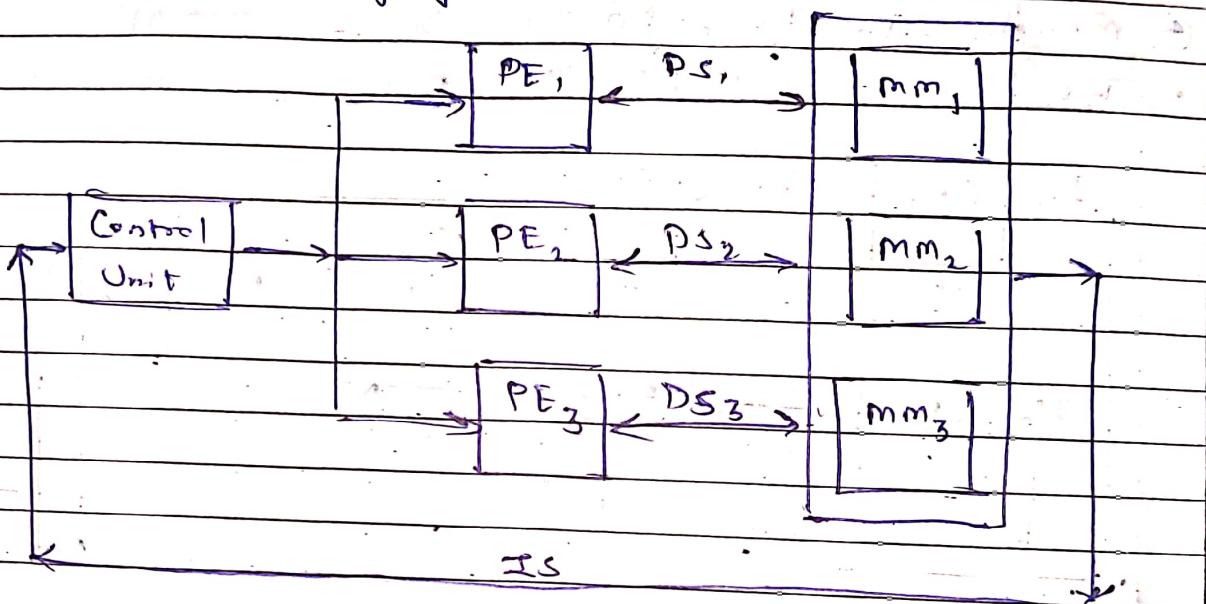
- In this case there is a single processor that executes one instruction at a time on single data stored in the memory.
- In fact, this type of processing can be said to be unit processing. Hence unit processors fall into this category.
- 



# Terna Engineering College, Navi Mumbai

## ② Single Instruction multiple Data (SIMD) :

- In this case, the same instruction is given to multiple processing element, but different data.
- This kind of system is mainly used when many data (Array of data) have to be operated with same operation.
- Vector processors and array processors fall into this category.

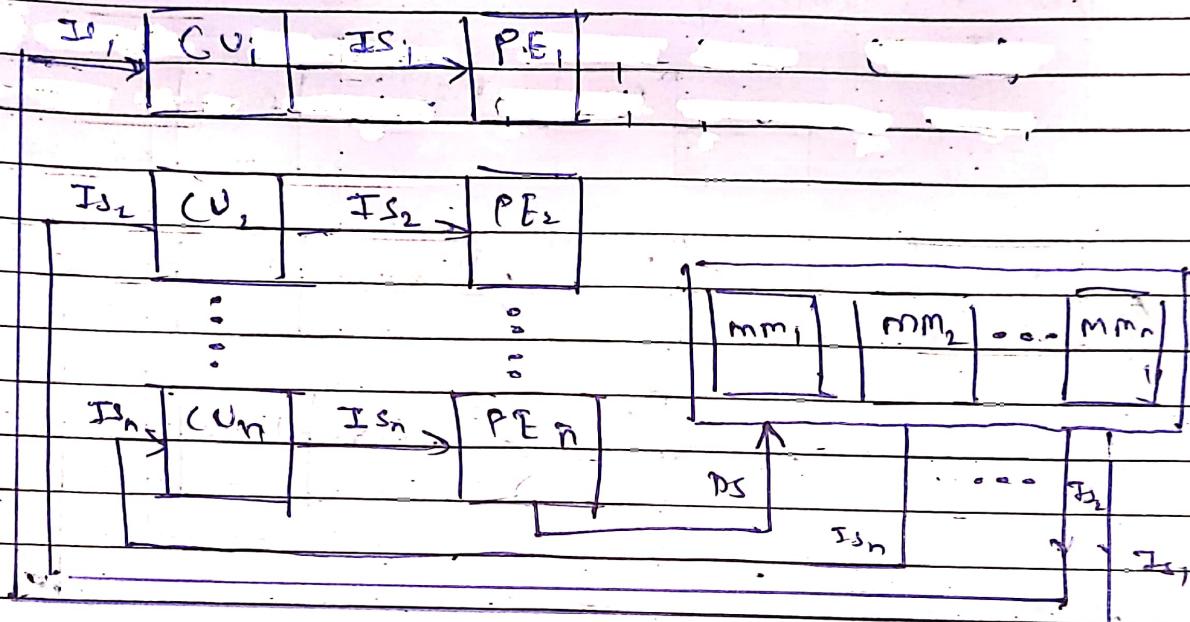


## ③ Multiple Instruction Single Data (MISD) :

- In case of MISD, there are multiple instruction streams and hence multiple control units to decode these instructions.
- Each control unit takes a different instruction from the different memory module in the same memory.
- The data stream is single. In this case the data is taken by the first processing element.
- This processing element performs an operation on the data given to it and forwards the result to the next processing element for further operation.

# Terna Engineering College, Navi Mumbai

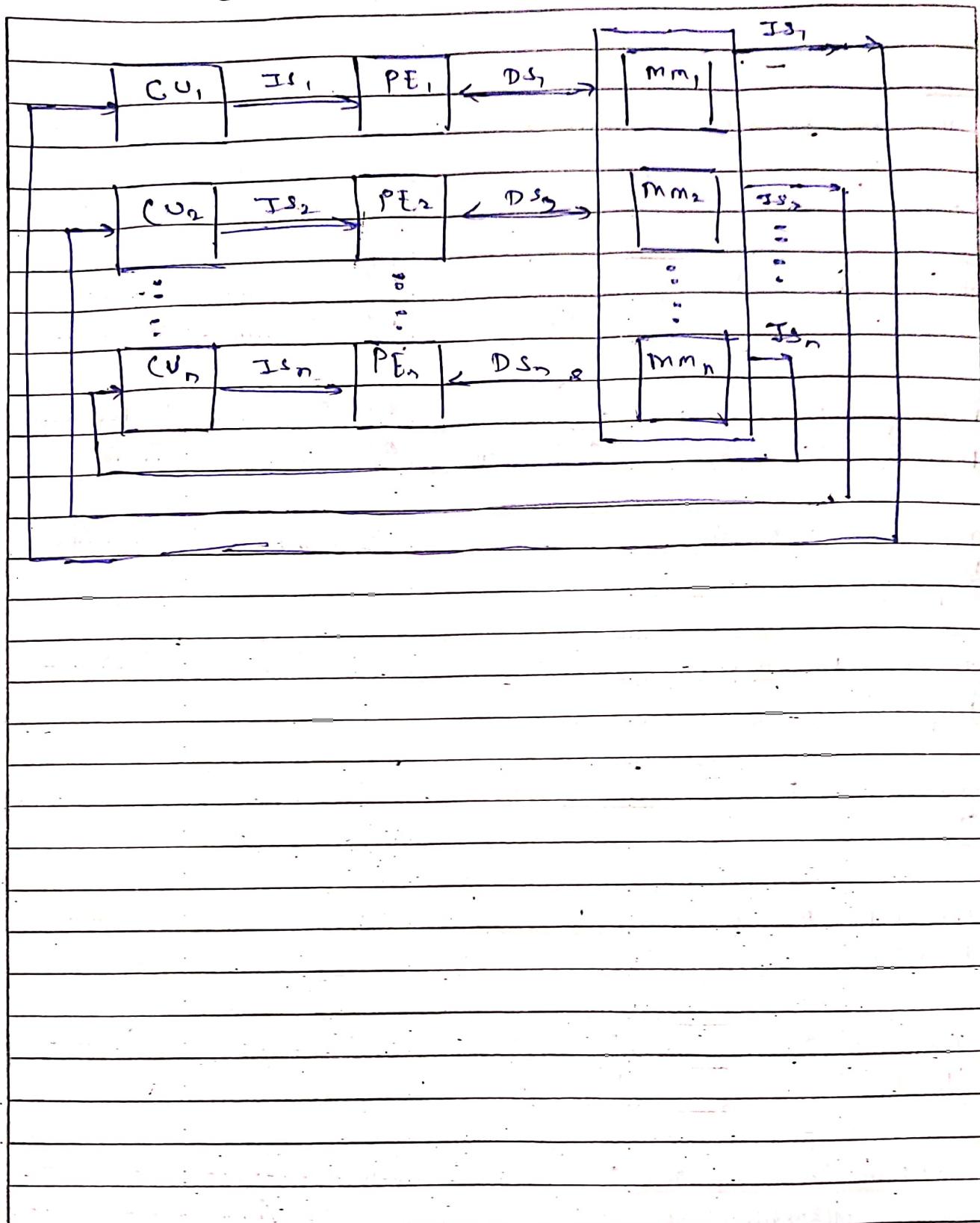
- This processing element performs a similar operation and so on the final result reaches back to the same memory module.
- This system is not used much but can be used in cases where a data has to undergo many computations to get the result for eg. to add two floating point numbers.



## (4) Multiple Instruction, multiple Data (MIMD)

- This is a complete parallel processing example. Here each processing element is having a different set of data and different instruction.
- Examples of this kind of system are SMP's (Symmetric Multiprocessors), clusters and NUMA (Non-uniform Memory Access).

# Terna Engineering College, Navi Mumbai



# Terna Engineering College, Navi Mumbai

Q9.

Ans:

Multicore processor architecture

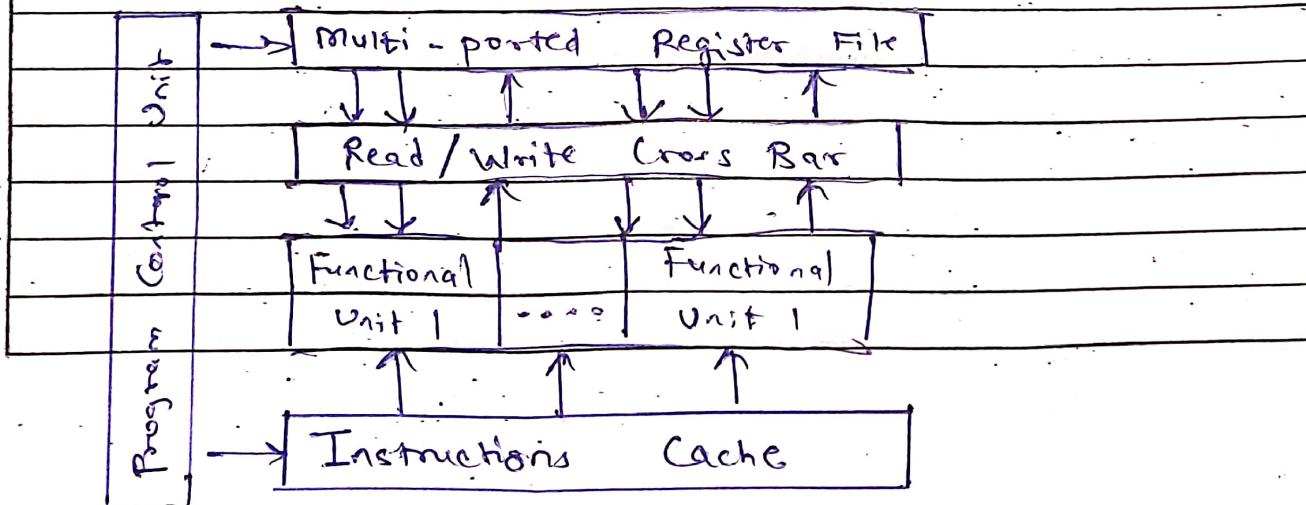
Features.	i3	i5	i7	i7 Extreme
① No. of cores	2 for desktop 2 for laptop	4 for desktop 2 for Laptop	4/6 for desktop 2/4 for Laptop	6 for Desktop 4 for mobile
② Processing threads	4 for desktop 4 for laptop	8 for desktop 4 for laptop	8/12 for desktop 4/8 for laptop	12 for desktop 8 for laptop
③ Maximum base clock frequency	3.4 GHz	3.4 GHz	3.2 GHz	3.3 GHz
④ Maximum turbo boot frequency	Not applicable	3.8 GHz	3.8 GHz	3.69 GHz
⑤ Maximum smart cache size	3 MB	6 MB	12 MB	15 MB
⑥ Intel turbo boost 2.0	Not present	Present	Present	Present
⑦ Intel Hyperthreading	Present	Present only in laptop processors	Present	Present
⑧ Best Desktop processor	Intel core i3-2130 (3.4 GHz, 3MB)	Intel core i5-2550K (3.4 GHz, 6MB)	Intel core i7-3930 (3.24 GHz, 612 MB)	Intel core i7-3770K (3.3 GHz, 15 MB)

# Terna Engineering College, Navi Mumbai

Q.10.

Ans:

- Very Long Instruction Word (VLIW) architecture in p-DSPs (Programmable DSP) increases the no. of instructions that are processed per cycle.
- It is concatenation of several short instructions and require many instruction unit running in parallel to carry out instructions in a single cycle.
- A language compiler or pre-processor separates program instructions into basic operations and places them into VLIW processor which then disassembles and transfers each operation to an appropriate execution unit.
- VLIW P-DSPs have no. of processor units (Data paths) i.e. they have a no. of ALUs, MAC units, shifters, etc.
- The VLIW is fetched from memory and is used to specify the operands and operations to be performed by each of the data paths.
- VLIW Architecture



# Terna Engineering College, Navi Mumbai

## Advantages of VLIW architecture

- Increased performance
- Potentially scalable; i.e. more execution units can be added and so more instructions can be packed into the VLIW instruction

## Disadvantages of VLIW architecture

- New programmer needed.
- Program must keep track of instruction scheduling
- Increased memory use
- High power consumption

# Terna Engineering College, Navi Mumbai

Q11.

Ans:

## Dynamic Instruction Scheduling

- It is widely used technique because of the speed up given by it.
- It is used in Pentium IV processor.
- Here the execution of the instruction of a program is done out-of-order i.e. not in the sequence as the instructions were written by the programmer.
- As and when the resources of an instruction is available, the execution of that instruction is done.
- If, for an instruction the resources are not available, it is kept in waiting state and the further instructions whose resources are available will be executed.
- This approach will have a problem.  
The logic implemented by the programmer will not be followed properly i.e. wrong sequence of instructions will be executed. The answer to that is, although the instructions are executed out of order but the write back is done in order and hence final result of the program is in sequence.
- The compiler is designed in such a way that while translating from high level language to machine language program, it detects data dependencies and re-orders the instructions.
- If necessary to delay the loading of the conflicting data it inserts No Operation Instruction (nop).

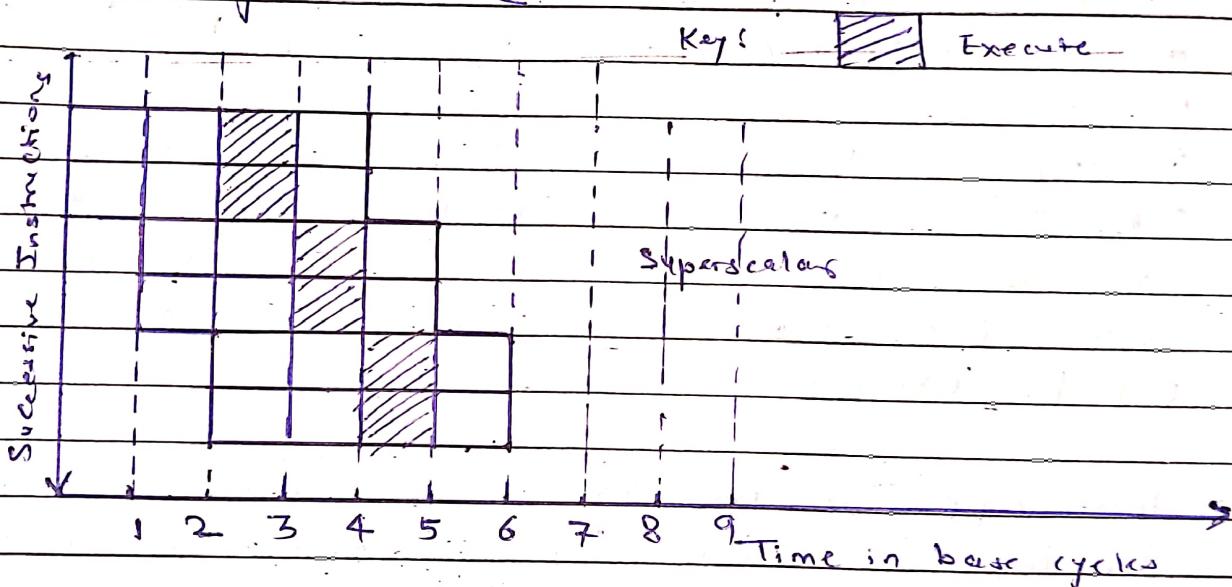
# Terna Engineering College, Navi Mumbai

Q12.

Ans:

## Pipelining in Super scalar Processors

- The pipelining is the most important representation of demonstrating the speed increase by the superscalar feature of the processor
- Timing diagram of the superscalar processor with degree  $m=2$



- Hence to implement multiple operations simultaneously we need to have multiple Execution unit to execute each instruction independently.
- The ID (Instruction Decode) and rename unit decodes the instruction and then by the use of register renaming avoids instruction dependency. The instruction window then takes the decoded instruction and based on some pairability rules, issues them to the respective execution units.

# Terna Engineering College, Navi Mumbai

- The instructions once executed move to the Retire and write back unit, wherein the instructions retire and result is written back to the corresponding destination

- Block diagram of typical superscalar processor.

