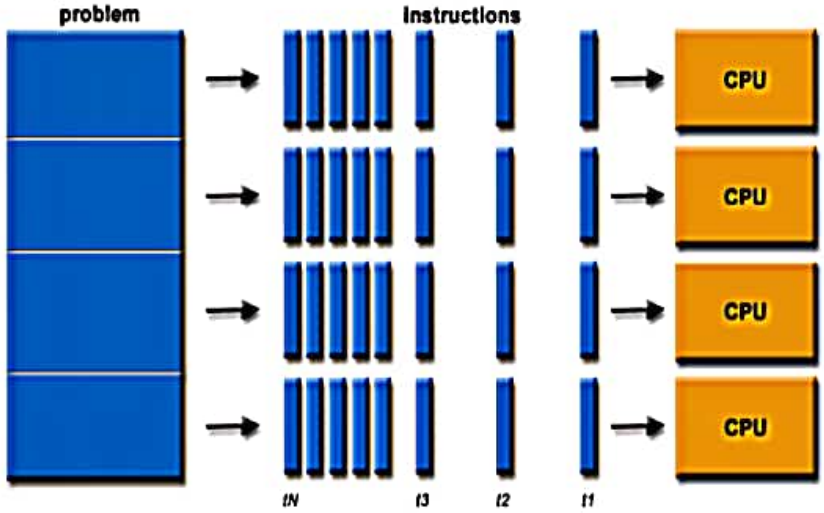
Unit-7

Parallel Processing

With the increased use of computers in every sphere of human activity, computer scientists are faced with two crucial issues today.

1.Processing has to be done faster like never before
2.Larger or complex computation problems need to be solved

- Parallel processing or computing is a form of computation in which many instructions are carried out simultaneously operating on the principle that large problems can often be divided into smaller ones, which are then solved concurrently (in parallel).
- Instead of processing each instruction sequentially a parallel processing system is able to perform concurrent data processing to achieve faster execution time



- Parallel Computing is used for the following reason
    - Saves time
    - Solve larger problems
    - Cost savings
    - Provide concurrency

It is broadly categorised into Four types

Bit level parallelism

When an 8-bit processor needs to add two 16-bit integers,it's to be done in two steps.

- The processor must first add the 8 lower-order bits from each integer using the standard addition instruction,
- Then add the 8 higher-order bits using an add-with-carry instruction and the carry bit from the lower order addition

Instruction Level Parallelism

The instructions given to a computer for processing can be divided into groups, or re-ordered and then processed without changing the final result. This is known as instruction-level parallelism i.e.,ILP.

1. e = a + b
2. f = c + d
3. g = e * f

Here, instruction 3 is dependent on instruction 1 and 2 . However,instruction 1 and 2 can be independently processed.

Task parallelism

Task Parallelism focuses on distribution of tasks across different processors. It is also known as functional parallelism or control parallelism

Data Parallelism

Data parallelism focuses on distributing the data across different parallel computing nodes. It is also called as loop-level parallelism.

Parallel processing can be achieved by using 3 technologies.

1.pipelining

2.Vector processing

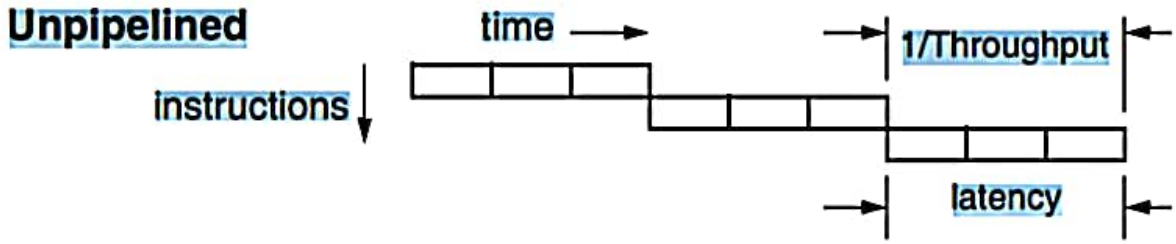
3.Array processing

Linear pipeline

The process of execution of instruction can be divided into 4 major steps.

1.Instruction Fetch(IF)
2.Instruction Decode(ID)
3.Operand fetch(OF)
4.Execute(EX)

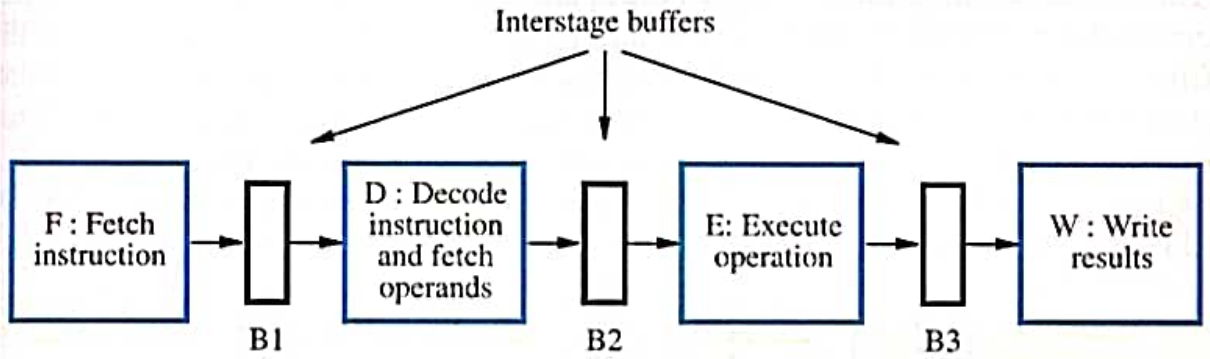
- During IF the instruction is fetched from main memory.
- During ID the operation is identified that is to be performed.
- During OF the operand is fetched from memory (if required).
- During EX the instruction is executed by ALU



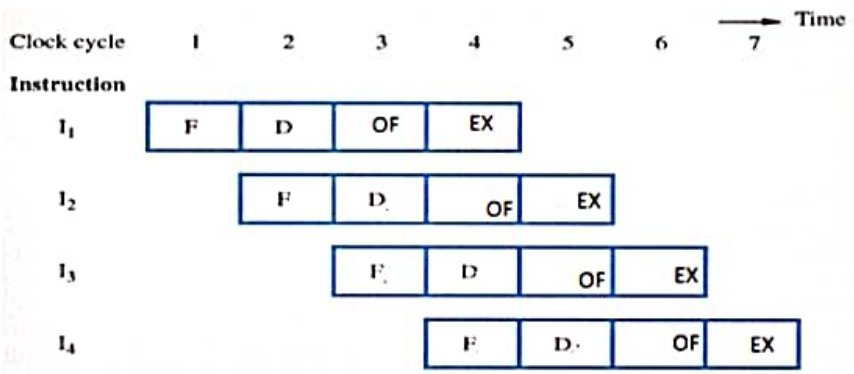
In unpipelined computer all these four steps must be completed before starting of the next instruction.

But in case of a pipelined computer successive instruction are executed in an overlapped fashion.

- It decomposes a sequential process into segments.
- It divide the processor into segment processors each one is dedicated to a particular segment.
- Each segment is executed in a dedicated segment-processor operates concurrently with all other segments.
- Information flows through these multiple hardware segments.



- Instruction execution is divided into k segments or stages
- Instruction exits pipe stage k-1 and proceeds into pipe stage k
- All pipe stages take the same amount of time; called one processor cycle
- Length of the processor cycle is determined by the slowest pipe stage
- There is a inter stage buffer between two successive stage.



FLYNN'S TAXONOMY

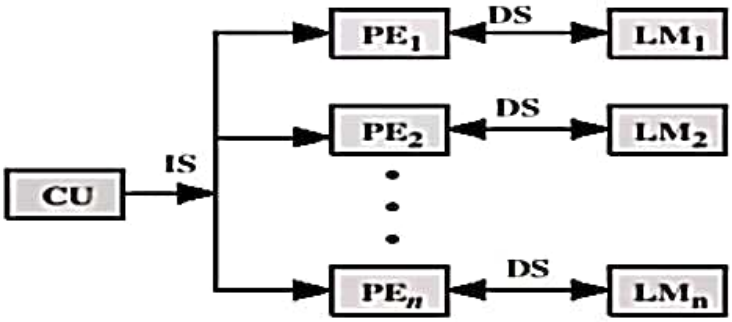
- In general, digital computers may be classified into four categories, according to the multiplicity of instruction and data streams.
- This scheme for classifying computer organizations was introduced by Michael J. Flynn. The essential computing process is the execution of a sequence of instructions on a set of data.
- The term stream is used here to denote a sequence of items (instructions or data) as executed or operated upon by a single processor.
- An instruction stream is a sequence of instructions as executed by the machine;
- A data stream is a sequence of data including input, partial, or temporary results, called for the instruction stream.Listed below are Flynn's four machine organizations:

• Single instruction stream single data stream (SISD)

SISD computer organization This organization represents most serial computers available today. Instructions are executed sequentially but may be overlapped in their execution stages.

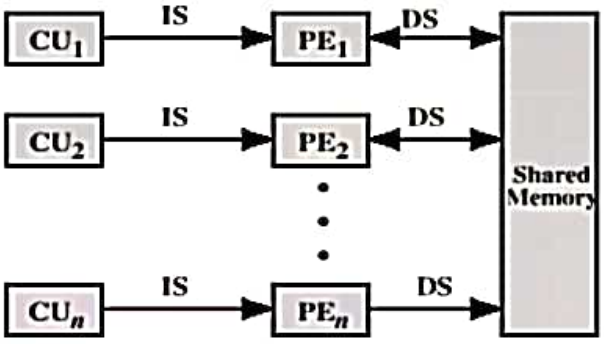
• Single instruction stream multiple data stream (SIMD)

SIMD computer organization In this organization, there are multiple processing elements supervised by the same control unit. All PE'(processing element) receive the same instruction broadcast from the control unit but operate on different data sets from distinct data streams.



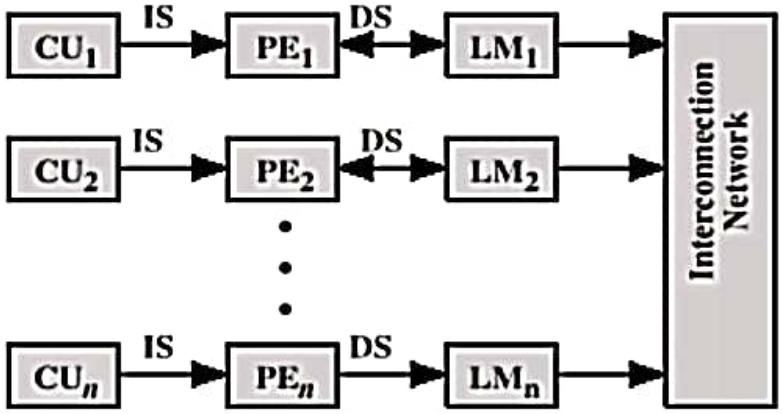
• Multiple instruction stream single data stream (MISD)

MISD computer organization There are n processor units, each receiving distinct instructions operating over the same data stream. The results (output) of one processor become the input (operands) of the next processor .This approach has no practical implementation.



- **Multiple instruction stream multiple data stream (MIMD)**

MIMD computer organization Most multiprocessor systems and multiple computer systems can be classified in this category. MIMD computer implies interactions among the n processors because all memory streams are derived from the same data space shared by all processors.



MULTIPROCESSOR

- A multiprocessor system is a interconnection of two or more cpus with memory & i\o equipment. The processor in multiprocessor can have either cpu or i\o processor.
- Computers are interconnected with each other by means of communication lines to form a computer network.The network consists of several autonomous computer that may or mynot communicate with each other.
- A multiprocessor system is controlled by one os that provides interaction between processor & all the components of the system cooperate in the solution of a problem.
- Multiprocessor are classified by the way their memory is organized.
    - ✔ Tightly coupled
    - ✔ Loosely coupled.

Tightly coupled

This systems contain multiple CPUs that are connected at the bus level. These CPUs may have access to a central shared memory (SMP or UMA), or may participate in a memory hierarchy with both local and shared memory (NUMA).

Loosely coupled

These system are based on multiple standalone single or dual processor or commodity computers interconnected via a high speed communication system .

Each PE has its own private local memory.the processor are tied together by a switching scheme designed to route information from one processor to another through a message passing system.