Cloud Computing

Session 2022-2023 (Odd Semester)

Dr. Prasenjit Chanak
Assistant Professor

E-mail: prasenjit.cse@iitbhu.ac.in

Department of Computer Science and Engineering
Indian Institute of Technology (BHU), Varanasi-221005

Syllabus

- **Introduction to Cloud Computing**: Overview of computing, Cloud computing, Properties, Characteristics & Disadvantages, Role of open standard
- **Virtualization:** Issues with virtualization, Virtualization technologies and architectures, Internals of virtual machine monitors/hypervisors, Virtualization of data centers, and Issues with Multi-tenancy.
- **Cloud Computing Architecture:** Cloud computing stack, Deployment Models
- **Service Management in Cloud Computing**
- **Resource Management and load balancing in Cloud Computing**
- **Data Management in Cloud Computing**
- **Cloud Security:** Infrastructure security, Data security and storage, Identity and access management, Access control, trust, reputation risk
- **Case study on open source and commercial clouds, cloud simulator**

Class Timings

Instructor: Dr. Prasenjit Chanak

Day	*Time*
Monday	11:00 am to 11:55 am
Tuesday	10:00 am to 10:55 am
Thursday	10:00 am to 10:55 am

Teaching Assistant (TA):

1. Archana Ojha (Email: archanaojha.rs.cse21@itbhu.ac.in)
2. Saurabh Kumar Srivastava (saurabhkrsrivastava.rs.cse21@itbhu.ac.in)

Evaluation Pattern

Evaluation Pattern	**Number of exam**	**Weightage (%)**
Project	01	15%
Labs (Programming Assignments)	-	15%
Assignments	-	10%
Mid semester	01	20%
End semester	01	30%
Attendance, sincerity etc		10%

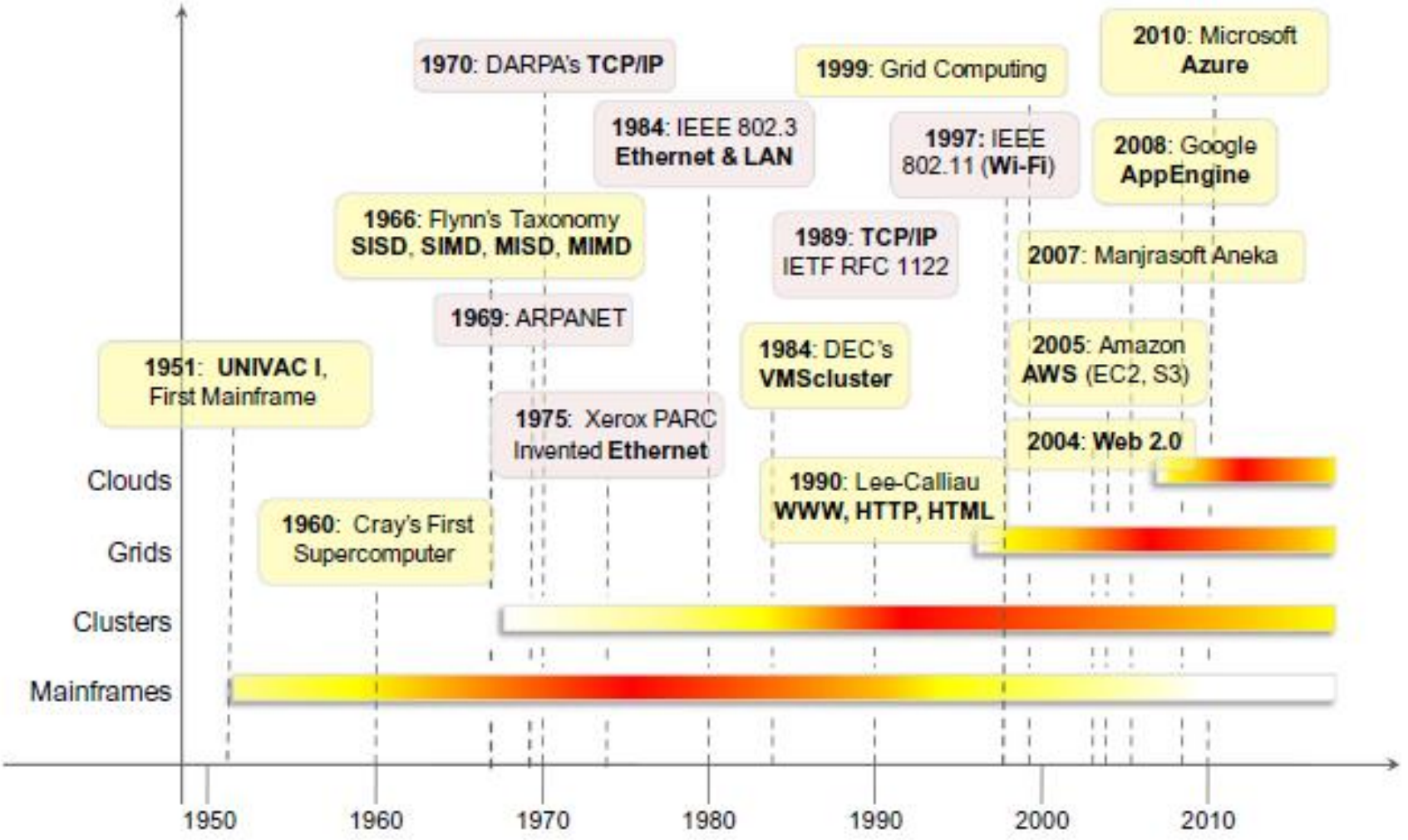
Text Books

1. Mastering Cloud computing, Rajkumar Buyya, Christian Vacchiola, S Thamarai Selvi, McGraw Hill, 2013.
2. Cloud Computing Principles and Paradigms, Rajkumar Buyya, James Broberg, Andrzej Goscinski, Wiley Publishers, 2011.
3. Cloud Computing Bible, Barrie Sosinsky, Wiley Publishers, 2010.
4. Cloud Computing : Web-based Applications that change the way you work and collaborate online, Michael Miller, Pearson Education, 2008.
5. Cloud Computing: A Practical Approach, Toby Velte, Antohy T Velte, Robert Elsenpeter, McGraw Hill, 2009.

Overview

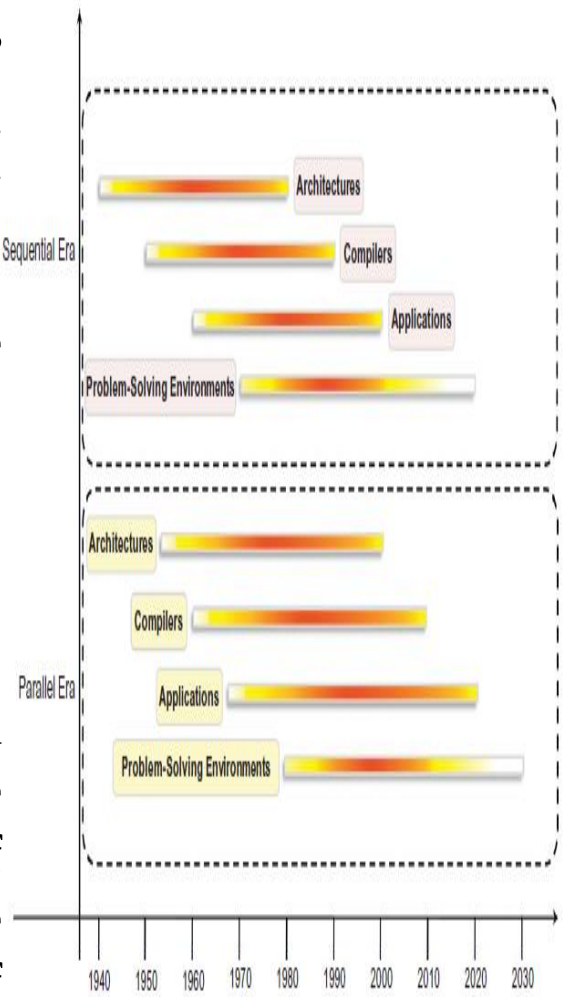
- Three major milestones have led to cloud computing evolution
  - **Mainframes:** Large computational facilities leveraging multiple processing units. Even though mainframes cannot be considered as distributed systems, they offered large computational power by using multiple processors, which were presented as a single entity to users.
  - **Clusters:** An alternative technological advancement to the use of mainframes and super computers.
  - **Grids**
  - **Clouds**

Mile Stones to Cloud computing Evolution



Eras of Computing

- Two fundamental and dominant models of computing are ***sequential*** and ***parallel***.
  - The sequential era began in the 1940s, and Parallel(and distributed)
  - computing era followed it within a decade.
- Four key elements of computing developed during three eras are
  - Architecture
  - Compilers
  - Applications
  - Problem solving environments
- The computing era started with development in hardware architectures, which actually enabled the creation of system software – particularly in the area of compilers and operating systems – which support the management of such systems and the development of applications



Principles of Parallel and Distributed Computing

- The term parallel computing and distributed computing are often used interchangeably, even though they mean slightly different things.
- The term parallel implies a tightly coupled system, where as distributed systems refers to a wider class of system, including those that are tightly coupled.
- More precisely, the term parallel computing refers to a model in which the computation is divided among several processors sharing the same memory.
- The architecture of parallel computing system is often characterized by the homogeneity of components: each processor is of the same type and it has the same capability as the others.

Contd…

- The shared memory has a single address space, which is accessible to all the processors.
- Parallel programs are then broken down into several units of execution that can be allocated to different processors and can communicate with each other by means of shared memory.
- Originally parallel systems are considered as those architectures that featured multiple processors sharing the same physical memory and that were considered a single computer.

Contd…

- Over time, these restrictions have been relaxed, and parallel systems now include all architectures that are based on the concept of shared memory, whether this is physically present or created with the support of libraries, specific hardware, and a highly efficient networking infrastructure.
- For example: a cluster of which of the nodes are connected through an InfiniBand network and configured with distributed shared memory system can be considered as a parallel system.

Contd…

- The term distributed computing encompasses any architecture or system that allows the computation to be broken down into units and executed concurrently on different computing elements, whether these are processors on different nodes, processors on the same computer, or cores within the same processor.
- Even though it is not a rule, the term distributed often implies that the locations of the computing elements are not the same and such elements might be heterogeneous in terms of hardware and software features.
- Classic examples of distributed computing systems are
  - Computing Grids
  - Internet Computing Systems

Elements of Parallel computing

- Silicon-based processor chips are reaching their physical limits. Processing speed is constrained by the speed of light, and the density of transistors packaged in a processor is constrained by thermodynamics limitations.
- A viable solution to overcome this limitation is to connect multiple processors working in coordination with each other to solve “Grand Challenge” problems.

Elements of Parallel computing

- The first step in this direction led
  - To the development of parallel computing, which encompasses techniques, architectures, and systems for performing multiple activities in parallel.
  - As discussed earlier, the term parallel computing has blurred its edges with the term distributed computing and is often used in place of later term.

What is Parallel Processing?

- Processing of multiple tasks simultaneously on multiple processor is called **parallel processing**
- The parallel program consists of multiple active processes (tasks) simultaneously solving a given problem
- A given task is divided into multiple subtasks using divide-and-conquer technique, and each one of them is processed on different CPUs
- Programming on multi-processor system using divide-and-conquer technique is called **parallel programming**

What is Parallel Processing?

- Many applications today require more computing power than a traditional sequential computer can offer.
- Parallel Processing provides a cost effective solution to this problem by increasing the number of CPUs in a computer and by adding an efficient communication system between them.
- The workload can then be shared between different processors. This setup results in higher computing power and performance than a single processor a system offers.

Parallel Processing influencing factors

- The development of parallel processing is being influenced by many factors. The prominent among them include the following:
- Computational requirements are ever increasing in the areas of both scientific and business computing. The technical computing problems, which require high-speed computational power, are related to life sciences, aerospace, geographical information systems, mechanical design and analysis etc.

Contd…

- Sequential architectures are reaching mechanical physical limitations as they are constrained by the speed of light and thermodynamics laws.
  - The speed which sequential CPUs can operated is reaching saturation point (no more vertical growth), and hence an alternative way to get high computation speed is to connect multiple CPUs (opportunity for horizontal growth).

Contd…

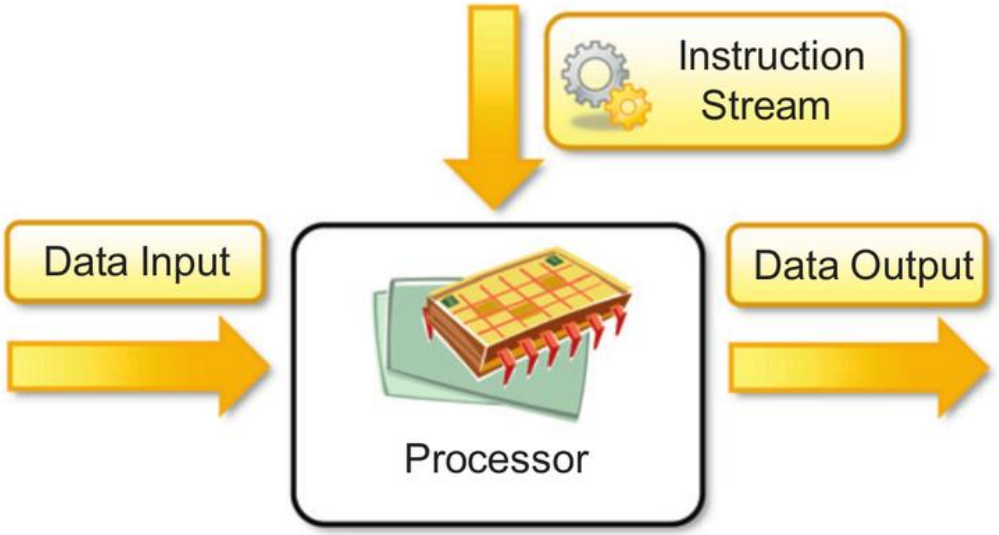
- Vector Developing such compiler technology is a difficult task. or processing works well for certain kinds of problems. It is suitable mostly for scientific problems (involving lots of matrix operations) and graphical processing. It is not useful for other areas, such as databases.
- The technology of parallel processing is mature and can be exploited commercially here is already significant R&D work on development tools and environments. Significant development in networking technology is paving the way for heterogeneous computing.

Hardware Architectures for Parallel Processing

- The core elements of parallel processing are CPUs
- Based on a number of instruction and data streams that can be processed simultaneously, computing systems are classified into the following four categories:
  - Single Instruction Single Data (SISD)
  - Single Instruction Multiple Data (SIMD)
  - Multiple Instruction Single Data (MISD)
  - Multiple Instruction Multiple Data (MIMD)
    - Shared Memory MIMD Machine
    - Distributed Memory MIMD Machine

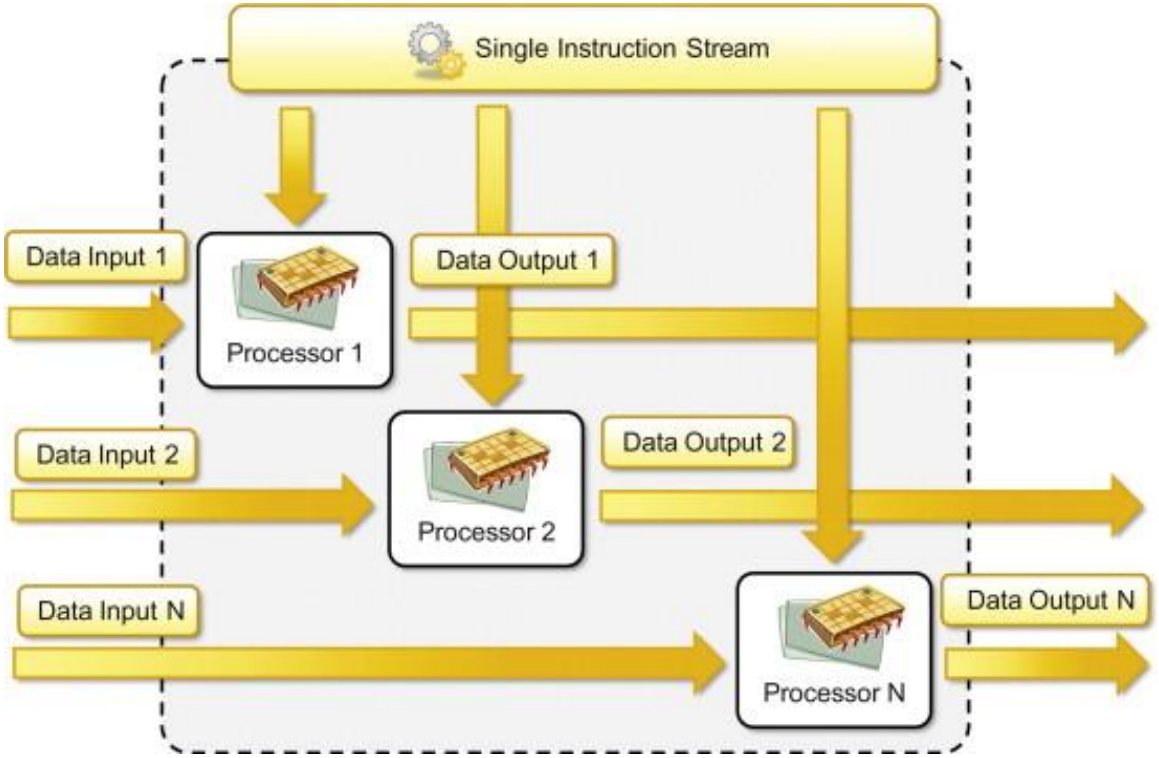
Single Instruction Single Data (SISD)

- In SISD, machine instructions are processed sequentially, and hence computers adopting this model are popularly called sequential computers



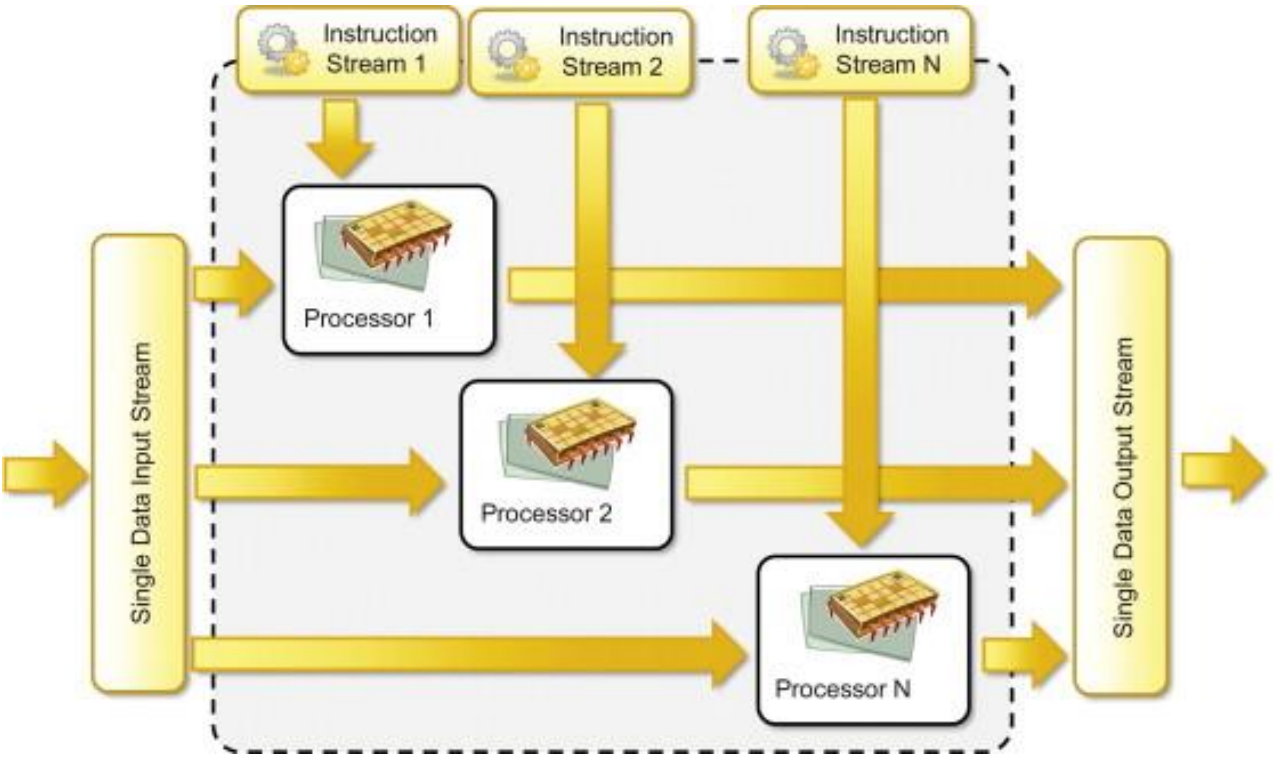
Single Instruction Multiple Data (SIMD)

- A SIMD computing system is a multiprocessor machine capable of executing the same instruction on all the CPUs, but operating on different data streams



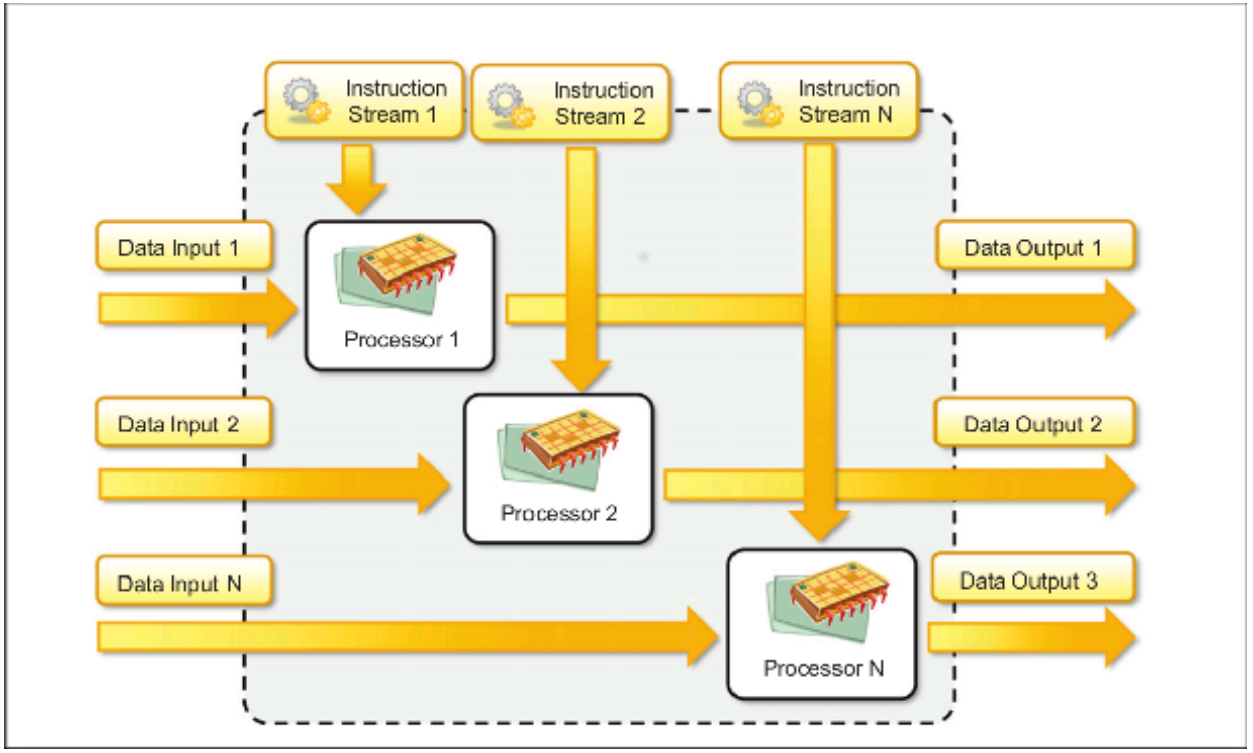
Multiple Instruction Single Data (MISD)

- A MISD computing system is a multiprocessor machine capable of executing different instructions on different CPSs, but all of them operating on the same data-set



Multiple Instruction Multiple Data (MIMD)

- A MISD computing system is a multiprocessor machine capable of executing multiple instructions on multiple data-sets



Approaches to Parallel Programming

- A wide variety of parallel Programming approaches are available. The most prominent among them are the following:
    - Data Parallelism
    - Process Parallelism
    - Farmer and Worker Model
- In case of data parallelism, divide-an-conquer technique is used to split data into multiple sets

Levels of Parallelism

- Level of parallelism are decided based on the lumps of code (grain size) that can be a potential candidate for parallelism
  - Large-grain (or task-level)
  - Medium-grain (or control-level)
  - Fine-grain (data-level)
  - Very-fine grain (multiple instruction issue)

Distributed Computing

- General definition proposed by Tanenbaum
  - *"A distributed system is a collection of independent computers that appears to its users as a single coherent system"*

- A definition proposed by Coulouris
  - *"A distributed system is one in which components located at networked computers communicate and coordinate their actions only by passing messages"*

Distributed Computing/System?

- Distributed computing
  - Field of computing science that studies distributed system.
  - Use of distributed systems to solve computational problems.
- Distributed system
  - Wikipedia
  - There are several autonomous computational entities, each of which has its own local memory
  - The entities communicate with each other by message passing.
- Operating System Concept
  - The processors communicate with one another through various communication lines such as high speed buses.