Introduction to Parallel Computing

Instructor

Dr. B Krishna Priya

Contents

- Introduction to Parallel computing
- Motivating Parallelism
- Scope of Parallelism
- Parallel Programming Platforms
- Implicit Parallelism: Trends in Microprocessor Architectures
- Limitations of Memory System Performance
- Dichotomy of Parallel Computing Platforms
- Physical Organization of Parallel Platforms

What is parallel computing

- Traditional software has been written for serial computing.
- To be run on single computer with single CPU
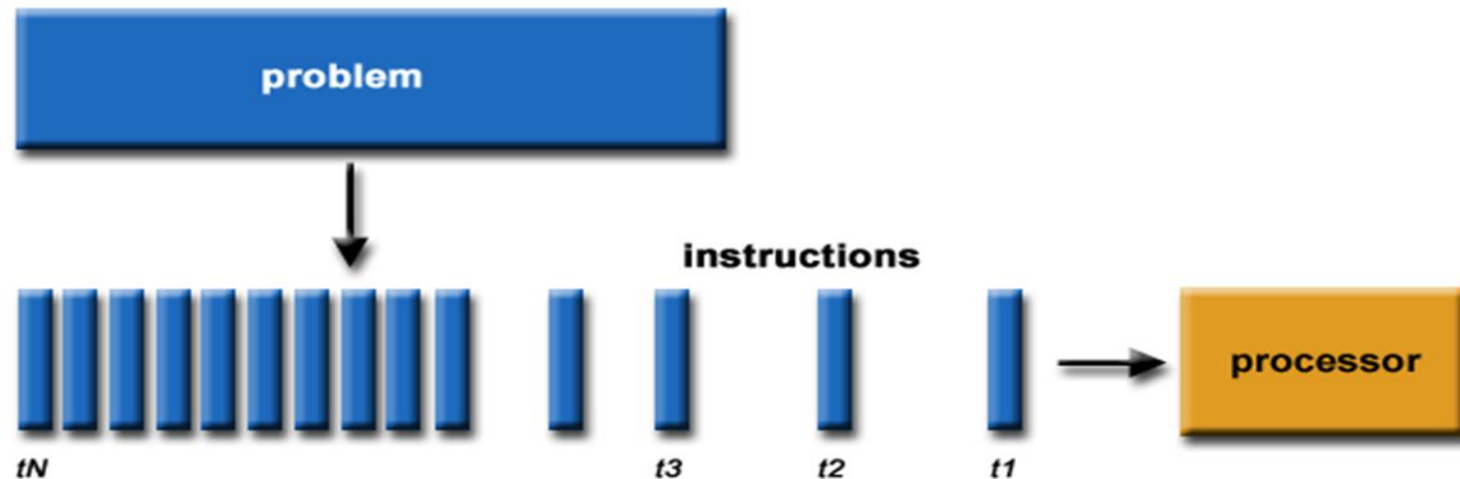
Limitations of Serial Computing

- Limits to serial computing - both physical and practical reasons pose significant constraints to simply building ever faster serial computers.

- Transmission speeds - the speed of a serial computer is directly dependent upon how fast data can move through hardware.

- Limits to miniaturization - processor technology is allowing an increasing number of transistors to be placed on a chip. However, even with molecular or atomic-level components, a limit will be reached on how small components can be.

- Economic limitations - it is increasingly expensive to make a single processor faster. Using a larger number of moderately fast commodity processors to achieve the same (or better) performance is less expensive.
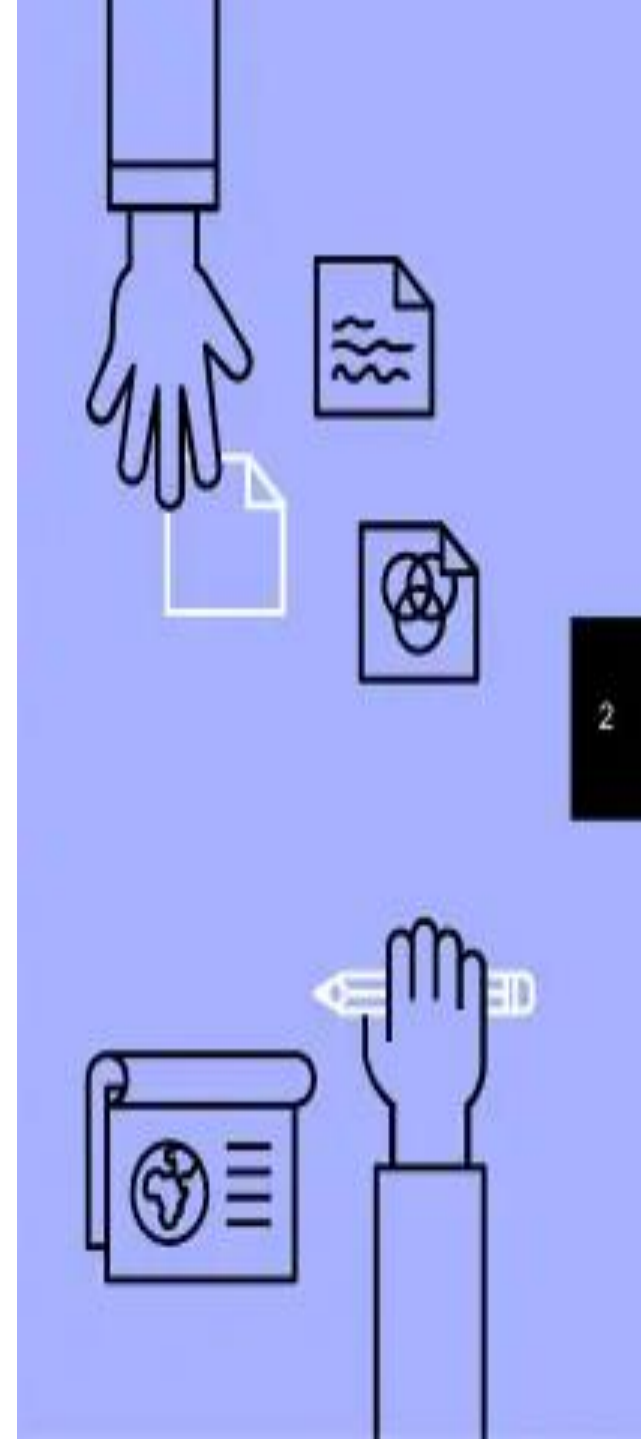
What is parallel computing

- In the simplest sense, parallel computing is the simultaneously use of multiple computing resource to solve computational problem.
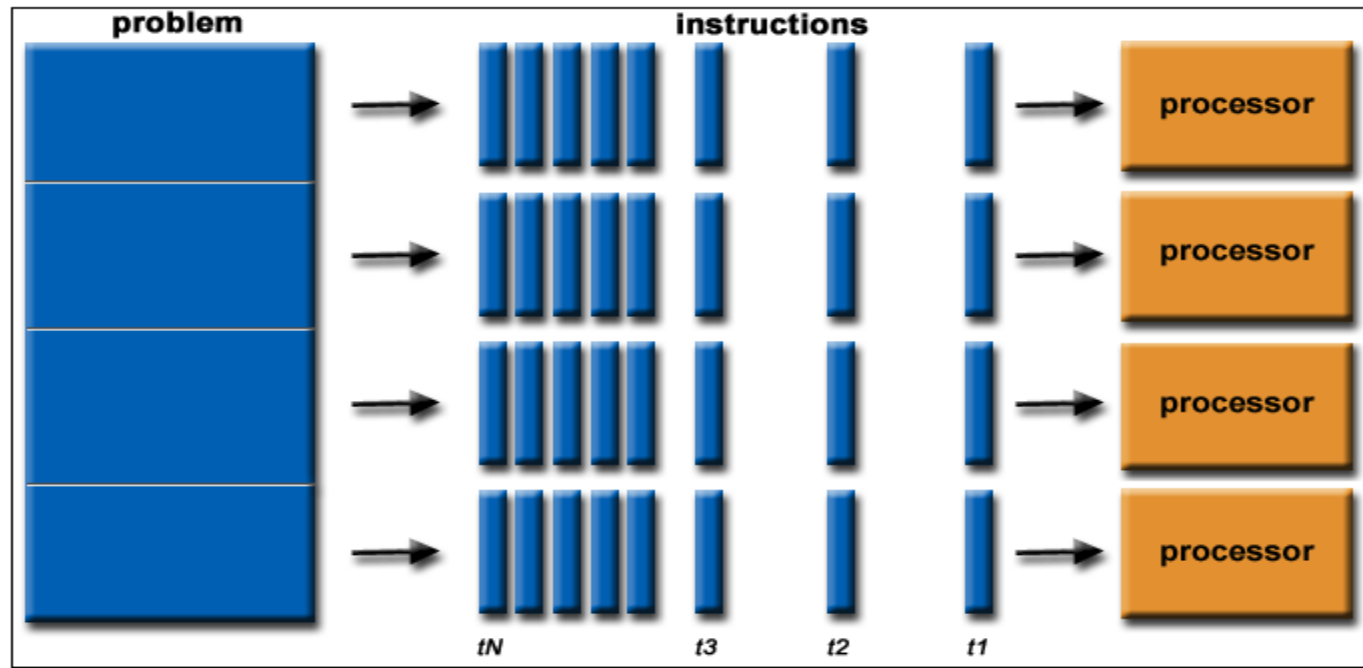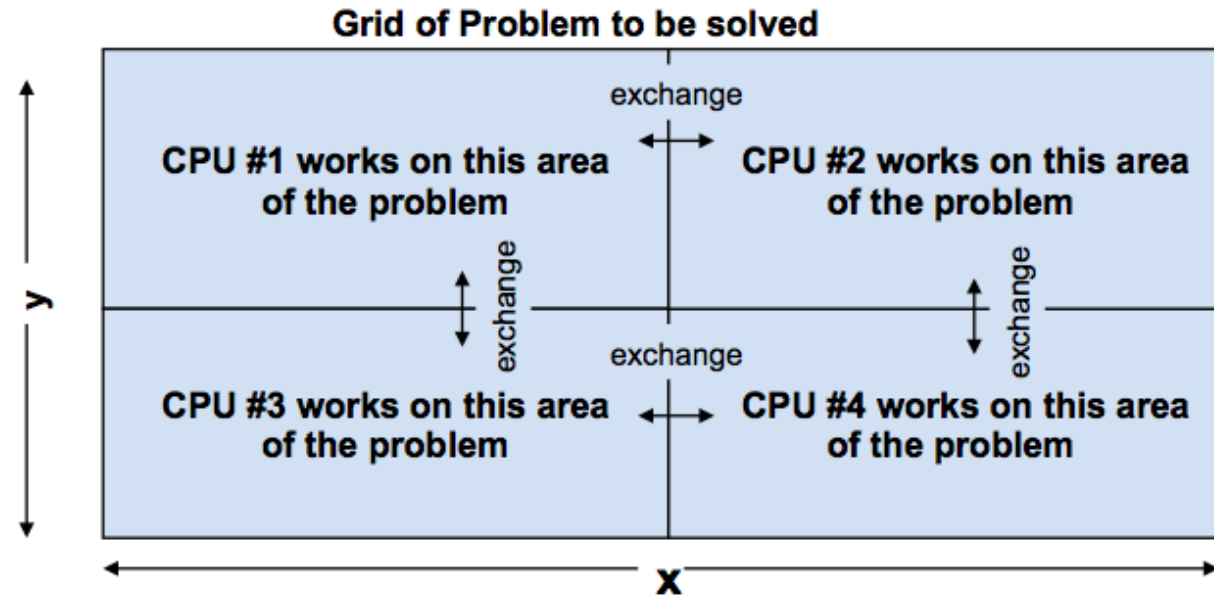
Parallel computing

- A parallel computing is a "Collection of processing elements that co-operate and communicate to solve large Problem"

- Basic Principle: A computation can be divided into smaller subproblem each of which can be solved simultaneously.

- Assumption: Parallel computing assumes that a parallel hardware is present at our disposal, which is capable of executing these computations in parallel.
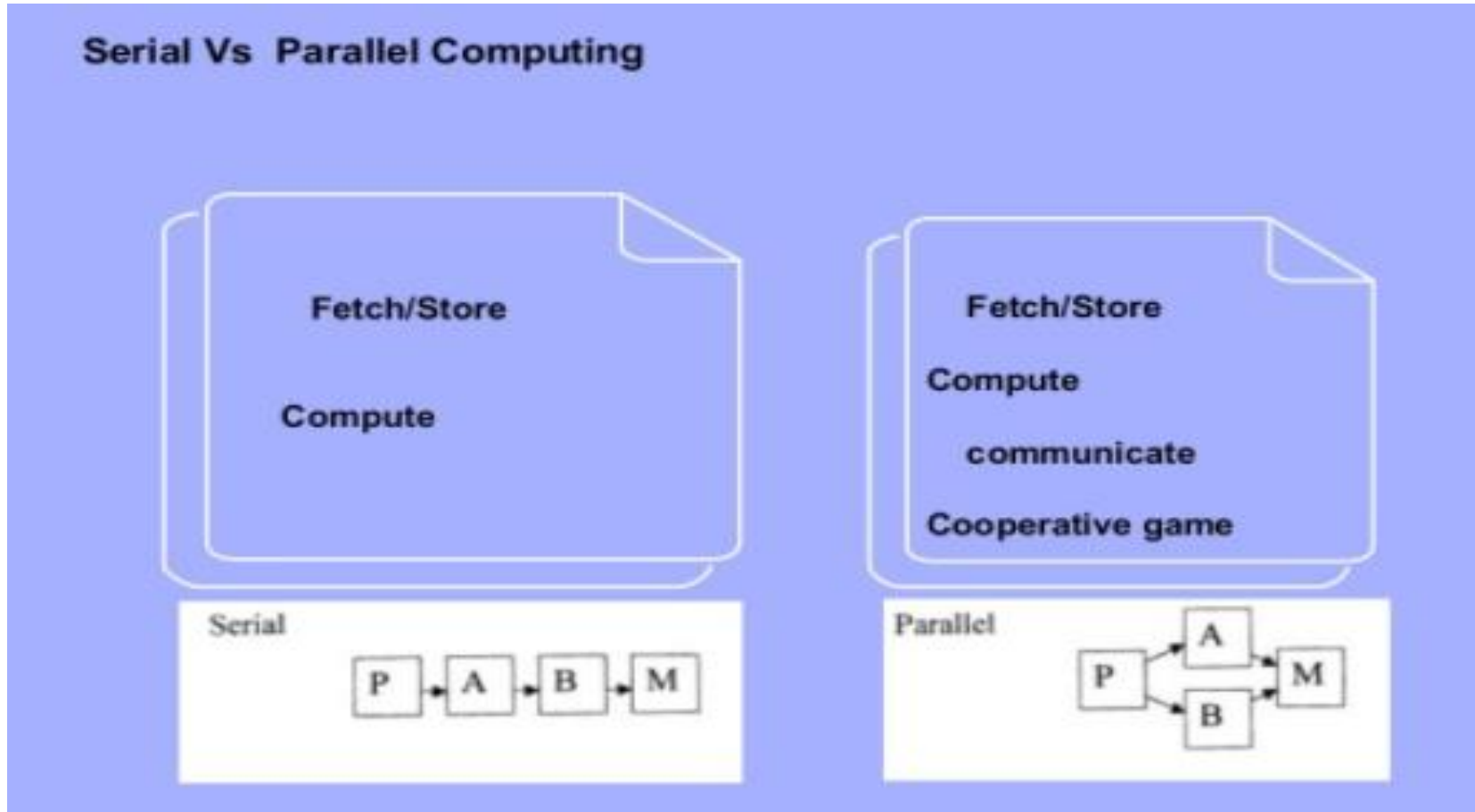
The concept of parallel computing

- One processor in your desktop or laptop can give you output in X hours

- X>10

- Why not use several (Y) processors, and get the results in X<1?

- Each processor works on one part of the problem
- Processors can exchange information with other processors

Grid of Problem to be solved

exchange

CPU #1 works on this area of the problem

exchange

CPU #2 works on this area of the problem

exchange

exchange

CPU #3 works on this area of the problem

exchange

CPU #4 works on this area of the problem

y

x

Serial vs Parallel Computing

Motivating Parallel Computing

The primary reasons for using parallel computing:

- Save time - wall clock time

- Solve larger problems

- Provide concurrency (do multiple things at the same time)

Motivating Parallel Computing Computational Power argument

Moore's law states [1965]:

- Moore law is a prediction formulated by Moore in 1965 about transistor density on integrated circuit chip(ICs).\

- According to Moore's Law, " the number of transistors used per square inch in the integrated circuits doubles about every 2 years."

- As a result the scale gets smaller and smaller and performance of microprocessor has enjoyed an exponential growth.

Technology Growth Graph
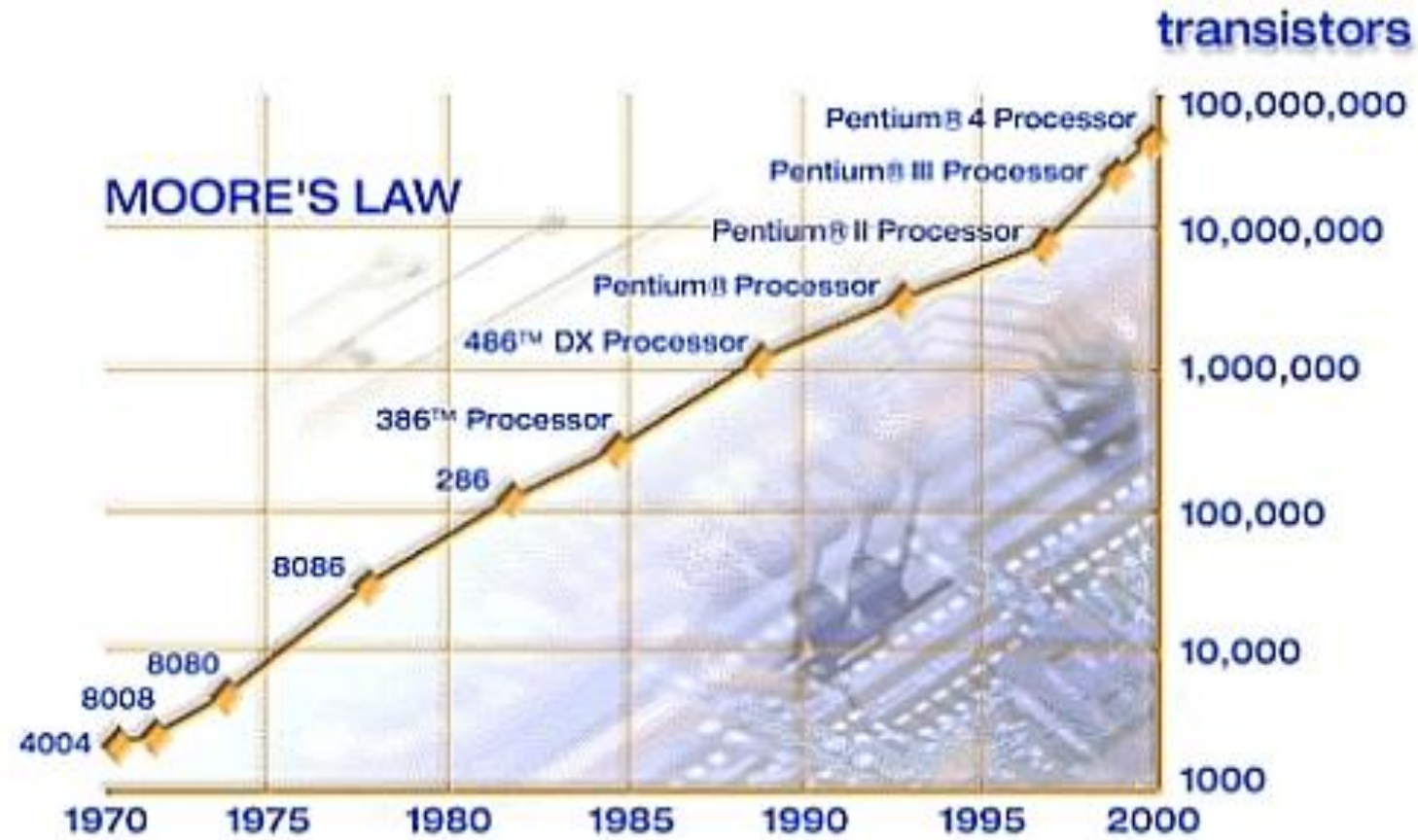
Motivating Parallel Computing
The Memory/Disk Speed Argument

- **While clock rates of high-end processors have increased at roughly 40% per year over the past decade, DRAM access times have only improved at the rate of roughly 10% per year over this interval.**

- **This mismatch in speeds causes significant performance bottlenecks.**

- **Parallel platforms provide increased bandwidth to the memory system.**

- **Parallel platforms also provide higher aggregate caches.**

- **Principles of locality of data reference and bulk access, which guide parallel algorithm design also apply to memory optimization.**

- **Some of the fastest growing applications of parallel computing utilize not their raw computational speed, rather their ability to pump data to memory and disk faster**.

Motivating Parallel Computing
The Data Communication Argument

- As the network evolves, the vision of the Internet as one large computing platform has emerged.

- In many other applications (typically databases and data mining) the volume of data is such that they cannot be moved.

- Any analyses on this data must be performed over the network using parallel techniques.

Scope of Parallel computin

- Parallelism finds applications in very diverse application domains for different motivating reasons.

- These range from improved application performance to cost considerations.
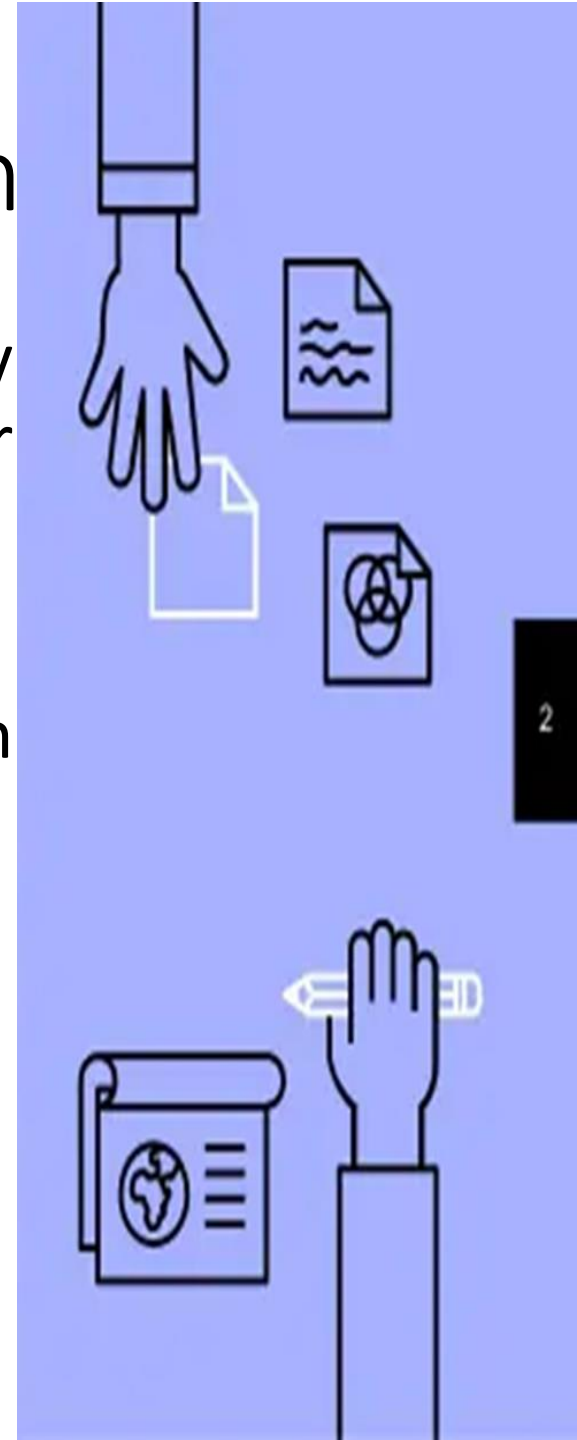
Applications

- Commercial computing
  - Weather forecasting
  - Remote sensors, image processing
  - Process optimization, operation resea
- Scientific and Engineering applicati
  - Computational chemistry
  - Molecular modeling
  - Structure mechanics
- Business application
  - E-governance
  - Medical imaging
- Internet applications
  - Internet servers
  - Digital libraries

HPC: What for?

European Commission

Weather, Climate & Ear

Bio/Life Sciences

New applications e.g. Health, Big data

Fundamental sciences: Physics, Chemistry, Material Sciences, Astrophysics

Industrial & Engineering