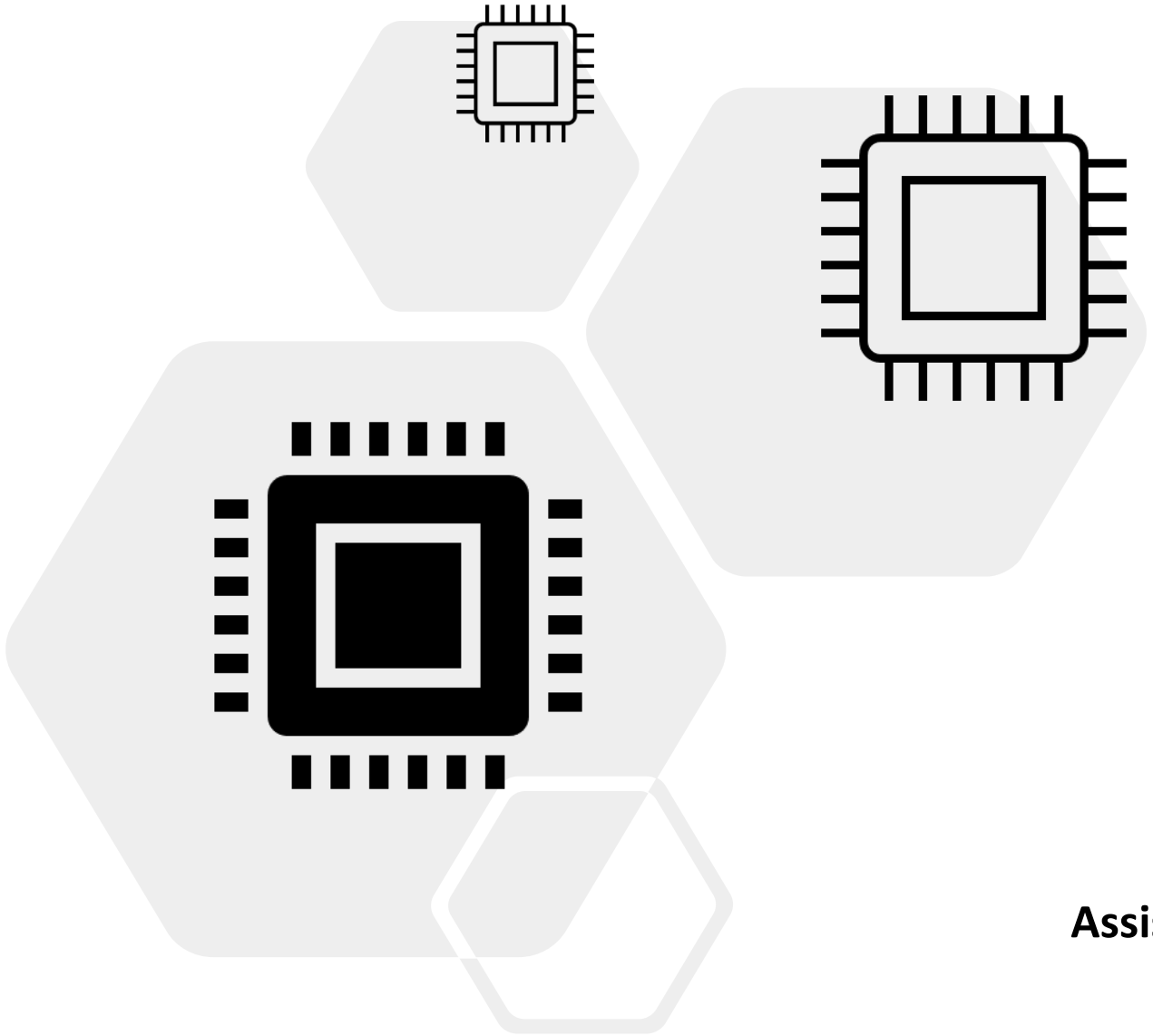
High Performance Computing

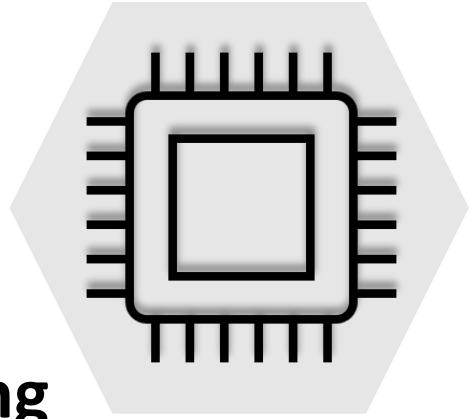


Introduction

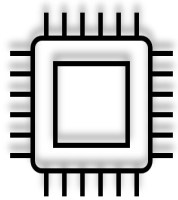
Introduction of Parallel Processing

Shafaque Fatma Syed
Assistant Professor - Dept. of Computer Engineering
A P Shah Institute of Technology, Mumbai

Topics to be discussed



- Difference between Serial Computing and Parallel Computing
- Fundamental limits on Serial Computing: Three “Walls” (Motivating Parallelism)
- Why is Parallel Computing Important?
- Who is Using Parallel Computing?(Scope of Parallel Computing)
- Why is Parallel Computing Hard?



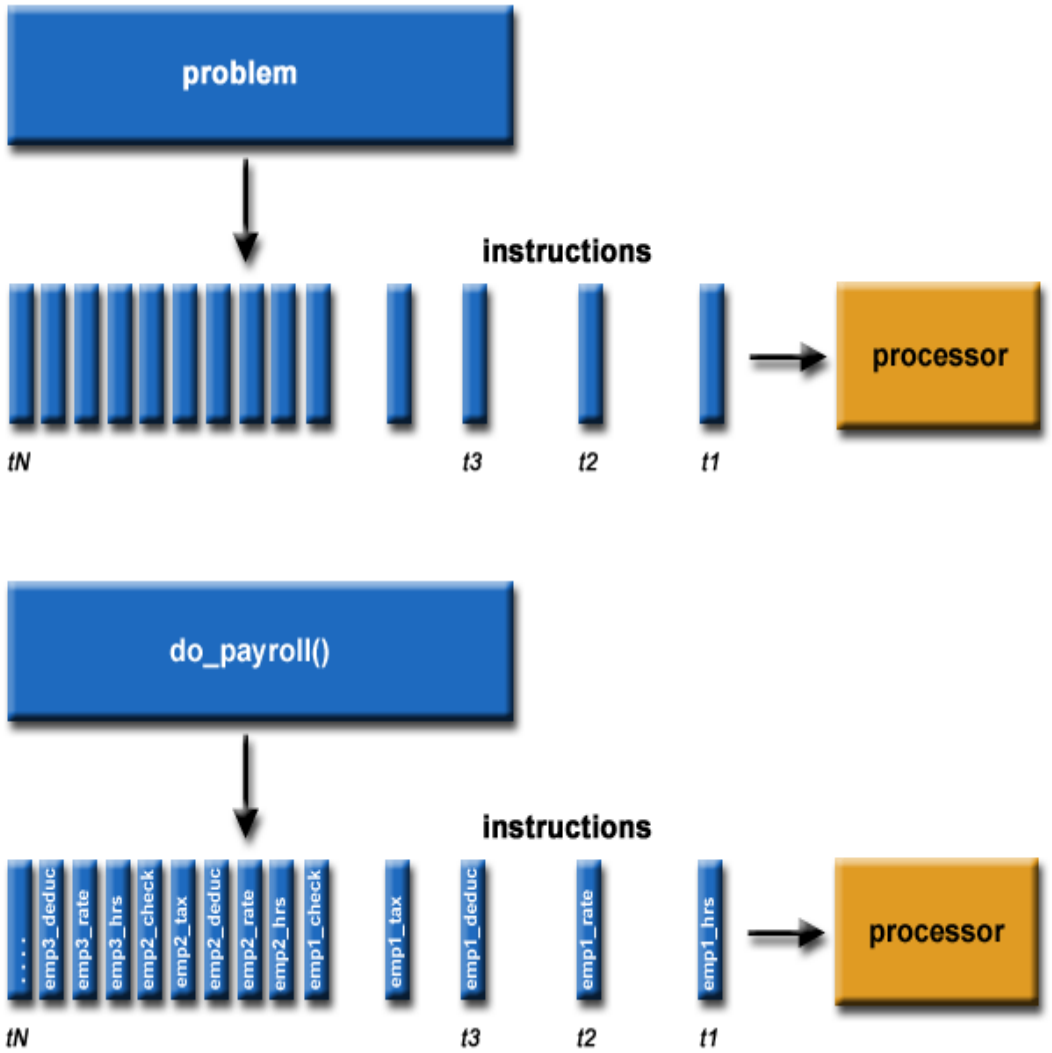
Let’s get clear with basic concepts

https://www.youtube.com/watch?v=RJNx90ODawk

Serial Computing

Traditionally, software has been written for ***serial*** computation.

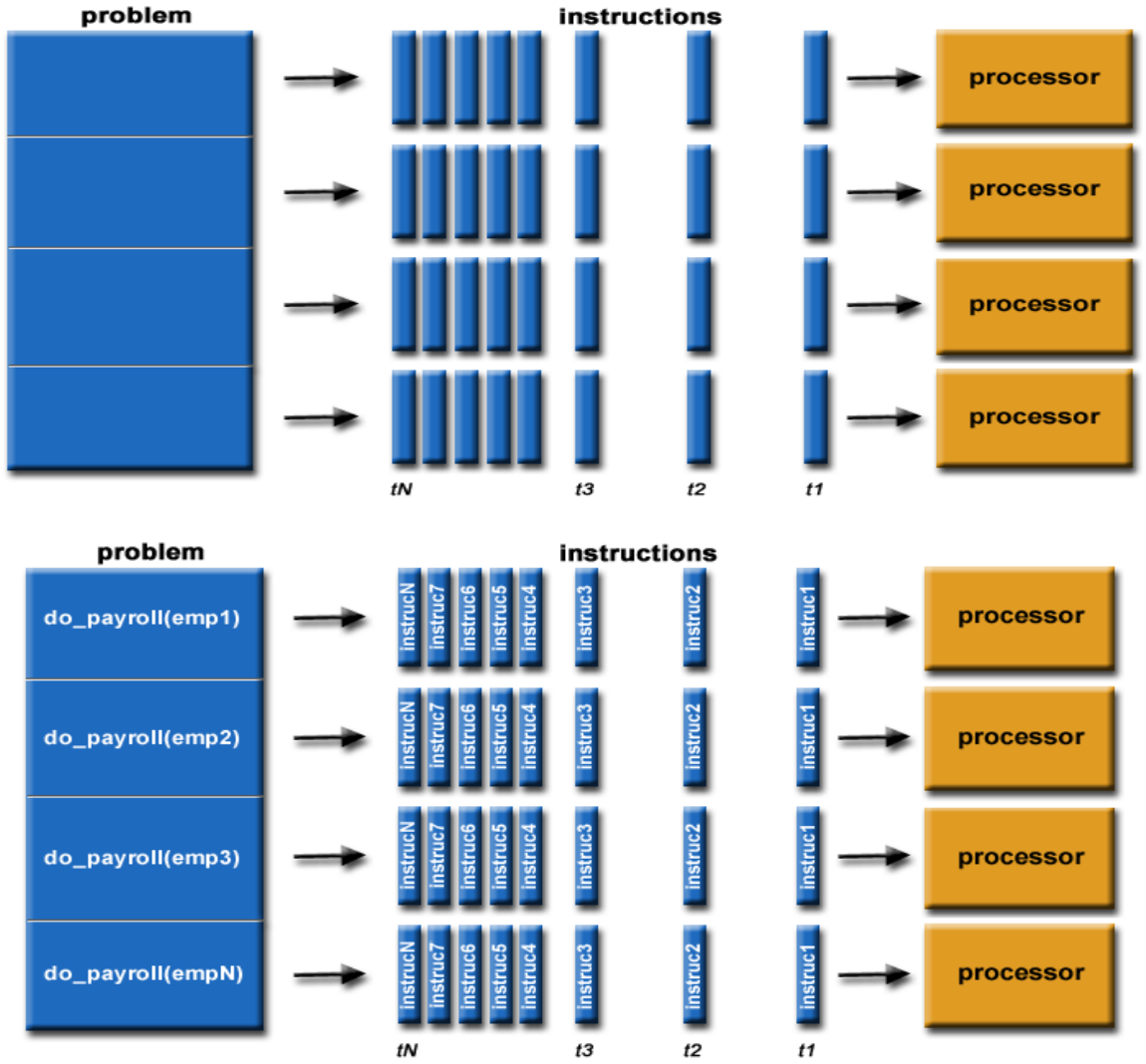
- A problem is broken into a discrete series of instructions
- Instructions are executed sequentially one after another
- Executed on a single processor
- Only one instruction may execute at any moment in time



Parallel Computing

Parallel computing is the simultaneous use of multiple compute resources to solve a computational problem.

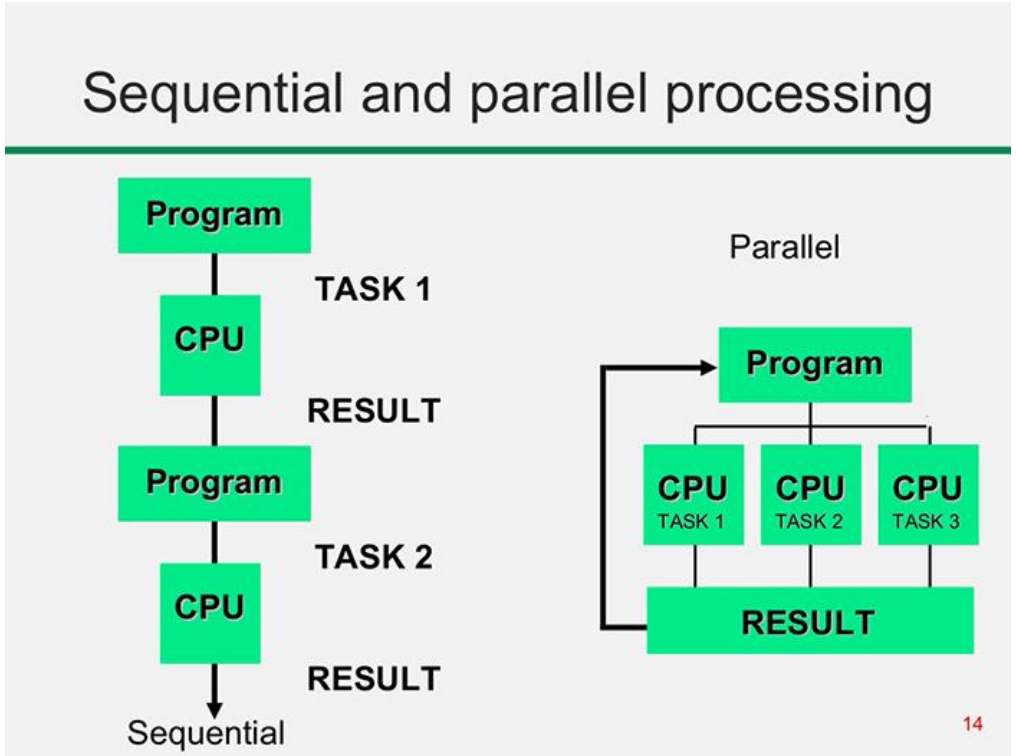
- A problem is broken into discrete parts that can be solved concurrently
- Each part is further broken down to a series of instructions
- Instructions from each part execute simultaneously on different processors
- An overall control/coordination mechanism is employed



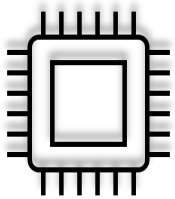
Difference between Serial and Parallel Processing

SERIAL PROCESSING VERSUS PARALLEL PROCESSING

SERIAL PROCESSING	PARALLEL PROCESSING
Type of processing in which one task is completed at a time and all the tasks are executed by the processor in a sequence	Type of processing in which multiple tasks are completed at a time by different processors
There is a single processor	There are multiple processors
Lower performance	Higher performance
Work load of the processor is higher	Work load per a processor is lower
Data transfers are in bit by bit format	Data transfers are in byte form (8 bits)
Requires more time to complete the task	Requires less time to complete the task
Cost is lower	Cost is higher



Fundamental limits on Serial Computing: Three "Walls"

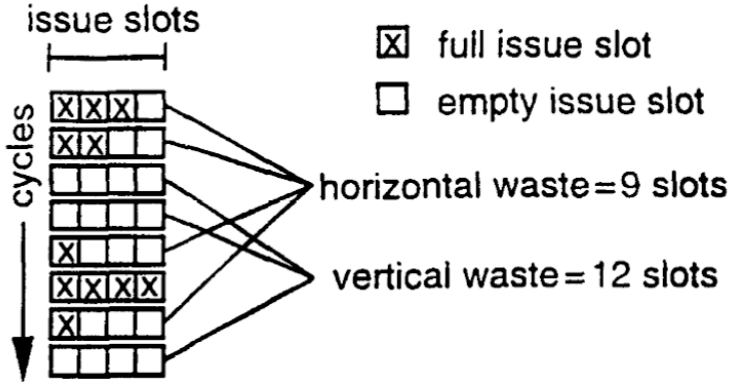


- **Power Wall (The Computational Power Argument – from Transistors to FLOPS):**

Increasingly, microprocessor performance is limited by achievable power dissipation rather than by the number of available integrated-circuit resources (transistors and wires).
Thus, the only way to significantly increase the performance of microprocessors is to improve power efficiency at about the same rate as the performance increase.

- **Frequency Wall(The Data Communication Argument):**

Conventional processors require increasingly deeper instruction pipelines to achieve higher operating frequencies.
This technique has reached a point of diminishing returns, and even negative returns if power is taken into account.



- **Memory Wall(The Memory/Disk Speed Argument):**

On multi-gigahertz symmetric processors --- even those with integrated memory controllers --- latency to DRAM memory is currently approaching 1,000 cycles.
As a result, program performance is dominated by the activity of moving data between main storage (the effective-address space that includes main memory) and the processor.

Why is Parallel Computing Important?

Imagine modeling these serially

Galaxy Formation | Planetary Movments | Climate Change

The Real World is Massively Parallel

- In the natural world, many complex, interrelated events are happening at the same time, yet within a temporal sequence.
- Compared to serial computing, parallel computing is much better suited for modeling, simulating and understanding complex, real world phenomena.

Example: Weather Prediction(one of the 1989 Grand Challenges to Computational Science Categories)

Atmosphere is divided into 3D cells ;Data includes temperature, pressure, humidity, wind speed and direction, etc

–Recorded at regular time intervals in each cell ;There are about 5×10^3 cells of 1 mile cubes.

Calculations would take a modern computer over 100 days to perform calculations needed for a 10 day forecast (Then)

Why is Parallel Computing Important?

SAVE TIME AND/OR MONEY:

- In theory, throwing more resources at a task will shorten its time to completion, with potential cost savings.
- Parallel computers can be built from cheap, commodity components.

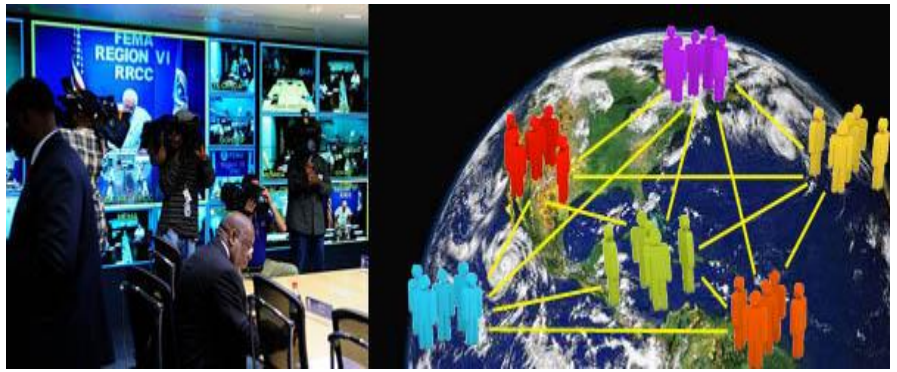
SOLVE LARGER / MORE COMPLEX PROBLEMS:

- Many problems are so large and/or complex that it is impractical or impossible to solve them using a serial program, especially given limited computer memory.
- Example:
  "Grand Challenge Problems" requiring petaflops and petabytes of computing resources.
  Web search engines/databases processing millions of transactions every second

Why is Parallel Computing Important?

PROVIDE CONCURRENCY:

- A single compute resource can only do one thing at a time. Multiple compute resources can do many things simultaneously.
- Example: Collaborative Networks provide a global venue where people from around the world can meet and conduct work "virtually".



TAKE ADVANTAGE OF NON-LOCAL RESOURCES:

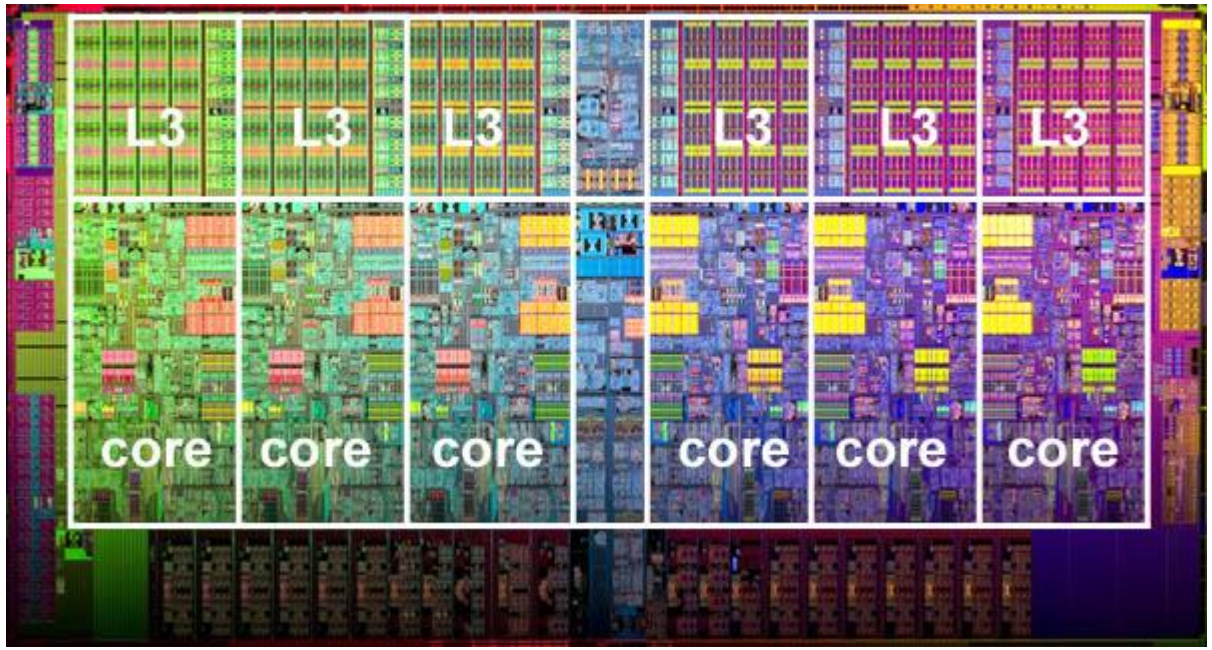
- Using compute resources on a wide area network, or even the Internet when local compute resources are scarce or insufficient.
- Example: SETI@home (setiathome.berkeley.edu) has over 1.7 million users in nearly every country in the world. (May, 2018).

SETI stands for the Search for Extra Terrestrial Intelligence and the project is dedicated to searching for patterns that may be signs of intelligent life amongst the mostly random mass of radio signals that reach the Earth from space. Each member of the project offers some of their computer's time to the cause.

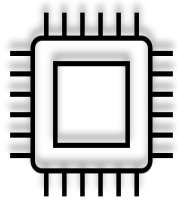
Why is Parallel Computing Important?

MAKE BETTER USE OF UNDERLYING PARALLEL HARDWARE:

- Modern computers, even laptops, are parallel in architecture with multiple processors/cores.
- Parallel software is specifically intended for parallel hardware with multiple cores, threads, etc.
- In most cases, serial programs run on modern computers "waste" potential computing power.



Who is Using Parallel Computing?(Scope of Parallel Computing)



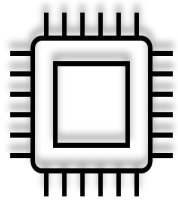
Applications in Engineering and Design:

- Design of airfoils (optimizing lift, drag, stability)
- Internal combustion engines (optimizing charge distribution, burn),
- High-speed circuits (layouts for delays and capacitive and inductive effects)
- Structures (optimizing structural integrity, design parameters, cost, etc.)
- Design of microelectromechanical and nanoelectromechanical systems (MEMS and NEMS)

Scientific Applications:

- Physics - applied, nuclear, particle, condensed matter, high pressure, fusion, photonics
- Bioscience, Biotechnology, Genetics
- Chemistry, Molecular Sciences
- Geology, Seismology
- Applications in astrophysics have explored the evolution of galaxies, thermonuclear processes,and the analysis of extremely large datasets from telescopes.
- Weather modeling, mineral prospecting, flood prediction, etc.,

Who is Using Parallel Computing?(Scope of Parallel Computing)



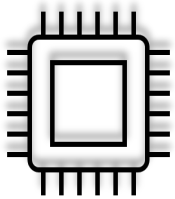
Commercial Applications:

- "Big Data", databases, data mining
- Web search engines, web based business services
- Financial and economic modeling
- Management of national and multinational corporations
- Advanced graphics and virtual reality, particularly in the entertainment industry
- Networked video and multimedia technologies
- Collaborative work environments

Applications in Computer Systems:

- network intrusion detection
- cryptography,
- Embedded systems
- Artificial Intelligence

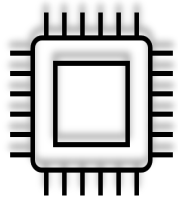
Why is Parallel Computing Hard?(Limitations)



- **Algorithm development is harder**
  - The algorithms must be managed in such a way that they can be handled in the parallel mechanism.
  - The algorithms or program must have low coupling and high cohesion. But it's difficult to create such programs.
  - complexity of specifying and coordinating concurrent activities

- **Software development is much harder**
  - lack of standardized & effective development tools, programming models, and environments
  - It addresses such as communication and synchronization between multiple sub-tasks and processes which is difficult to achieve.

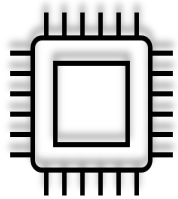
- **Rapid pace of change in computer system architecture**
  - today's hot parallel algorithm may not be suitable for tomorrow's parallel computer!
  - More technically skilled and expert programmers can code a parallelism based program well.



You have completed this topic, you should be able to:

Explain why is parallel processing important ?

Differentiate between Serial and parallel processing.



Food for thought!!!

Concurrency versus Parallelism:

Are they the same?

Which one is better and why?

What are their applications?

Watch these videos :

https://www.youtube.com/watch?v=oV9rvDllKEg

and https://www.youtube.com/watch?v=FChZP09Ba4E

References Used

- Ananth Grama, Anshul Gupta, George Karypis, Vipin Kumar , —Introduction to Parallel Computing,Pearson Education, Second Edition, 2007.

- https://hpc.llnl.gov/training/tutorials/introduction-parallel-computing-tutorial