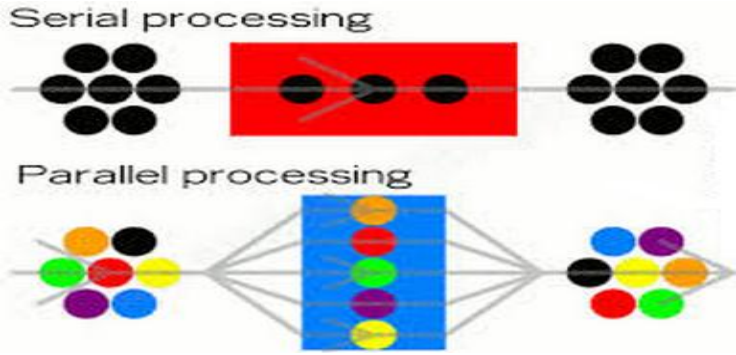
Parallel Processing

By

Fahad Layth Malallah

Parallel Processing Introduction



4th grade, Computer Science
Cihan University
2nd Semester, 2014-2015.

Reference Books:

1- F. B. Moreshwer, **Parallel Computing**, Addison-Wesley Publishing Company, 2008.

2- M. Morris Mano, **Computer System Architecture** ,3rd edition.

3-Hesham El-Rewini, Mostafa Abd-El-Barr, **Advanced Computer Architecture And Parallel Processing**, Weily, 2005.

4- David A. Patterson, John L. Hinnessy, **Computer Organization and design: the hardware / Software interface**. 3rd edition, Elsevier, 2005.

Syllabus:

1. **Introduction:**
   - 1.1- Basic Computer.
   - 1.2- Parallel Computing.
   - 1.3- Applications & Importance.
   - 1.4- Some General Parallel Terminology.
2. **Architecture Classifications:**
   - 2.1- Flynn's Classification.
   - 2.2- Shore's Classification.
   - 2.3- Feng's Classification.
   - 2.4- Handler's Classification.
3. **Parallel Computing Architecture:**
   - 3.1- Introduction.
   - 3.2- Parallelism types.
   - 3.3- Multiprocessors.
     - 3.2.1-Synchronous Multiprocessors(Array processors).
     - 3.2.1- Asynchronous (Conventional) Multiprocessors.

Syllabus:

3.4- Memory architectures:

3.4.1- Shared Memory (Tightly Coupled).

A- Uniform Memory Access (UMA).

B- Non-uniform Memory Access (NUMA).

3.4.2- Distributed Memory (Loosely Coupled).

A-Grid.

B-Cluster.

3.4.3- Hybride Shared and Distributed Memory.

4. Performance of Parallel Processing:

4.1- Speedup & Efficiency.

4.2- Amdahl's Law.

4.3- Minsky's Conjecture.

4.4- Gustafson's Law.

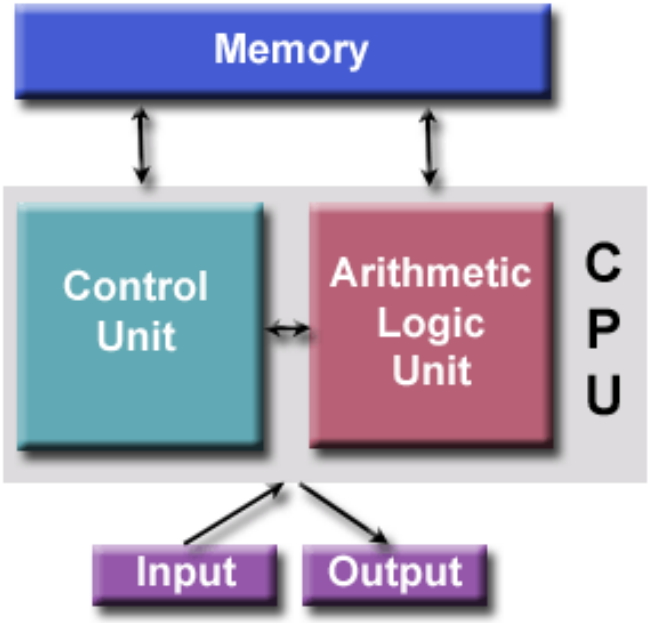
Syllabus:

5. **Pipeline Processing:**
   - 5.1- Pipeline Definition.
   - 5.2- Pipeline Performance in terms of Parallel Processing.
   - 5.3- Arithmetic Pipeline.
     - Floating Point Adder-Subtractor Architecture.
     - Floating Point Adder-Subtractor Performance.
   - 5.4- Instruction Pipeline.
     - RISC Instruction Architecture.
     - RICS Instruction Performance.
6. **Interconnection Network and Classification.**
   - 6.1- Direct Connection Networks
   - 6.2-Indirect Connection Networks .
     - 6.2.1- Busses .
     - 6.2.2- Multistage networks.
     - 6.2.3- Crossbar switches.
7. **Load Balancing.**

❑1.1 Basic Computer

Von Neumann Architecture

•First computer architecture model was invented by Von Neumann, Hungarian mathematician John von Neumann.
•A von Neumann computer uses the stored-program concept.
• CPU executes a stored program.



John von Neumann circa 1940s (Source: LANL archives)

❑1.1 Basic Computer

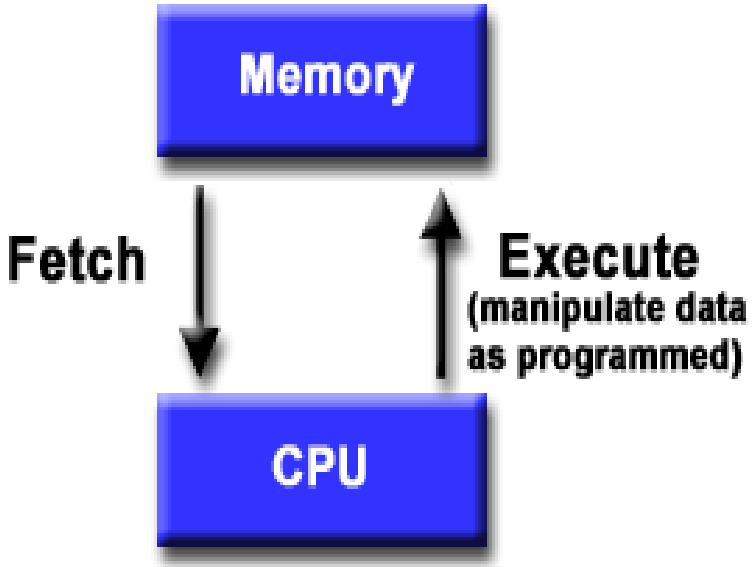
•Basic design: memory is used to store both: program and data instructions.

• Program instructions are coded data which tell the computer to do something (ex: add+, subtract-, compare ><, etc).

•Program→ many tasks→ many Instructions.

•Data is simply information to be used by the program.

•A central processing unit (CPU) gets instructions and/or data from memory, decodes the instructions and then ***sequentially*** performs them.



❑1.1 Basic Computer

- Uniprocessors must stop getting faster due to limit of high speed achieved, 2.5GHz, 3.3 GHz.
- Increasing more speed, results more power consuming and higher temperature of the device.
- Need faster speed and better performance. What do we do?.

Parallel Processing is the solution.

Single Processor computer architecture work details with below link:
https://www.youtube.com/watch?v=cNN_tTXABUA

❑1.2- Parallel Computing.

- Definition: It is a collection of processing elements that cooperate and communicate to solve large problems fast.
- or it is doing more than one thing at a time.
- Example: multiprocessors (multicore) computer.

❑1.2- Parallel Computing.

- Questions about parallel computers:
  1- How large a collection?
  2- How powerful are processing elements?
  3- How do they cooperate and communicate?
  4- How are data transmitted?
  5- What type of interconnection?

❑1.2- Parallel Computing.

Goals of parallel computing are:

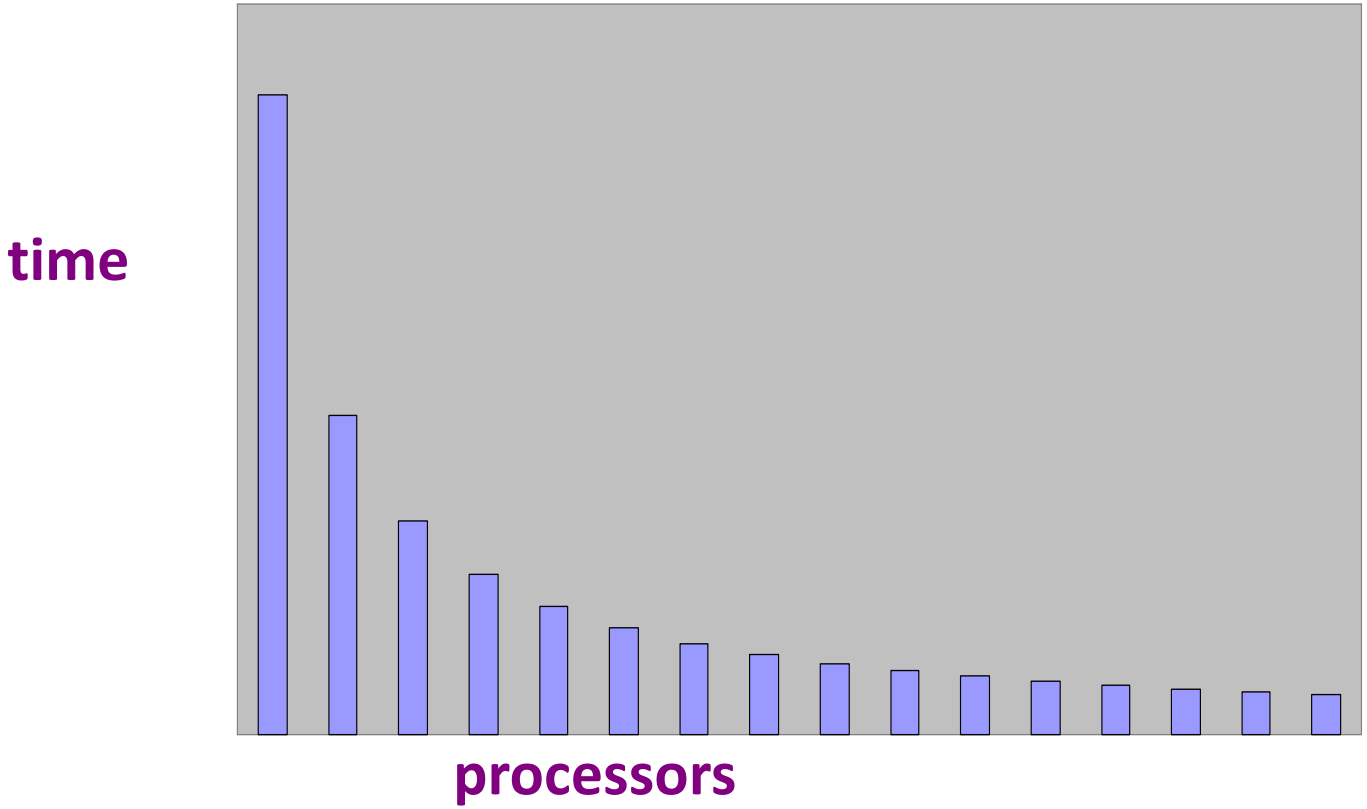
1. Improve performance: Execution time or task throughput

1. Reduce power consumption
   (4N units at freq F/4) consume less power than (N units at freq F).

3. Improve cost efficiency and reduce complexity

4. Improve dependability: Redundant execution in space

❑1.2- Parallel Computing.



❑1.2- Parallel Computing.

- **Level of Parallelism:**

1- Bit level parallelism: 1970 to ~1985

- 4 bits, 8 bit, 16 bit, 32 bit microprocessors
- 2004→ 64-bit.

2- Instruction level parallelism (ILP):
~1985 through today

- Pipelining

3- Data level parallelism:

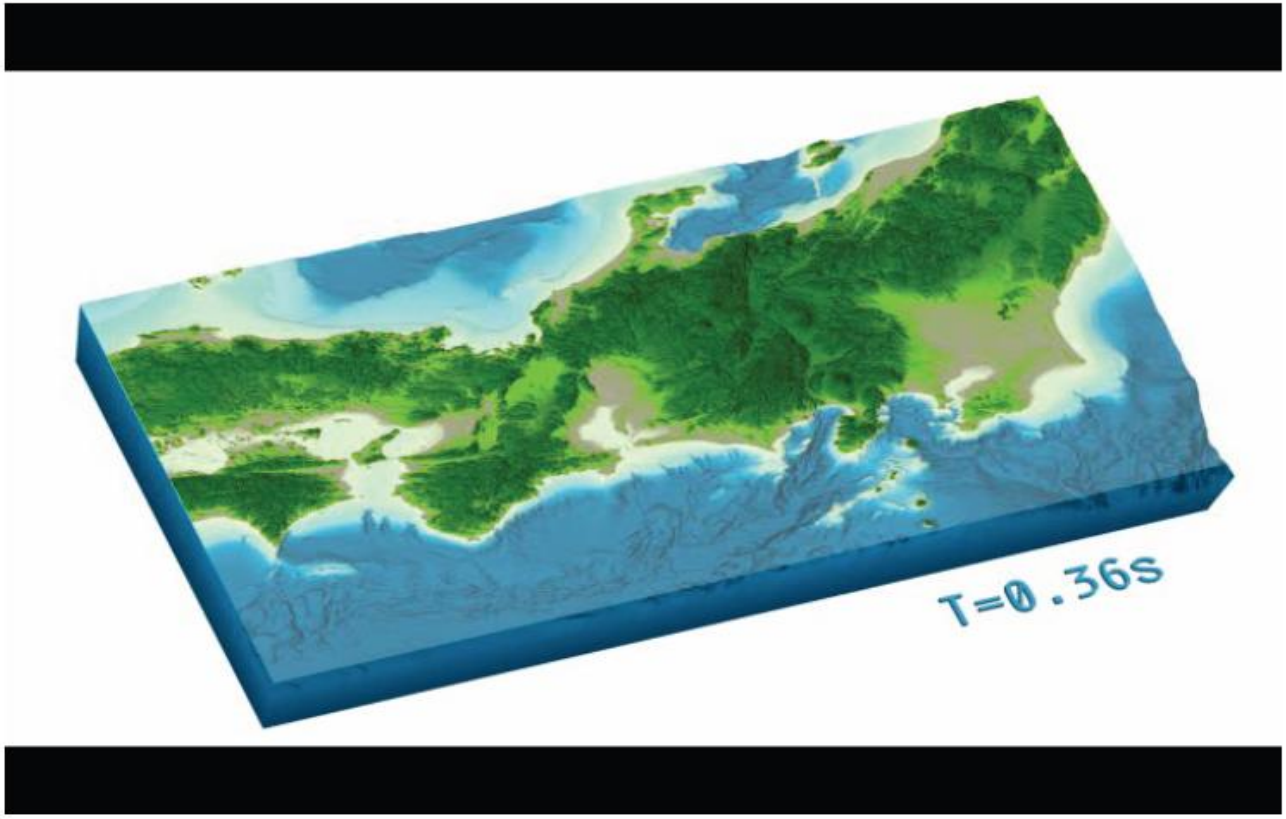
- Different piece of data can be operated on in parallel.
- Systolic arrays.

4- Process Level or Thread level parallelism:

- Desktop dual processor (PC).

❑1.3- Applications & Importance

Earthquake Simulation



Earthquake Research Institute, University of Tokyo
Tonankai-Tokai Earthquake Scenario
Photo Credit: The Earth Simulator Art Gallery, CD-ROM, March 2004

❑1.3- Applications & Importance

Galaxy Formation

Planetary Movments

Climate Change

Rush Hour Traffic

Plate Tectonics

Weather

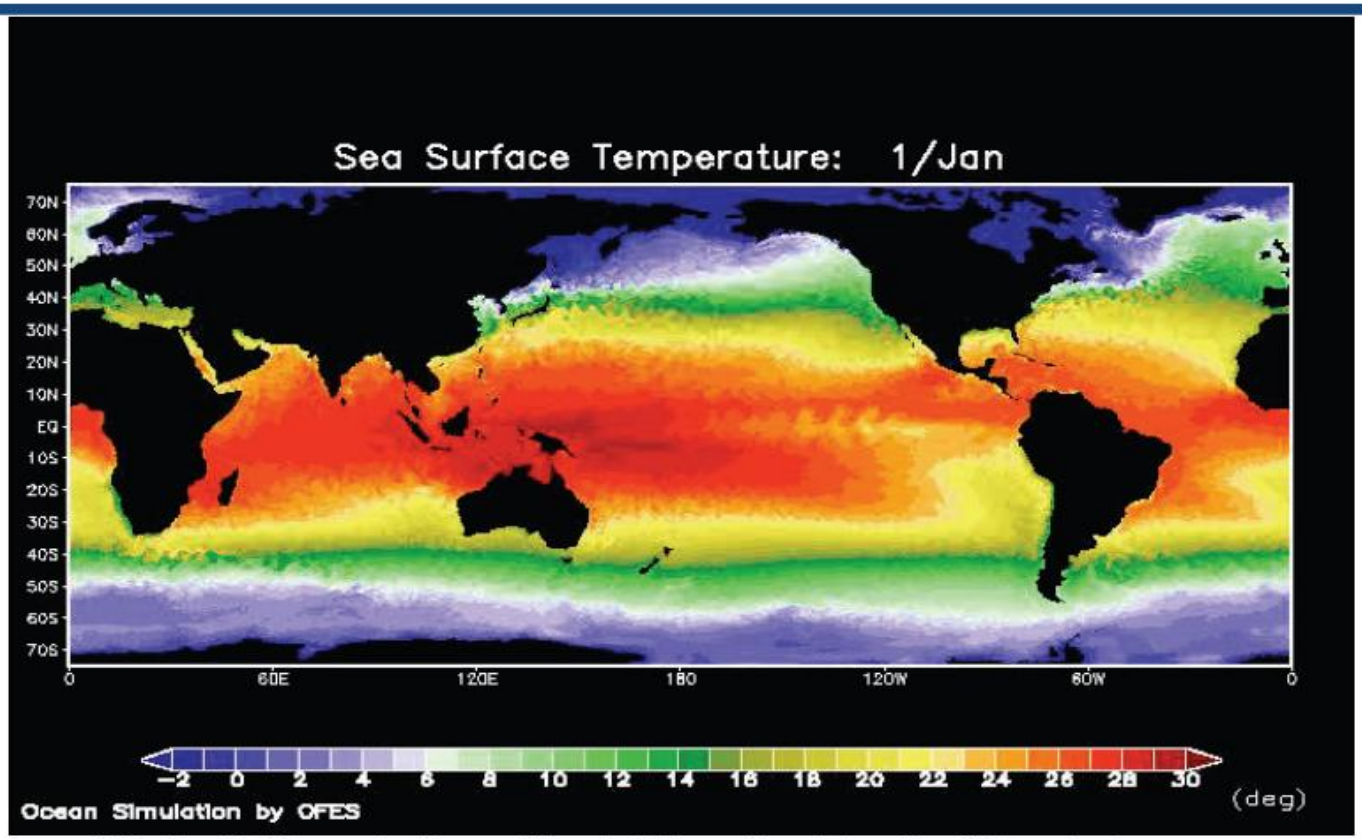
Auto Assembly

Jet Construction

Drive-thru Lunch

❑1.3- Applications & Importance

Ocean Circulation Simulation



Ocean Global Circulation Model for the Earth Simulator
Seasonal Variation of Ocean Temperature
Photo Credit: The Earth Simulator Art Gallery, CD-ROM, March 2004

❑1.3- Applications & Importance

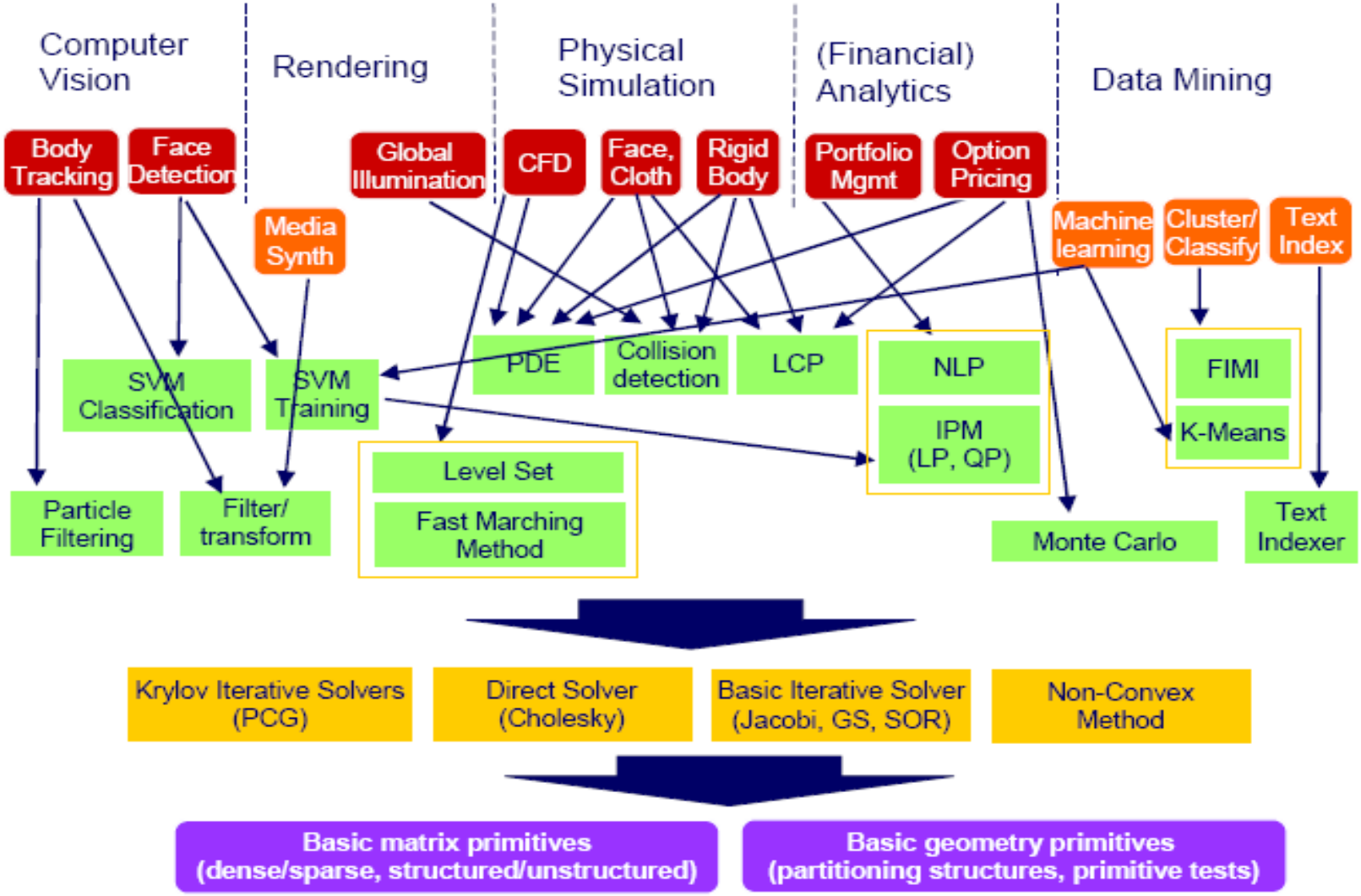
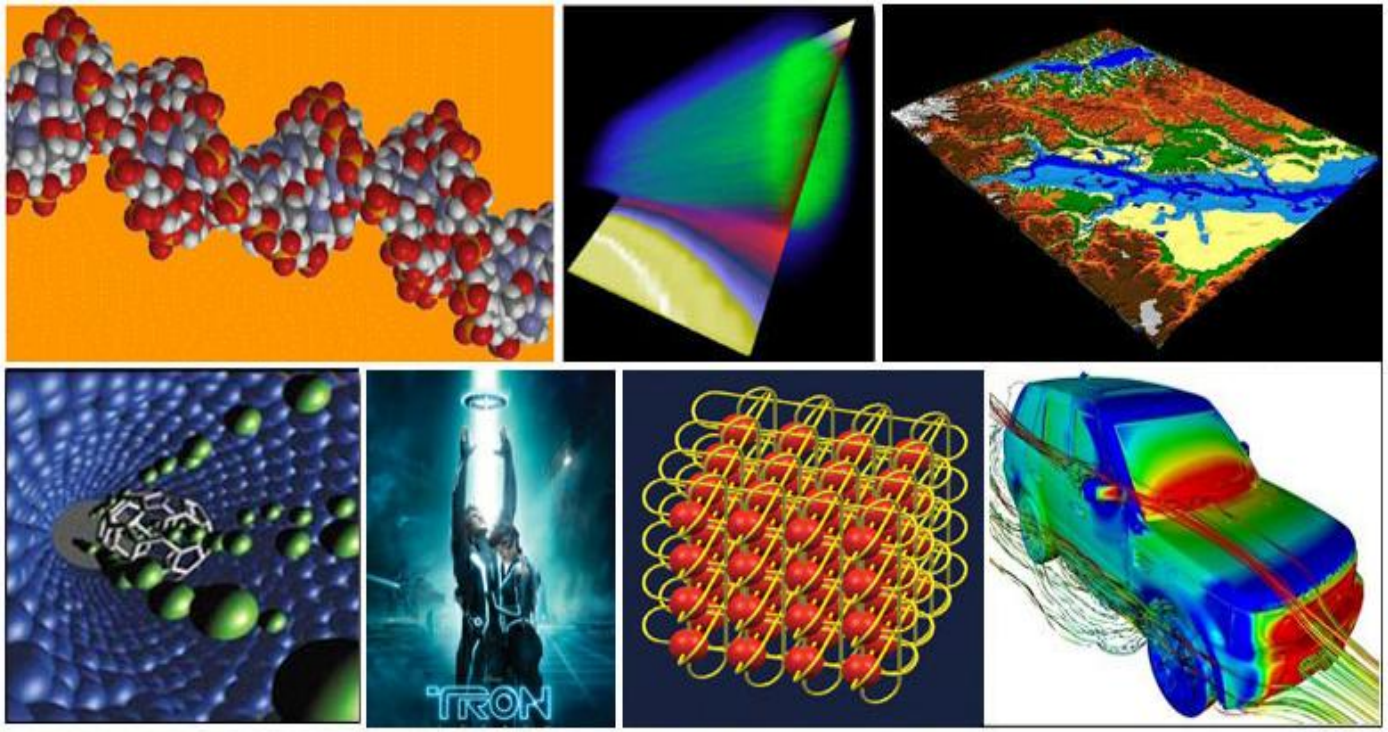


Figure 5. Intel's RMS and how it maps down to functions that are more primitive. Of the five categories at the top of the figure, Computer Vision is classified as Recognition, Data Mining is Mining, and Rendering, Physical Simulation, and Financial Analytics are Synthesis. [Chen 2006]

❑1.3- Applications & Importance

•**Who is using Parallel computing:** 1- Industrial and communication.
2- Science.

•Databases, data mining
•Oil exploration
•Web search engines, web based business services
•Medical imaging and diagnosis.

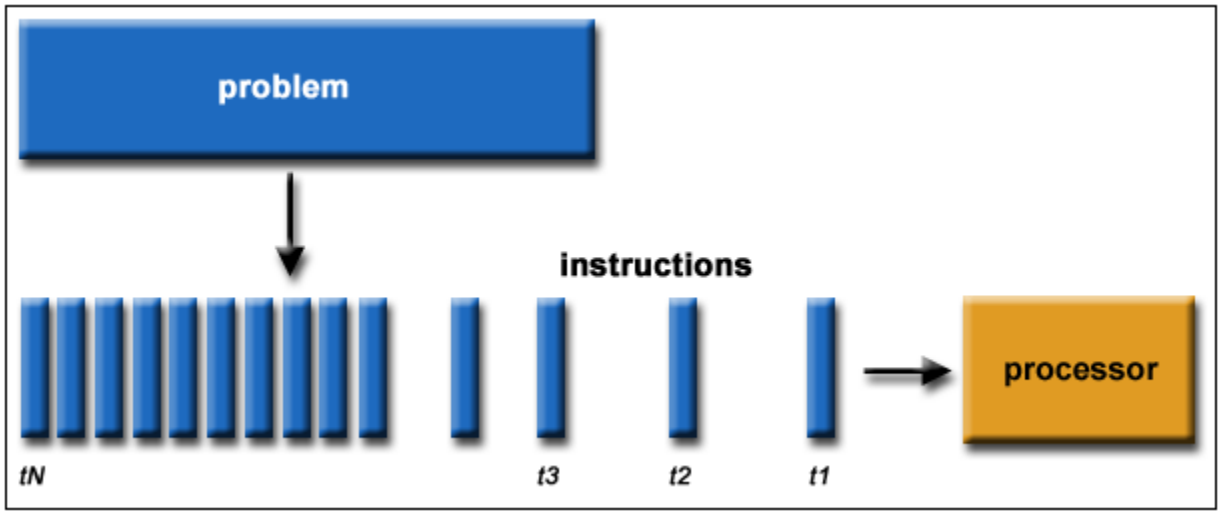


❑1.4- Some General Parallel Terminology

•**Task** : A logically discrete section of computational work. A task is typically has set of instructions that is executed by a processor.

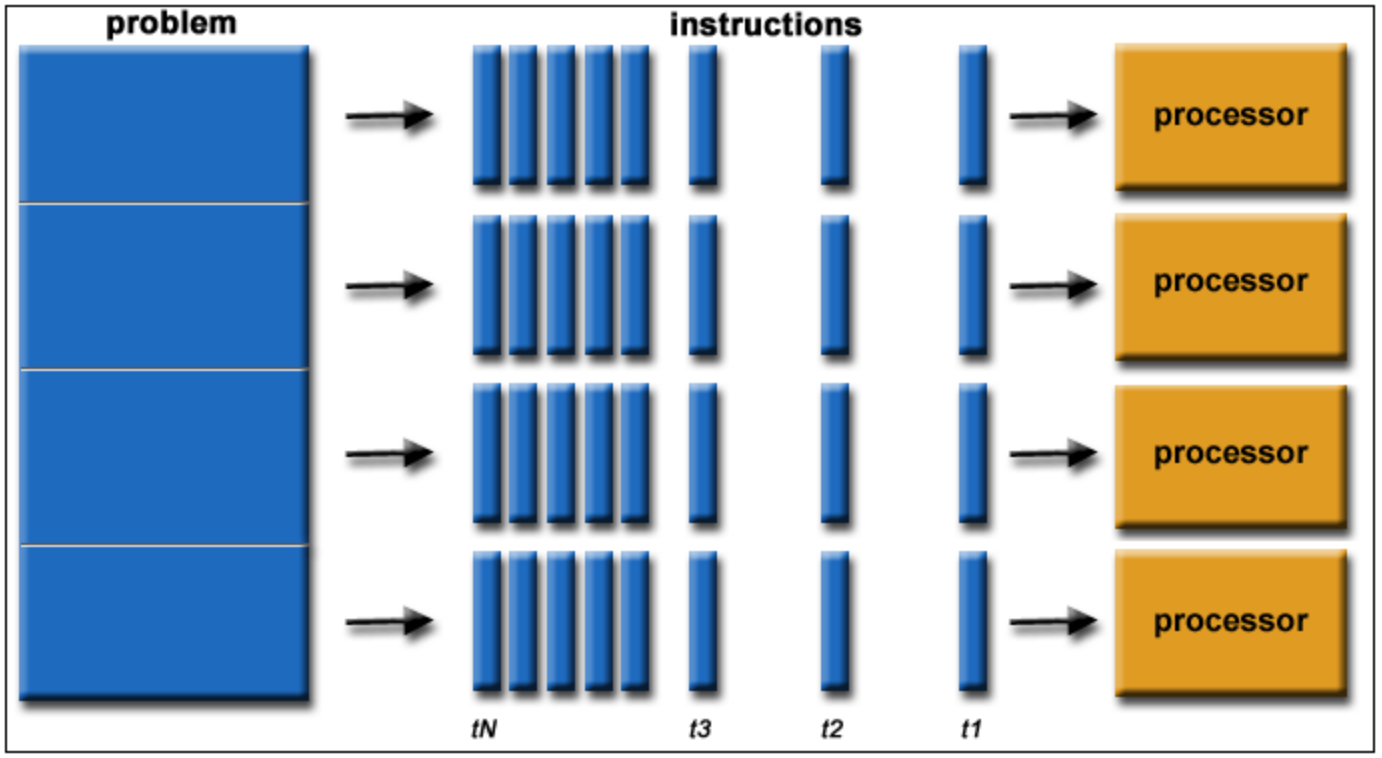
•**Parallel Task**: A task that can be executed by multiple processors safely (yields correct results).

•**Serial Execution**: Execution of a task sequentially, one statement at a time. In the simplest sense, this is what happens on a one processor machine.



❑1.4- Some General Parallel Terminology

•**Parallel Execution:**
Execution of a program by more than one task, with each task being able to execute the same or different statement at the same moment in time.



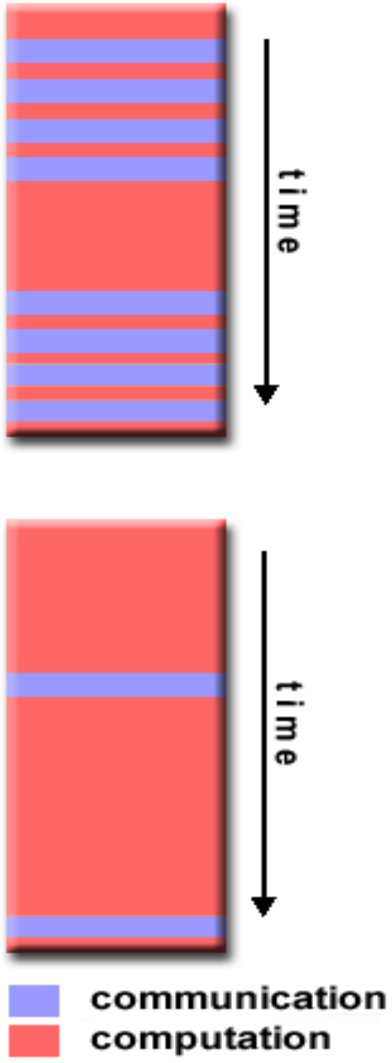
❑1.4- Some General Parallel Terminology

•**Granularity**

In parallel computing, granularity is a qualitative measure of the ratio of computation to communication. Types:

1- Coarse: relatively large amounts of computational work are done between communication events.

2- Fine: relatively small amounts of computational work are done between communication events.



❑Exercices

1. What is granularity and what are its types? Explain with drawing.
2. Define the following: Task, Instruction, parallel processing.
3. What are the main purpose of the parallel computers?.
4. Processor execution are two types? What are they?, and what is the difference between them illustrated the answer with diagrams.
5. List 10 applications require parallel computer.
6. What are the level of parallelism.
7. Who is Von Neumann?, what did he invent?, draw it?. Is it still using now?, Why?.