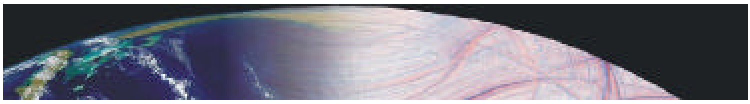
SPICE Research and Training workshop
Venice September 25– October 2, 2004



# Parallel Architectures and Parallel Programming Paradigms: an overview

Giovanni Erbacci
High Performance Systems Department
CINECA



# Agenda

- Introduction to Computational Sciences
- Evolution of HPC
- Overview of parallel architectures
- Parallel Programming Paradigms

- HPC -Europa

# Modern Science

Continuous Interaction between

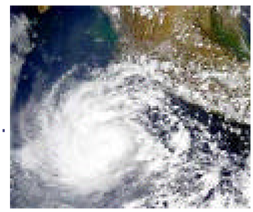
Experiments and Theory

Models <=> Simulation

Computers play a fundamental role

Computer numerical simulation allows *ad libitum* extension of time and space scale

- Macroscopic scale: astrophysics...
- Microscopic scale: biology, nuclear physics...





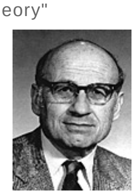


# What is Computational Science?

- The science of using computers for problem solving
- Diverse problems such as protein synthesis, lattice structures, weather modeling, geophysics, etc.
- Extend traditional definition to include problems such as stock market analysis, economics, psycology, cultural heritages applications, virtual archaeology

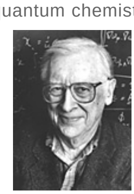
**The Nobel Prize in Chemistry 1998**

"for his development of the density-functional theory"



**Walter Kohn**

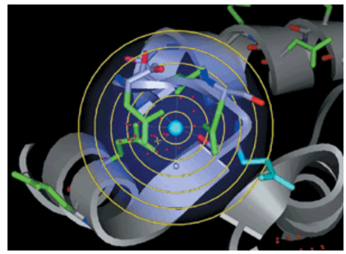
"for his development of computational methods in quantum chemistry"



**John A. Pople**

## How is different from Computer Science?

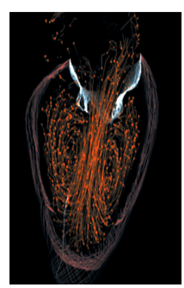
- The focus in computer science is the computer itself. The issues dealt with are related to speeding up the computers;
  - Making them easier to use
  - Finding fast algorithms for core problems (sorting, searching, etc.)
- In computational science the focus is on using computers for solving scientific problems.





## What makes Computational Science Difficult?

- Interdisciplinary nature: people from computer science, mathematics, pure sciences, and other have to work together
- The core techniques in problem solving draw heavily on mathematical sciences
- Many problems in computational science require the use of High Performance Computing
- Grand Challenge problems represent a subset of problems in scientific computing
- The diverse background of the community makes it difficult to disseminate results



# What are Grand Challenge Problems?

"...fundamental problems in science or engineering with potentially broad social, political, and/or scientific impact, that could be advanced by applying high performance computing resources"

Identified thrust areas for grand challenge problems:

- Electonic structure of materials
- Turbolence
- Genome sequencing and structural biology
- Global wether and climate
- Speech and language
- Drug design
- Vision and cognition
- Fundamental properties of matter (QCD)
- Pollution dispersion
- Earth Sciences.

**Kenneth. Wilson**

Nobel Prize in Physics 1982

*"for his theory for critical phenomena in connection with phase transitions"*



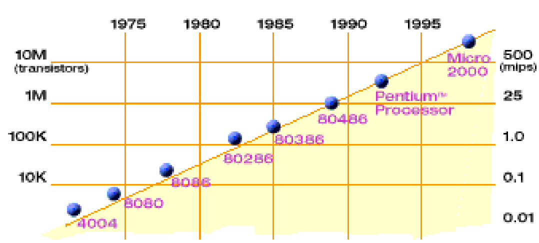
# Powerful Computers

- Increase processor speed
  - Performances affected also by other factors (bandwidth, latency,...)

Moore Law

The number of transistors per sqare inch in microprocessors doubles every 18 months.

Gordon Moore, co-founder of Intel, 1965

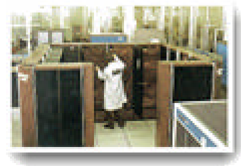



- Increase the degree of parallelism

CINECA Consorzio Interuniversitario

# HPC in ‘80

Scalar systems
Vector systems
Vector-augmented mainframes
Shared memory parallel systems



Vector Compilers:

Focus on loop optimisation

MegaFlop/s era



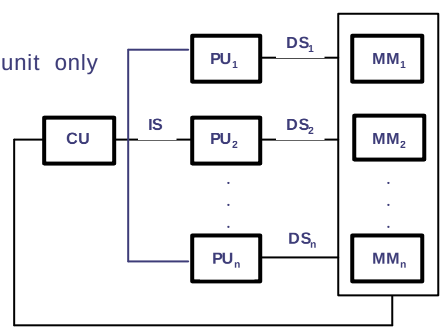
CINECA Consorzio Interuniversitario

# Syncronous Parallelism

SIMD systems have a single control unit only

- A single instruction simulateously operates on multiple data

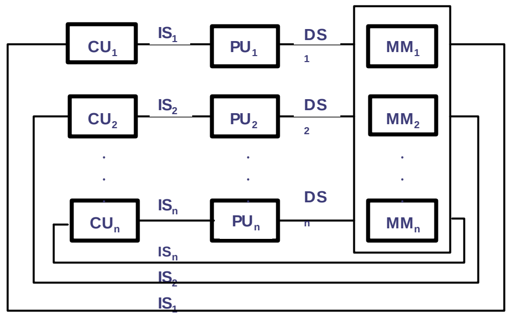
Array processor
Vector systems




**CU** Control Unit
**PU** Processing Unit
**MM** Memory Module
**DS** Data stream
**IS** Instruction Stream

# Asyncronous Parallelism

- Different processors interpretes multiple instructions operating on different data
- Multiprocessor version of the SIMD class
- From *linked main frame* computer to big micro-processor cluster

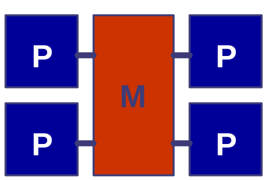



**CU** Control Unit
**PU** Processing Unit
**MM** Memory Module
**DS** Data stream
**IS** Instruction Stream

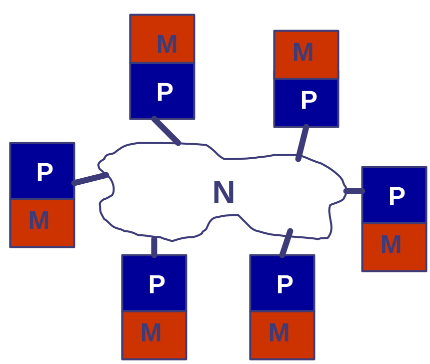




# Memory Classification




**Shared Memory System**




**Distributed Memory System**

CINECA Consorzio Interuniversitario

# Evolution of Hardware

- Parallel computers offer magnitudes higher performance enabling solution of hitherto unrealizable problems
- The microprocessor revolution puts tremendous performance on the desktop. In addition to problem solving, this facilitates visualisation and representation of complex data
- Advances in network technology enables the use of clusters of workstation as parallel platforms
- Advances in symmetric multiprocessing put the power of expensive vector processors in cheap high end clusters of SMP



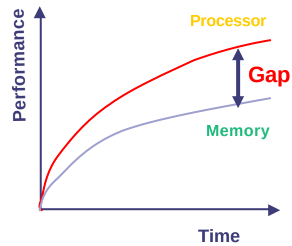
CINECA Consorzio Interuniversitario

# Evolution of Hardware (the flip side)

- Memory speeds have not kept up with processor speeds (deep cache hierarchies). This causes a large gap in peak and realizable performance of conventional computers : hand-tuning necessary even for uniprocessors
- Parallel computers are still hard to program (push for standardisation)
- Complex architectures require more complicated algorithms and programs
- Solution : identify key computational kernels , use handcoded libraries fo these
- Removes the burden of tuning programs from application engineers

# HPC today

- MPP Systems
- ccNUMA
- Clusters of SMP
- Beowulfs
- Grid Computing



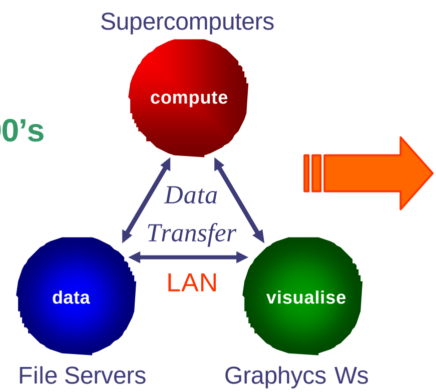

**Interconnection Network**
**Bandwidth**
**Latency**





# High Performance Computing: 90's

**90's**

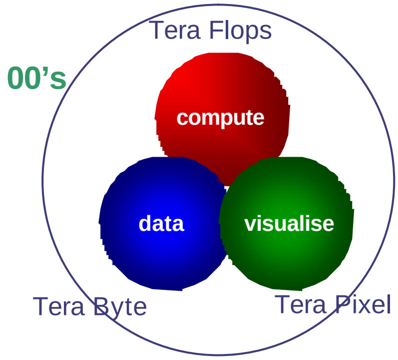



HPC is rapidly evolving

Dramatic change, driven by:

- the invasion of commodity components
- Internet Technology

## High Performance Computing: today




- Big Power
- Big Data
- Big Insight

Training
Co-operation
Maintenance

More integration:
Logically integrated not necessarily physically integrated:
- Grid Computing

It is impossible to maintain competitiveness without continuous innovative efforts

Focus on Applications



## Top 500

http://www.top500.org/

The TOP500 project was started in 1993 to provide a reliable basis for tracking and detecting trends in high-performance computing. Twice a year, a list of the sites operating the 500 most powerful computer systems is assembled and released. The best performance on the **Linpack** benchmark is used as performance measure for ranking the computer systems. The list contains a variety of information including the system specifications and its major application areas.

The LINPACK Benchmark was introduced by Jack Dongarra.
The LINPACK Benchmark is to solve a dense system of linear equations.
For the TOP500, is used that version of the benchmark that allows the user to scale the size of the problem and to optimize the software in order to achieve the best performance for a given machine.
This performance does not reflect the *overall performance* of a given system, as no single number ever can. It does, however, reflect the *performance of a dedicated system for solving a dense system of linear equations.*
Since the problem is very regular, the performance achieved is quite high, and the performance numbers give a good correction of peak performance.

# Top 500: number 1 and 2 (June '04)

**EARTH SIMULATOR**, Yokohama, Japan

Rmax: 35.86 TFlops

- **5,120 (640 8-way nodes) 500 MHz NEC CPUs**
- **8 GFLOPS per CPU (41 TFLOPS total)**
- **2 GB (4 512 MB FPLRAM modules) per CPU (10 TB total)**
- **shared memory inside the node**
- **640 × 640 crossbar switch between the nodes**
- **16 GB/s inter-node bandwidth**
- **20 kVA power consumption per node**

**Intel Itanium 2 Tiger4** , Lawrence Livermore Nat. Lab.

Rmax:19.940 TFlops

| | |
|---|---|
| Nodes | 1024 |
| Processors per node (Itanium2) | 4 |
| Total number of processors | 4096 |
| Processor speed (GHz) | 1.4 |
| Theoretical system peak performance (TFlop/s) | 22.938 |
| Memory/node (GB) | 8 |
| Total memory (TB) | 8.2 |
| Type of memory | SDRAM |
| Total disk space | 75 GB (local) 151 TB (global) |



# CINECA


To advance the computational infrastructure at the service of science

**CINECA operates the most advanced high-performance computing center in Italy**

Consortium of 24 Universities and National Research Council

Founded in 1969 as a computing center open to public and private research activities

# IBM SP4

IBM SP Power4 (SP4)
2.7 TFlops peak performance

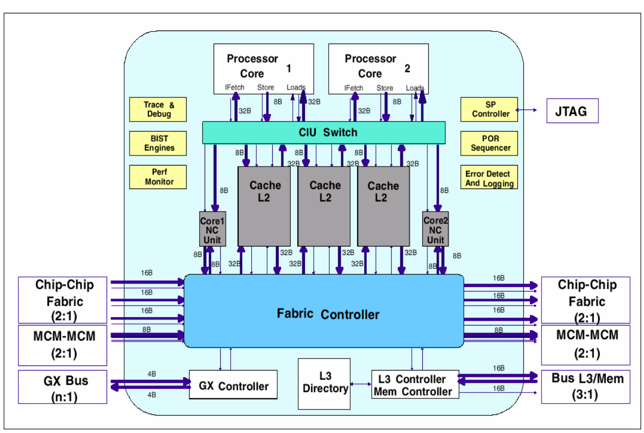
- 512 CPUs Power4 1.3GHz
- 32 CPUs per node (16 nodes)
- 64 GB (128) main memory per node
- 3 cache levels
- GPFS, global filesystem
- Interconnection network: switch2

Colony → Federation

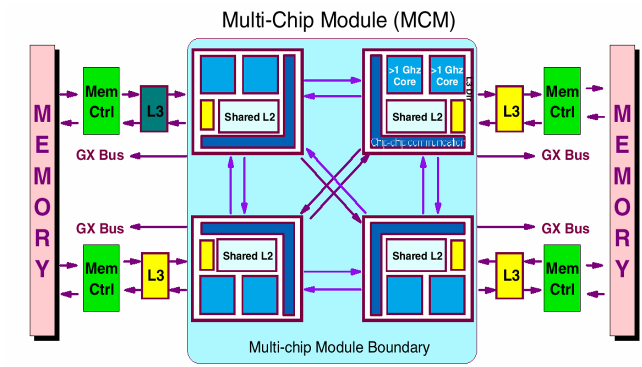




# Chip Power 4



# Multi Chip Module








# Cluster Linux

| | |
|---|---|
| Compute Node | IBM linux Cluster 1350 |
| Processor (PE) | Intel Xeon Pentium IV 3.055 GHz |
| Number of PEs | 512, 256 nodes with 2 proc. |
| L2 Cache | 512KB per processor |
| DRAM | 512 Gbytes (2GB/node) |
| Disk space | 5.5 TB |
| Peak performance | 3.1 Tflop/s |

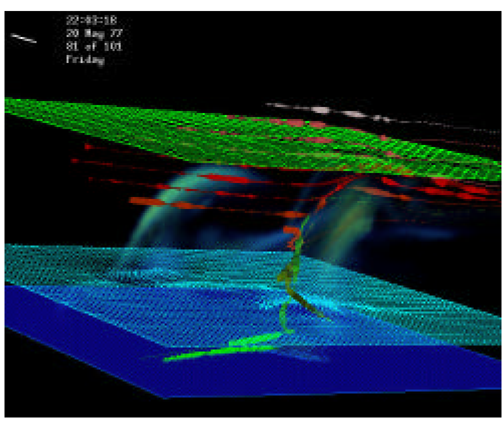
Switch Myrinet IPC network

# The Virtual Theatre





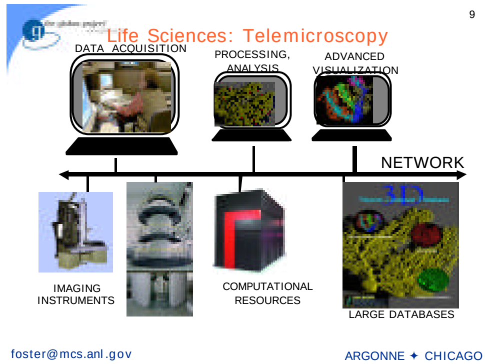
# Meteorological data



# Toward Grid Computing

## Grid computing

Networked infrastructure for advanced research that can seamlessly connect distributed teams to the high-end computing systems, instruments, advanced simulation and visualization software, and sensor networks they need to work collaboratively on data- and computation-intensive problems.



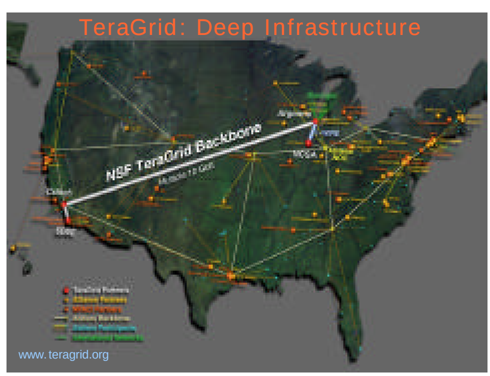





# Toward Grid Computing / 1

## TeraGrid

- multi-year effort to build and deploy the world's largest, fastest, distributed infrastructure for open scientific research.
- 20 Tflops of computing power distributed at five sites
- facilities capable of managing and storing nearly 1 Pbyte of data
- high-resolution visualization environments
- toolkits for grid computing
- tightly integrated and connected through a 40 Gbits/s network



# Paallel Programming Models

Parallel Programming Paradigms:

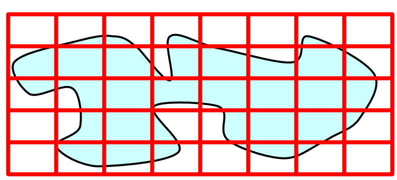
- shared Memory:
  - C, C++, Fortran90 & directives (OpenMP)
- distributed Memory:
  - C, C++, Fortran90 & libraries (MPI)

- Mixed Mode Parallelism

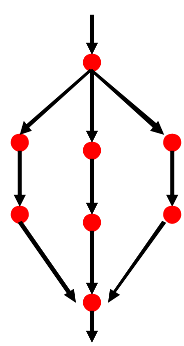


# Parallel Methods

Data Parallelism
- Data distribution
- Neighbour communication (ghost cells)



Control Parallelism (Functional Parallelism)
- distribute the different functions
- load balancing problems



## Functional or Data Parallelism?

Partition by task (functional parallelism)

- o each process performs a different "function"
- o identify functions, then data requirements
- o commonly programmed with message-passing

Partition by data (data parallelism)

- o each process does the same work on a unique piece of data "owner computes"
- o data placement is critical
- o more scalable than functional parallelism
- o programmed with HPF or message-passing





## Evolution of Software

- Problem solving environments: Integrated visualisation and problem solving
- A PSE is an integrated hardware-software system providing all the computational faiclities required for solving a target class of applications
- Availability of standardized libraries for a variety of problems (BLAS, LAPACK, LINPACK, ...)
- Greatly increased software quality (integrated libraries and standardised interfaces allow the programmers to write code at a much higher level)

## Evolution of Algorithms

- Faster machines pose additional challenges for existing algorithms (finer finite element meshes, impact of finite machine precision on convergence, etc...)
- New problems drive development of newer algorithms (hierarchical methods for n-body methods, linear system solver such as GMRES, Fast Fourier Transforms...)
- Focus on:
  - error analysis
  - exact solutios and expansions
  - uniqueness proofs
  - theorems



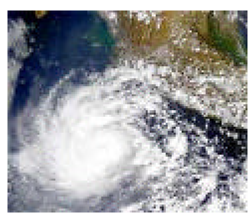
## Focus on Applications

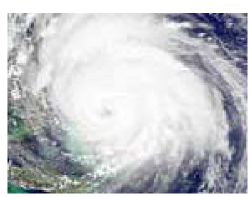
- Scientific problems
- Industrial problems
- Business problems

- Astrophysics
- Meteo-climatology
- Biology
- . . . .
- Earth Sciences
- Design and synthesis of new materials
- Oil industry
- Management and mining of vast collections of data

**Think Big**

**Think Parallel!!!**





# An example: ....SPICE

## Theory and applications of acoustic (elastic, seismic) wave propagation

- entering a new era in fields such as seismology, oceanography, meteorology, acoustics, engineering, material sciences, medical sciences
- computational techniques in combination with parallel computer architectures allow the simulation of the complete three-dimensional phenomena of wave propagation for realistic complex structures with unprecedented detail

The reverse processes will experience a big improvement in resolution and accuracy over the next decade:

- imaging of the Earth's internal structure,
- physical description of hydrocarbon reservoirs,
- monitoring of zones of weakness in constructions,
- characterization of earthquake rupture processes, etc

determination of global Earth structure;
quantitative estimation of shaking hazard;
characterization and monitoring of reservoirs;
understanding the structure and processes inside volcanoes;
simulating the physical processes of earthquake rupture;
characterizing the small-scale properties of rocks.

....Pflop/s era





# HPC-Europa

Pan-European Research Infrastructure on High Performance Computing

It is a consortium

- six leading HPC infrastructures
- five centres of excellence

Goals

- transnational access to HPC systems
- provision of a suitable computational environment
- to allow the European researchers to remain competitive with teams elsewhere in the world.
- the deployment and operation of a "virtual global infrastructure" for HPC

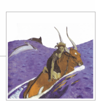
## HPC-Europa

**CINECA, Italy**
Consorzio Interuniversitario di Calcolo, Bologna
**EPCC, United Kingdom**
Edinburgh Parallel Computing Centre
**CEPBA, Spain**
Centro Europeo de Paralelismo de Barcelona
**HLRS, Germany**
High Performance Computing Centre, Stuttgart
**IDRIS, France**
Institut du Développement et des Ressources en Informatique Scientifique, Paris
**SARA, Netherlands**
Stichting Academisch Rekencentrum Amsterdam
**PSNC, Poland**
Poznan Supercomputing and Networking Centre
**PARALLAB, Norway**
HPC Laboratory, Bergen
**TCD, Ireland**
Trinity Centre for HPC, Dublin
**CASPUR, Italy**
Consorzio Interuniversitario, Roma
**NTUA, Greece**
National Technical University of Athens



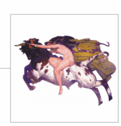

## HPC-Europa
Pan-European Research Infrastructure on High Performance Computing

NETWORKING ACTIVITIES



Joint Research and Networking actions
- to foster a culture of cooperation to generate critical mass for computational activities,
- to drive new advances in HPC allowing so to better structure the European Research Area.
- to provide new tools for better handling HPC resources

RESEARCH ACTIVITIES



CINECA Consorzio Interuniversitario

HPC-Europa

Pan-European Research Infrastructure on High Performance Computing

TRANSNATIONAL ACCESS

The core project is the **Transnational Access activity to HPC Infrastructures**

- Access to state of the art HPC infrastructures for European researchers
- Provision of an advanced computational environment
- HPC consultancy from experienced staff
- Collaboration with scientists working in related fields at a local research institute
- Travel costs, subsistence expenses and accomodation

All fields of computational sciences are involved

800 researchers will benefit from these Access visits in 2004-07



CINECA Consorzio Interuniversitario

HPC-Europa

Pan-European Research Infrastructure on High Performance Computing

TRANSNATIONAL ACCESS

## Who can ask for access

European researcher who need HPC resources and want to spend a working period in a foreign European country collaborating with local researchers in their research field and experts in supercomputig support

## How long you can stay abroad:

From one to twelve weeks

## When can you apply

Continuous Call for proposals over 4 years!

Evaluations are every 3 months

International Scientific selection Panel

**http://www.hpc-europa.org**

Grazie!