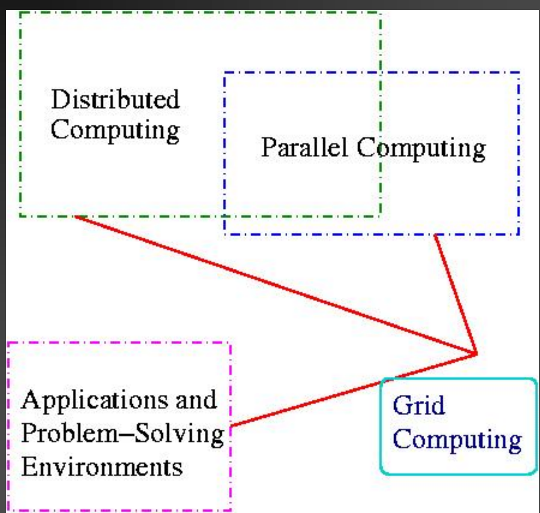
Grid Computing

José Cardoso Cunha
Dep. Informática
CITI – Centre for Informatics and Information Technologies
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa

1. Motivation
2. Concept of a Grid
3. Grid Architecture
4. Applications and User Profiles
5. Research Directions
6. Portuguese Efforts
7. Conclusions

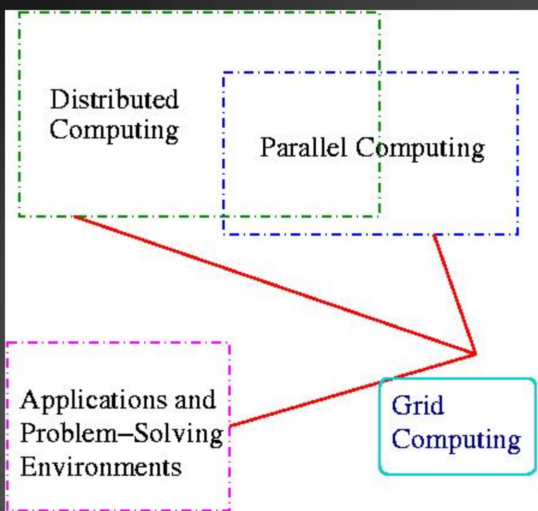
Genesis of Grid Computing



Distributed Computing

- Physically distributed computations and data
- Local (LAN) or large scale (WAN)
- Geographical distribution
  - Users and access sites
  - Processing sites and data archives
- Availability and Reliability
  - Fault tolerance
  - Replication of hardware and software
- Goals:
  - Adapt to geographical application distribution
  - Provide appropriate levels of transparency

Genesis of Grid Computing



Parallel Computing

- Computer System Architectures: 1980s-90s
  - Supercomputers
  - Shared / Distributed memory multiprocessors
  - LANs and Clusters of PCs
- Parallel Programming requires:
  - Decompose application in parts
  - Launch tasks in parallel processes
  - Plan the cooperation between tasks
- Goal: to reduce execution time, compared to sequential execution
- Quite a difficult task!

Developing Parallel Applications

- Costs of task decomposition and cooperation depend critically on the system layers:
  - Application
  - Algorithm
  - Programming Language
  - Operating System
  - Computer Architecture
- How to evaluate the overall result?
  - Correctness
  - Performance
- Long term research on Models, Tools and Environments

Factors affecting Performance

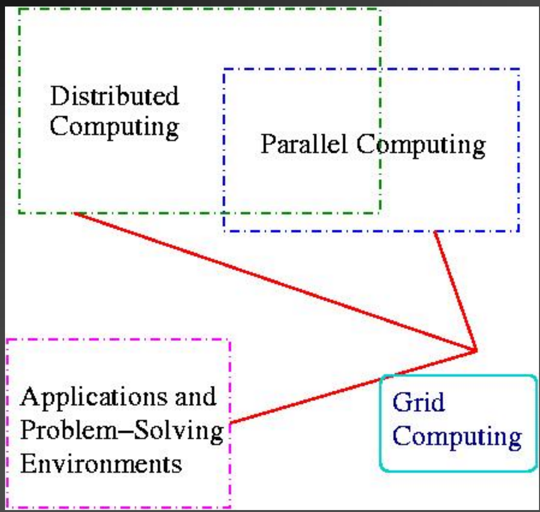
- Memory access vs CPU times
- Shared memory access conflicts
- Task and data distribution
- Sequential code and I/O
- Process management overheads
- Communication delays
- Synchronization
- Processor load unbalanced

They have a combined global effect.

Reasons to exploit Parallelism

- Why to develop parallel applications?

Genesis of Grid Computing



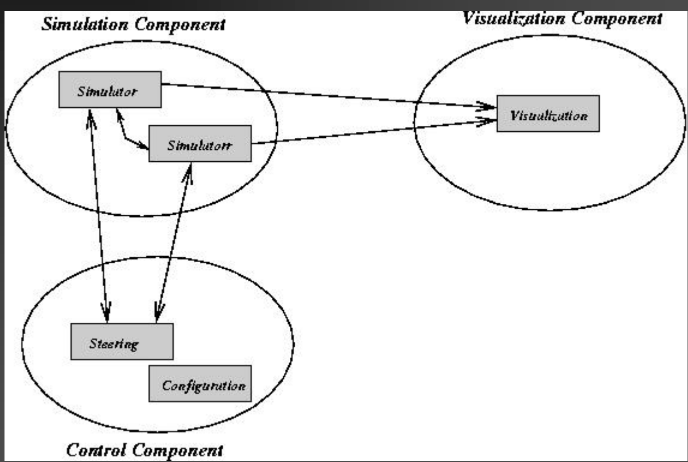
Examples of Application Areas

- Science and Engineering
  - Fluid Dynamics
  - Particle Systems in Physics
  - Weather Forecast and Climate
- Simulation of VLSI systems
- Parallel Databases
- Artificial Intelligence
- ...

Some Application Characteristics

- Complex models – simulations
- Large volumes of input / generated data
- Difficult interpretation and classification
- High degree of User interaction:
  - Offline / online data processing / visualization
  - Distinct user interfaces
  - Computational steering
- Multidisciplinary:
  - Heterogeneous models / components
  - Interactions among multiple users, collaboration
- Require parallel and distributed processing

A Parallel / Distributed Application

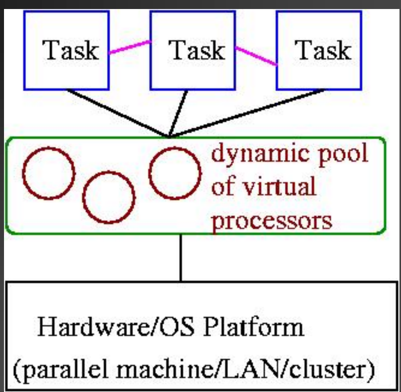


Heterogeneous Components

- Sequential, Parallel, Distributed Problem Solvers (simulators, mathematical packages,etc.)
- Tools for data / result processing, interpretation and visualization
- Online access to scientific data sets and databases
- Interactive (online) steering

- Can be mapped onto a parallel and distributed platform e.g. Based on PVM or MPI

Parallel Virtual Machine (PVM)



Typical Cycle of User Activities

1. Problem specification
2. Configuration of the environment:
   - Component selection (simulation, control, visualization) and configuration
3. Component activation and mapping
4. Initial set up of simulation parameters
5. Start of execution, possibly with monitoring, visualization and steering
6. Analysis of intermediate / final results

Problem-Solving Environments (PSE)

- Integrated environments for solving a class of related problems in an application domain:
  - Easy-to-use by the end-user
  - Based on state-of-the-art algorithms
- An old idea:
  - Examples:
    - MatLab, Mathematica
    - For standalone and local use

PSE Impact

- Several fully developed PSE in the Industry, e.g. Automotive, Aerospace
- Many applications in Science and Engineering:
  - Design optimization
  - Application behavior studies
  - Rapid prototyping
  - Decison support
  - Process control
- Emerging areas: Education, Environment, Health, Finance
- A new profile of end-user, beyond the scientist and engineer

PSE Functionalities

- Support for problem specification
- Support for resource management
- Support for execution services

Distributed PSE

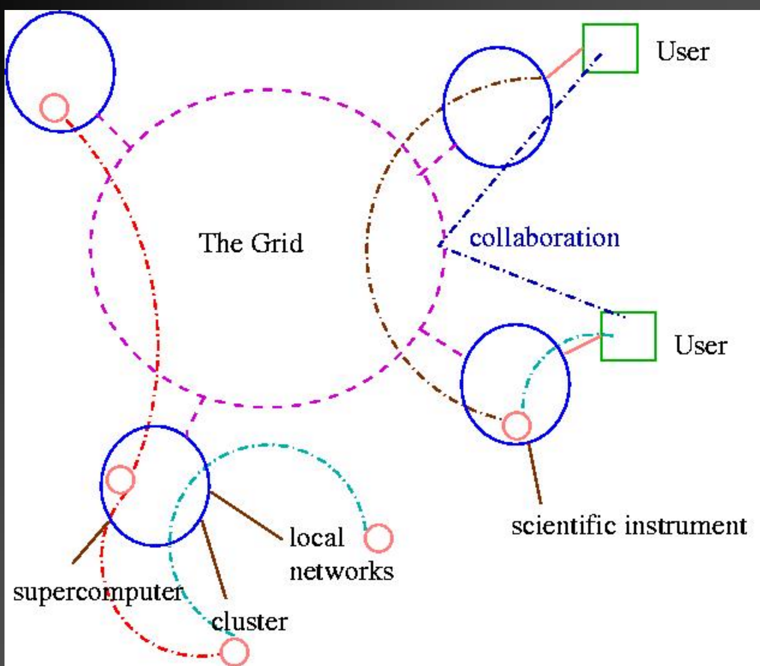
- Integrates heterogeneous components into an environment
- Transparent access to distributed resources
- Collaborative modeling and simulation
- Web-accessed

An Example - NetSolve

- A client-server system for remote solutions of complex scientific problems:
  - On request: performs computational tasks on a set of servers
  - Based on agents or resource brokers
- Access to languages C, Fortran, MatLab, Mathematica
- Application Service Provider: supports the resources for a particular set of services

Motivations for Grids

- Enable ´´heavy´´ applications in science and engineering
  - Complex simulations with visualization and steering
  - Access and analysis of large remote datasets
  - Access to remote data sources and special instruments (satellite data, particle accelerators)
- distributed in wide-area networks, and
- accessed through collaborative and multi-disciplinary PSE, via Web Portals.



Concept of a Grid

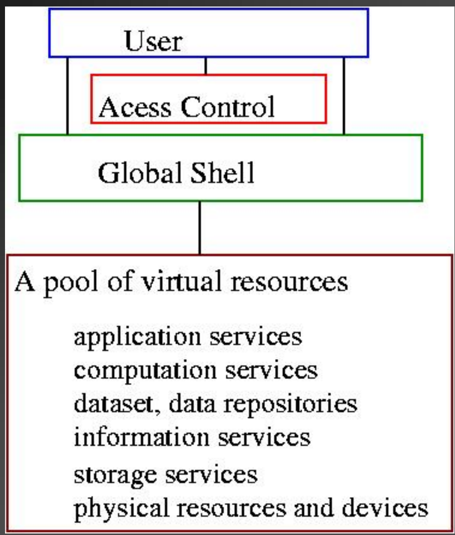
- Gathers a large diversity of distributed physical resources:
  - supercomputers and parallel machines
  - clusters of PCs
  - massive storage systems
  - databases and data sources
  - special devices

Concept of a Grid

Access is globally unified through virtual layers:

- solve new or larger problems by aggregating available resources
- access a large diversity of computation, data and information services
- enable coordinated resource sharing and collaboration across virtual organizations

Concept of a Grid



Grids are very complex systems

- Aim at providing unifying abstractions to the end-user
- Large-scale universe of distributed, heterogeneous, and dynamic resources
- Critical aspects:
  - Distributed
  - Large-scale
  - Multiple administrative domains
  - Security and access control
  - Heterogeneity
  - Dynamic

Main goals

- Towards uniform and standard large-scale computing environments
- Virtual resources:
  - Transient: to support experiments
  (computation, data, scientific instruments)
  - Persistent
  (databases, catalogues, archives)
  - Collaboration spaces

Applications and User Profiles

- Computational Grids:
  - provide a single point of access to a high-performance computing service
- Scientific Data Grids:
  - Access large datasets with optimized data transfers and interactions for data processing
- Virtual Organizations:
  - Access to virtual environments for resource sharing, user interaction and collaboration
- Information and Knowledge services:
  - Access large geographically distributed data repositories, e.g. for data mining applications

Data Grids

- EU DataGrid project:
  - Large-scale environment for accessing and analysing large amounts of data:
    - High energy physics, Biology, Earth observation
  - Petabytes of data (1 000 000 Giga)
  - Thousands of researchers
  - Scalable storage of datasets: replicated, catalogued, distributed in distinct sites

Virtual Organizations

- Resource sharing and collaboration between dynamically changing collections of individuals and organizations
- E.g. Consortium of companies collaborating in a design of a new product
    - Sharing design data, Collaborative simulations, etc
- E.g. Scientists collaborating in common experiments via a distributed virtual laboratory

Layers of a Grid Architecture

- User Interfaces, Applications, PSEs
- Development Tools and Environments
- Grid Middleware: Services and Resource Management
- Heterogeneous Resources and Infrastructure

Elements of a Grid Architecture

- User interfaces and grid portals
- Applications and PSEs
- Development environments and tools
- Grid middleware:
  - Resource management and scheduling
  - Information registration and discovery
  - Authentication, Security
  - Storage access, and communication
- Heterogeneous and physical resources, and network infrastructure

Example – The Globus toolkit

- Grid middleware: Provides secure and uniform access to remote computation and storage resources
- Used in most ongoing grid projects

Ongoing efforts

- Ongoing research on the Grid:
  - On the core grid middleware
  - On the application tools and environments
  - On the integration of grid systems
  - On the applications

Dimensions

- Resource management
  - Configuration of parallel and distributed virtual machines
  - Resource discovery, scheduling, and reservation
  - Quality of Service

Further Research

- How to specify, compose, develop, and understand dynamic distributed large-scale applications: models, languages, and tools
- Coordination models
  - Dynamic change of application structure, interaction patterns and operation modes
- Strategies for adaptive resource scheduling
- New problem-solving strategies

Portuguese Efforts

- Ongoing initiatives:
  - LIP, ADETI, and Universities
- Recent National Meeting promoted by FCCN – Fundação para a Computação Científica Nacional
- Plans for cooperation at a national scale

Conclusions

- Grid Computing:
  - Aims at some hard (impossible?) to achieve goals
  - It poses many challenges
  - It is already driving significant research and development efforts that will have great impact upon many areas

Grids

- The Electrical Power Grid
  - Simple local interface
  - Transparency
  - Pervasive access
  - Secure
  - Dependable
  - Efficient
  - Inexpensive
- The Computational and Data Grid:
  - Not really true (yet!?)