

The DEISA Project and its Joint Research Activities

Hermann.Lederer@rzg.mpg.de

Garching Computing Centre of the Max Planck Society

<http://www.rzg.mpg.de>





FP6 EU Integrated Infrastructure Initiative (I3) Project

Contract no. 508803

Project start: May 1, 2004

Project duration: 5 years



Project Aim

Advancement of Computational Sciences
in Supercomputing in Europe

Project Direction

Europe-wide collaboration of major supercomputer centres
for a joint deployment and operation of a distributed Terascale
supercomputer infrastructure

Approach

Proposed infrastructure: tight coupling of large supercomputers
Deployment of a virtual (distributed) supercomputing platform
in multi-cluster-mode



FI



I



D

Initially participating sites



D

UK

UK



NL



F



Stepwise approach

- Start with four homogeneous platforms
(„Proof of Concept Sites“)
- Inclusion of further homogeneous platforms
- Extension of the homogeneous multi cluster platform towards a larger, heterogenoeus environment



Project Core Activities

- *Network Activities*

Organisation in analogy to a virtual supercomputer centre

- *Service Activities*

Integration of leading grid technologies

- *Joint Research Activities*

Inclusion of leading applications in computational sciences

Involvement of leading computational scientists in Europe



Co-scheduling and Meta-Computing

Problems:

- co-scheduling, as provided by Globus, is not yet mature enough to be massively used in national supercomputing centres without disrupting the efficiency of national services.
- Communication latencies across Europe are of the order of 10-100 milliseconds, instead of a few microseconds (“speed of light not fast enough”)

Consequences

These latencies are not acceptable for demanding, tightly coupled parallel applications which require high performance computing.

DEISA does not rely on meta-computing for massive science production. Instead, it adopts a simple, innovative operational model that relies on communication bandwidth to make the latency problem irrelevant.

Co-scheduling and meta-computing will only be used in selected cases of weakly coupled distributed applications.



The DEISA Operational Model

- Global management of a “resource pool”, that will allow:
 - The availability of an integrated supercomputing environment for trans-national collaborations, providing global data management through global file systems.
 - Implementation of job migration across sites (transparent to end users) as a way of releasing significant resources in one particular site for demanding DEISA applications. We are load balancing computational throughput at a European scale.
 - Support of distributed applications.
- With this operational model, the European super-cluster is not very different from a huge local supercomputing cluster (which will be partitioned anyhow because of fault tolerance)



Data Management with Global File systems (GFS)

- Global file systems allow direct access to European-wide / world-wide distributed data in a secure and comfortable way
 - > Enormous facilitation of data management in a virtual (distributed) environment
- Fundamental technology choice of the DEISA consortium:
Deployment of global file systems at the continental level in the DEISA multi-cluster distributed platform.
 - > GFS “symmetric” with respect to all computing platforms,
data and executable files need not to be transferred by user action.
 - > Job migration among sites facilitated

**GFS necessary – but not sufficient – to provide a single system image
of a clustered computing platform**



Global File systems: GPFS

- Fully distributed software providing a “Parallel” global file system, strongly oriented towards performance and scalability.
- Adapts and takes full advantage of underlying high speed network
- Efficiently operating on AIX national clusters, and on Linux clusters
- DEISA requires, in phase 1, only fine tuning of existing installed software, plus the next release of AIX. TeraGrid is using a Linux extension of GPFS produced by IBM.
- We have started to coordinate actions with TeraGrid on GPFS.



Global File systems: AFS

- Client-server Global File System with solid security model
- Open source
- Integrated HSM feature (with mr-afs)
- Interoperates with Globus
- Expertise available at RZG

- Interest for DEISA
 - Solid option for global data management at the “external” level and possibly – if performance is validated – at the “core” level.



Heterogeneous Extension of Core Platform

- Fundamental strategic choice: no permanent commitment to any particular technology or vendor.
- This is an operation system (OS) issue, not a hardware issue.
- The only OS whose survival seems guaranteed in the long turn is Linux.
- Heterogeneous extension means integration of Linux systems
- Technology options:
 - AVAKI DataGrid
 - IBM GPFS for Linux, if it works on non-IBM Linux systems
 - AFS
New technology from RZG allows to combine advantages of GPFS and AFS
 - NFS4
 - lustre
 - other future options emerging



DEISA and Grid technologies

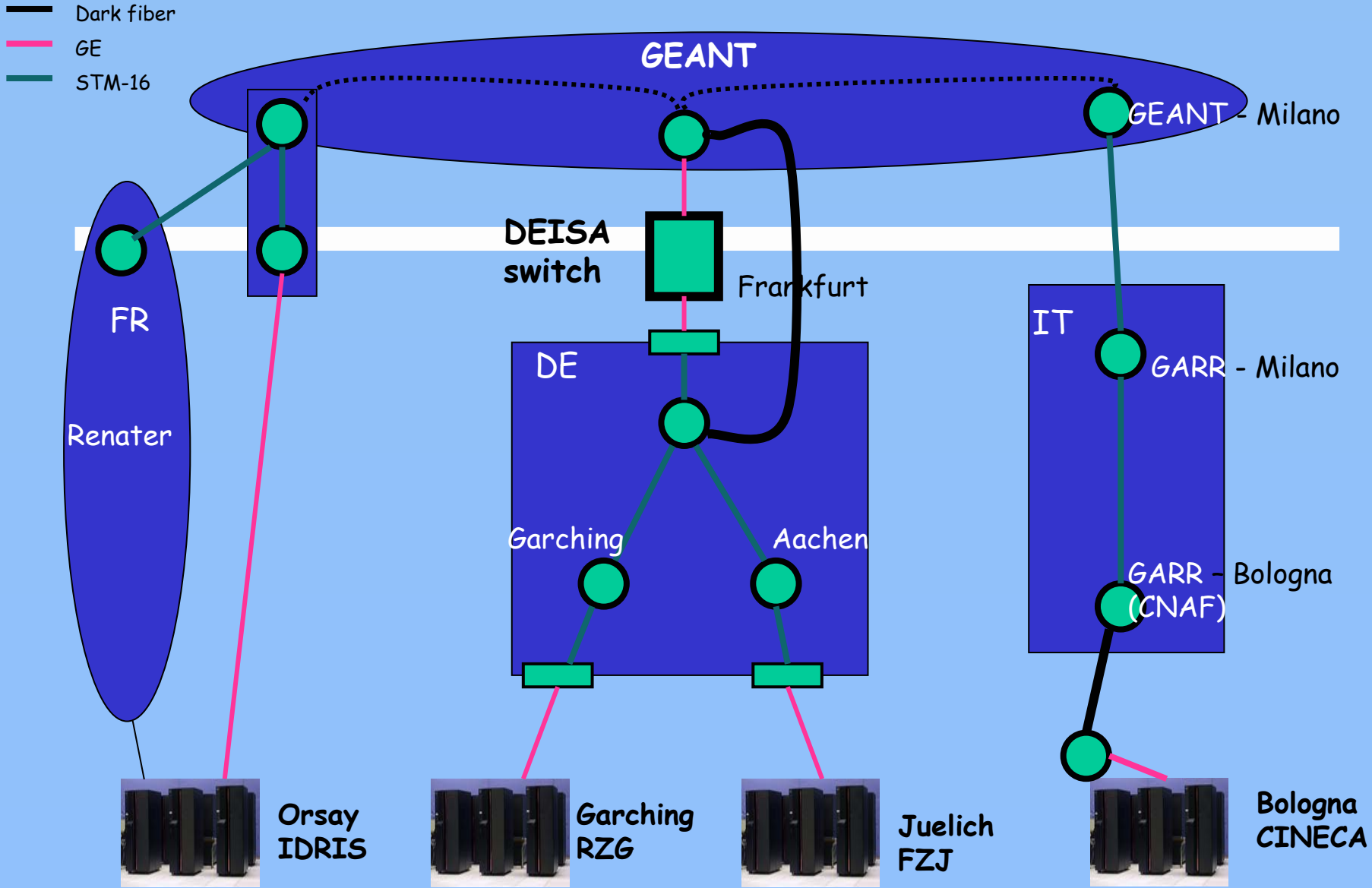
- Large scale integration of IT systems is, for DEISA, primarily a strategic issue
 - DEISA is technology neutral. Technology choices follow from their capability to adapt to a pre-established operational model, and to provide real services to end users.
 - Three criteria for technology choices:
 - The necessity of disposing as soon as possible of a stable and reliable European production platform (the “core” platform). *Grid technologies work in the background through global file systems and multi-cluster batch managers.*
 - The necessity interfacing this platform with other systems in Europe (the “outer” Grid environment). *Deployment of traditional Grid environments, like the Globus Toolkit or Unicore.*
 - The preparation of their future evolution by the integration of new Grid technologies as they reach maturity (*a specific JRA activity*).
-



- Virtual private network
- In phase with GEANT development
- Connection through GEANT, with national access via NRENs
- Realisation via Premium IP Services / Priority Routing
- Start with 1 Gb/s connection per “proof of concept” site
- Extension to 10 Gb/s after GEANT upgrade in 2005/2006



DEISA network / Initial Phase



Categories of Service Activities

SA1 – Network Operation and Support.

SA2 – Data Management with Global file systems.

SA3 – Resource Management.

SA4 – Applications and User Support.

SA5 – Security.



Joint Research Activities

Key Aspects

- Inclusion of leading applications in computational sciences
- Involvement of leading computational scientists in Europe
- Synchronisation with SA activities



Categories of Joint Research Activities

JRA1 – Materials Sciences	(RZG)
JRA2 – Cosmology	(EPCC)
JRA3 – Plasma Physics	(RZG)
JRA4 – Life Sciences	(IDRIS)
JRA5 – Industry (CFD)	(CINECA)
JRA6 – Coupled Applications	(IDRIS)

*Migration and deployment of leading applications
in these scientific areas*

JRA7 – Access to Resources in Heterogeneous Environments (EPCC)



The overall guidance of this JRA
especially with its relevance
for the scientific community
is in the hands of
Prof. Michele Parrinello.



Joint Research Activities – Materials Sciences

Involvement of leading computational scientists in Europe

Prof. Michele Parrinello

ETH Zuerich, CH

Prof. Matthias Scheffler

Fritz-Haber-Institute, Berlin, Germany

Prof. Karlheinz Schwarz

Technical University of Vienna, Austria

Prof. Peter Coveney

University College London, UK

Prof. Kurt Kremer

Max-Planck-Institute for Polymer Research,
Mainz, Germany

Prof. Risto Nieminen

Helsinki University of Technology, Finland



*Grand challenge:
to span time and length scales of many orders of magnitude*

Microscopic level:

typical bond distance between atoms is of the order of Angstroms, atoms vibrate with a frequency of approx. 10^{13} 1/s.

Macroscopic level:

phenomena and applications of practical interest occur on a time scale of seconds, system sizes of microns or larger.



Inclusion of leading applications in computational sciences

Already selected applications:

CPMD	www.cpmd.org	CH
WIEN2K	www.wien2k.at	A, D
Espresso	www.espresso.mpg.de	D
Set of codes from UK RealityGrid	www.realitygrid.org	UK

Starter: CPMD



Joint Research Activities - Materials Sciences

Involvement of leading computational scientists in Europe

Principal Investigator: Prof. Michele Parrinello, ETH Zuerich, CH

Proposal

Ab-initio base crystal structure prediction in a GRID environment

Approach:

Mixed quantum/classical procedure that combines the CPMD program with one hyper cluster method generally used for crystal structure prediction. Implementation and validation of the method on a subset of the Cambridge Structural Database for molecular solids.



Joint Research Activities - Materials Sciences

Involvement of leading computational scientists in Europe

Principal Investigators:

Prof. Matthias Scheffler, Fritz-Haber-Institute, Berlin, Germany

Prof. Karlheinz Schwarz, Technical University of Vienna, Austria

Proposal:

Bridging Time and Length Scales in Materials Science and Bio-Physics:

Studies of surface processes, such as oxidation catalysis, as well as studies of biopolymer structures.

Approach:

Application of hybrid methods that combine density functional theory (DFT) with statistical mechanics and thermodynamics. Basis for the atomistic, first-principles description to time and length scales of the order of seconds and micrometers. Novel APW+lo method recently implemented in the WIEN2k code from the Technical University of Vienna combined with parallelization on each level of the multi-scale modelling, new classes of problems will be tractable for the first time.



Joint Research Activities - Materials Sciences

Involvement of leading computational scientists in Europe

Principal Investigator: Prof. Peter Coveney, University College London

Proposal:

Reality Grid Project in DEISA environment

Application of high performance codes, including lattice gas, lattice Boltzmann and molecular dynamics, inter alia, for the simulation and modelling of complex condensed matter structures and biomolecules over multiple length and time scales as well as the discovery of new materials.

Approach

The RealityGrid project is a major U.K. e-Science research partnership involving seven U.K. universities and as many non-academic affiliates.

The aim of RealityGrid is to use grid technology to closely couple high performance computing, high throughput experiment and visualization. There are already capabilities to perform computational steering in an OGSA-compliant framework, running over Unicore and Globus



Joint Research Activities - Materials Sciences

Involvement of leading computational scientists in Europe

Principal Investigator:

Prof. Kurt Kremer, Max-Planck-Institute for Polymer Research, Mainz, Germany

Proposal

Non-equilibrium and multiscale molecular dynamics simulations of macromolecules with new code package ESPRESSO

Approach

With the development of our molecular dynamics package Espresso we intend to be able to attack problems where alternative approaches, such as lattice gas models, are doomed to failure due to the lack of detail on the small scale. Espresso supports coarse grained force fields on various levels which simplify the atomistic structure as much as preferred and thus facilitates the computation of huge problem sizes. E.g., tens if not hundreds of millions of particles are needed to be able to understand complicated effects, such as the development of multilamellar vesicles in the presence of a shear field. Although the overall length scale of those vesicles reaches the order of micrometers, the molecular structure on the length scale of Ångströms immensely influences the result.



Joint Research Activities - Materials Sciences

Involvement of leading computational scientists in Europe

Principal Investigator:

Prof. Risto Nieminen, Helsinki University of Technology, Finland

Project:

Multiscale modelling and simulation of new materials, their physical and chemical properties, their processing, and of structures made thereof, especially those relevant for nano- and microtechnologies.

Code: CPMD



Joint Research Activities – Plasma Physics

Involvement of leading computational scientists in Europe

Consortium established with principal investigators:

Prof. Karl Lackner

Max Planck Institute of Plasma Physics, Garching

Prof. Jürgen Nührenberg

Max Planck Institute of Plasma Physics, Greifswald

Prof. Laurent Villard

Centre de Recherches en Physiques des Plasmas, EPFL-Lausanne



Joint Research Activities – Plasma Physics

Involvement of leading computational scientists in Europe

Consortium includes scientists from:

- Max Planck Institute of Plasma Physics, Germany
- Centre de Recherches en Physiques des Plasmas, EPFL-Lausanne, Switzerland

European collaborations exist with:

- Observatory Paris-Meudon, France
- UKAEA, Culham, UK
- Risø National Laboratory, Denmark
- Centre for Interdisciplinary Plasma Science, Germany
- Research Centre, Jülich, Germany
- Christian-Albrechts-University at Kiel, Germany

World wide collaborations exist with institutes from

USA: Princeton Plasma Physics Laboratory, Princeton University, University of Maryland, General Atomics, MIT Plasma Science & Fusion Center;

Japan: National Institute for Fusion Science

Australia: Australia National University



Joint Research Activities – Plasma Physics

One of the grand challenges of physics: theoretical analysis of turbulence

Turbulence responsible for energy and particle transport in magnetically confined plasmas, one of the key elements determining the minimum size of self-sustained burning plasmas.

Magnetic field in plasma: turbulent structures become highly anisotropic

Plasma turbulence involves spatial scales ranging from the electron gyro-radius to the plasma cross-section and time scales ranging from approximately the ion gyration time to the confinement time.

More complexity is added to the system: a simple one-fluid description of the plasma is not adequate and has to be substituted by a more complex two fluid model with kinetic corrections, or even a fully kinetic model.



Inclusion of leading applications in computational sciences

List of important codes contains already 10 codes with world-wide relevance

From this list codes will be selected for development to grid- computing capability. One will of course be also open for new developments.

Starter: ORB code family, ORB code first
(from CH – D - collaboration)

ORB applies an electrostatic Monte Carlo (particle-in-cell) model to solve the gyrokinetic equations for the study of transport-related instabilities and turbulence in toroidal magnetic confinement devices.





DEISA

-

A major step
towards
European
Supercomputing