

DEISA : towards cooperative extreme computing in Europe

V. Alessandrini
www.deisa.org

The DEISA Consortium



WWW.DEISA.ORG

epcc

sara

lrz

HLRIS

European Centre for Medium-Range Weather Forecasts

Forschungszentrum Jülich
in der Helmholtz-Gemeinschaft

CINECA
Consorzio Interuniversitario

RZG IPP

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

CSC

Partners

- Institut du Développement et des Ressources en Informatique Scientifique, France
- Forschungszentrum Jülich, Germany
- Rechenzentrum Garching of the Max Planck Society, Germany
- Consorzio Interuniversitario, Italy
- Edinburgh Parallel Computing Centre, UK
- SARA Computing and Networking Services, The Netherlands
- Finnish Information Technology Center for Science, Finland
- European Centre for Medium-Range Weather Forecasts, UK
- High Performance Computing Center, Germany
- Leibniz Computing Centre of the Bavarian Academy of Sciences and Humanities, Germany
- Barcelona Supercomputing Center (BSC), Spain

DEISA objectives



- *To enable Europe's terascale science by the integration of Europe's most powerful supercomputing systems.*
- *Enabling scientific discovery across a broad spectrum of science and technology is the only criterion for success*
- **DEISA is an European Supercomputing Service built on top of existing national services.**
- **DEISA deploys and operates a persistent, production quality, distributed supercomputing environment with continental scope.**

Basic requirements and strategies for the DEISA research Infrastructure



- **Fast deployment of a persistent, production quality, grid empowered supercomputing infrastructure with continental scope.**
- **European supercomputing service built on top of existing national services requires reliability and non disruptive behavior.**
- **User transparency**
- **Application transparency**
- **Top-bottom approach: technology choices follow from the business and operational models of our virtual organization. DEISA technology choices are fully open.**

The DEISA supercomputing Grid: a layered infrastructure



- **Inner layer: a distributed super-cluster resulting from the deep integration of similar IBM AIX platforms** at IDRIS, FZ-Julich, RZG-Garching and CINECA (phase 1) then CSC (phase 2). It looks to external users as a single supercomputing platform.
- **Outer layer: a heterogeneous supercomputing Grid:**
 - IBM AIX super-cluster (IDRIS, FZJ, RZG, CINECA, CSC) close to 24 Tf
 - BSC, IBM PowerPC Linux system, 40 Tf
 - LRZ , Linux cluster (2.7 Tf) moving to SGI ALTIX system (33 Tf in 2006, 70 Tf in 2007)
 - SARA, SGI ALTIX Linux cluster, 2.2 Tf
 - ECMWF, IBM AIX system, 32 Tf
 - HLRS, NEC SX8 vector system, close to 10 Tf

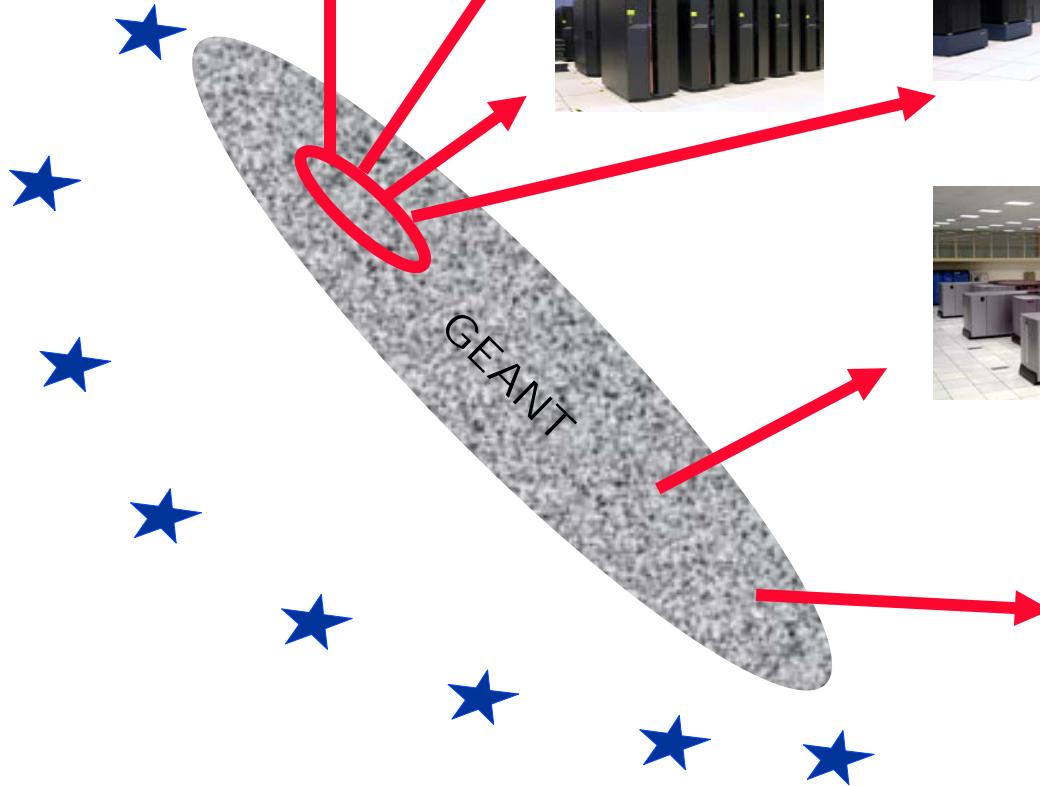
THE DEISA SUPERCOMPUTING GRID



AIX distributed super-cluster

Vector systems (NEC, ...)

Linux systems (SGI, IBM, ...)



DEISA technology leaders

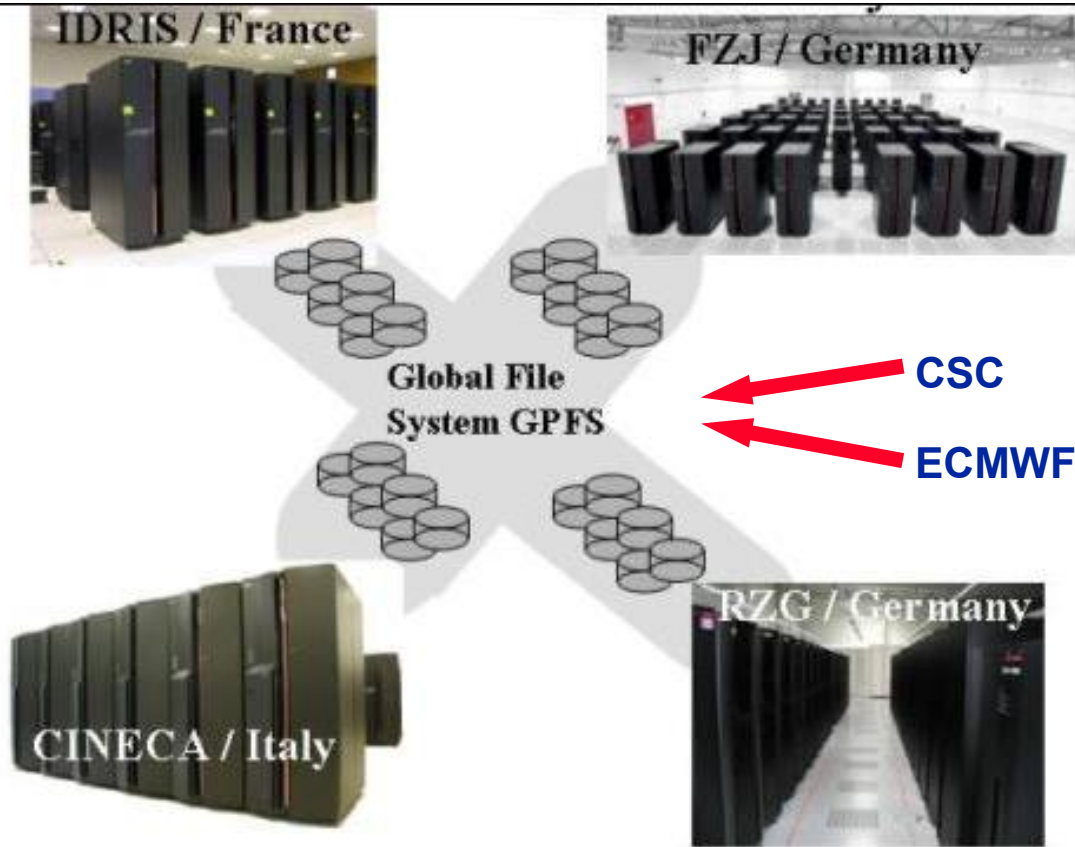


- **Deployment and operation of Grid infrastructure:**

- **Network operation and support** **Ralph Niederberger, FZJ-JULICH**
- **Global File Systems** **Stefan Heinzl, RZG-GARCHING**
- **Middleware** **Paolo Malfetti, CINECA**
- **Applications and User Support** **Denis Girou, IDRIS**
- **Security** **Jules Wolfrat, SARA**

- **DEISA specific middleware R&D:**

- **Heterogeneous Resource Management** **Terry Sloan, EPCC**



Services:

High performance datagrid via GPFS
Access to remote files use the full available network bandwidth

Job migration across sites
Used to load balance the global workflow when a huge partition is allocated to a DEISA project in one site

Common Production Environment

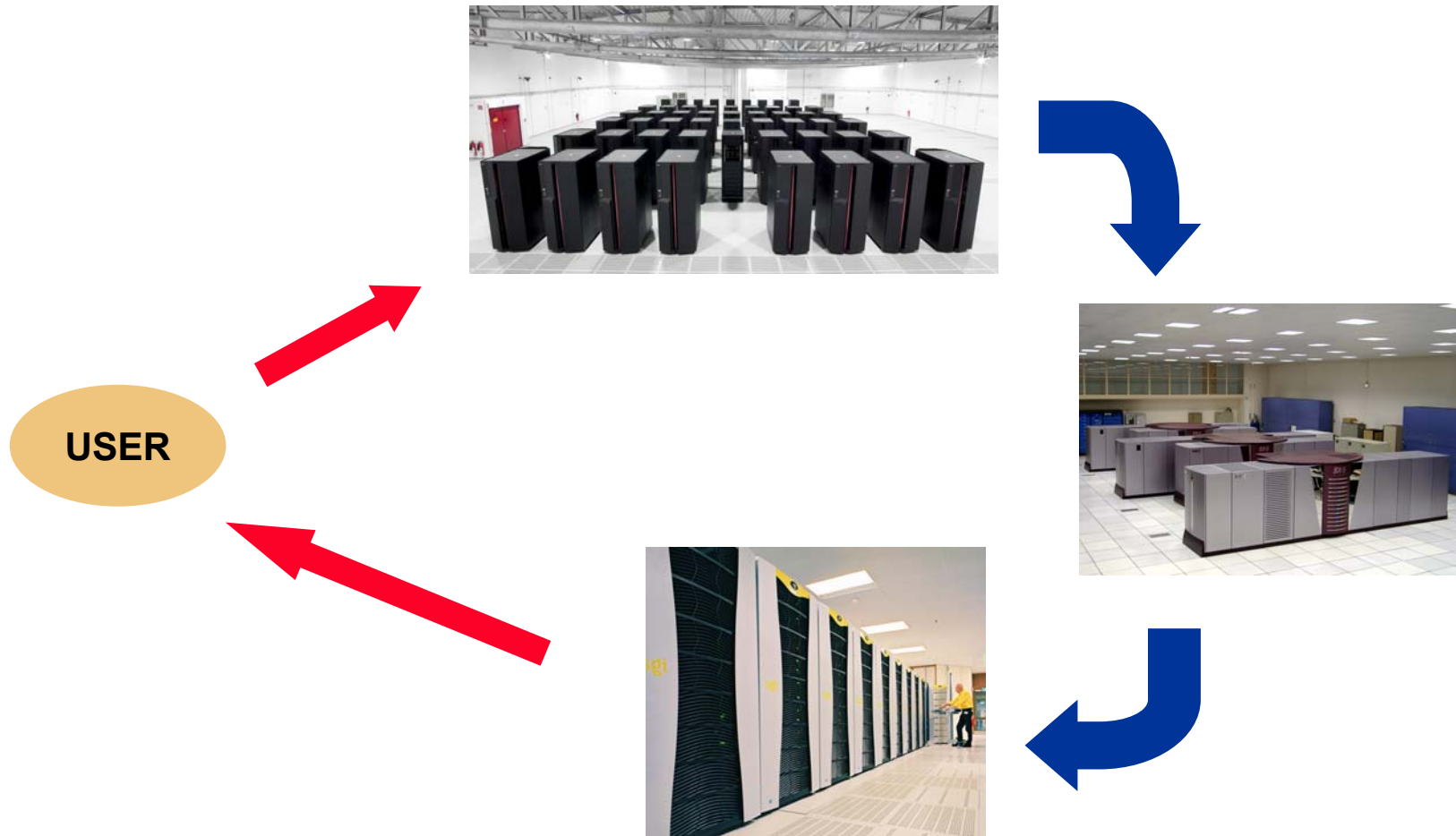
Full production status of dedicated 1 Gb/s network
GPFS : Full production IDRIS and RZG, FZJ and CINECA to follow immediately
JOB MIGRATION: test status in all sites

DEISA Supercomputing Grid services

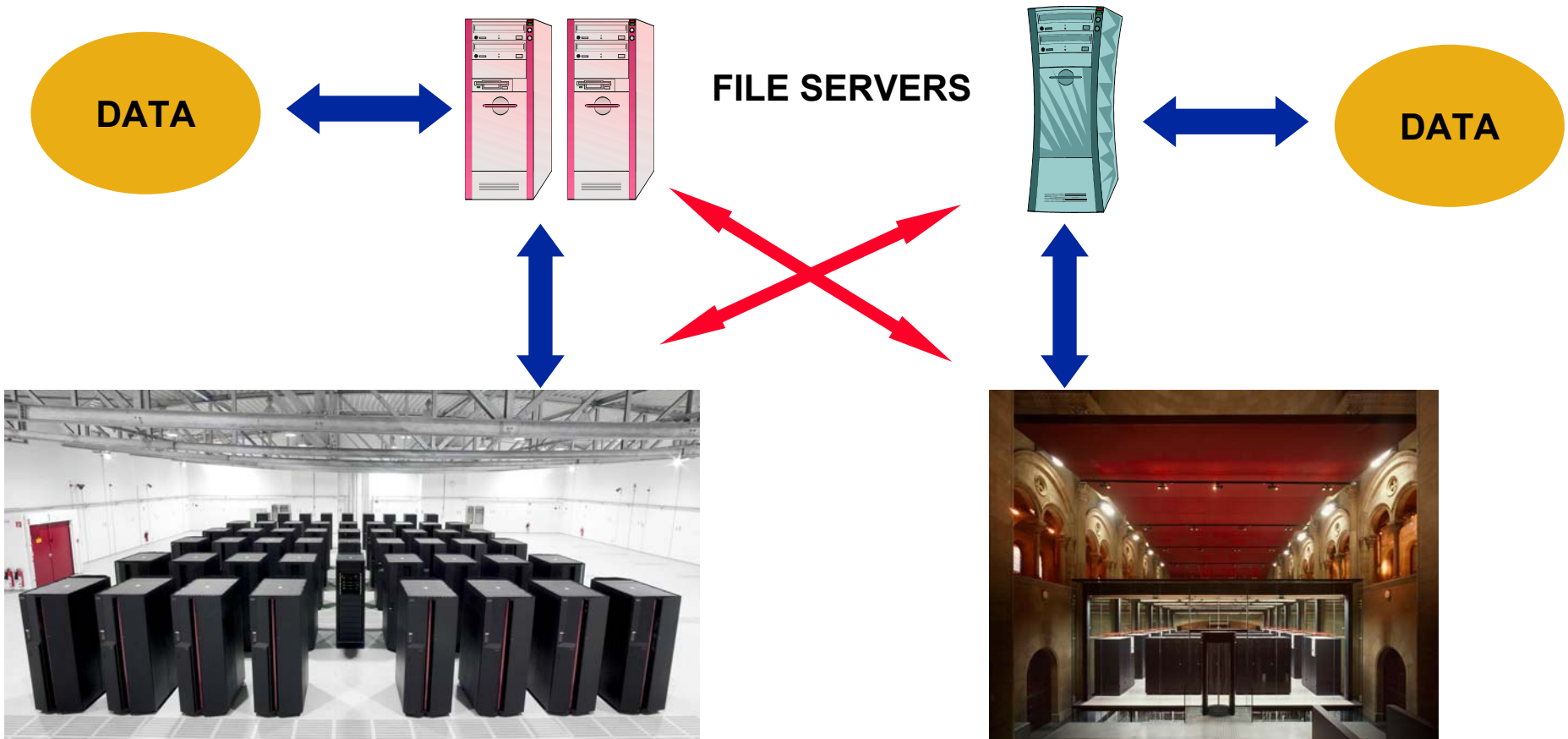


- **Workflow management:** based on UNICORE plus further extensions and services coming from DEISA's JRA7 and other projects (UniGrids, ...)
- **Global data management:** a well defined architecture implementing extended global file systems on heterogeneous systems, fast data transfers across sites, and hierarchical data management at a continental scale.
- **Co-scheduling,** needed to support Grid applications running on the heterogeneous environment.
- **Science Gateways and portals:** specific Internet interfaces to hide complex supercomputing environments from end users, and facilitate the access of new, non traditional, scientific communities.

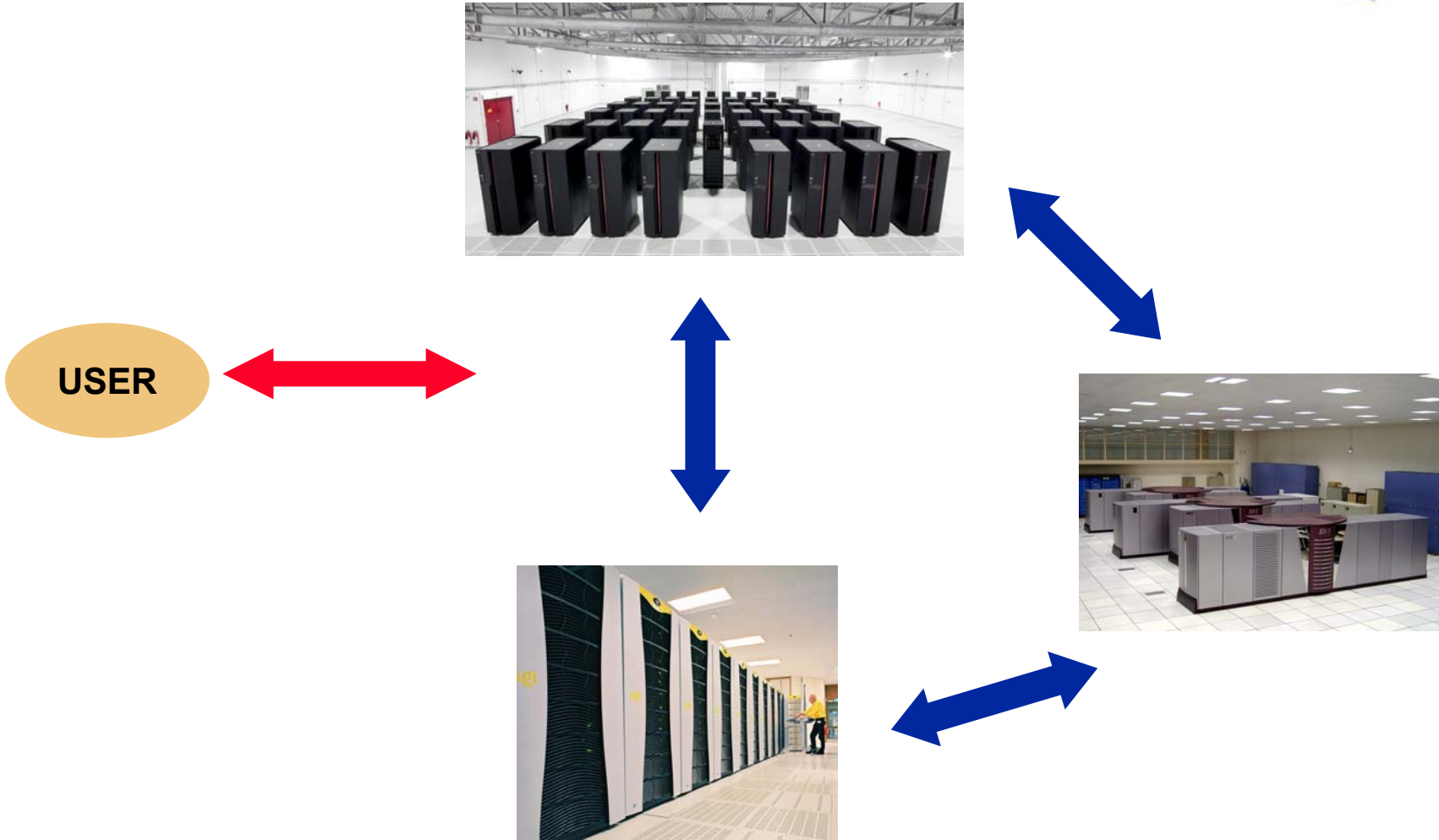
Workflow applications (UNICORE)



Global data management



Co-scheduling for coupled applications



Portals (Science Gateways)



- Same concept as TeraGrid's Science Gateways
- Needed to enhance the outreach of supercomputing infrastructures
- Hiding complex supercomputing environments from end users, providing discipline specific tools and support, and moving in some cases towards **community allocations**.
- There is already work done by DEISA on Genomics and Material Sciences portals
- Intense brainstorming on the desing of a global strategy, if possible interoperable with TeraGrid's Science Gateways

Enabling science



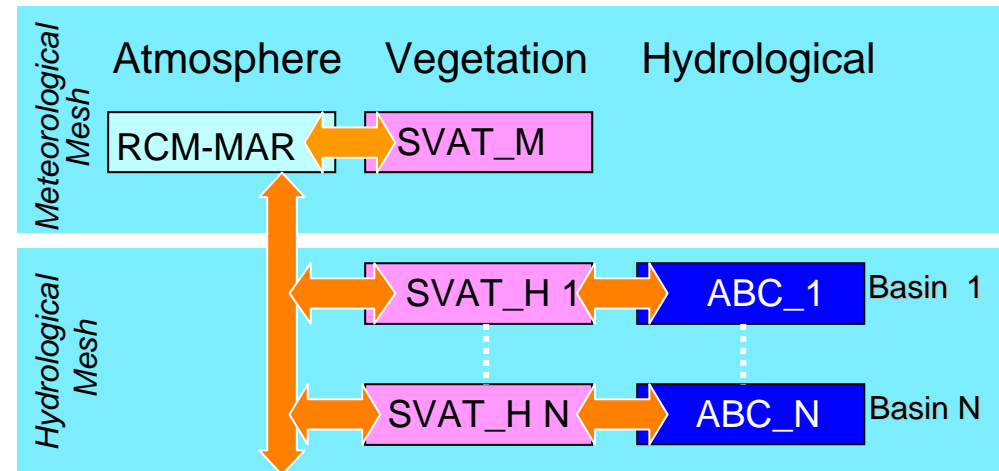
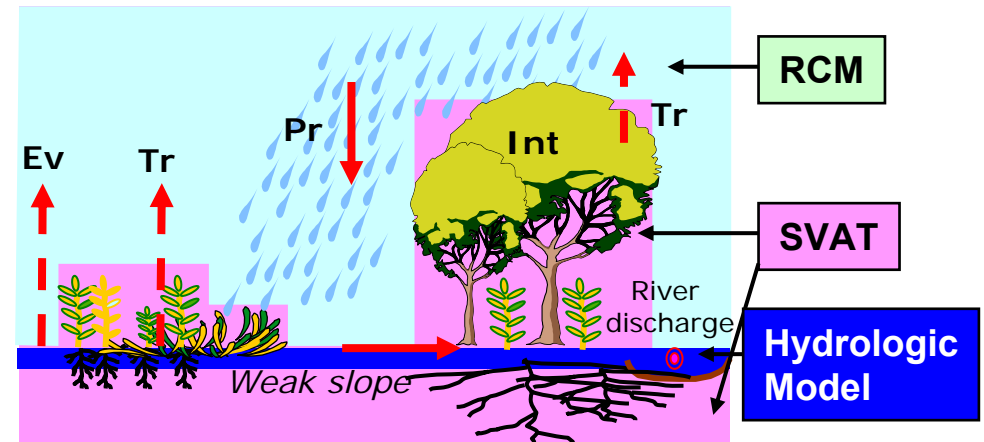
- Initial, « early users » program: a number of Joint Research Cctivities integrated in the project from the start.
- Moving towards « exceptional users » : the DEISA Extreme Computing Initiative

Activity	Scientific program	Partners	Leader
JRA1	Enabling Material Sciences CPMD codes, portals	RZG	<u>Hermann Lederer,</u> RZG
JRA2	Computational environment for applications in Cosmology	EPCC	Gavin Pringle, EPCC
JRA3	Enabling the TORB Plasma Physics code	RZG	<u>Hermann Lederer,</u> RZG
JRA4	Life sciences: genomic and <u>eHealth Applications</u>	IDRIS, (BSC)	VA , IDRIS -> BSC
JRA5	CFD in the automobile industry	CINECA, CRI	<u>Roberto Tregnago,</u> CRI
JRA6	<u>Coupled applications: Astrophysics,</u> Combustion, <u>Environment</u>	IDRIS (HLRS)	<u>Gilles Grasseau,</u> IDRIS

EARLY USERS PROGRAM

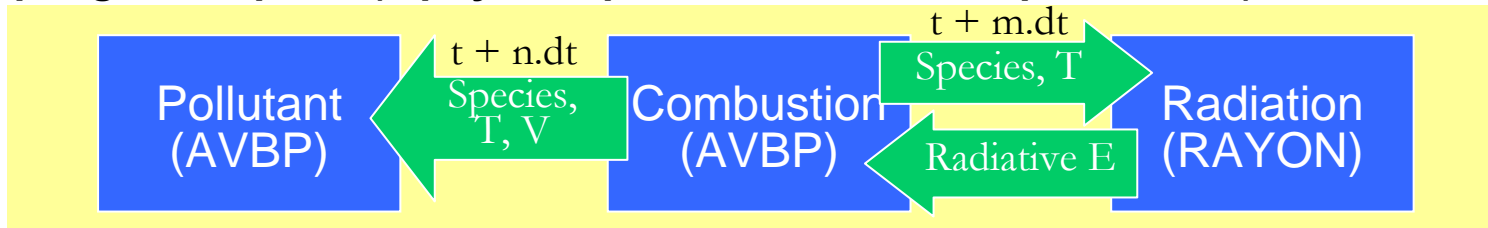
Coupled applications : Environment

- Leaders: Michel Vauclin and Christophe Messager (LTHE)
- Evaluate the importance of the water cycles between:
 - the atmosphere (RCM),
 - the soil / vegetation (SVAT),
 - the hydrological basins (hydrologic model)
- over the West Africa.
- Add easily new basins — the coupling architecture is modular and extensible.
- Collaboration with the international AMMA project (Africa Monsoon Multidisciplinary Analysis) by adding a new basin (Ouémé) and a new hydrologic model (dedicated to humid catchments).

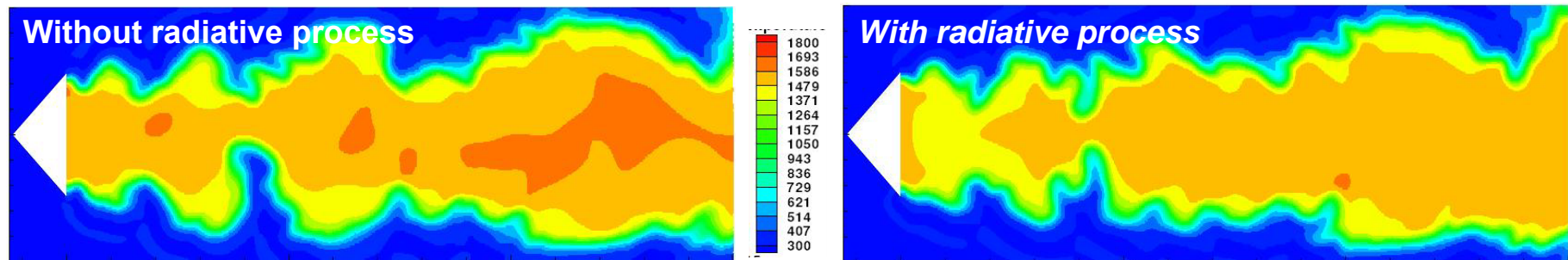


Coupled applications: combustion

- Leader: Denis Veynante (EM2C)
- Develop and optimize the efficiency of the combustion and reduce pollutant emissions in industrial systems (engines, energy production, industrial furnaces, ...)
- Take account of the radiative process in the combustion (rarely considered in previous works)
- Coupling description (3 physical phenomena → 3 coupled codes):



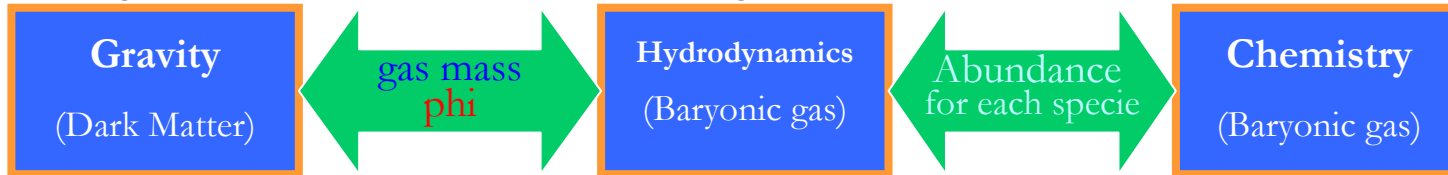
- First simulations about the impact of radiative process on the flame behaviour



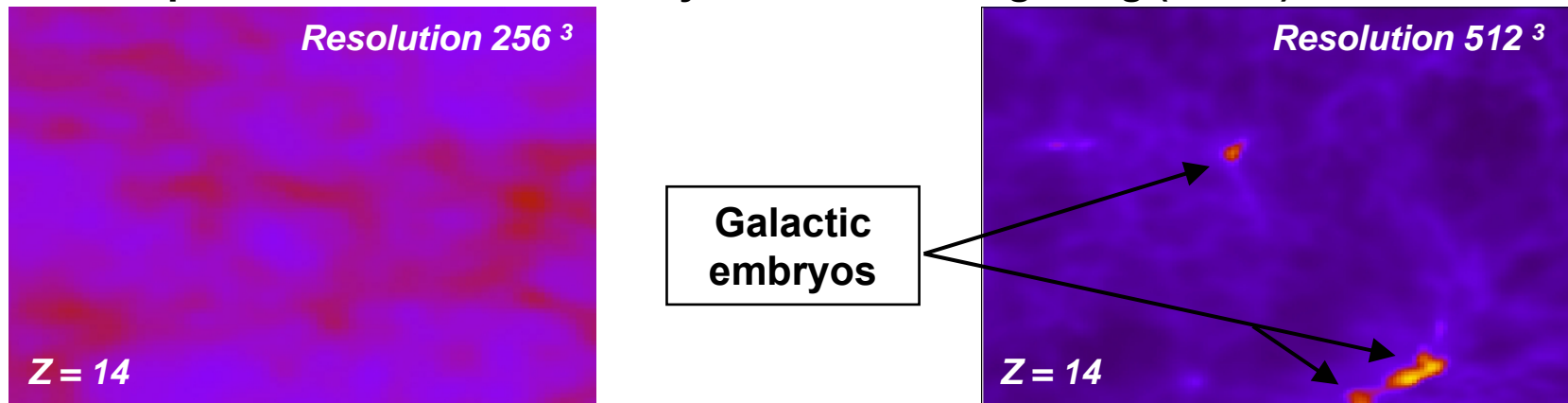
The temperature field is largely modified: the temperature decreases and the field is more homogeneous when the radiative process runs

Coupled applications: astrophysics

- Leader: Jean-Michel Alimi (LUTH)
- Modeling of the galaxy formation requires to take account of many physical processes.
- 3 main physical phenomena are currently considered → 3 coupled codes:



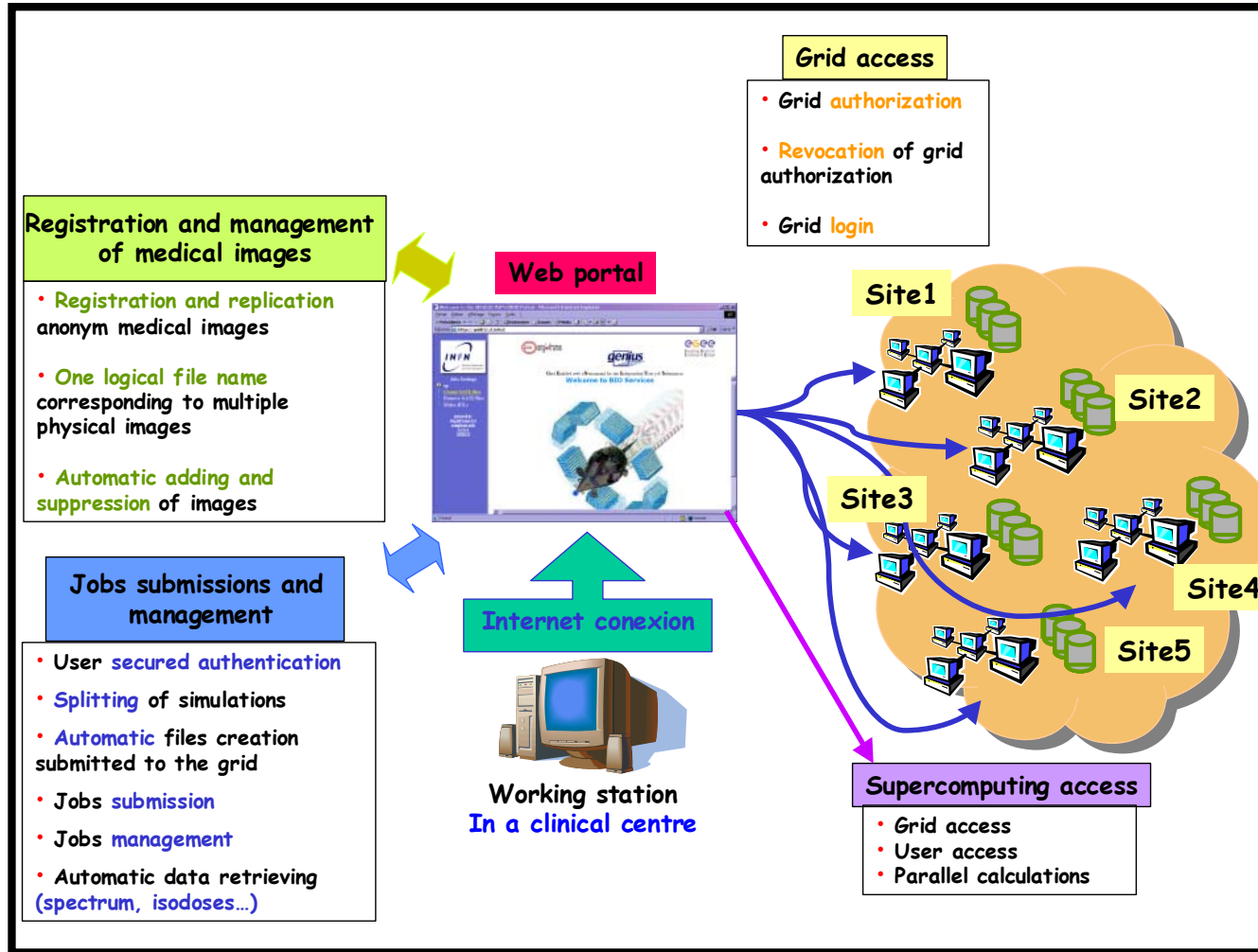
- Impact of the resolution in the process of galaxy formation
Gas temperature at 240 millions of years after the Big Bang ($z = 14$):



The high temperature zones (galactic embryos) appear later (470 millions of years after the Big Bang — $Z=10$) in the 256^3 resolution.

- Genomics computational environment on AIX super-cluster: databases, fine tuning of multithreaded BLAST.
- InfoBioGen portal: transparent rerouting of jobs to the DEISA platform.
- Leading applications
 - **INSERM: *Identification of new human mitochondrial proteins*** (ready for production)
 - **INRA : *Large scale microbial genome reannotation***
 - **BSC: *Prediction of protein interactions***
- Deployment, with BSC, of a high performance *heterogeneous* platform for bioinformatics applications

Radiation Therapy planning (joint application with EGEE)



The Extreme Computing Initiative



- Identification, deployment and operation of a number of « flagship » applications in selected areas of science and technology
- Applications must rely on the DEISA Supercomputing Grid services (application profiles have been clearly defined). They will benefit from exceptional resources from the DEISA pool.
- Applications are selected on the basis of scientific excellence, innovation potential, and relevance criteria.
- **European call for proposals:** April 1st -> May 30, 2005

Evaluation and allocation of DEISA resources



- National evaluation committees evaluate the proposals and determine priorities.
- On the basis of this information, the DEISA consortium examines how the applications map to the resources available in the DEISA pool, and negotiates internally the way the resources will be allocated and the final priorities for projects.
- Exceptional DEISA resources will be allocated – as in large scientific instruments – at well defined time windows (to be negotiated with the users).

Facing the « last mile problem »



- Creation, in March 2005, of the **Applications Task Force** (ATASKF), to support the Extreme Computing initiative.
- The ATASKF carries out a prospective action with the European scientific community, to support the design on new, leading applications.
- The ATASKF provides guidance to find the best fit between the user requirements and the distributed supercomputing environment
- The ATASKF determines the actions to be taken to move applications into the « extreme computing » domain.
- We have today (May 9) **25 solid Extreme Computing proposals**, and **10 more proposals** are expected before the ends of the month.

The ATASKF



- IDRIS – France
- FZJ-Juelich
- RZG – Garching
- CINECA, Italy
- EPCC, UK
- CSC, Helsinki
- ECMWF, UK
- LRZ, Munich
- BSC, Barcelona
- HLRS, Stuttgart

Pierre Francois Lavallée

Norbert Attig

Hermann Lederer (leader)

Chiara Marchetto

David Henty

Juha Fageholm

Umberto Modigliani

Mattias Brehm (co-leader)

Felipe Lozano

Mattias Mueller (co-leader)

DECI proposals - 1



- COSMOLOGY
- CINECA – Output simulation, calibration and data analysis for the instrument LFI of Planck project (INAF + Int.coll, 100.000 cpu h)
- EPCC – Evolution of Univers from Big Bang (Durham+VIRGO, 500.000 h)
- FZJ – Simulation of gravitational systems (Heidelberg, 250.000 h)
- IDRIS – Galaxy formation at high redshift in a cosmological frame (HORIZON project, IAP, LUTH, CEA, CNRS, LERMA 1.000.000 h)
- IDRIS – Lyman-alpha high resolution simulation (LUTH, 1500 PEs)

DECI proposals - 2



- ASTROPHYSICS

- IDRIS – Simulation of turbulent, active and rotating stars (LUTH – CEA – CNRS, 200.000 cpu hs)
- LRZ – Large eddy simulation of turbulence in galaxy cluster cores and star forming regions (Wurzburg Univ, 500.000 hs)
- RZG – Supernova research: 3D simulations of a relativistic AGN jet (MPI Garching + Valencia Univ, 500.000 hs)
- RZG – Realistic simulation of solar supergranularity (MPI + Stanford Univ, 300.000 hs)
- CINECA – Stellar system dynamics describes as self gravitating systems of N bodies interacting through newtonian potentials (Rome Univ, 250.000 hs)

DECI proposals - 3



- **MATERIALS SCIENCES**
- RZG – Catalysis by Ab-initio Metadynamics in Parallel (ETH – Zurich, 300.000 hs)
- RZG – Material functions in heterogeneous and environmental catalysis (MPI + Berlin Univ)
- RZG – Polymer simulations (MPI + Mainz Univ, 300.000 hs)
- RZG – Many particle dynamics in finite atomic systems (MPI + CNRS)
- CSC – High-k nanotubes FETs (Helsinki Univ, 100.000 hs)
- FZJ – Simulating complex molecular many body systems with CPMD (Univ. Bochum, 300.000 hs)
- CINECA – Surface structure and growth of deposited metal clusters on thin films (IPCF-CNR, 400.000 hs)

DECI proposals - 4



- **PLASMA PHYSICS**

- RZG – Gyrokinetic turbulence simulations (MPI-IPP + Lausanne)
- RZG – Magnetohydrodynamic turbulence (MPI-IPP + CNRS-Meudon, 350.000 hs)

- **QUANTUM CHROMODYNAMICS**

- FZJ – Lattice QCD Simulations (DESY, 350.000 hs)
- CSC – Hot and dense nuclear matter in particle physics and astrophysics (Univ. Oulu and Helsinki + CERN, 300.000 hs)

- **BIO-INFORMATICS**

- LRZ – Computation of a comprehensive tree of life from genetic information (TU Munich, Lehrstuhl für Rechner-technik, 300.000 hs)

DECI proposals - 5



- **CFD**
- IDRIS – Coupling of combustion and thermal radiation models (Ecole Centrale + CNRS, 500.000 hs)
- CINECA – Direct simulation of Lagrangian evolution of passive and heavy particle in a highly resolved turbulent field (IAC-CNR + Int. Coll, 400.000 hs)

- **OCEANOGRAPHY**
- IDRIS – High resolution modelling of the variability of the global ocean (Proj Drakkar, CNRS-LEGI, LPO, LODIC, 200.000 hs)

- **QUANTUM CHEMISTRY**
- IDRIS – Highly accurate potential surfaces for quantum dynamics simulations of triiodide-I3 (Univ Strasbourg, Amsterdam, Duseldorf,)



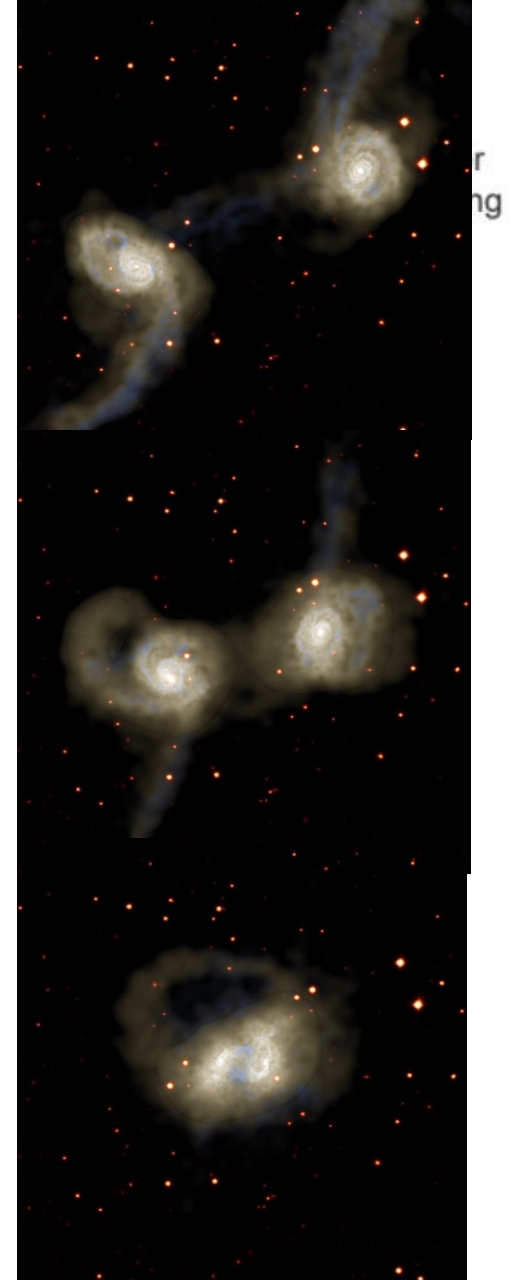
The Virgo Consortium

Simulating cosmic structure formation

Since 1996, the fore-front international consortium carrying out leading-edge simulations of the growth of structure in the Universe

Principal sites:

Max-Planck Institute for Astrophysics (D)
Institute for Computational Cosmology (UK)
Institute for Astronomy, Edinburgh (UK)
MacMaster University (CAN)
University of Victoria (CAN)
University of Michigan (US)



Extreme computing demands of the VIRGO Consortium



Theoretical input for the International Virtual Observatory, a global petabyte grid of observed and simulated data

The Millennium Run: the largest ever N-body simulation of cosmic evolution

>20 TB data output from a runs using 512 PEs
with 1 TB main memory, 10^{10} particles

Grid access and query enablement as a work package within the EURO-VO initiative

ESO (European Southern Observatory)

ESA (European Space Agency)

INSU/CDS (F) PPARC /Astrogrid (UK)

INAF/DRACO (I)

Max-Planck Society/BMBF/GAVO (D)

INTA/SVO(E) NOVA (NL)

Extreme computing runs proposed by VIRGO

As of June 2005

Ultra high resolution cosmological halos

Determining the nature of dark matter

10^9 particles, 200 GB main memory, 300 000 cpu-h

2H 2005

Hydrodynamical structure formation in the Universe

Precision measurements of cosmic parameters

10^{10} particles, 2 TB main memory, 300 000 cpu-h

2006/2007

Hydrodynamical structure formation in the Universe

Following the formation of galaxies

10^{11} particles, 20 TB main memory

Proposals from Materials Science

M. Parrinello, ETH Zurich

Studies of Catalysis by Ab-initio Metadynamics in Parallel

Combined approach using CPMD and **new application Quickstep** with which all-electron DFT calculations are feasible using an extension of the GPW method, the so-called Gaussian augmented and plane waves (GAPW)

Support of Quickstep within DEISA is in preparation

Proposals from Materials Science

M. Scheffler, Fritz-Haber-Institute, Berlin

***First principles statistical mechanics for material functions
in heterogeneous and environmental catalysis***

using a powerful new approach

“The ability to predict and modify the rate-determining steps in chemical reactions would be a boon in designing better catalysts. Technical innovations in computer simulations bring that goal closer. Writing in Physical Review Letters, Reuter et al. [1] herald a powerful new approach that should ultimately help make prediction possible.”

NEWS and VIEWS, NATURE VOL 432, 18 NOV 2004

Proposals from Materials Science

K. Kremer, Max Planck Institute for Polymer Research, Mainz

***proposes several projects using different applications,
among them***

Study of Long Range interactions in fluid membranes

Approach by new package ESPResSO

Extensible Simulation Package for Research on Soft matter

Support of ESPResSO within DEISA is in preparation

Proposals from Plasma Physics

Strong European Efforts to contribute to
Numerical Tokamak and to Numerical Stellarator

Newest developments will allow turbulence
simulations in Tokamak research, and will also
be useful for Stellarator

Additional developments will allow methodologically
complementary turbulence simulations for Tokamak

Principle Investigators in DEISA JRA Plasma Physics:

K. Lackner, Max-Planck Institute of Plasma Physics, Garching
J. Nuehrenberg, Max-Planck Institute of Plasma Physics, Greifswald
L. Villard, Centre des Recherches de Physiques des Plasmas, Lausanne)

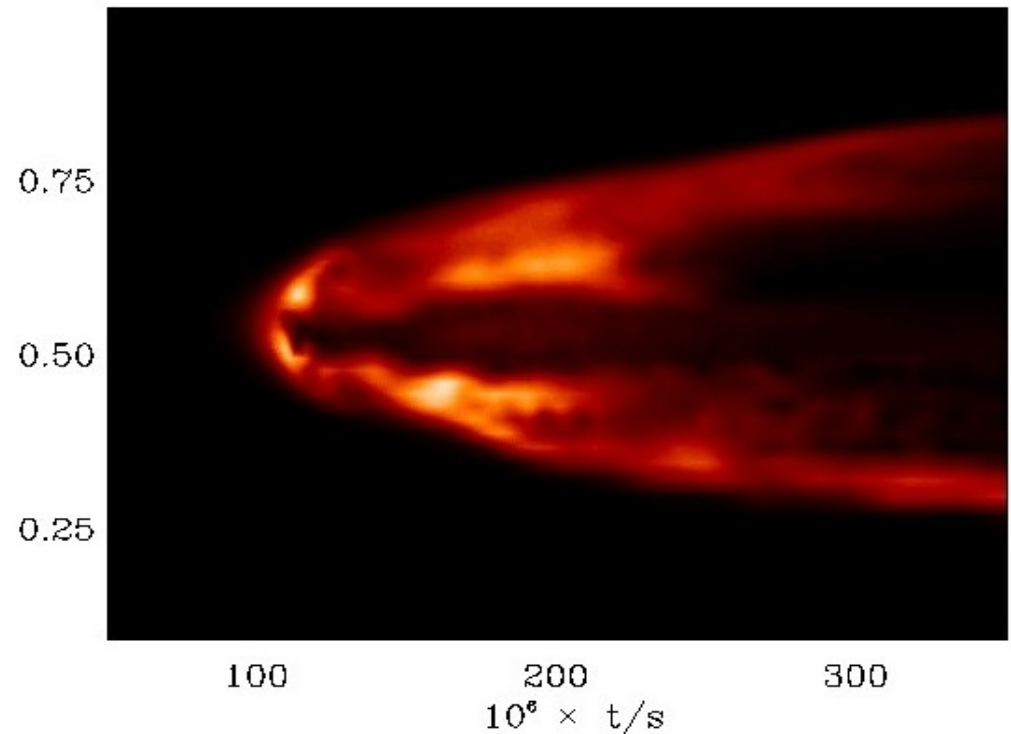
Proposals from Plasma Physics

Extreme Gyrokinetic Turbulence Simulations

The gyrokinetic simulation model gives a quantitative correct description of a large variety of plasma physics effects.

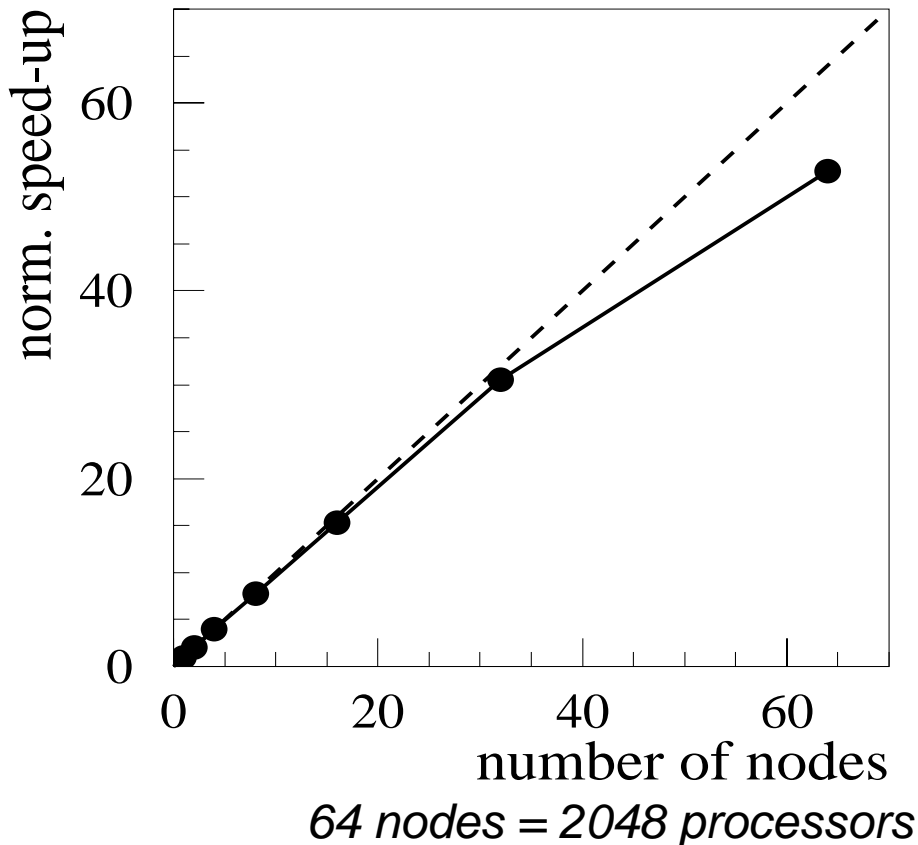
The excellent physical properties of this approach result, especially for turbulence simulations, in very demanding computational efforts.

Time development of the
turbulent radial heat flux
Code: TORB



Proposals from Plasma Physics

Extreme Gyrokinetic Turbulence Simulations



The nonlinear particle-in-cell code TORB uses a Monte Carlo particle approach to simulate the time evolution of turbulent field structures in fusion plasmas

(J. Nuehrenberg, IPP, Greifswald & L. Villard, CRPP, Lausanne)

Within DEISA, TORB has been improved for extreme scalability at IBM system at ECMWF:

On 2048 procs:

Speedup = 1680 Parallel efficiency = 82%

Sustained performance = 1.3 TF

Conclusions



- **DEISA adopts Grid technologies to integrate national supercomputing infrastructures, and to provide an European Supercomputing Service.**
- **Service activities are supported by the coordinated action of the national center's staffs. DEISA operates as a virtual European supercomputing centre.**
- ***The big challenge we are facing is enabling new, first class computational science.***
- ***Integrating leading supercomputing platforms with Grid technologies may help to create a new research dimension in Europe.***