Distributed European Infrastructure for Supercomputing Applications

DEISA eScience Applications

Hermann Lederer
Garching Computing Centre of the Max Planck Society
www.rzg.mpg.de

Distributed European Infrastructure for Supercomputing Applications

The DEISA Project

FP6 EU Integrated Infrastructure Initiative (I3) Project

Contract no. 508803

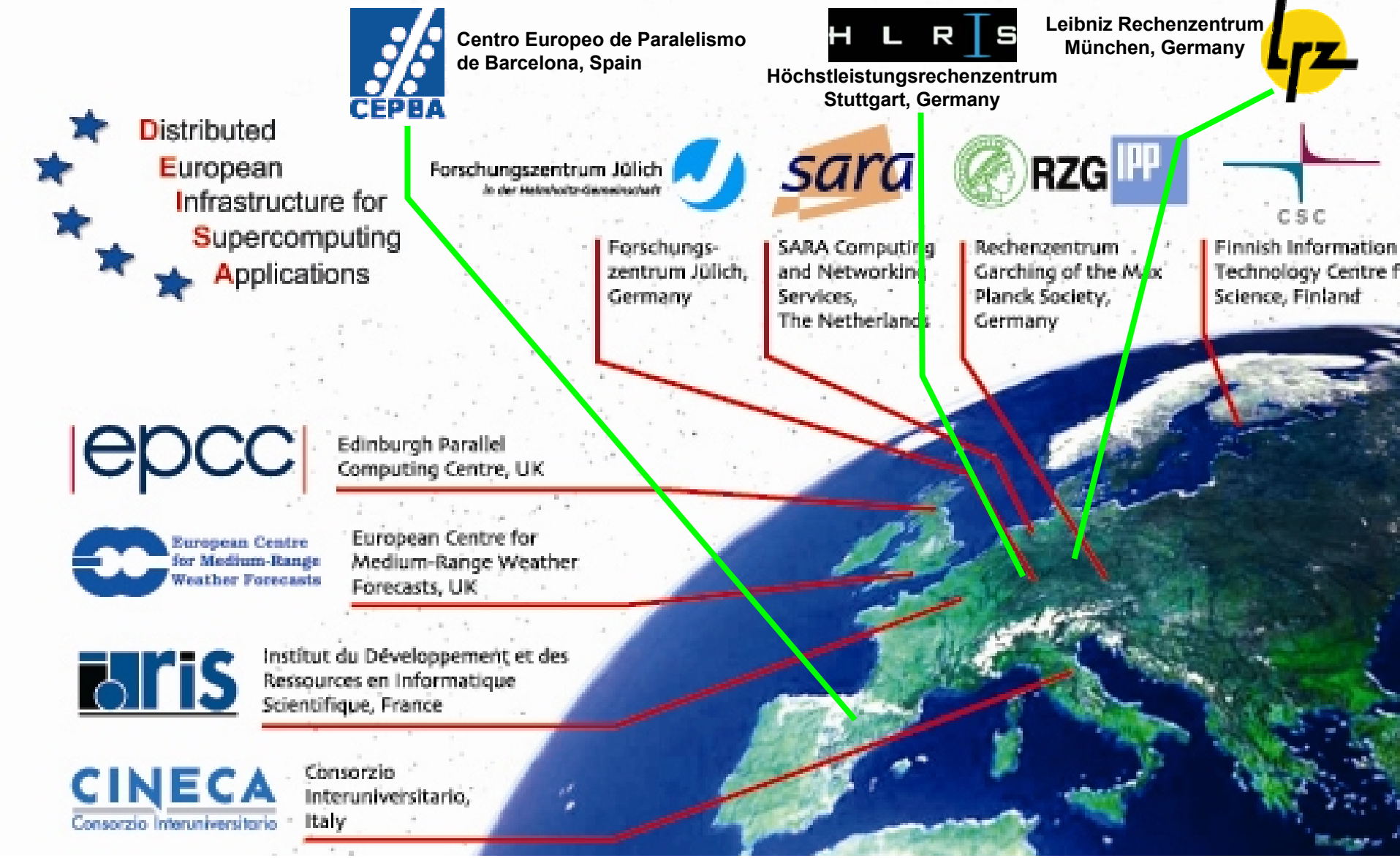
Project start: May 1, 2004
Project duration: 5 years

Distributed European Infrastructure for Supercomputing Applications

DEISA Objectives

- *To enable scientific discovery across a broad spectrum of science and technology*
- *Advancement of computational sciences in supercomputing in Europe*

- *Collaboration of leading supercomputing centres in Europe*
- *Joint deployment and operation of a persistent, production quality, distributed supercomputing environment with continental scope.*
- *European supercomputing service built on top of existing national services*
- *Integration of Europe's most powerful supercomputing systems.*



Supercomputing Centres in DEISA

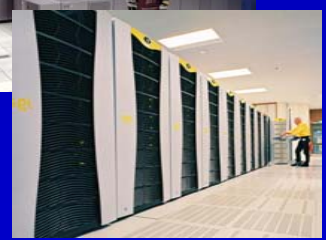
Distributed European Infrastructure for Supercomputing Applications

THE DEISA SUPERCOMPUTING GRID

AIX distributed super-cluster

GEANT

Vector systems (NEC, …)



Linux systems (SGI, IBM, …)

Distributed European Infrastructure for Supercomputing Applications

System Integration, deployment and operation of the infrastructure

- SA in Networking
- SA in Global File Systems
- SA in Middleware and Ressource Management
- SA in User Support
- SA in Security
- JRA in Heterogeneous resource management (JRA7)

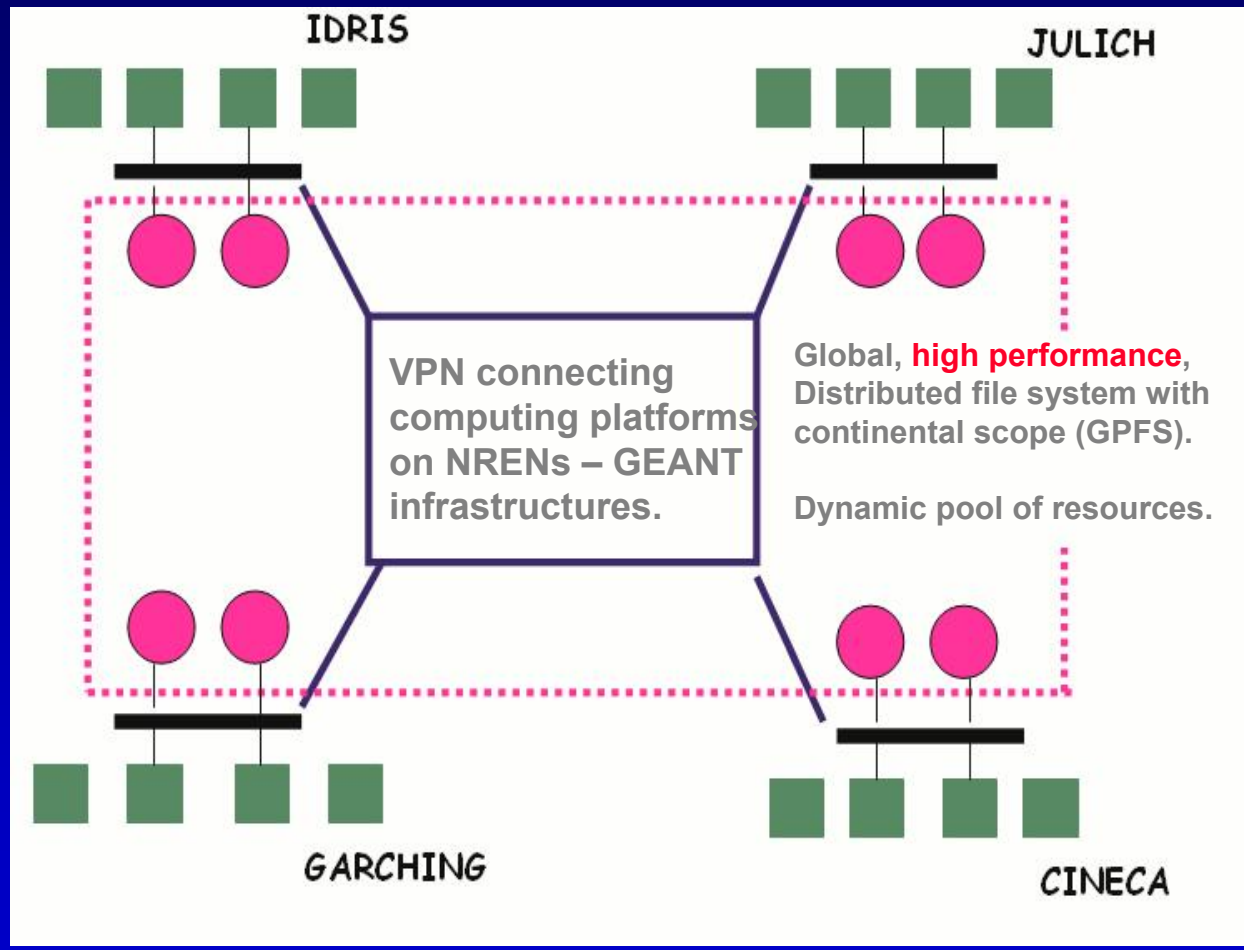
Distributed European Infrastructure for Supercomputing Applications

Network

Virtual private network via GEANT, national access via NRENs

Start with 1 Gb/s connection for 4 sites in July 2004

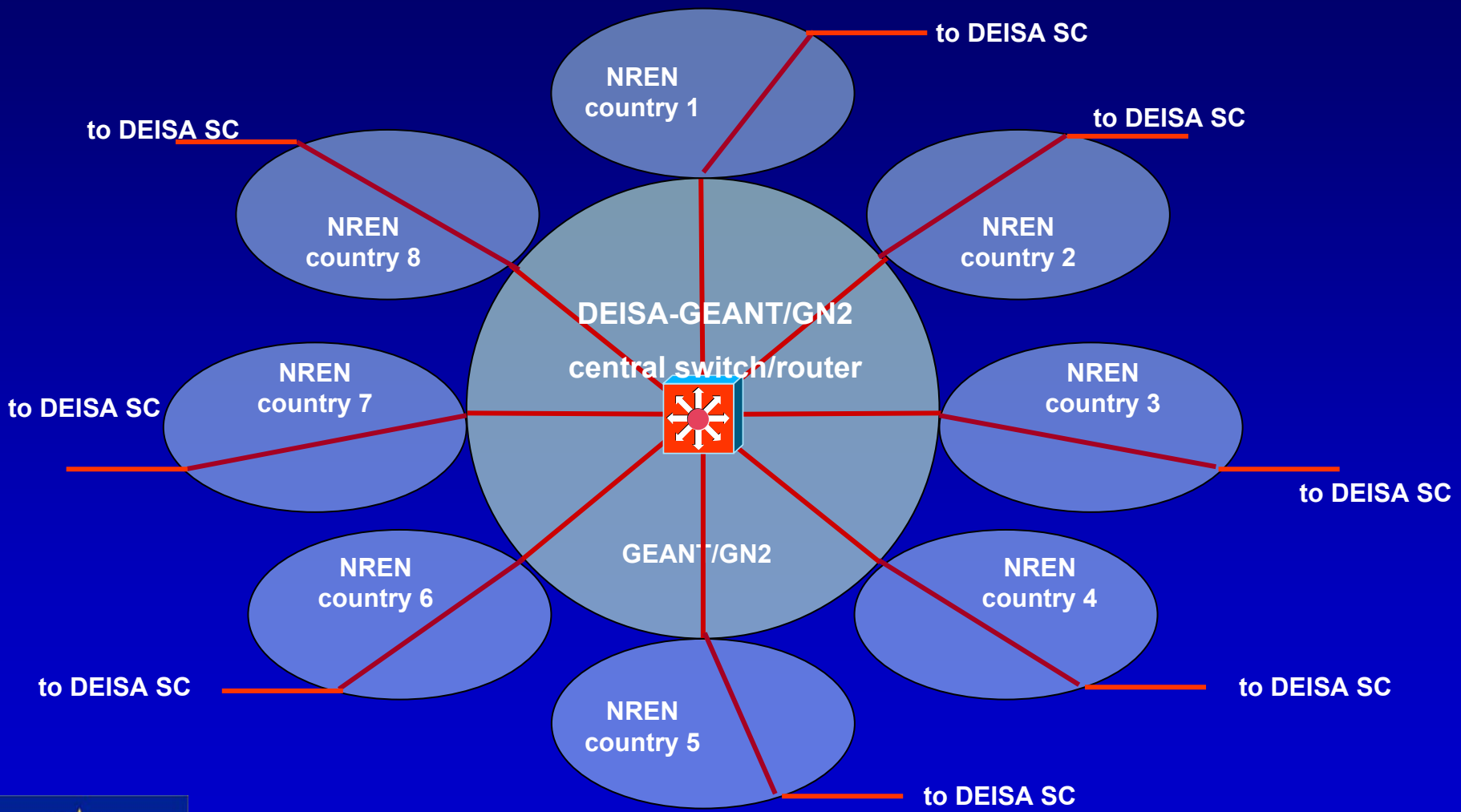
Extension to 10 Gb/s in 2006 in phase with GEANT2



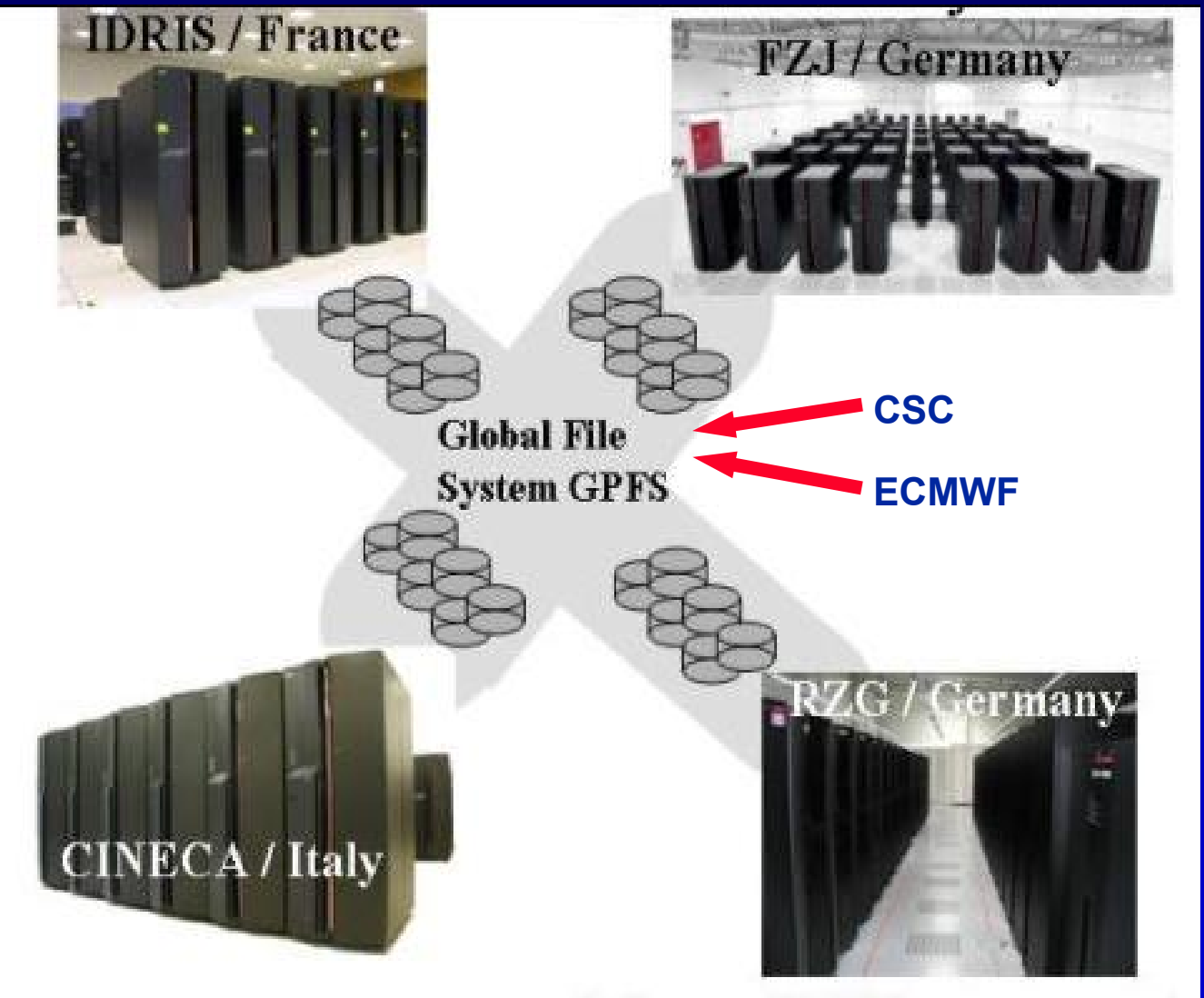
The DEISA super-cluster (phase 1)

Distributed European Infrastructure for Supercomputing Applications

Logical view of DEISA network



Global File System



First global file system wordwide in production environments at continental scope

AIX SUPER-CLUSTER JUNE 2005

Distributed European Infrastructure for Supercomputing Applications

GPFS demo at European Scale *(June 16, 2005, Paris)*

1) A 256 processor job is running at RZG (Germany). The data for this run have been read from IDRIS (France), where hey have been produced by a former job (No ftp, transparent for the user as in a local cluster)

2) The application writes the restart data to the Jülich GPFS file system and result data to RZG

3) The results are interpreted and displayed in CINECA, reading the data from RZG (Germany)

Distributed European Infrastructure for Supercomputing Applications

eScience Applications

Applications – what the infrastructure is being built for: scientific research via applications

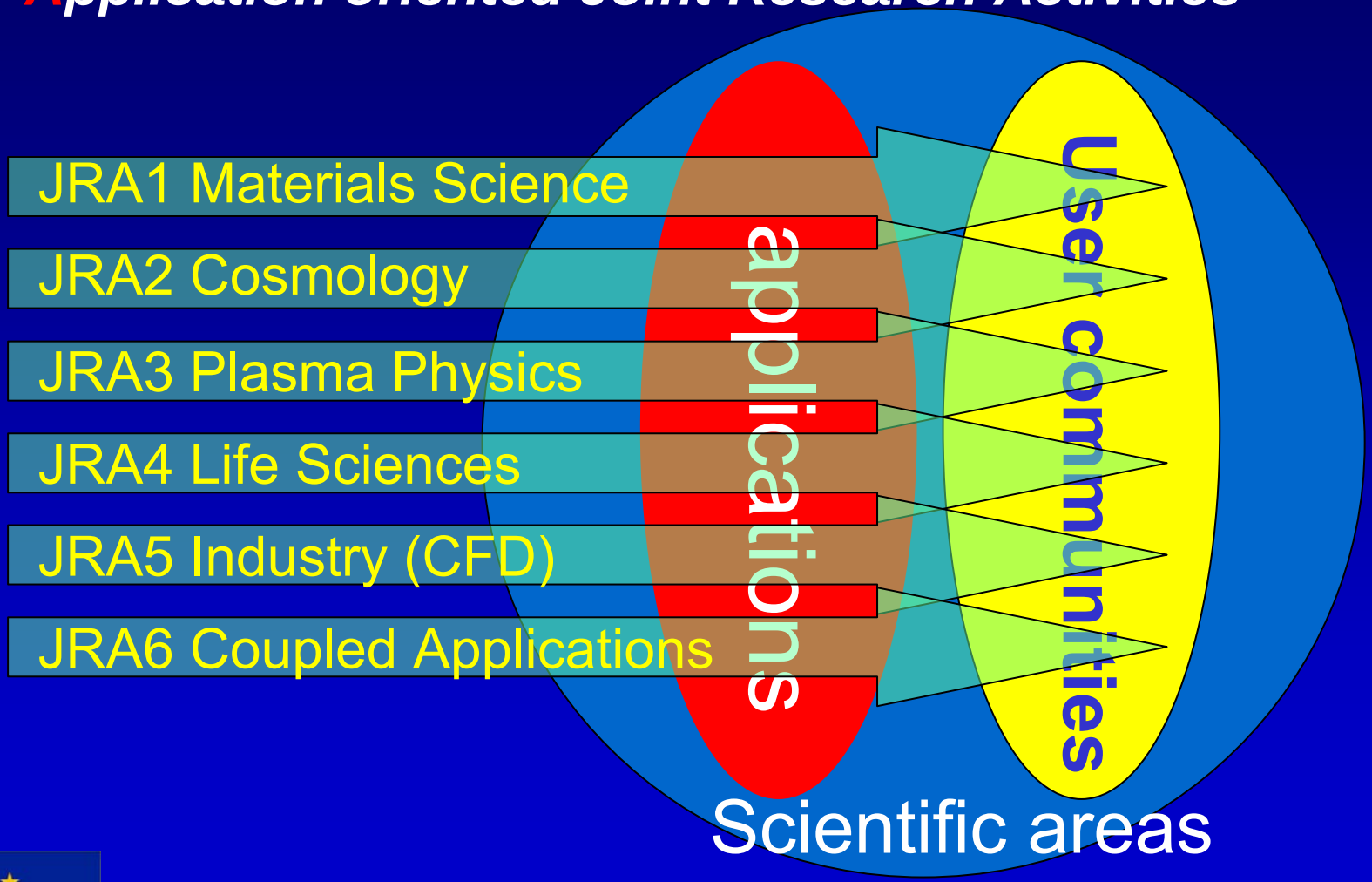
Applications – interfaces *to* the scientific communities and *among* scientists

Applications – driving forces during infrastructure build-up phase (and expansion phases)

Distributed European Infrastructure for Supercomputing Applications

eScience Applications

Application oriented Joint Research Activities



Distributed European Infrastructure for Supercomputing Applications

eScience Applications

Application oriented Joint Research Activities used as "vehicles" for the different scientific disciplines

- to build pioneer applications for the new infrastructure
- to establish contacts to the scientific user communities and explore their needs
- to pave the way for big projects
- to guide/accompany the build-up phase of the infrastructure

Distributed European Infrastructure for Supercomputing Applications

eScience Applications

JRA1 Materials Science

Scientists from

ETHZ, Zurich

Fritz-Haber-Institute, Berlin

University College London

Max Planck Institute for Polymer Research, Mainz

CPMD

WIEN2K

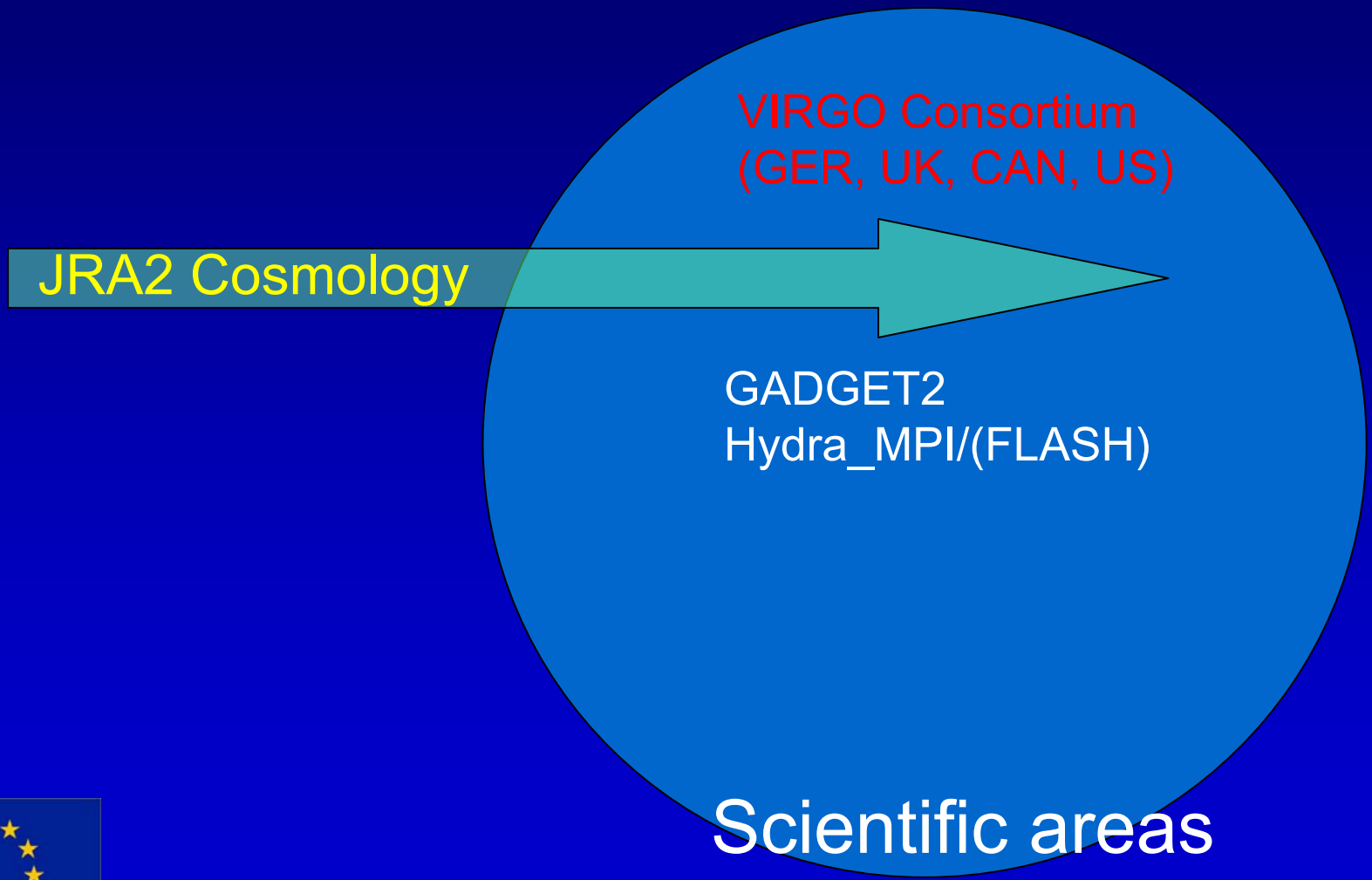
ESPREsSo

(NWCHEM)

Scientific areas

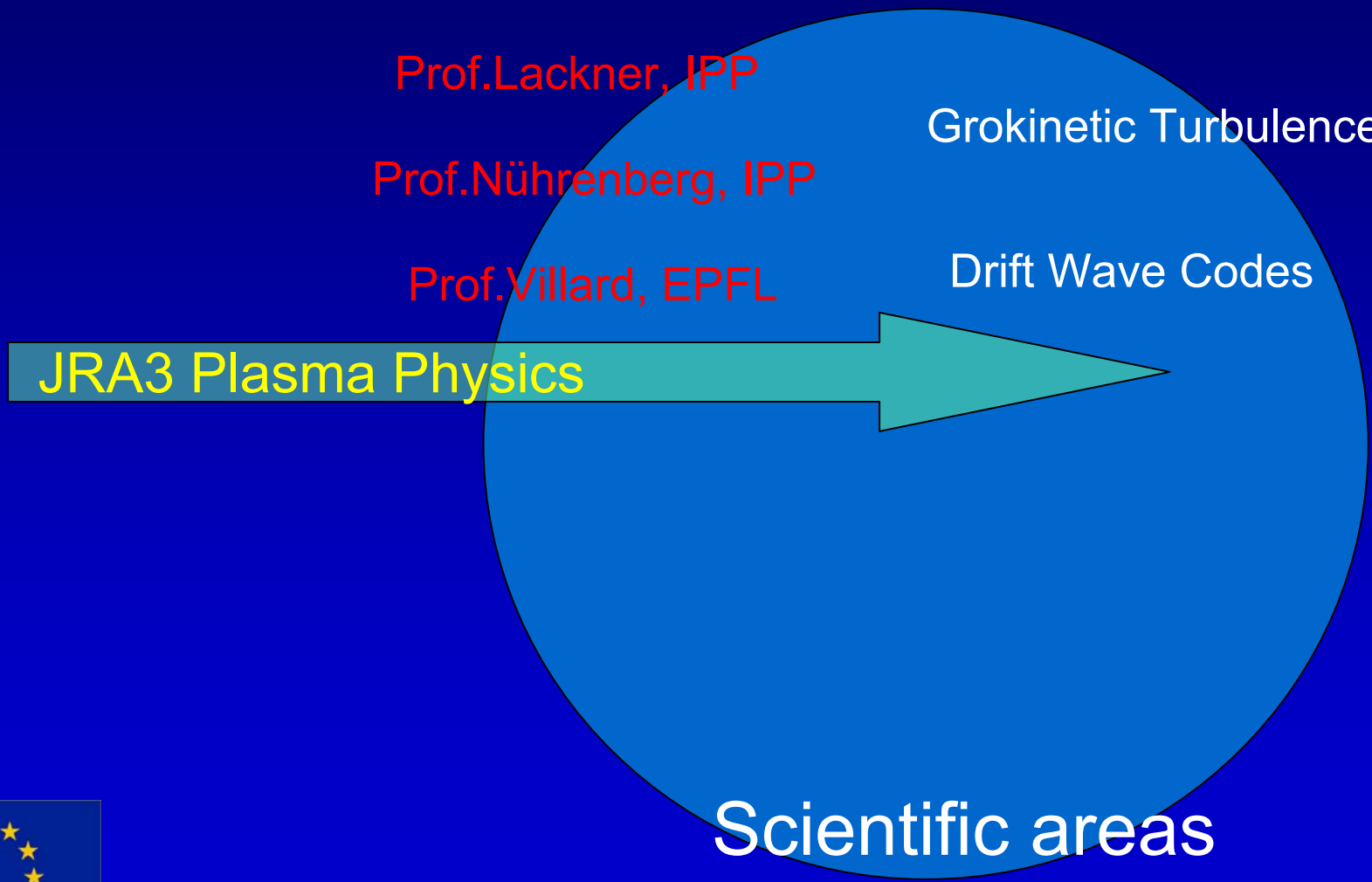
Distributed European Infrastructure for Supercomputing Applications

eScience Applications



Distributed European Infrastructure for Supercomputing Applications

eScience Applications



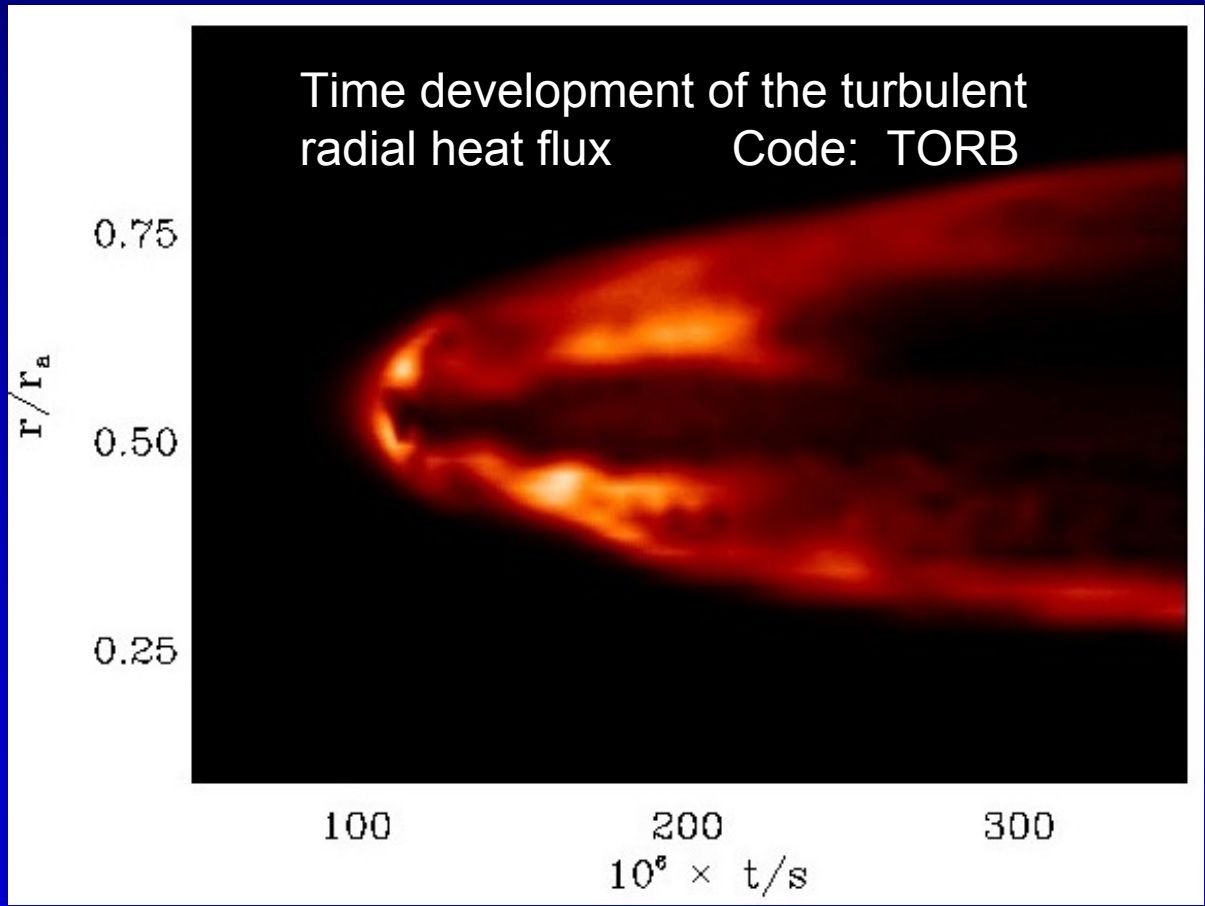
Distributed European Infrastructure for Supercomputing Applications

Extreme Gyrokinetic Turbulence Simulations

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.

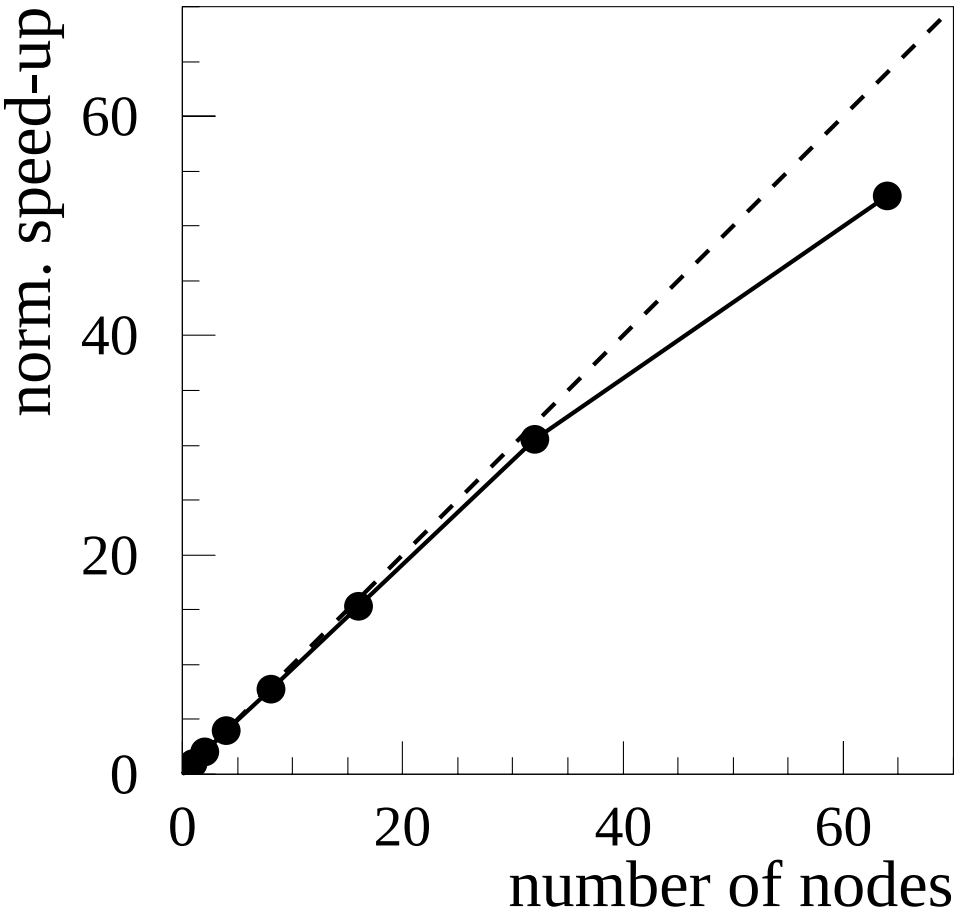
The gyrokinetic simulation model gives a quantitative correct description of a large variety of plasma physics effects.

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



Distributed European Infrastructure for Supercomputing Applications

Extreme Gyrokinetic Turbulence Simulations



64 nodes = 2048 processors

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:

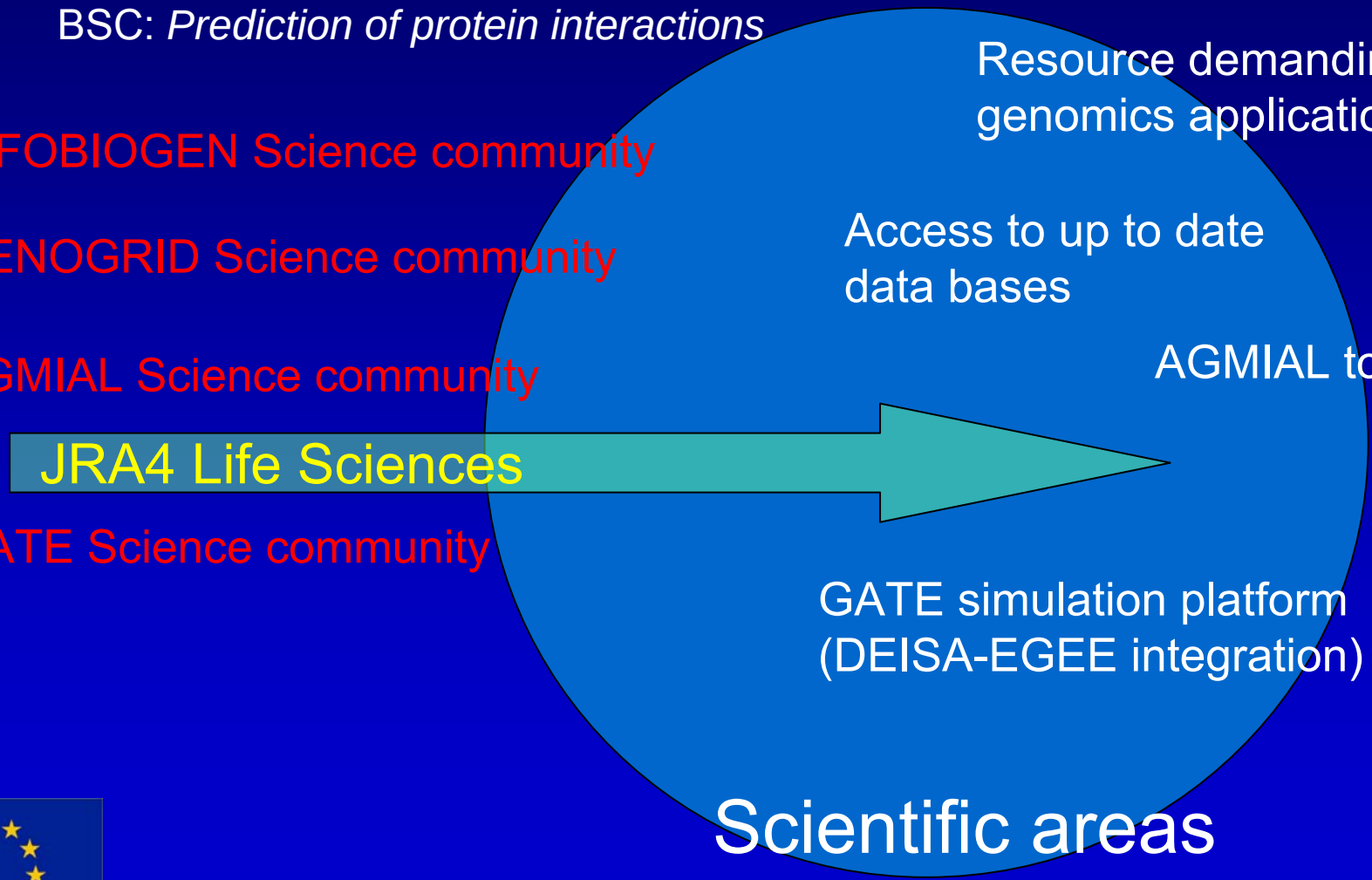
On 2048 procs:

Speedup = 1680
Parallel efficiency = 82%
Sustained performance = 1.3 TF

Distributed European Infrastructure for Supercomputing Applications

eScience Applications

INSERM: *Identification of new human mithocondrial proteins*
INRA : *Large scale microbial genome reanotation*
BSC: *Prediction of protein interactions*



Distributed European Infrastructure for Supercomputing Applications

Radiation Therapy planning

Grid access

- Grid authorization
- Revocation of grid authorization
- Grid login

Registration and management of medical images

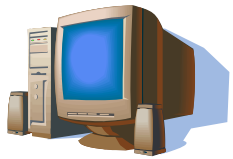
- Registration and replication anonym medical images
- One logical file name corresponding to multiple physical images
- Automatic adding and suppression of images

Web portal

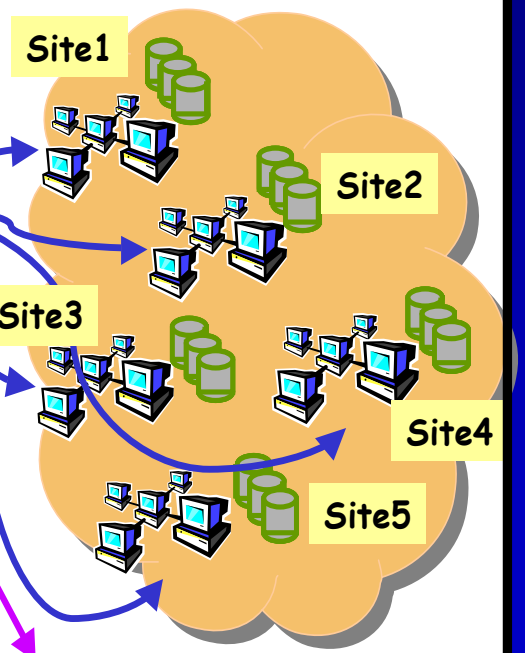
Jobs submissions and management

- User secured authentication
- Splitting of simulations
- Automatic files creation submitted to the grid
- Jobs submission
- Jobs management
- Automatic data retrieving (spectrum, isodoses…)

Internet conexion



Working station
In a clinical centre



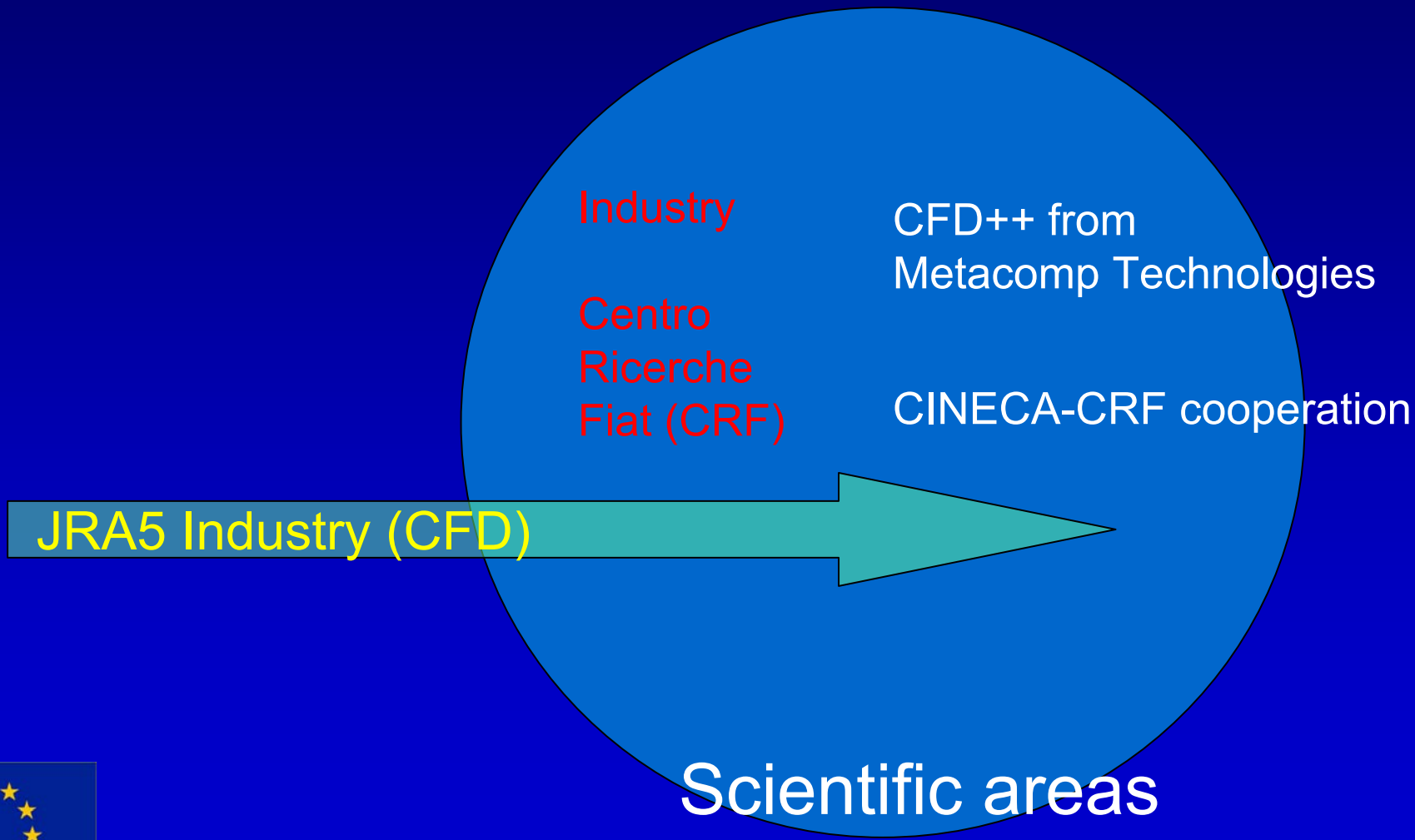
Supercomputing access

- Grid access
- User access
- Parallel calculations

DEISA Joint Application with EGEE

Distributed European Infrastructure for Supercomputing Applications

eScience Applications



Distributed European Infrastructure for Supercomputing Applications

eScience Applications

Environmental science
Prof. Vauclin, Dr. Messager,
LTHE + LGGE, HSM, IRM,
LEGOS,LEGI, U Lancaster

Coupled application from project "Hydrological Cycles over West African Continent" (HYCYMAC)

Combustion
Prof. Veynante, EM2C
+ CERFACS/IMFT

Large Eddy Simulations of turbulent combustion including pollutant species prediction and radiative heat transfer
- combustion/pollution coupling
- combustion/radiative coupling

Astrophysics
Prof. Alimi, LUTH
+ U Elche(ES) + CEA

Gravity-Hydrodynamics-Chemistry coupling for star formation

JRA6 Coupled Applications

Scientific areas

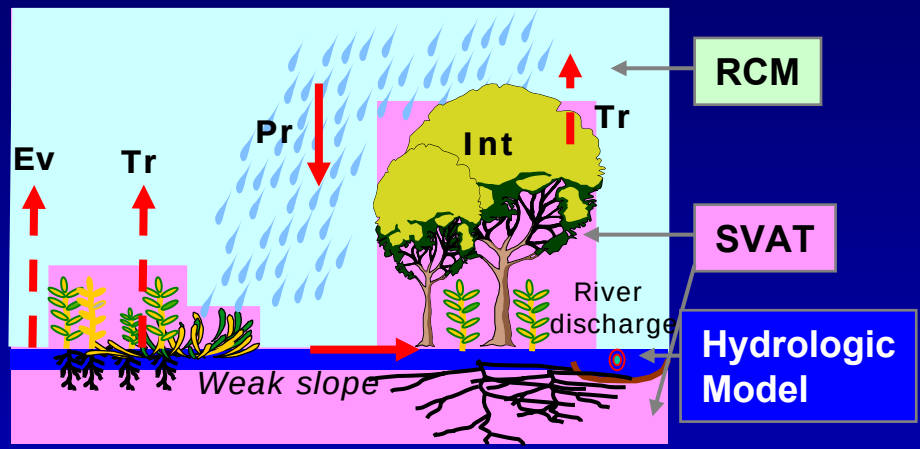
Distributed European Infrastructure for Supercomputing Applications

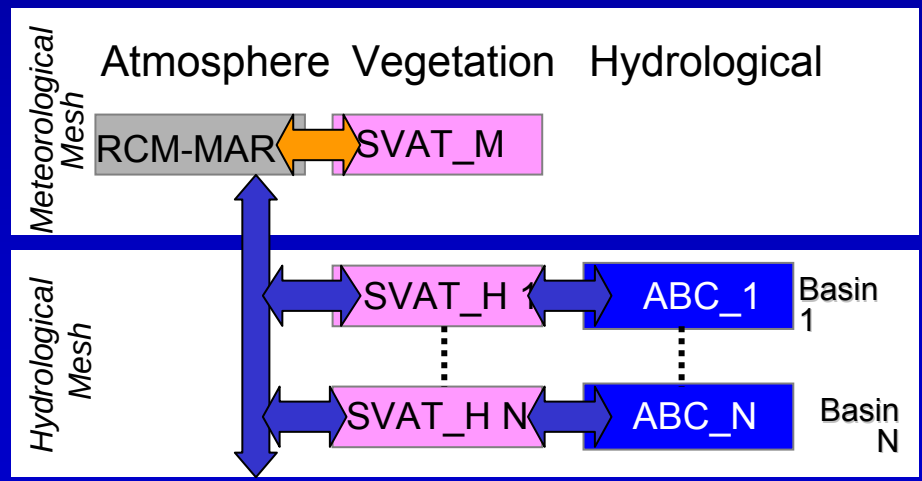
Coupled applications : Environment

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).

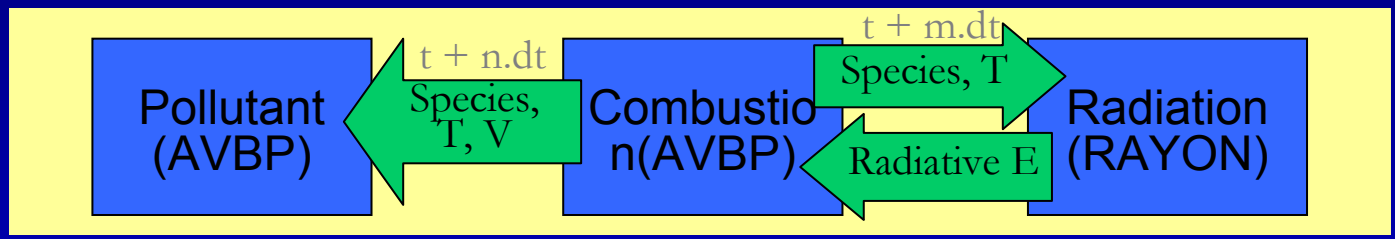




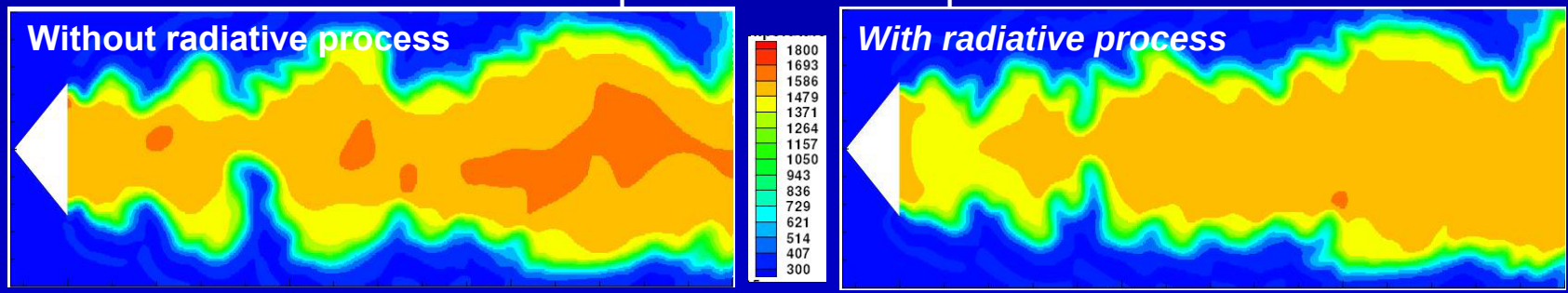
Distributed European Infrastructure for Supercomputing Applications

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

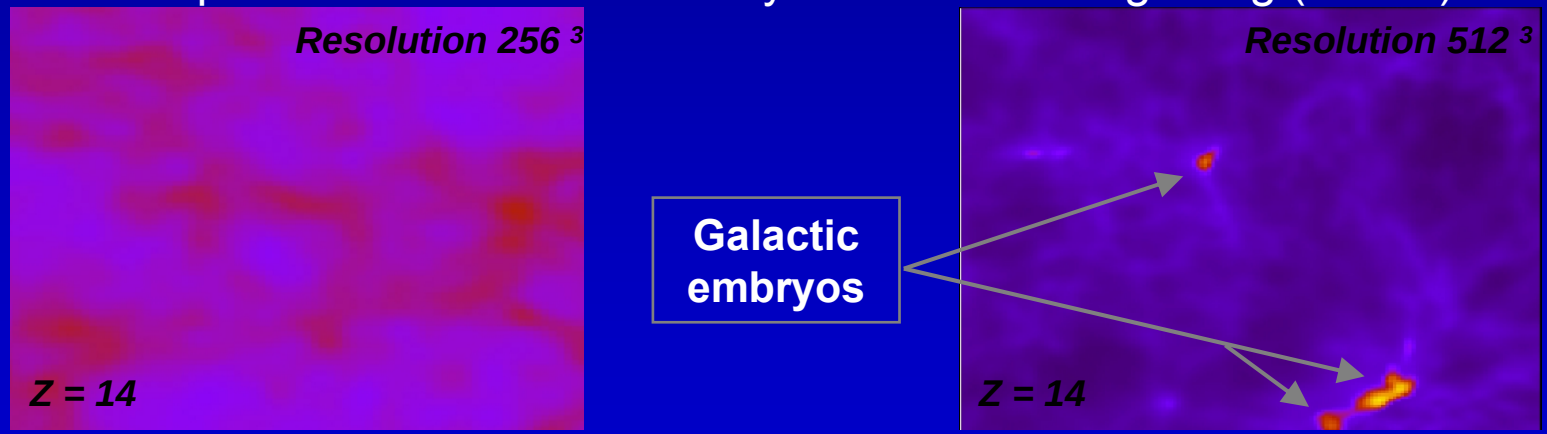
Distributed European Infrastructure for Supercomputing Applications

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.

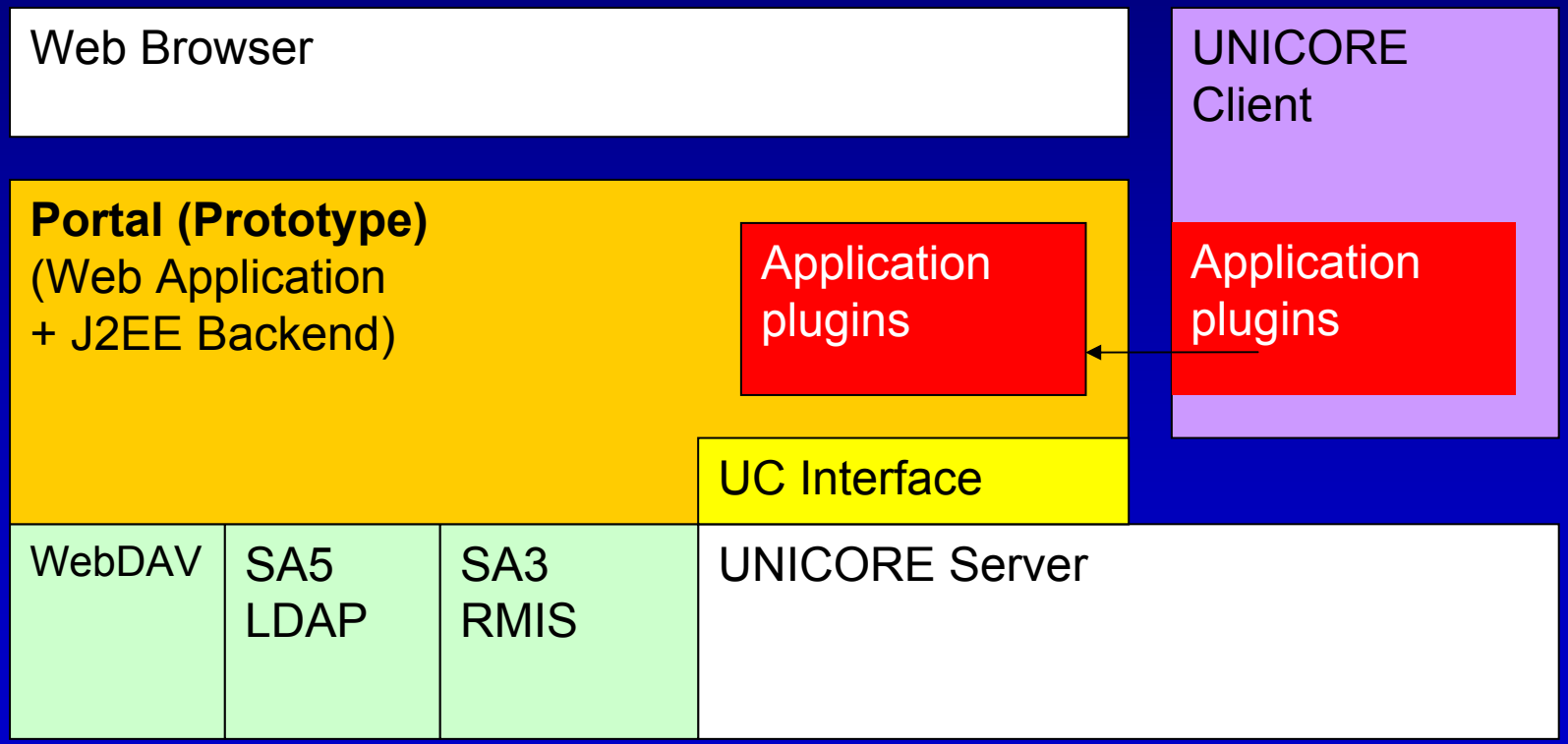
Distributed European Infrastructure for Supercomputing Applications

Portals / Science Gateways

- Hiding complex supercomputing environments from end users, providing discipline specific tools and support, and moving in some cases towards community allocations for anonymous users.
- Needed to enhance the outreach of supercomputing infrastructures
- There is already work done by DEISA on Genomics and Material Sciences portals
- Similar concepts as TeraGrid's Science Gateways

Portals / Science Gateways

Example: Materials science / application-plugins



DEISA

DISTRIBUTED EUROPEAN INFRASTRUCTURE FOR SUPERCOMPUTING APPLICATIONS

Alexandria

The Science Gateway

Welcome

to the DEISA's Application Portal of the Joint Research Activities (JRAs)

- **Material Sciences (JRA1)**

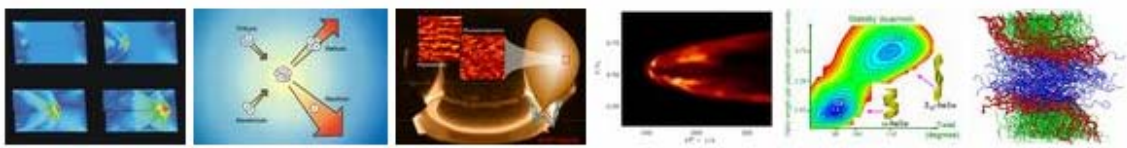
  Physical, chemical, and biological processes for many problems in computational physics, biology, and materials sciences span length and time scales of many orders of magnitude. For example, on the microscopic level, the typical bond distance between atoms is of the order of Angstroms (the lattice constant). More ...

- **Plasma Physics (JRA3)**

  Research on magnetic confinement fusion has undergone large changes during the last decade, moving away from the semi-empirical, predominantly experiment-driven approach to one accompanied and supported in all areas by first-principle based modelling. This development has been particularly dramatic in the area of turbulence. More ...

Access via this Portal Application or the UNICORE client

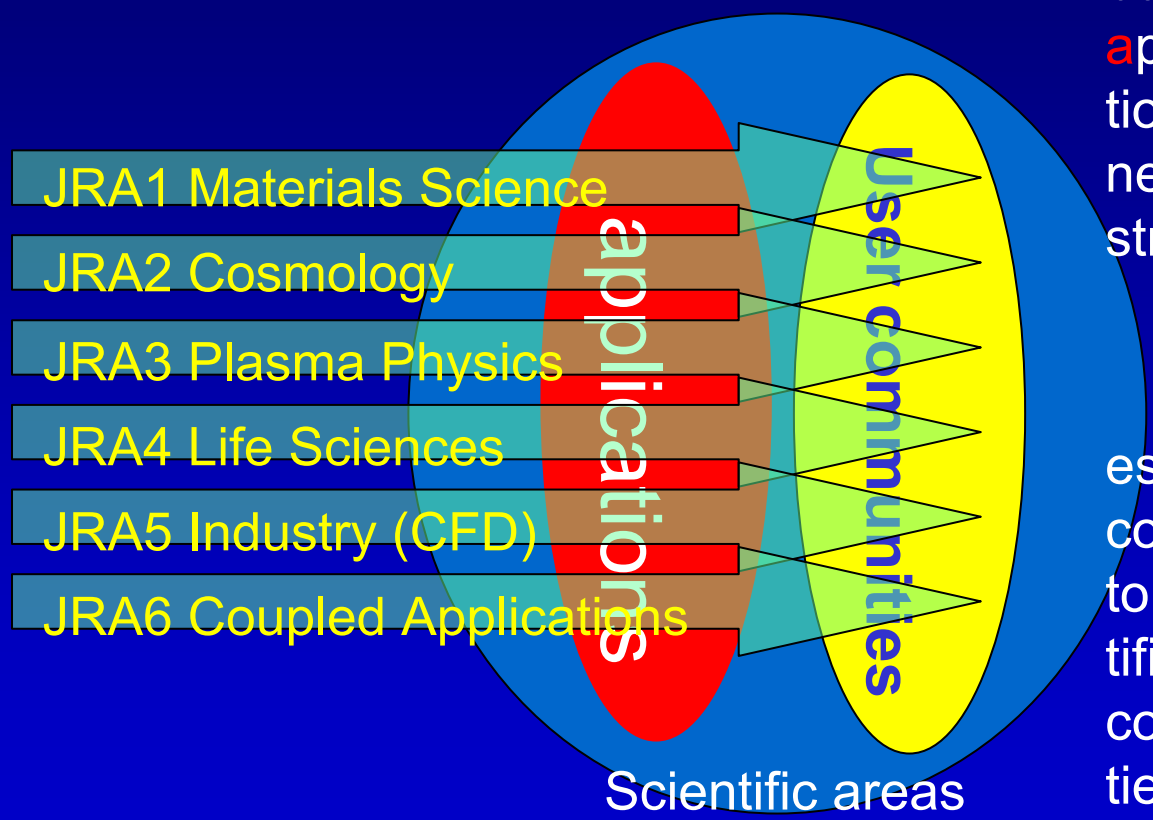
The middleware infrastructure provided to submit job to the DEISA environment is based on UNICORE.



Distributed European Infrastructure for Supercomputing Applications

eScience Applications

Application oriented Joint Research Activities



build pioneer applications for the new infrastructure

establish contacts to the scientific user communities and explore their needs

pave the way for big European projects

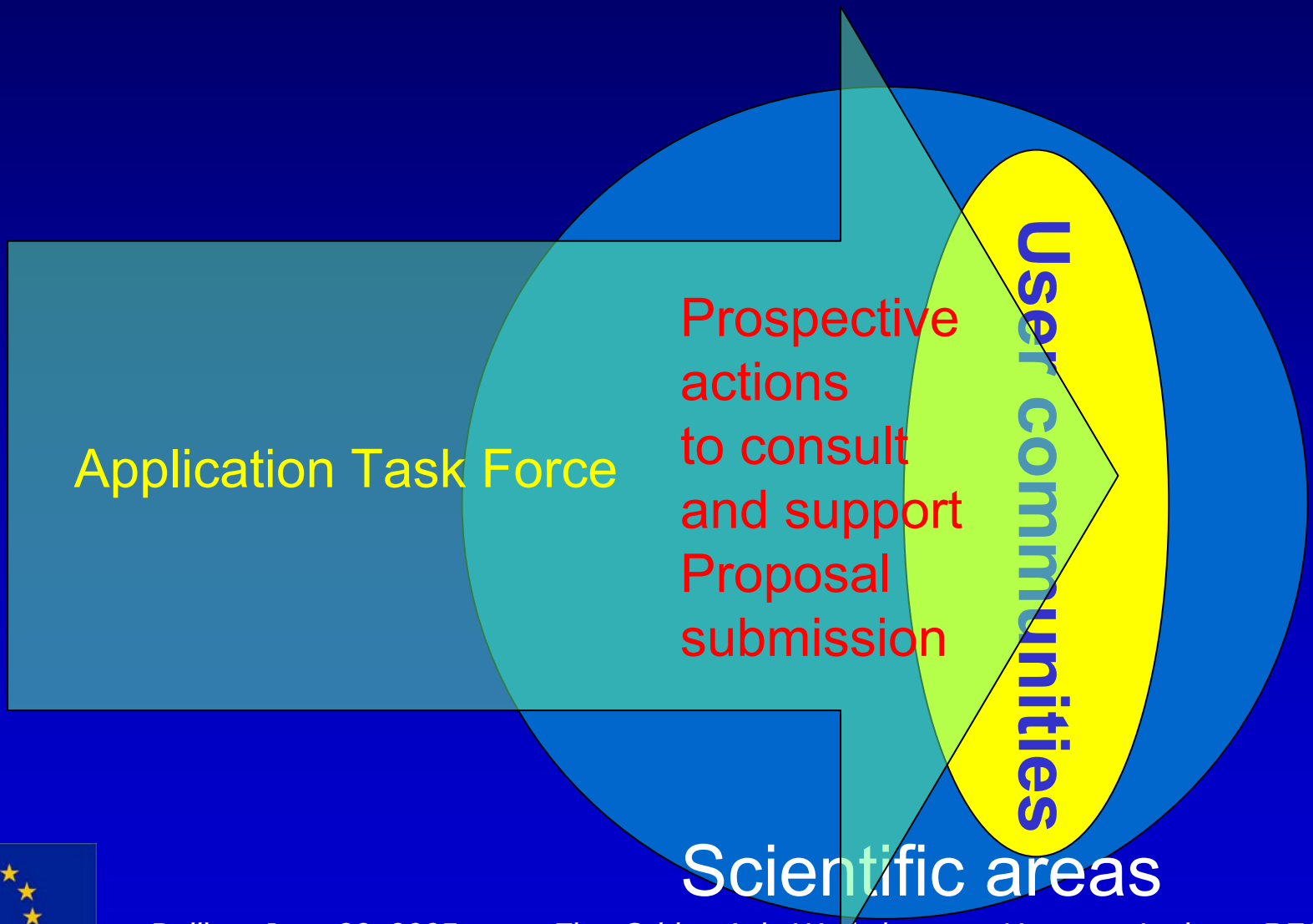
Distributed European Infrastructure for Supercomputing Applications

DEISA Extreme Computing Initiative DECI

- 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. 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

Distributed European Infrastructure for Supercomputing Applications

DEISA Extreme Computing Initiative DECI



Distributed European Infrastructure for Supercomputing Applications

DEISA Extreme Computing Initiative DECI

Call for Expressions of Interest / Proposals in April and May 2005

More than 50 proposals by May 30, 2005 from all scientific areas

Requested CPU time: more than 30 million cpu-h

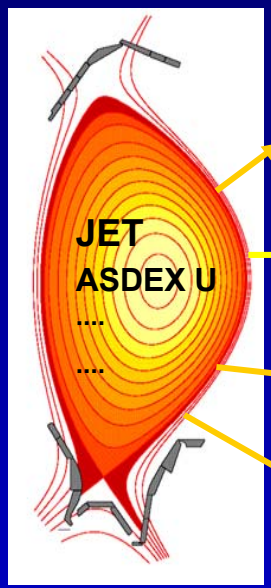
European countries involved:

Finland, France, Germany, Greece, Hungary, Italy, Netherlands, Russia, Spain, Sweden, Switzerland, UK

Distributed European Infrastructure for Supercomputing Applications

Extreme computing by German and Swiss theoretical Plasma Physicists

Fusion Theory (and Physics) Roadmap (courtesy of K. Lackner, IPP)

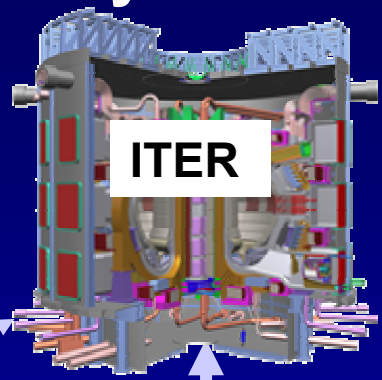


linear & nonlinear MHD instabilities)

fast particle driven instabilities

turbulent transport

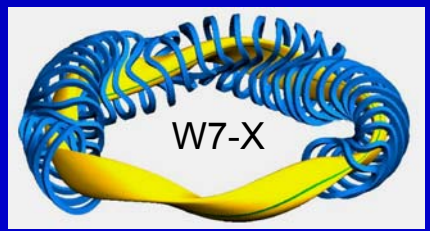
plasma-wall interaction:



DEMO

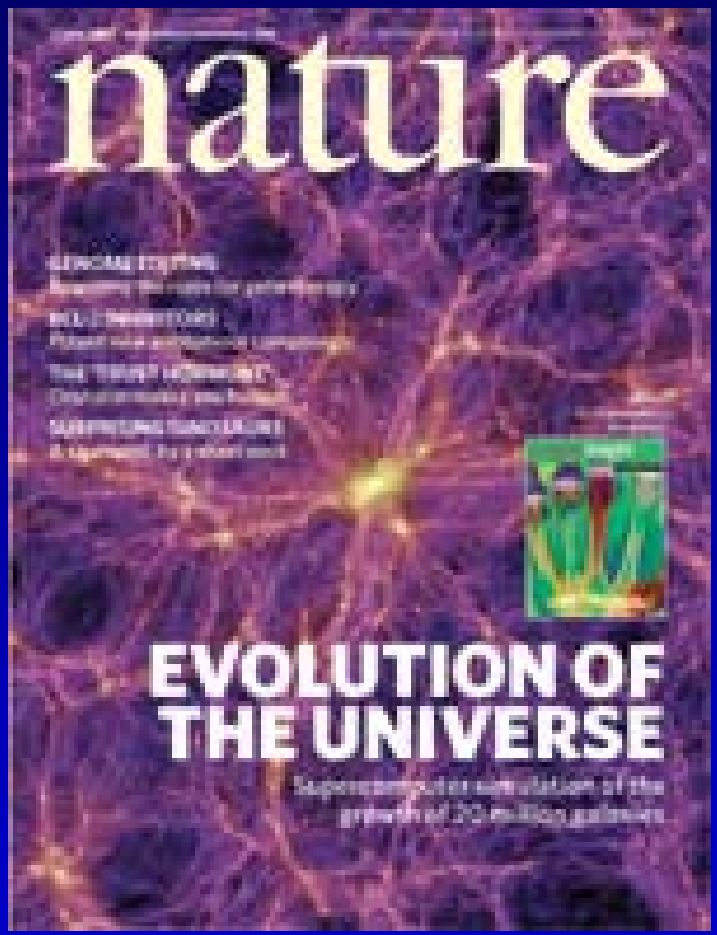
Numerical TOKAMAK

Numerical STELLAR.



Distributed European Infrastructure for Supercomputing Applications

Extreme computing by the VIRGO Consortium (GER, UK, CAN, US)



Volume 435, Number 7042
pp537-712 2 June 2005

Simulating cosmic structure formation

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

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

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