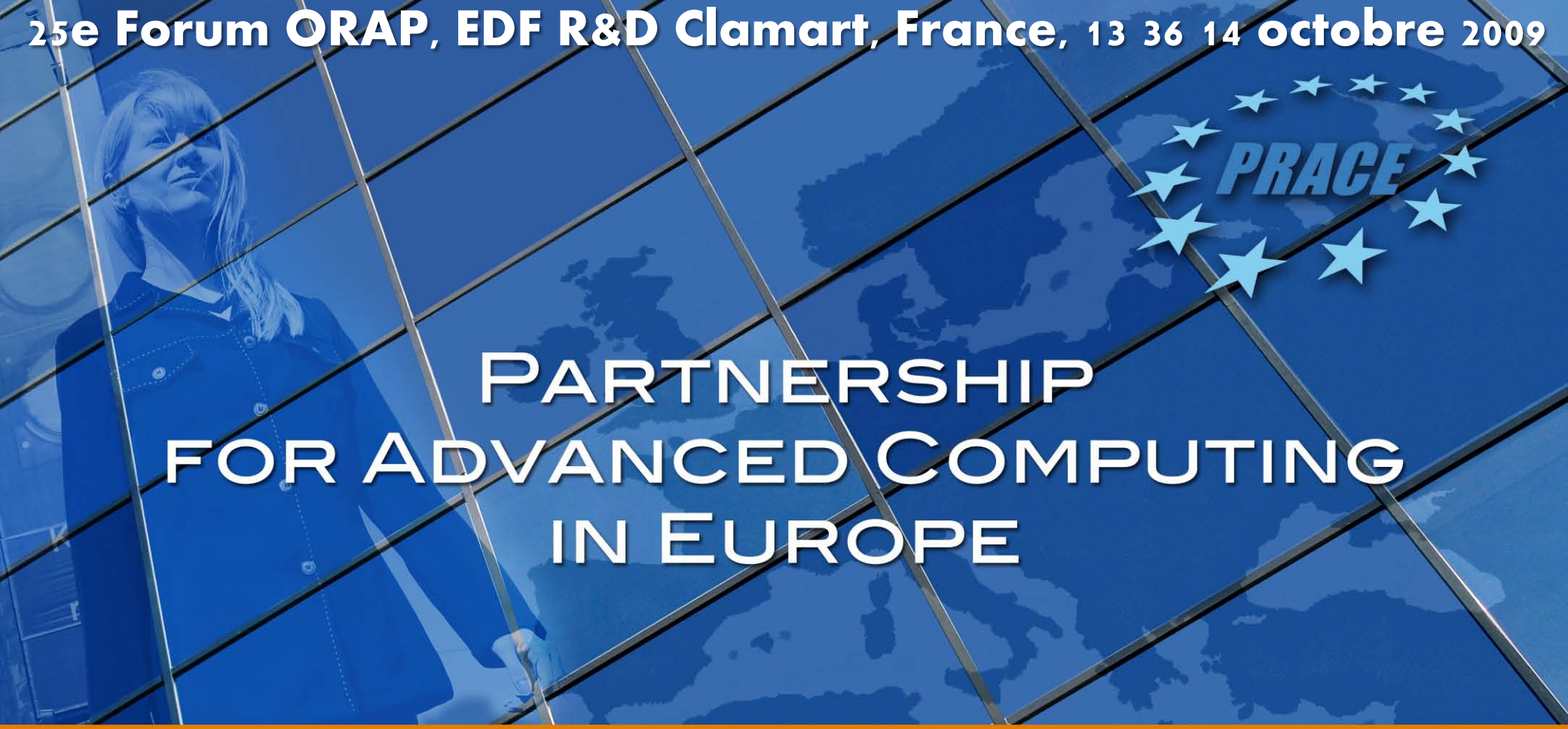
25e Forum ORAP, EDF R&D Clamart, France, 13 36 14 octobre 2009



Creating a European Supercomputing Infrastructure

Thomas Lippert, Forschungszentrum Jülich, Germany

CAPACITIES

e-infrastructure

Regards
from

Prof. Dr.
Achim Bachem

Coordinator
of the
PRACE Project

Chair of
Research Centre
Jülich

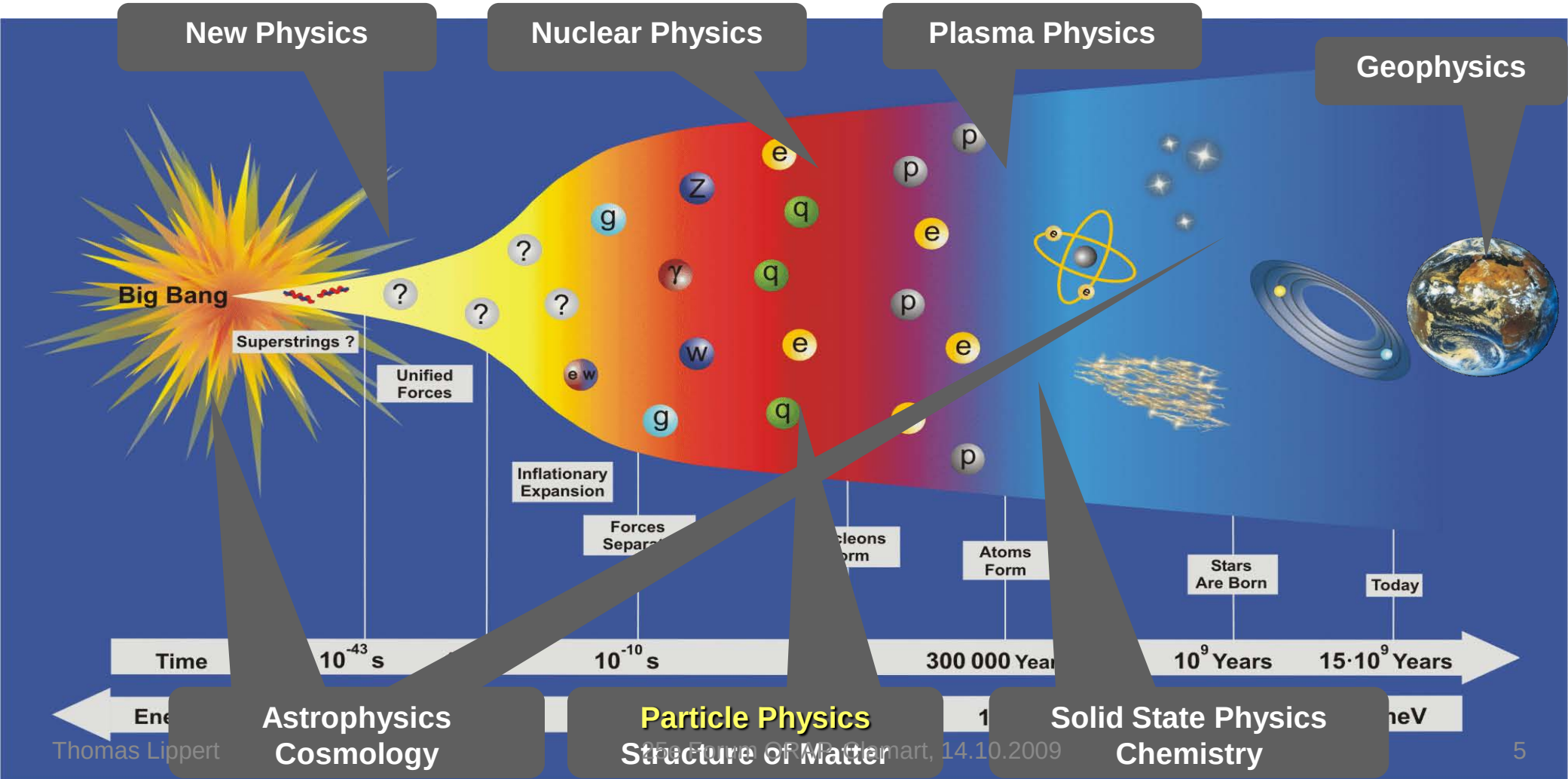
Outline

- A European SC RI: Why?
- What is PRACE?
- Where do we stand?
- What comes next?
- Questions

Outline

- A European SC RI: Why?
- What is PRACE?
- Where do we stand?
- What comes next?
- Questions

Supercomputing Drives Basic Sciences

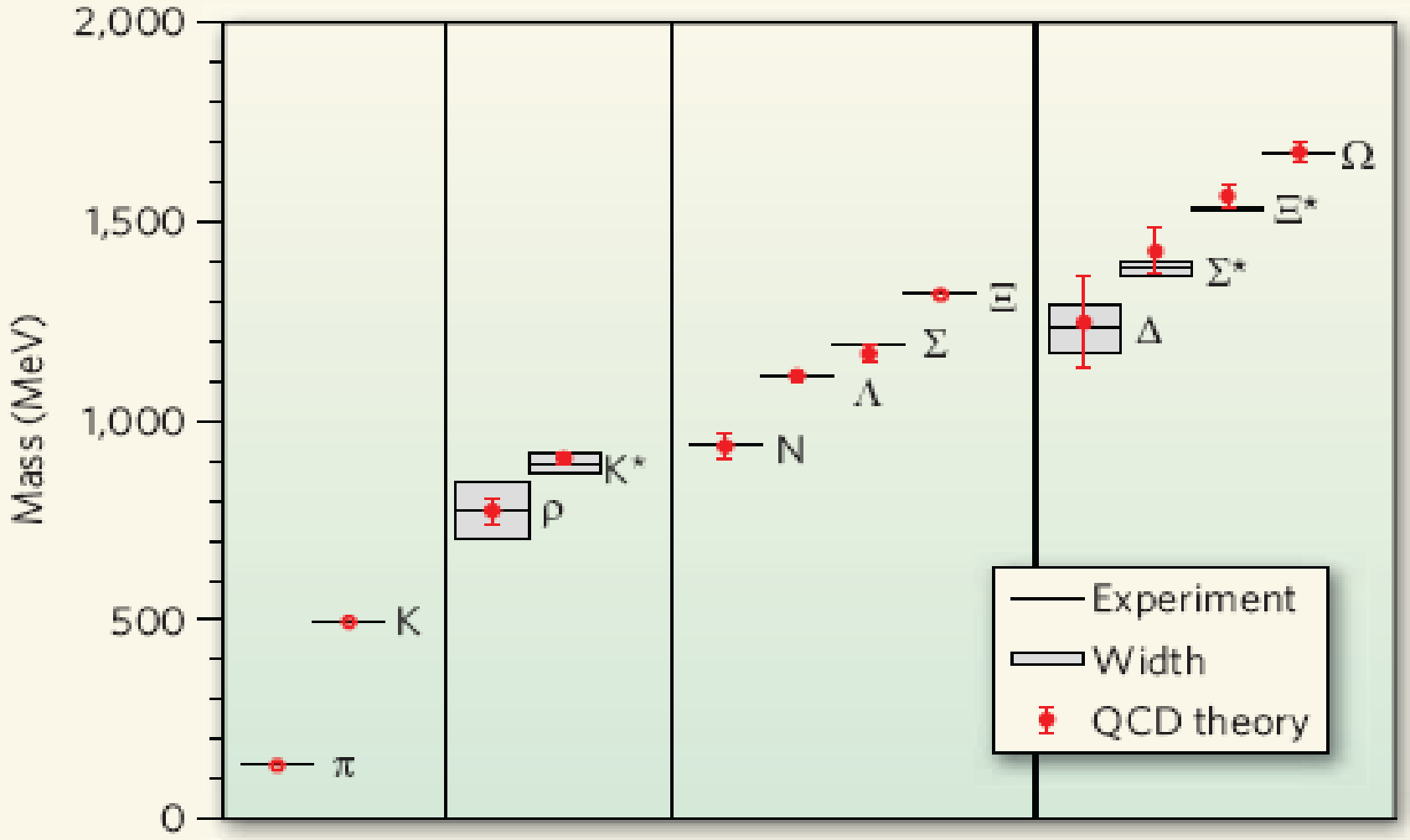


10 Breakthroughs of the Year 2008 SCIENCE VOL 322:

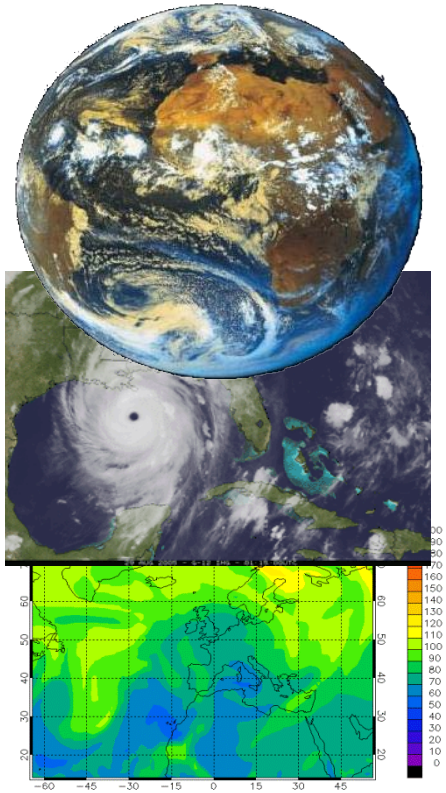
Proton's Mass 'Predicted' [Fodor et al. (F,G,Hu)]

STARTING FROM A THEORETICAL DESCRIPTION OF ITS INNARDS,

physicists precisely calculated the mass of the proton and other particles made of quarks and gluons. The numbers aren't new; experimenters have been able to weigh the proton for nearly a century. But the new results show that physicists can at last make accurate calculations of the ultra-complex strong force that binds quarks….

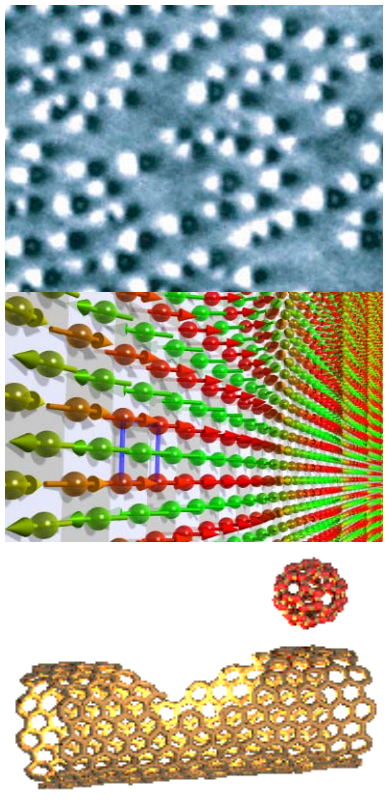

2,000
1,500
1,000
500
0
Mass (MeV)
π
K
ρ
K*
N
Λ
Σ
Ξ
Δ
Σ*
Ξ*
Ω
Experiment
Width
QCD theory

Supercomputing Drives Applied Science

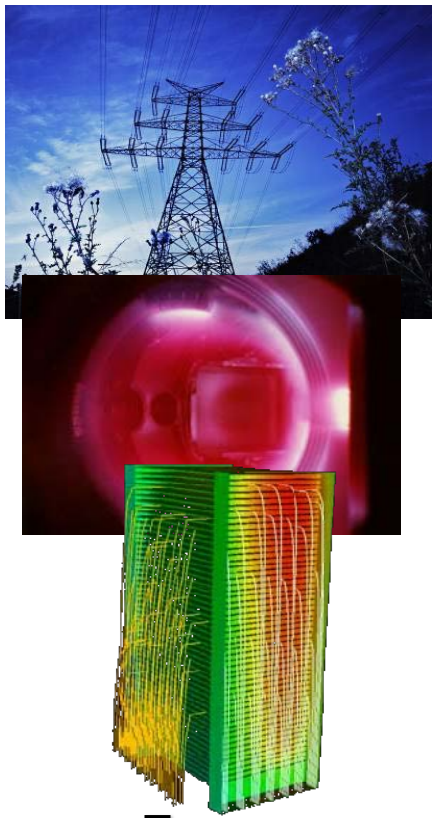


Environment
Weather/ Climatology
Pollution / Ozone Hole

Ageing Society
Medicine
Biology



Materials/ Inf. Tech
Spintronics
Nano-science

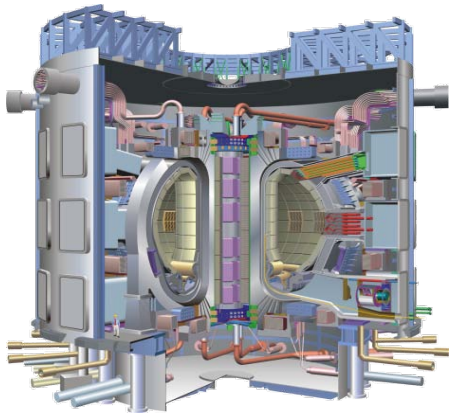


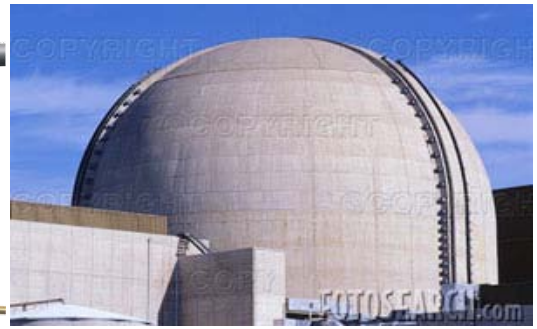
Energy
Plasma Physics
Fuel Cells

Supercomputing Drives Engineering and Business Competitiveness

- Reducing design costs by virtual prototyping:

  → faster time to market



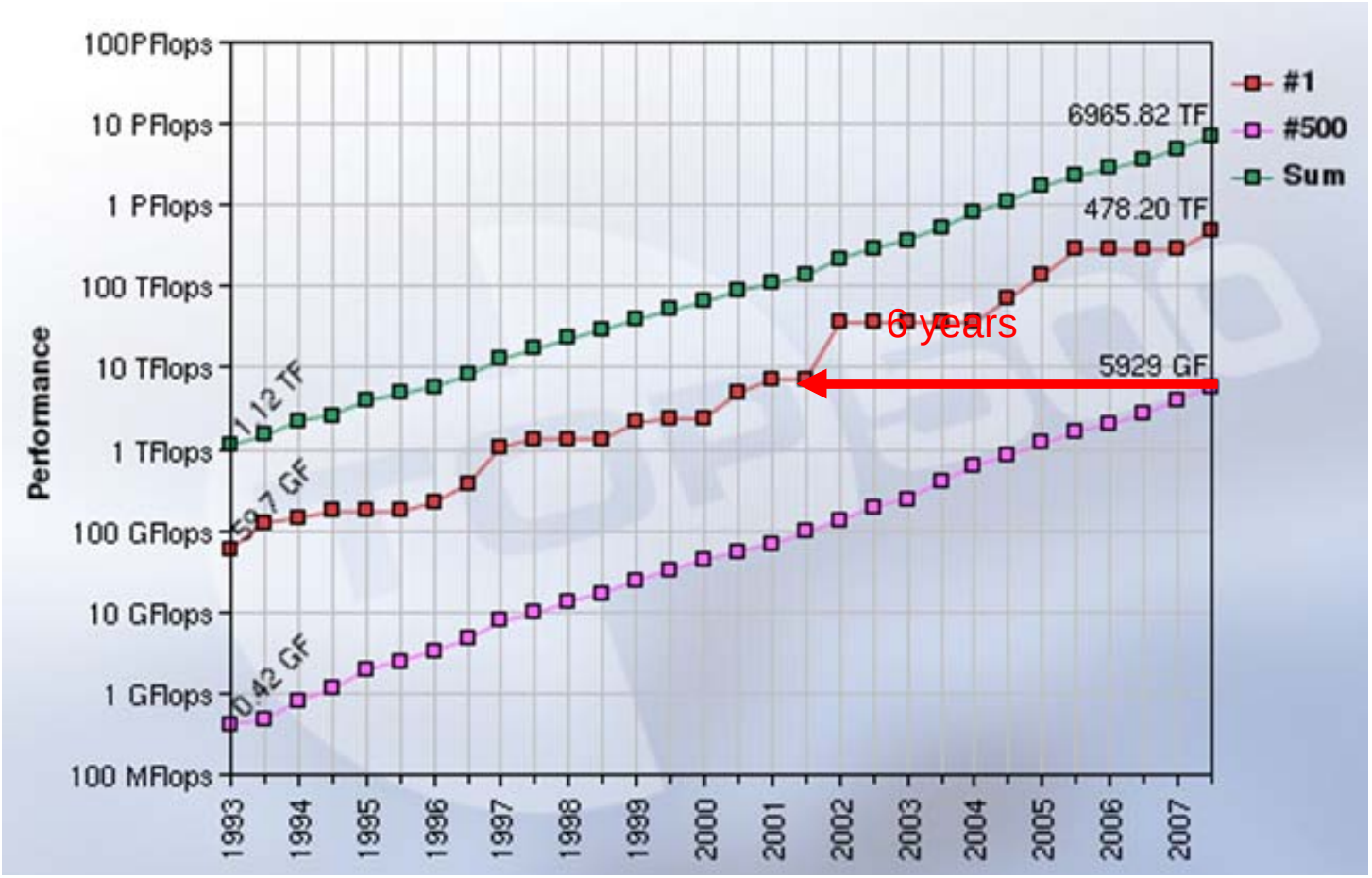


- Allowing investigations where economics or ethics preclude experimentation

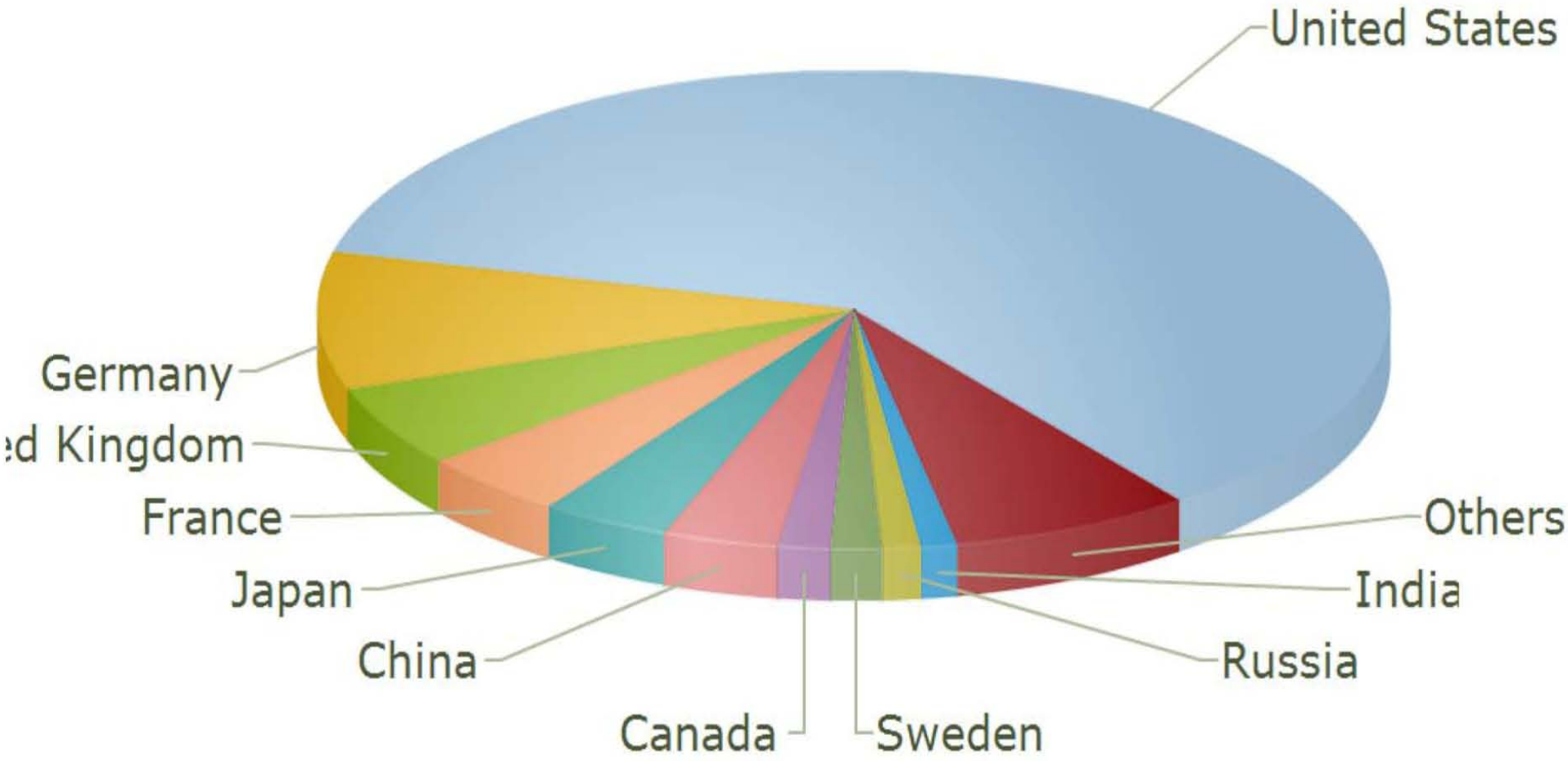
  → imperative of supercomputing

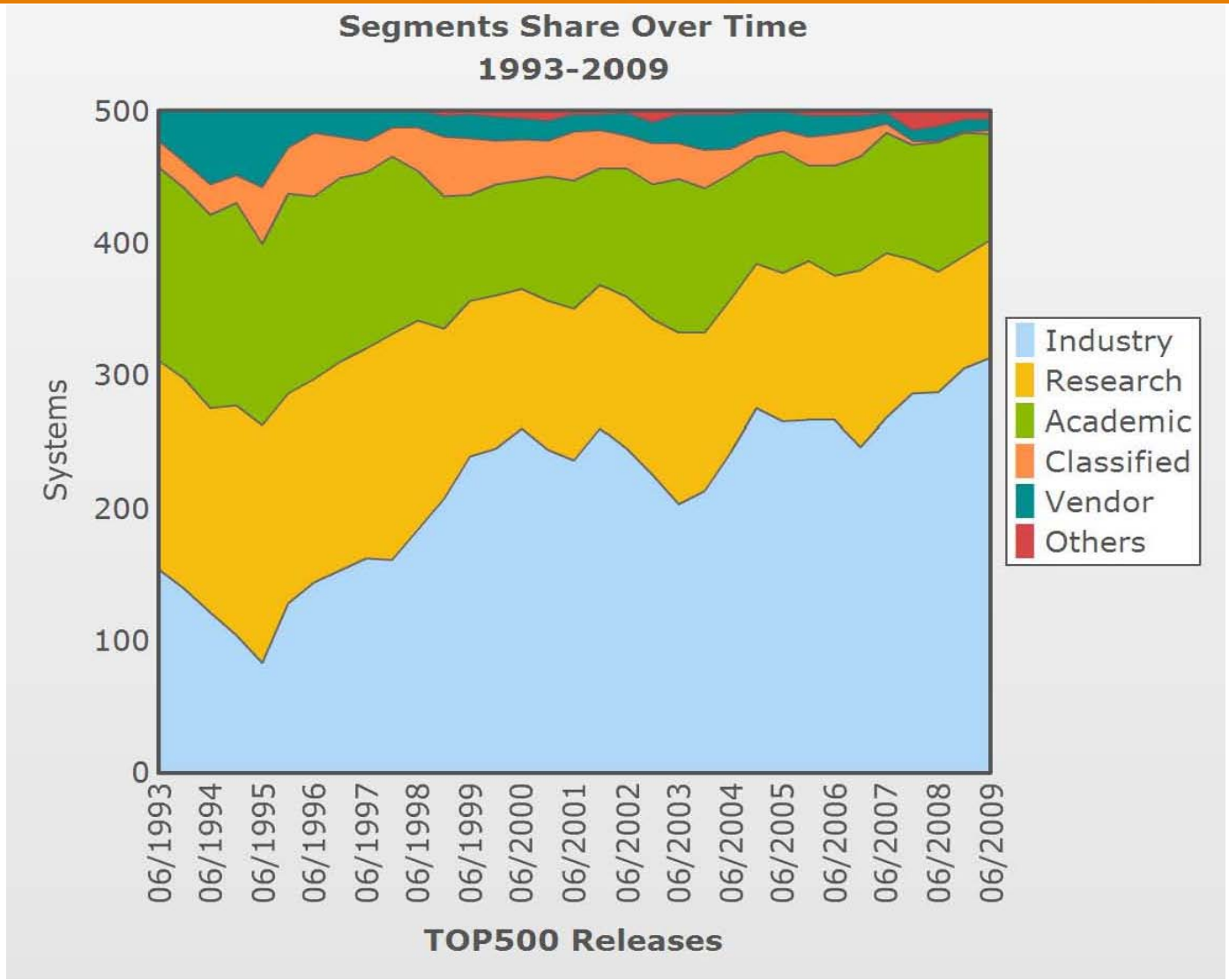
Scale length of computational advantage between #1 and #500

Top 500 list 11/07



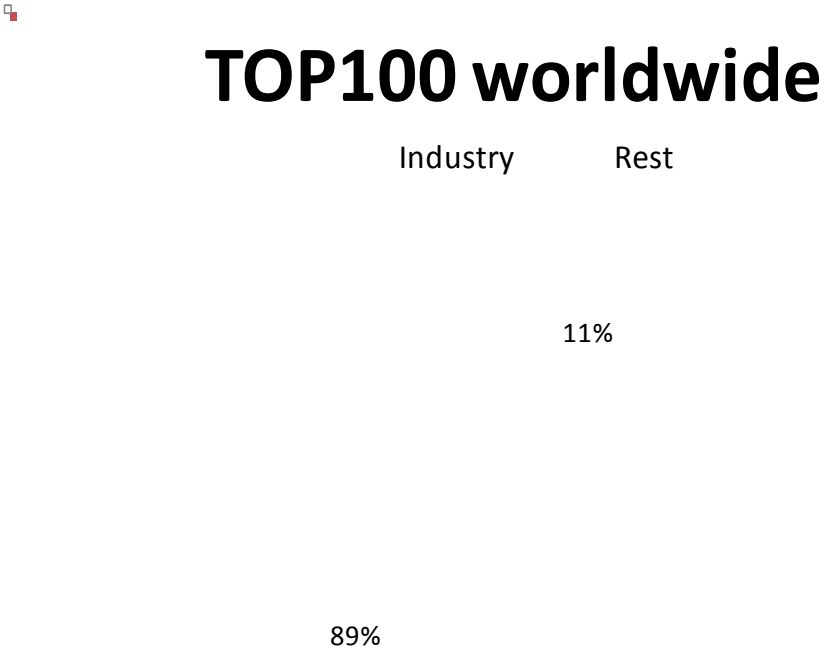
Performance per Country (TOP500 6/2009)

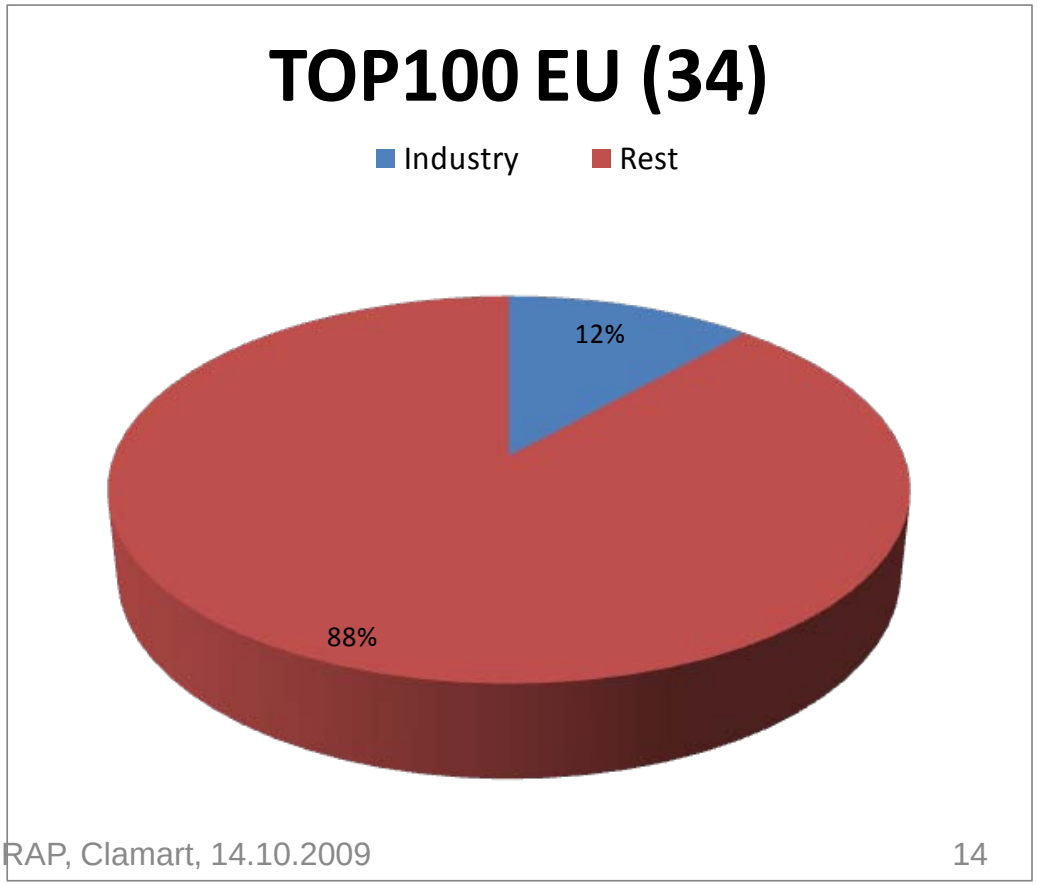



Segments Share Over Time
1993-2009
500
400
300
200
100
0
Systems
06/1993
06/1994
06/1995
06/1996
06/1997
06/1998
06/1999
06/2000
06/2001
06/2002
06/2003
06/2004
06/2005
06/2006
06/2007
06/2008
06/2009
TOP500 Releases
Industry
Research
Academic
Classified
Vendor
Others

Numbers of Systems in Industry

Numbers of Systems in Industry from TOP100





Outline

- A European SC RI: Why?
- What is PRACE?
- Where dowe stand?
- What comes next?
- Questions

Preparation (2004-2006): The Scientific Case

- **Weather, Climatology, Earth Science**
- **Astrophysics, Elementary particle physics, Plasma physics**
- **Material Science, Chemistry, Nanoscience**
- **Life Science**
- **Engineering**

  Airplane/helicopter simulation, biomedical flows, gas turbines, combustion engines, forest fires, virtual power plant, etc.

Vision

- Achieve EU leadership in public and private research
- Provide world-class HPC systems for word-class science

… and Mission

- Create a leading, persistent high-end HPC infrastructure
  - Deploy 3 – 5 systems of the highest performance level (tier-0)
  - Provide world-class support and training
- Implement pan-European Peer-Review procedure

Tier-0: JUGENE@FZJ – #3 TOP500, #1 in Europe

1st PRACE system

IBM Blue Gene/P

72 racks, 294912 cores

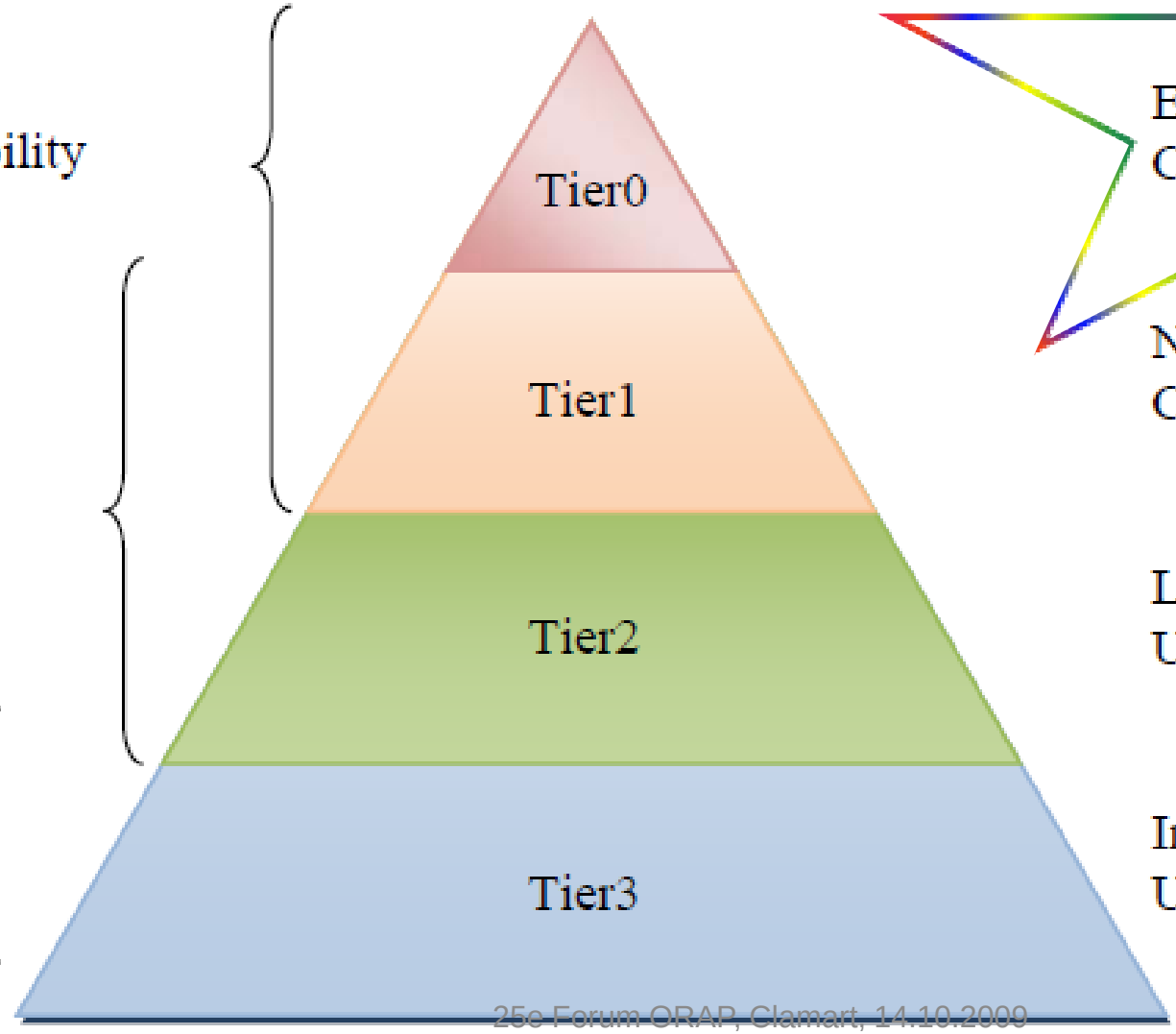
1 Petaflop/s peak

144 Tbyte memory

6 Pbyte disks

25 PByte tape capacity

Highest scalability



HPC Capability Resources

Capacity Resources

Desktop Resources

European Top Capability

National and Regional Centres

Local Facilities at Universities and Labs

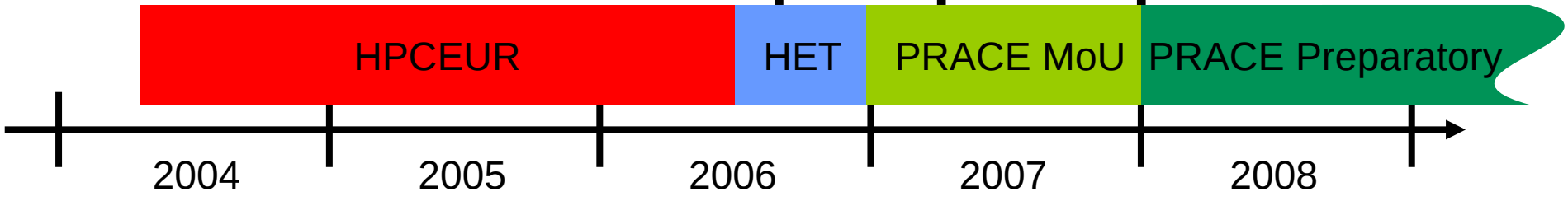
Individual Terminals of Users

Tier-1: JuRoPA + HPC-FF – #10 in TOP500, #2 in Europe

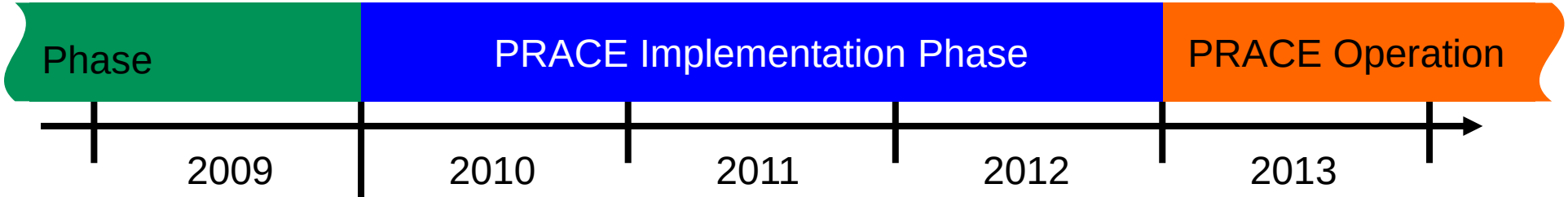
Cluster computer
Bull NovaScale R422-E2
1080 nodes, 8640 cores
101 TF peak, Intel Nehalem
24 GB memory
Infiniband QDR (Mellanox)
ParaStation Cluster-OS
HPC for Fusion

Cluster computer
SUN-blades
2208 nodes, 17664 cores
207 TF peak, Intel Nehalem
48 GB memory
Infiniband QDR (SUN M9)
ParaStation Cluster-OS
General Purpose HPC

PRACE History and first steps

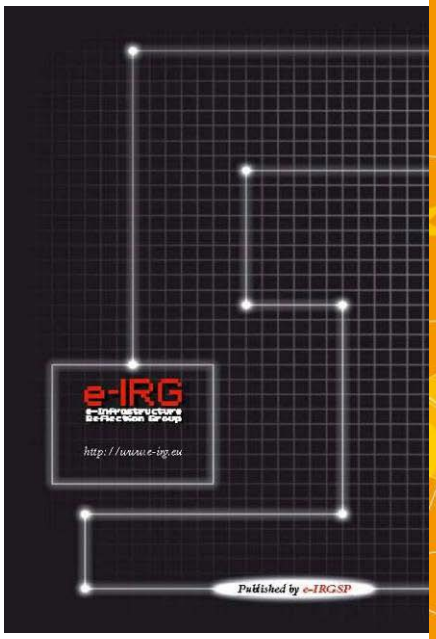


EU-Grant: INFSO-RI-211528, 10 Mio. €



Foreseen: PRACE Tier-0 centres providing HPC-capability service in a legal entity

The HPC-RI is an item on the ESFRI Roadmap



The European Roadmap for Research Infrastructures is the first comprehensive definition at the European level

Research Infrastructures are one of the crucial pillars of the European Research Area

A European HPC service – impact foreseen:

- strategic competitiveness
- attractiveness for researchers
- supporting industrial development

The PRACE Initiative (MoU)

- 2007: MoU by 15 European member states
- 2008: F, D, E, NL, UK reconfirmed their commitment for EU HPC Research Infrastructure
- 2009: Italy became a Principal Partner
- 2009: 4 new European member states have joined the PRACE initiative
- Bulgaria and Czech Republic joined

The PRACE Project

EU is funding the PRACE Preparatory Phase Project **(Grant: INFSO-RI-211528)**

- Partners: 16 from 14 countries
- Duration: 1/2008 – 12/2009
- Budget: 20 M€, EC: 10 M€
- Kickoff: January 29-30, 2008

PRACE Project Organization

EU-Preparatory Project Organization

- WP1 Management
- WP2 Organizational concept
- WP3 Dissemination, outreach and training
- WP4 Distributed computing
- WP5 Deployment of prototype systems
- WP6 Software enabling for prototype systems
- WP7 Petaflop/s systems for 2009/2010
- WP8 Future Petaflop/s to Exaflop/s technologies

Outline

- A European SC RI: Why?
- What is PRACE?
- Where do we stand?
- What comes next?
- Questions

After the First Successful Year:

- PRACE is a collaborative achievement of over 250 persons at the 16 partner sites
- Project review March 5-6, 2009, in Brussels

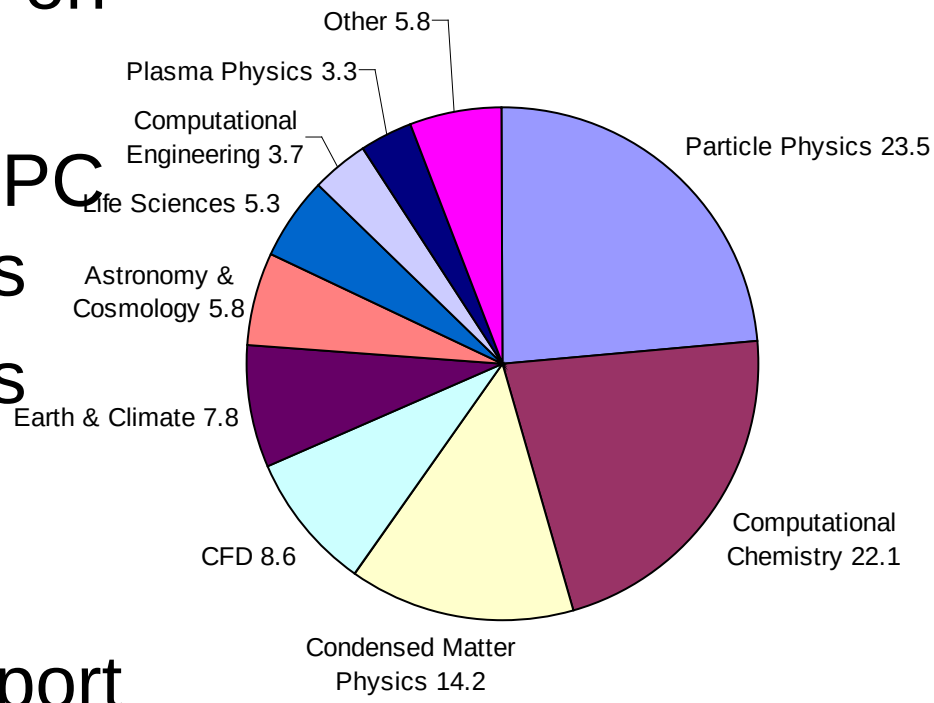
“The project made very good progress“

Selected Results and Highlights

- Applications
- Systems/Architectures
- Training and Outreach

Categorisation of Applications

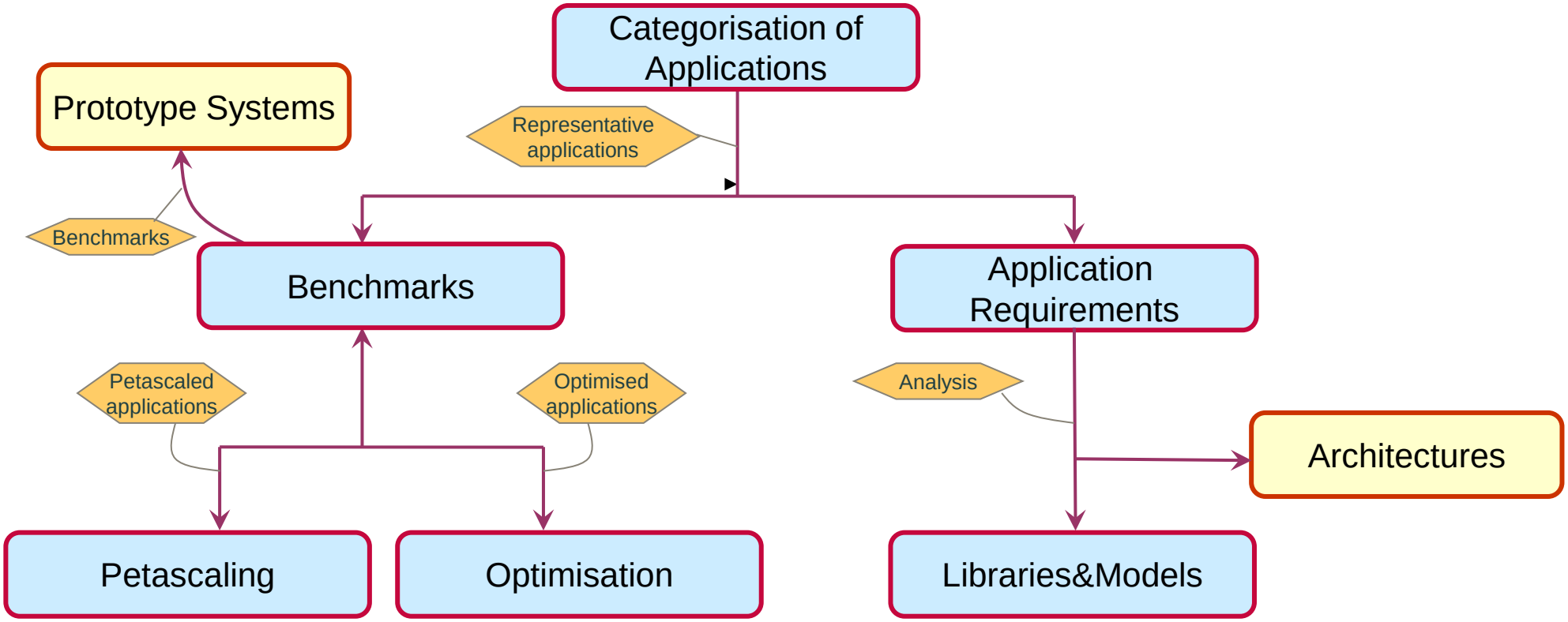
- Benchmark applications based on European HPC usage
- Surveys of PRACE partners' HPC systems and major applications
- 24 systems and 69 applications
- Quantitative basis for selecting representative applications
- Disseminated as Technical Report



Representative Benchmark Suite

- Set of applications benchmarks to be used in the procurement process for Petaflop/s systems
- 12 core applications, plus 8 additional applications
  - *Core:* NAMD, VASP, QCD, CPMD, GADGET, Code_Saturne, TORB, ECHAM5, NEMO, CP2K, GROMACS, N3D
  - *Additional:* AVBP, HELIUM, TRIPOLI_4, PEPC, GPAW, ALYA, SIESTA, BSIT
- Each application will be ported to appropriate subset of prototypes
- Synthetic benchmarks for architecture evaluation
  - Computation, mixed-mode, IO, bandwidth, OS, communication
- Applications and Synthetic benchmarks integrated into JuBE
  - **Ju**elich **B**enchmark **E**nvironment

Software Enabling for Petaflop/s Systems (WP6)



Mapping Applications to Architectures

Code	**MPP** (i.e. BlueGene L/P or CRAY XT4/5)	**Thin node clusters** (i.e. Bull INCA or SGI ICE)	**Fat node clusters** (i.e. Bull MESCA, SGI UltraViolet or IBM Power6)	**Vector systems** (NEC SX8-9, Cray X2	**Accelerated systems** (i.e. scalar or vector + GPU, FPGA or Clearspeed).	**Accelerated systems – Cell based** (i.e. Roadrunner, Maricell)
NAMD				E	E	E
CPMD						E
VASP						E
QCD				E	E	E
GADGET						
Code Saturne				E	E	E
TORB						E
NEMO						
ECHAM5						
CP2K	E					
GROMACS					E	E
N3D		E	E			
AVBP	E					
HELIUM						
TRIPOLI 4	E		E			
GPAW						
ALYA						E
SIESTA						E
BSIT						E
PEPC				E	E	E

Table 4 : application mapping to Petaflop/s systems archite

E = estimated

Prototypes for Petaflop/s systems in 2009/2010

IBM BlueGene/P (FZJ)
01-2008 / 06-2009

IBM Power6 (SARA)
07-2008

Cray XT5 (CSC)
11-2008

IBM Cell/Power (BSC)
12-2008

NEC SX9, vector part (HLRS)
02-2009

Intel Nehalem/Xeon (CEA/FZJ)
06-2009

Procurement Strategy, Cost Estimates

- Analysis of European procurement proced. completed
- Work in progress:
  - Definition of general procurement process
  - Definition of selection and evaluation criteria
  - Evaluation process for offers by vendors
- Market watch: what will it take to be in the Top 5 / Top 10 in 2010 … 2011 ?
- TCO estimates based on market survey, vendor input and partners' experience, repeated annually
  - Overall goals are consistent with committed funding of 80 to120 M€/a

Prototypes for Systems beyond 2010

Sites	Hardware/Software	Porting effort
CEA "*GPU/CAPS*"	1U Tesla Server T1070 (CUDA, CAPS, DDT) Intel Harpertown nodes	"Evaluate GPU accelerators and GPGPU programming models and middleware." (e.g., *pollutant migration code* (ray tracing algorithm) to CUDA and HMPP)
CINES-LRZ "*LRB/CS*"	Hybrid SGI ICE2 /UV /Nehalem-EP&Nehalem-EX /ClearSpeed / Larrabee	**Gadget**,SPECFEM3D_GLOBE, RaXml, Rinf, RandomAccess, ApexMap, Intel MPI BM
CSCS *"UPC/CAF"*	Prototype PGAS language compilers (CAF + UPC for Cray XT systems)	"The applications chosen for this analysis will include some of those already selected as **benchmark codes**"
EPCC "*FPGA*"	Maxwell – FPGA prototype (VHDL support & consultancy + software licenses (e.g., Mitrion-C))	"We wish to port several of the **PRACE benchmark codes** to the system. The codes will be chosen based on their suitability for execution on such a system."

Prototypes beyond 2010 (cont'd)

Sites	Hardware/Software	Porting effort
FZJ (BSC) "*Cell & FPGA interconnect*"	eQPACE (PowerXCell cluster with special network processor)	Extend FPGA-based interconnect beyond QCD applications.
LRZ "*RapidMind*"	RapidMind (Streaming Processing Programming Paradym) X86, GPGPU, Cell	ApexMap, Multigrid, FZJ (QCD), CINECA (linear algebra kernels involved in solvers for ordinary differential equations), SNIC
NCF "*ClearSpeed*"	ClearSpeed CATS 700 units	Astronomical many-body simulation, Iterative sparse solvers with preconditioning, finite element code, cryomicrotome image analysis
CINECA	I/O Subsystem (SSD, Lustre, pNFS)	-
KTH	AMD Istanbul Cluster (HP)	Energy effcient system with standard hardware

Outreach and Education

Industry seminars:

1st Seminar Sept. 3, 2008 Amsterdam, Netherlands
2nd Seminar Sept. 7-8, 2009 Toulouse, France

Summer & winter schools:

Stockholm, Athens

PRACE booths:

ISC, ICT, SC,

PRACE booth at SC08

1st Industry seminar, 3.9.2008

PRACE website

ICT 2008, PRACE-Booth

PRACE Summer School Stockholm

PRACE Winter Schoolat the OTE academy, Athens 26-29.8.2009

Industry seminars

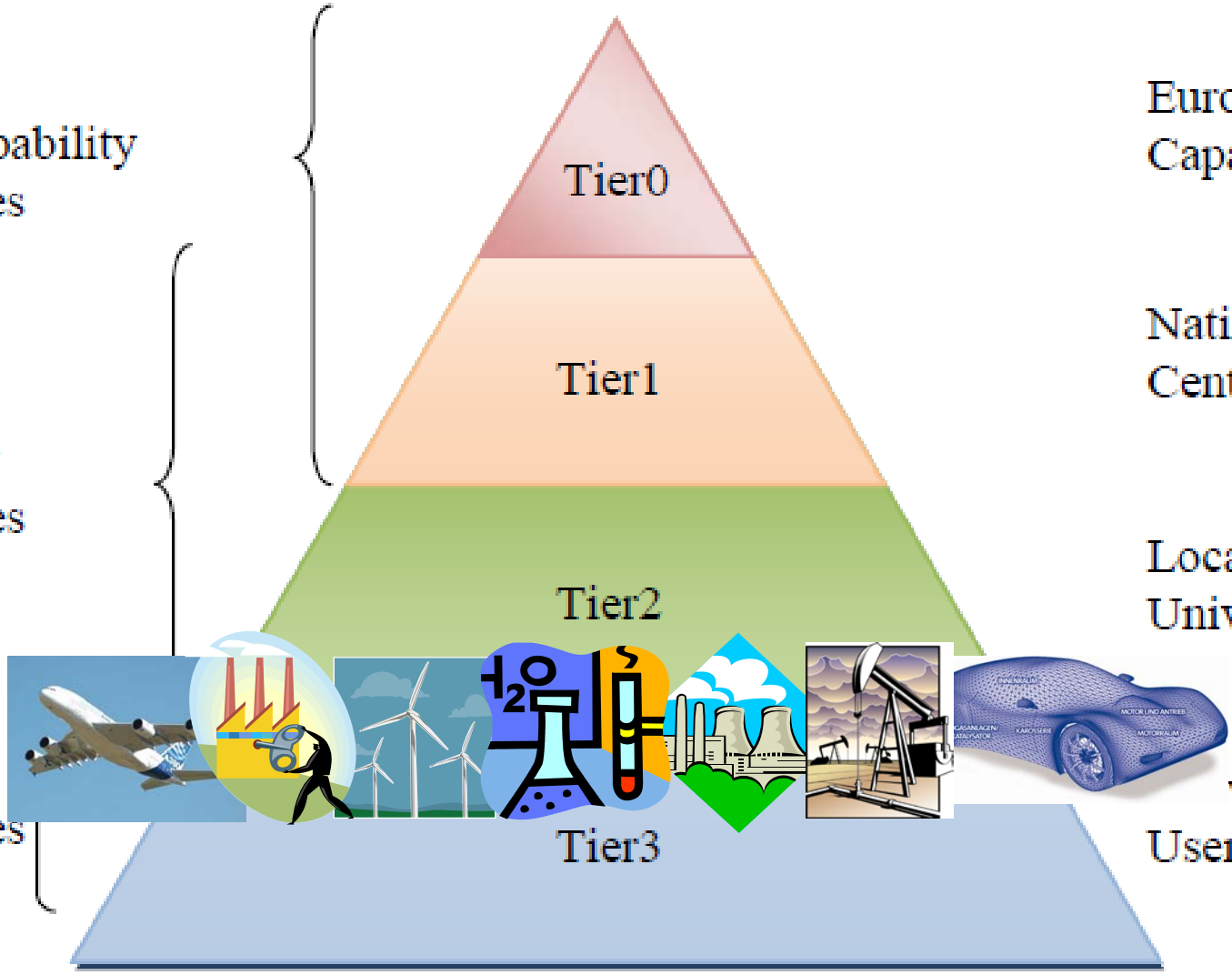
(1st August 31, 2008, Amsterdam, 2nd: September 8, 2009 Toulouse)

- **Goals**
  - Understanding industrial needs and expectations
  - Options for industry involvement in PRACE
- **Motivation**
  - HPC is a key technology in the 21st century and science
  - Industry needs independent access to HPC in Europe
  - PRACE will foster industrial application on top-end
  - PRACE pushes HPC technology development in EU

HPC Capability Resources

Capacity Resources

Desktop Resources



European Top Capability

National and Regional Centres

Local Facilities at Universities and Labs

vidual Terminals of Users

Outline

- A European SC RI: Why?
- What is PRACE?
- Where do we stand?
- What comes next?
- Questions

Next Steps

- Contracts for the legal entity are in final negotiation, signature planned in December 2009
- Temporary seat in Lisbon
- PRACE Tier-0 Infrastructure will become operable during first half of 2010
  - Small HQ in charge of organisational, financial tasks and Peer Review
  - Tier-0 services for the Scientific Community provided by the “hosting members” of the legal entity.

Next Steps II

- PRACE will organize platform for pan-European Tier-1 access as pioneered by DEISA
- Implementation of PRACE and further development will be major task in the next years: co-funded by the EC
  - By the way: Have to find a new name because of legal issues
- **→ 1st PRACE Implementation Phase Project in preparation**
  - **2nd and 3rd will follow**

Accessing the future PRACE RI

Disclaimer: this is work in progress

Access Model

- Based on peer-review: “Best systems for best science”
- Three types of resource allocations
  - Test / evaluation access
  - Project access – for a specific project, grant period ~ 1 year
  - Programme access – resources managed by a community
- Free-of-charge for European scientific communities

Accessing the future PRACE RI cont.

Disclaimer: this is work in progress

Access Models for Commercial users (< 10%)

- Peer Review based, free of charge if results are made public (similar to US INCITE programme)
- Paid Access

Funding of the Infrastructure

- Mainly national funding through partner countries
- European contribution
- Monitor relation between contribution and usage

How can users get involved ?

- **EU country:** **Join PRACE Initiative!**
- **International:** **Cooperation of partner organizations is welcome!**
- **Porting of codes to the PRACE Prototypes**
  - Prototypes are mainly be used project-internally, but
  - Prototypes are also made available to selected users for testing/porting purposes using a light-weight peer-review process
- **Participation in the PRACE training events**
  - 5 Code Porting Workshops so far
  - PRACE booth at SC'09

Summary

In its first 18 months, the project had …

- major achievements in all areas
- raised significant awareness with all stakeholders
- reconfirmed the commitment of the Governments

PRACE is well prepared to master the future challenges towards a European RI

PRACE All-Hands meeting February 2009

PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE