PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

PRACE – A Mid-Term Update

Dietmar Erwin, Forschungszentrum Jülich
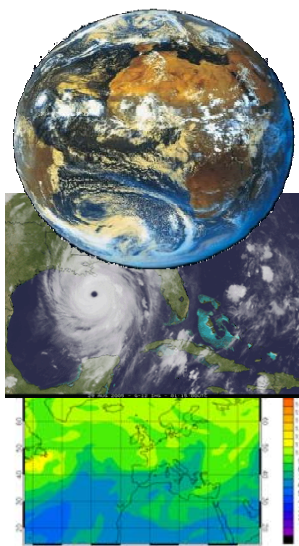ORAP, Lille, March 26, 2009

Outline

- What is PRACE
- Where we stand
- What comes next
- Questions

Outline

- What is PRACE
- Where of we stand
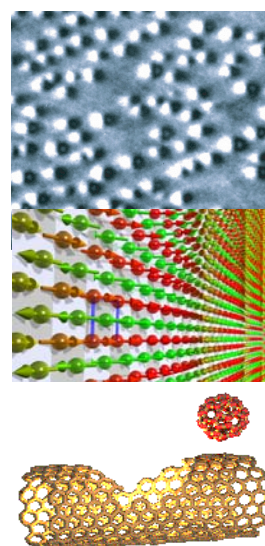- What comes next
- Questions
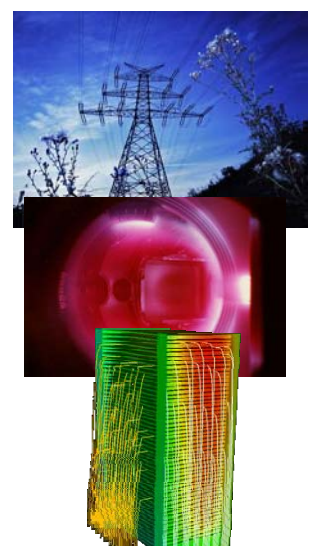
3

Supercomputing Drives Science through Simulation

Environment
Weather/ Climatology
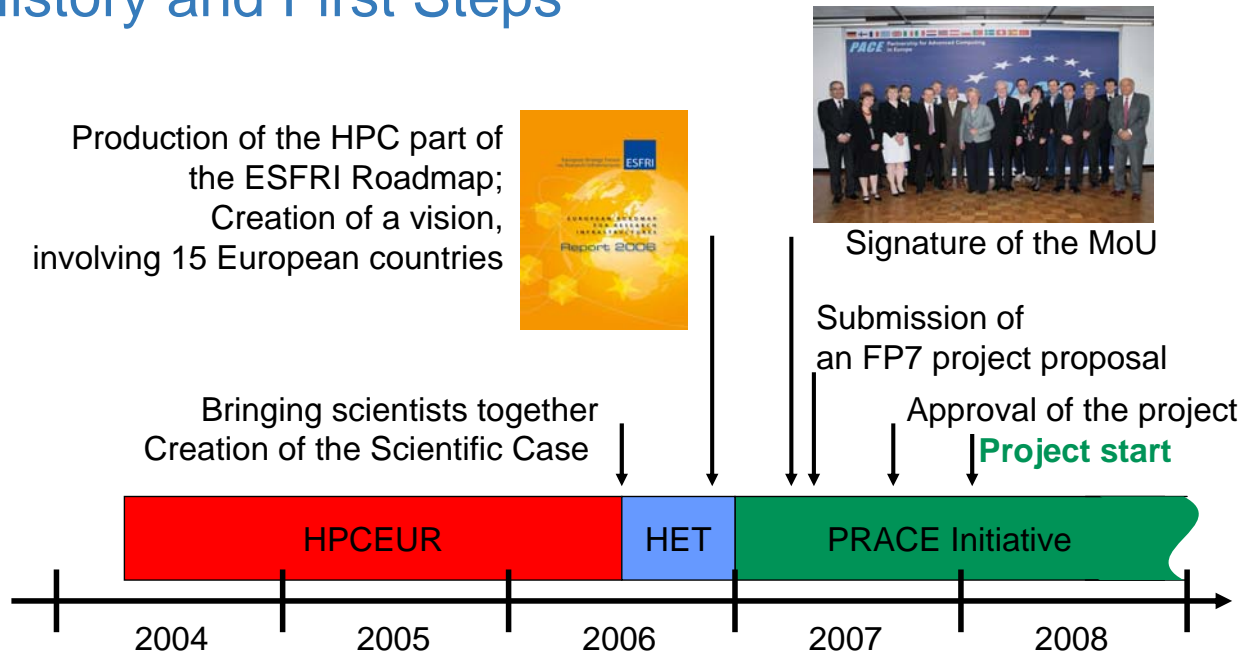Pollution / Ozone Hole

Ageing Society
Medicine
Biology

Materials/ Inf. Tech
Spintronics
Nano-science

Energy
Plasma Physics
Fuel Cells

4

History and First Steps

Production of the HPC part of
the ESFRI Roadmap;
Creation of a vision,
involving 15 European countries

Signature of the MoU

Submission of
an FP7 project proposal

Bringing scientists together
Creation of the Scientific Case

Approval of the project
Project start

| HPCEUR | HET | PRACE Initiative |

2004   2005   2006   2007   2008

HET: The Scientific Case

- Weather, Climatology, Earth Science
  - degree of warming, scenarios for our future climate.
  - understand and predict ocean properties and variations
  - weather and flood events
- Astrophysics, Elementary particle physics, Plasma physics
  - systems, structures which span a large range of different length and time scales
  - quantum field theories like QCD, ITER
- Material Science, Chemistry, Nanoscience
  - understanding complex materials, complex chemistry, nanoscience
  - the determination of electronic and transport properties
- Life Science
  - system biology, chromatin dynamics, large scale protein dynamics, protein association and aggregation, supramolecular systems, medicine
- Engineering
  - complex helicopter simulation, biomedical flows, gas turbines and internal combustion engines, forest fires, green aircraft,
  - virtual power plant
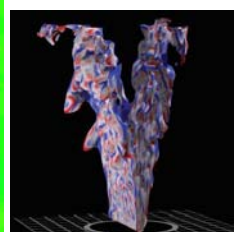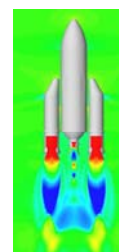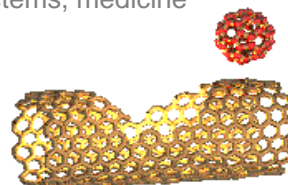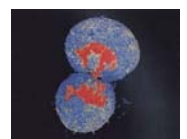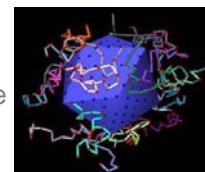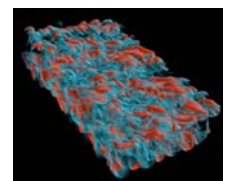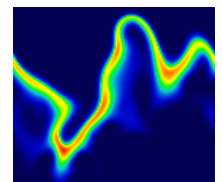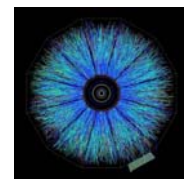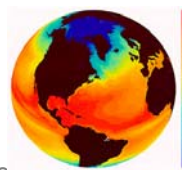
First success: HPC in ESFRI Roadmap

European Strategy Forum on Research Infrastructures — ESFRI

EUROPEAN ROADMAP FOR RESEARCH INFRASTRUCTURES
Report 2006

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

Second success: The PRACE Initiative

- Memorandum of Understanding signed by 15 States in Berlin, on April 16, 2007

- France, Germany, Spain, The Netherlands, UK committed funding for a European HPC Research Infrastructure (LoS)

New:

Third success: The PRACE Project

EU approved the PRACE Preparatory Phase Project
(Grant: INFSO-RI-211528)

- 16 Partners from 14 countries
- Project duration:
  January 2008 – December 2009
- Project budget: 20 M € ,
  EC funding: 10 M €
- Kickoff: Jülich, January 29-30,
  2008

e-infrastructure

9

PRACE Objectives in a Nutshell

- Provide world-class systems for world-class science
- Create a single European entity
- Deploy 3 – 5 systems of the highest performance level (tier-0)
- Ensure diversity of architectures
- Provide support and training

PRACE will be created to stay

Outline

- What is PRACE
- **Where we stand**
- What comes next
- Questions

After the First Successful Year:

- Project review March 5-6, 2009, in Brussels
- A full day of presentation of the project results of year1
- Coordinator and all work packages presented
- PRACE is the collaborative achievement of over 250 persons at the 16 partner sites
- Expertise includes a Chairman of the Board, (coordinator), executives, scientists, programmers, legal experts, …

"The project made very good progress in many areas"

PRACE Work Packages

- 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 technologies

Selected Results and Highlights of 2008

- Applications
- Systems/Architectures
- Training and Outreach

Software Enabling for Petaflop/s Systems
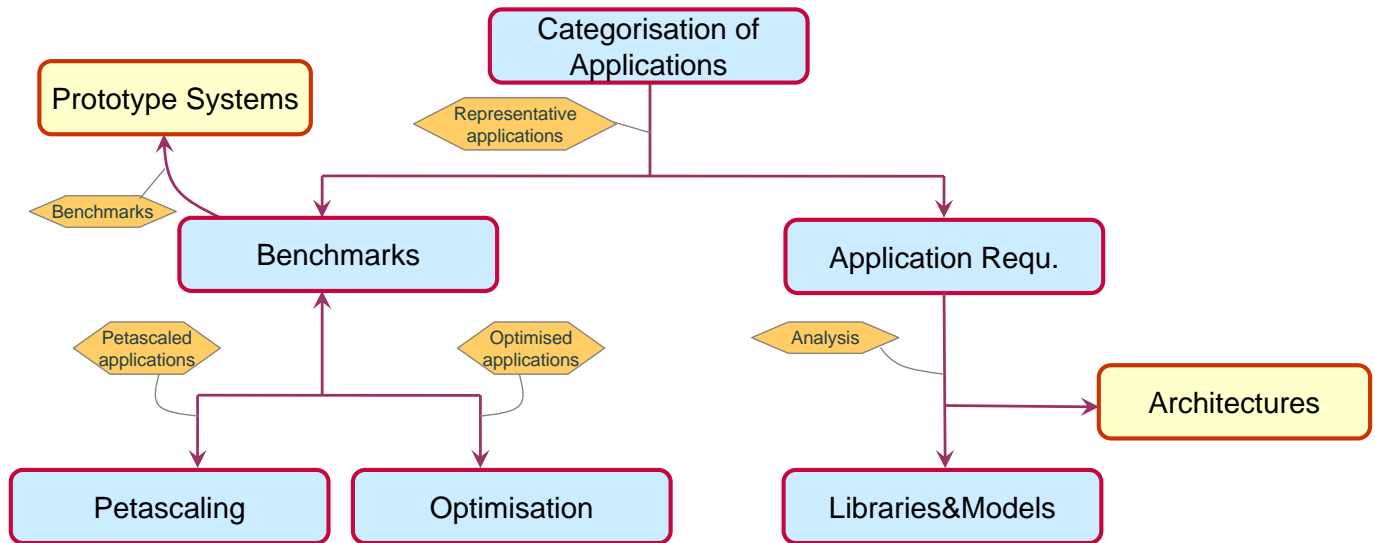
Categorisation of Applications

- Benchmark applications should be *representative* of European HPC usage
- We conducted surveys of PRACE partners' HPC systems and major applications
  - Collecting various interesting data for 24 systems and 69 applications
- Quantitative basis for selecting representative applications
- Disseminated as Technical Report

Other 5.8
Plasma Physics 3.3
Computational Engineering 3.7
Life Sciences 5.3
Astronomy & Cosmology 5.8
Earth & Climate 7.8
CFD 8.6
Condensed Matter Physics 14.2
Particle Physics 23.5
Computational Chemistry 22.1
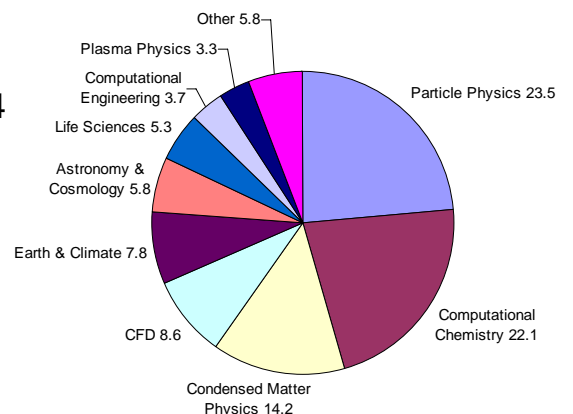
Application Requirements

- Analysis of representative applications
  - Ported to relevant architectures
- Result: Quantitative data from initial porting
- Supplemented by user survey
  - Sent to Top 10 users in each PRACE country
- Questions covered
  - The user
  - Usage patterns
  - HPC infrastructure
  - Upcoming algorithms
- Analysed almost 70 responses from these major users
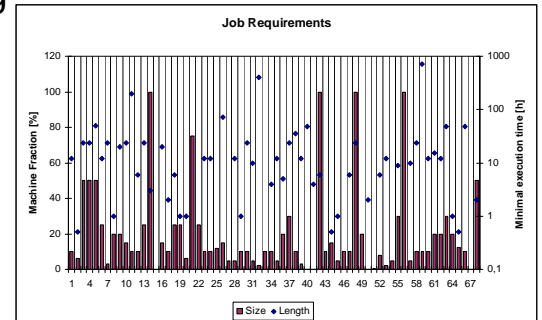
17

Representative Benchmark Suite

- Defined a 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

18

Mapping Applications to Architectures

- Identified affinities and priorities
- Based on the application analysis - expressed in a condensed, qualitative way
  - Need for different "general purpose" systems
  - There are promising emerging architectures
- Will be more quantitative after benchmark runs on prototypes

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 architec

E = estimated

Selected set of Prototypes

- Process
  - Call
  - evaluation
  - selection
  - Approval by EC
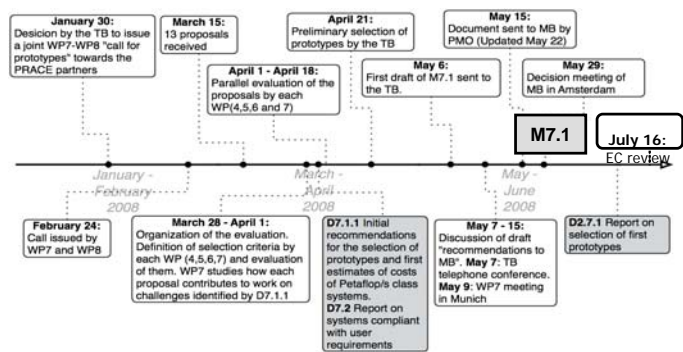
- Funding 2.2 M€
  (50% of cost in average)

| Site	Architecture Vendor/Technology	
FZJ Germany	MPP IBM BlueGene/P	
CSC-CSCS Finland+Switzerland	MPP Cray XT5/XTn - AMD Opteron	
CEA-FZJ France+Germany	SMP-TN Bull et al. Intel Xeon Nehalem	
NCF Netherlands	SMP-FN IBM Power 6	
BSC Spain	Hybrid – fine grain IBM Cell + Power6	
HLRS Germany	Hybrid – coarse grain NEC Vector SX/9 + x86	

Installed prototypes

IBM BlueGene/P (FZJ)
01-2008

IBM Power6 (SARA)
07-2008

Cray XT5 (CSC)
11-2008
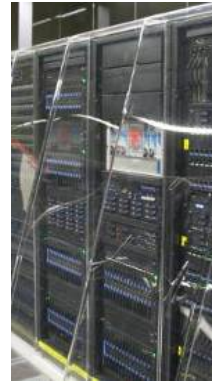
IBM Cell/Power (BSC)
12-2008

NEC SX9, vector part (HLRS)
02-2009

Intel Nehalem/Xeon (CEA/FZJ):
expected installation date 4/5-2009

Summary of current prototype status

milestone	IBM BlueGene/P at FZJ	IBM Power6 at SARA	Cray XT at CSC	IBM Cell/Power at BSC	NEC SX9/x86 at HLRS	Intel Nehalem/Xeon at CEA/FZJ
system installed	yes	yes	yes	yes	partly (vector)	no
system in production	yes	yes	yes	yes	partly (vector)	no
technical assessment	yes	nearly	yes	started	no	no
evaluation of communication and I/O infrastructure	yes	nearly	yes	started	no	no
evaluation and benchmarking of user applications	started	started	started	started	started	no

Web site and the dissemination channels

- The PRACE web presence with news, events, RSS feeds etc. http://www.prace-project.eu
- Alpha-Galileo service: 6500 journalists around the globe: http://www.alphagalileo.org

- Belief Digital Library
- HPC-magazines
- PRACE partner sites, top 10 HPC users

The PRACE website, www.prace-project.eu

PRACE Dissemination Package

- PRACE WP3 has created a dissemination package including templates, brochures, flyers, posters, badges, t-shirts, USB-keys, badges etc.

The PRACE logo

PRACE USB-key

Heavy Computing 10^15: the PRACE t-shirt

PRACE booth at ISC, ICT, SC

- PRACE had an exhibition booth at
  ISC'08 Dresden, D
  SC08   Austin, Texas, US
  ICT 2008, Lyon, F

PRACE booth at SC08

PRACE booth at ICT 2008

Selected Events   Industry Seminar

PRACE Award
ISC'08
Best student
paper on peta-
scaling to
Dominik
Göddeke,
Dortmund

Industrial Competitiveness: Europe goes HPC / Attendeed companies

---

Training: Survey of HPC education and training needs

- The Top 10 users at each participating PRACE member site were invited to participate in completing the survey.
- The data was obtained from the most comprehensive evaluation of user training requirements
- Over 90% of respondents believed they would benefit from formal training in the following areas: performance optimization, debugging tools and techniques, code testing and compiler optimisations;
- Over 90% of users considered that there is an important need for improved HPC training programmes
- Request to use survey from Hong Kong and US

HPC

TeraGrid

Training: summer School

- PRACE Petascale Summer School, August 26-29, Stockholm, Sweden;
- It attracted 31 students, representing all PRACE member countries plus South Africa. In the anonymous feedback answered by 26 students, the School received excellent grades;
- Training material published on the PRACE website.
- Special access to large European systems - BG/P,Cray XT5, IBM Power6, and CELL - was offered to the students in conjunction with lectures about programming models, hardware, development tools and optimisation/debugging

PRACE Summer School took place at PDC, Stockholm

2009 training: Winter School

- PRACE Winter School, February 9-13, 2009, Athens, Greece;
- 48 registered attendees (78 applications);
- Training material is published on the PRACE web site.

The PRACE Winter School
at the OTE academy, Athens

HPC training and HPC training events on the PRACE website

Outline

- What is PRACE
- Where we stand
- **What comes next**
- Questions

**PRACE
All-Hands meeting
February 2009**

PARTNERSHIP
FOR ADVANCED COMPUTING
IN EUROPE

The next important challenges

Get the initial commitments by Government converted into signed contracts

Governance document

Report on Operating model

Doc on funding agreement

selection of next production systems

Draft contract

Peer review process

PRACE CONTRACT

Month 15 *Month 18* *Month 24*

Prototypes: Change Status to Green

milestone	IBM BlueGene/P at FZJ	IBM Power6 at SARA	Cray XT at CSC	IBM Cell/Power at BSC	NEC SX9/x86 at HLRS	Intel Nehalem/Xeon at CEA/FZJ
system installed	yes	yes	yes	yes	partly (vector)	no
system in production	yes	yes	yes	yes	partly (vector)	no
technical assessment	yes	nearly	yes	started	no	no
evaluation of communication and I/O infrastructure	yes	nearly	yes	started	no	no
evaluation and benchmarking of user applications	started	started	started	started	started	no

Prototypes: Change Status to Green

milestone	IBM BlueGene/P at FZJ	IBM Power6 at SARA	Cray XT at CSC	IBM Cell/Power at BSC	NEC SX9/x86 at HLRS	Intel Nehalem/Xeon at CEA/FZJ
system installed	yes	yes	yes	yes	yes	yes
system in production	yes	yes	yes	yes	yes	yes
technical assessment	yes	yes	yes	yes	yes	yes
evaluation of communication and I/O infrastructure	yes	yes	yes	yes	yes	yes
evaluation and benchmarking of user applications	yes	yes	yes	yes	yes	yes

Refine Costs Analysis / Estimates

- Initial: simplified scenarios, using market survey, vendor input and partners' experience
- Updated and refined iteratively
- Consistent with the a priori PRACE initiative estimates in 2007

	Low consumption system	Cluster based on commodity components (thin nodes)	Clusters of fat nodes
Power Consumption	1 to 2 MW	3 to 4 MW	5 to 6 MW
Total floor space	600 m2	800 m2	1000 m2
Supercomputer	10 to 30 M€	40 to 60 M€	60 to 80 M€
Acquisition cost	24 M€ (11 to 38 M€)	59 M€ (44 to 75 M€)	82 M€ (66 to 100 M€)
Running cost (for 5 years)	33 M€ (19 to 53 M€)	59 M€ (38 to 89 M€)	84 M€ (55 to 127 M€)
Full cost (5 years)	56 M€ (30 to 91 M€)	118 M€ (82 to 164 M€)	167 M€ (121 to 227 M€)

Table 2: Examples of cost analysis estimates for 1 Petaflop/s (peak) systems in 2010

Market watch – Top 500 watch (Petascale worldwide…)

	Top 10	Top 5
Late 2009	0.5 Pflops	1 Pflops
Late 2010	1 Pflops	2 Pflops
Late 2011	2 Pflops	5 Pflops

Applications: Petascaling and Optimisation

Petascaling

– Mixed-mode parallelisation
– Load balancing
– Minimisation of communication overheads
– Parallel I/O
– Checkpointing

Optimisation

– Optimising serial performance
  • CPU
  • Memory
– Optimising both for general-purpose architectures and specialised architectures
– Algorithmic optimisations

PRACE will disseminate
best practice *in these areas*

Libraries and Programming Models

- Classification of Benchmark applications
- Current programming models
  - MPI, OpenMP, mixed-mode, …
- PGAS and other future programming models
- Accelerator Languages
  - CUDA, RapidMind, openCL …
- Petascale libraries

Future PRACE Events

- Second scientific seminar in 11.-13. May 2009 in Amsterdam in collaboration with DEISA2 (DEISA PRACE Symposium)
- The Second industry seminar Toulouse, 7-8th September organised by GENCI & GAUSS: covers also small and medium size enterprises
- 2nd PRACE Award at ISC09, Hamburg
- Exhibition booths at major events: ISC09 (June), SC09 (November)
- Five code porting and optimization workshops Finland, Sweden (2), Poland, Switzerland

- Additional training on GPU programming April, France
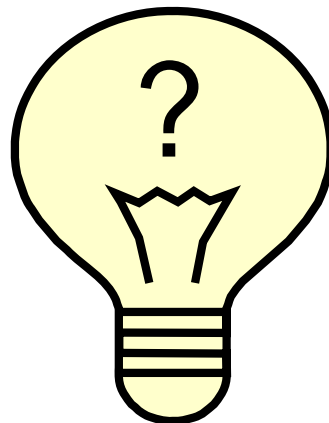
Summary

In 2008, 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 challenges of 2009

Thank you

Special thanks to all project collaborators for the content of the presentation