PRACE Overview

OASIS User Meeting, Toulouse, May 26, 2009

X. Delaruelle, CEA, PRACE WP4 co-leader

CAPACITIES

e-infrastructure

Outline

- What is PRACE
- Where we stand
- Questions

Outline

- What is PRACE
  - Excerpts from A. Bachem's talk at DEISA-PRACE Symposium, May 11, Amsterdam http://www.prace-project.eu/documents/deisa-prace-symposium-presentations/AB_Ver%20PRACE_DEISA%20Amsterdam.pdf
- Where of we stand
- Questions

HPC-Service 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

PRACE – an unusual ESFRI-list item

- There is no specific scientific community for the usage of HPC: it's **"the"** scientific community in total. Thus there is no special scientific community which "fights" for "money".
- There is not one infrastructure, not one experiment or machine to build and to operate (like e.g. ITER, FAIR or XFEL), but several independent machines each with possibly different architecture, (currently) national funding, build and operated by a European member state.
- Current Governance-Structure: usually the operators are in the "driver seat" and not the scientific user.

➔ **PRACE is challenged to reconsider this "culture"**

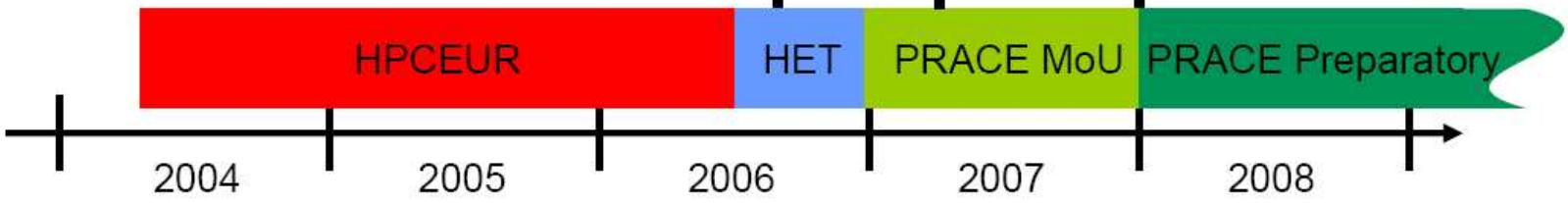
PRACE: An EU-Project and an European Initiative (MoU)

- 2007 April, 16: Memorandum of Understanding signed by 15 European member states in Berlin
- 2008: France, Germany, Spain, The Netherlands and UK reconfirmed their commitment for establishing an European HPC Research Infra-structure
- 2009: 4 new European member states have joined the PRACE initiative and more are interested

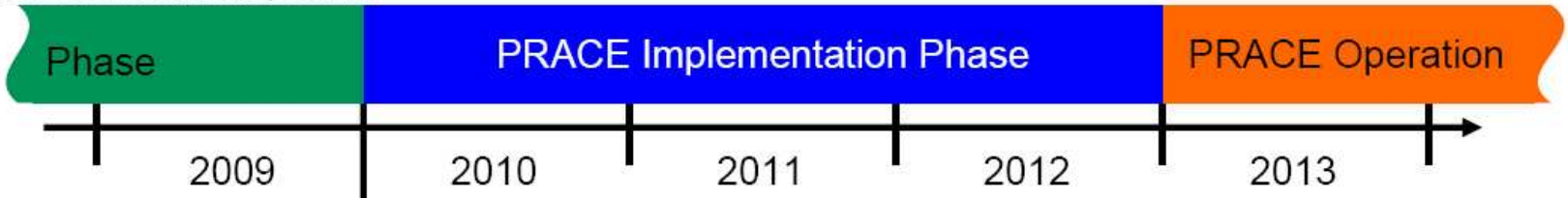
Europe needs access to peta- and exascale computing power in all of its member states

- There is no sustained HPC service beyond the national scale (excepts topical centres like ECMWF)
- Building the European Research Area: Access to HPC power must be organized on an European level:
  - Either via national funding and an appropriate compensation key (e.g. Just-Retoure) or
  - European funding e.g. via FP8
- We need the skills for peta- and exascale computing of all member states and not only of those which can afford to build up its own peta- and exascale installation for national use.

PRACE Timeline



EU-Grant: INFSO-RI-211528, 10 Mio. €



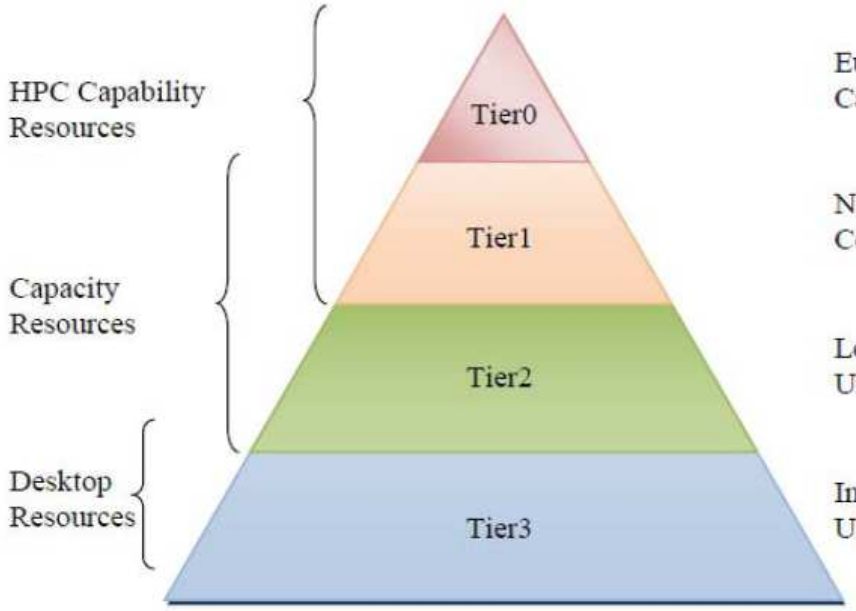
PRACE

Forseen: > 3 PRACE centres providing HPC-capability service in a legal entity

PRACE progress

- Collaborative achievement of over 250 persons at 16 European partner sites
- EU-Preparatory Project INFSO-RI-211528, 20 Mio. €, 50% EU-Funding
  - 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

WP2: Organizational concept



European Top Capability

National and Regional Centres

Local Facilities at Universities and Labs

Individual Terminals of Users

stakeholders

HPC vendors, Academic community, software companies…

Outline

- What is PRACE
- Where we stand
- Questions

Selected Results and Highlights of 2008

- Training and Outreach (WP3)
- Applications (WP6)
- Systems/Architectures

WP3: Outreach and Education

Industry seminars:

1st Seminar Sept. 3, 2008 Amsterdam, The Netherlands
2nd Seminar Sep. 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

Importance of training & education

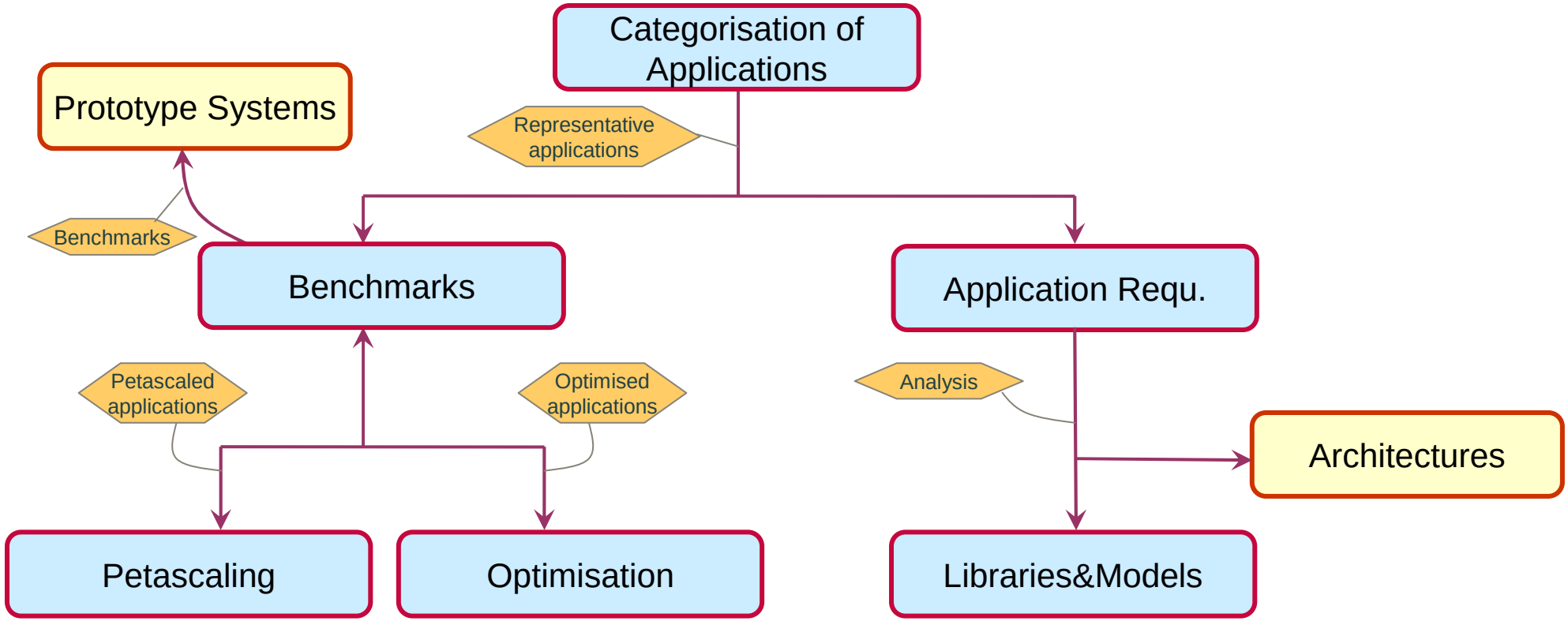
- More Schools to come
  - During 2009 and beyond

  + CEA/EDF/INRIA Summer Schools (& others)
- Training portal project
  - Cf Tim Stitt, CSCS – S. Requena, GENCI

Selected Results and Highlights of 2008

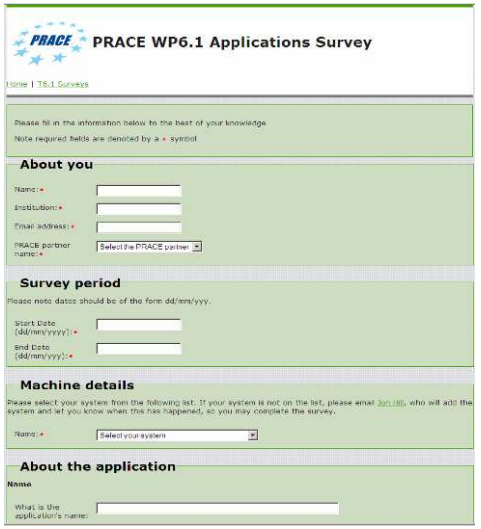
- Training and Outreach (WP3)
- Applications (WP6)
- Systems/Architectures

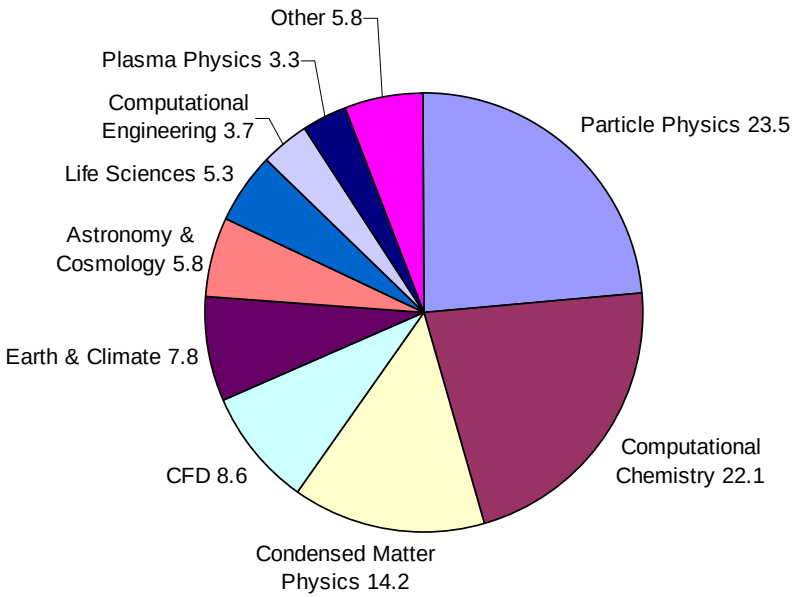
Software Enabling for Petaflop/s Systems = WP6



Categorisation of Applications

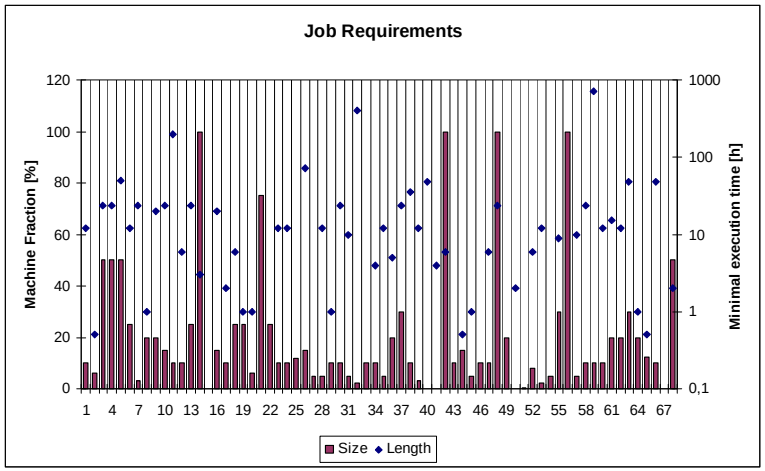
- Benchmark applications should be *representative* of European HPC usage
- We conducted surveys of PRACE partners' HPC systems and major applications
  - Collecting various interesting data for 24 systems and 69 applications
- Quantitative basis for selecting representative applications
- Disseminated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





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



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

Selected Results and Highlights of 2008

- Training and Outreach (WP3)
- Applications (WP6)
- Systems/Architectures… & prototypes

Prototypes

- Selection of systems or components to assess existing of future technologies and architectures
- EC + partners' funding, installation 2008+early 2009 so as to be useful during PRACE project
- WP7 = systems (available at full scale in 2010)
  Either:
  - **Access to existing production systems for scaling (e.g. BG, Juropa cluster)**
  - **Extension to existing system with latest components available (e.g. XT5)**
  - **New medium-size systems with latest processors (e.g. SX/9; Xeon 5500)**

  => Span all promising architectures available 2009-2010 (first round of acquisitions)
- WP8 = components (for supercomputers after 2010)

Installed prototypes WP7-WP5

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

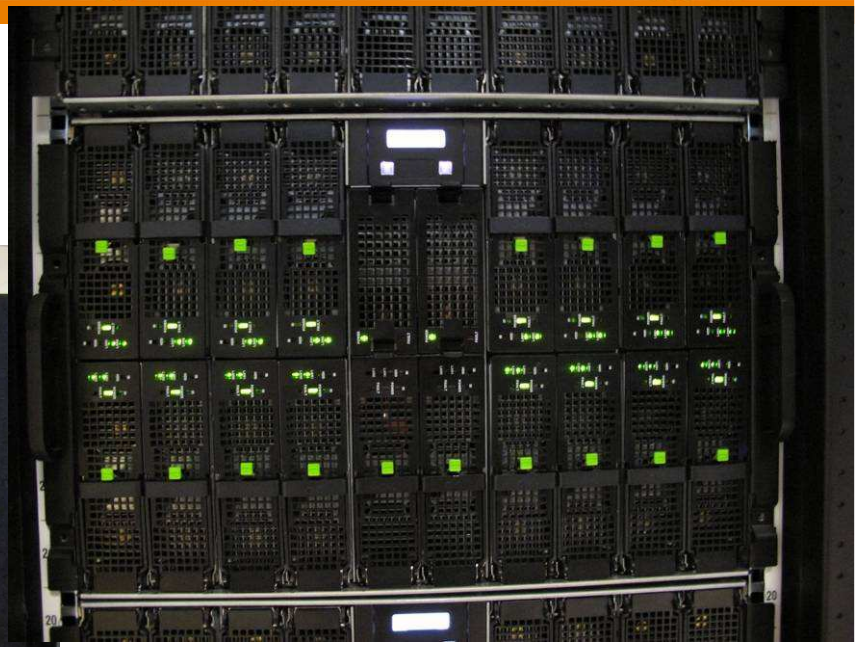
+ Intel Nehalem/Xeon 5500
04-2009 at CEA (Bull)
06-2009 at FZJ

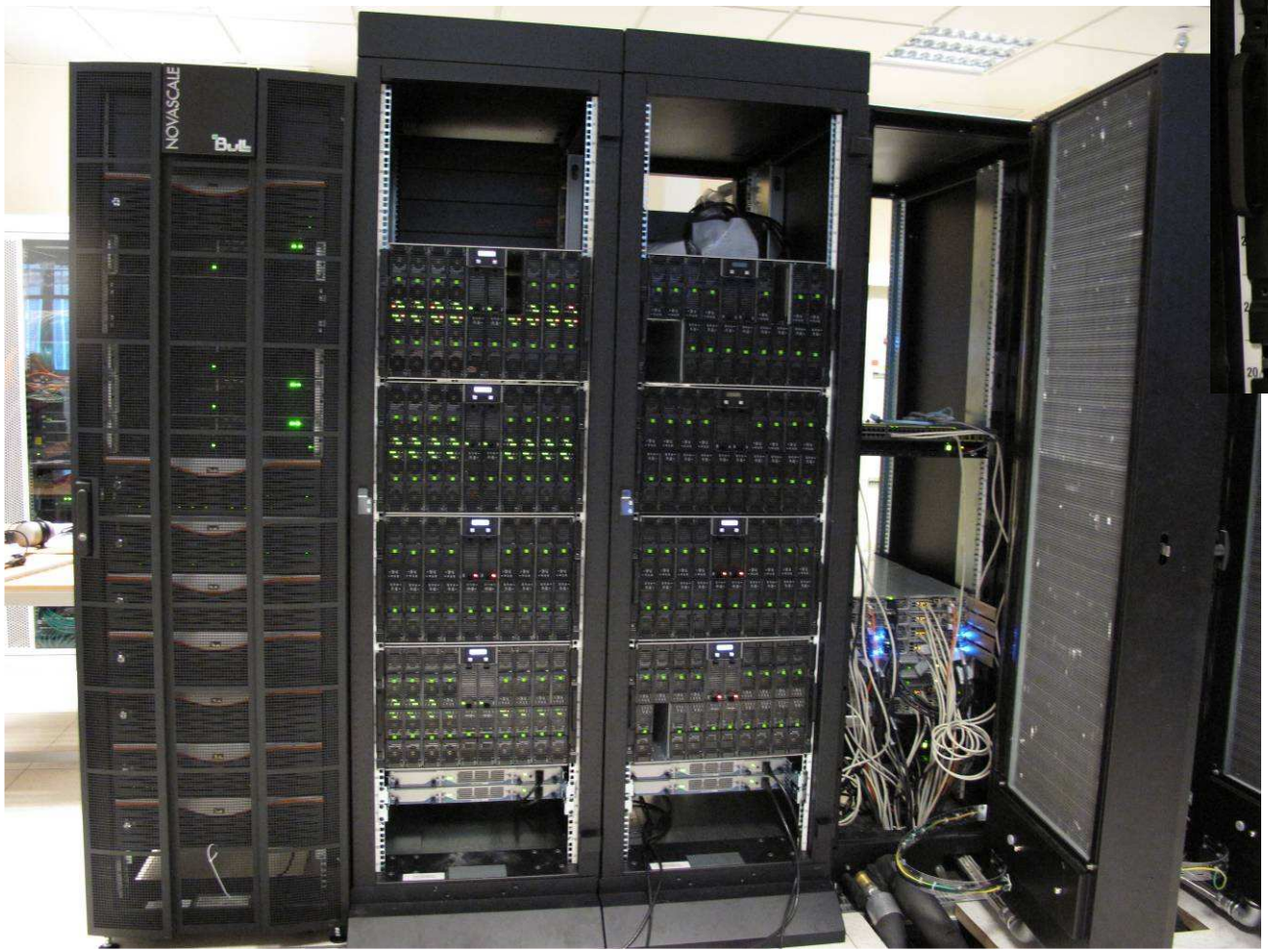
WP7-WP5 prototypes status

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	nearly	nearly
system in production	yes	yes	yes	yes	nearly	nearly
technical assessment	yes	nearly	yes	nearly	started	started
evaluation of communication and I/O infrastructure	yes	nearly	yes	nearly	started	started
evaluation and benchmarking of user applications	started	started	started	started	started (vector)	no

Last update: May 2009

Bull INCA at CEA – coming soon…





Prototype access procedure: go and try them!

- http://www.prace-project.eu/prototype-access

PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

FORM FOR REQUESTING PROTOTYPE ACCESS

A. GENERAL INFORMATION

Project name	
Research field*	

* *Please choose one of the four areas: 1) Astronomy, Aerospace and Earth Sciences; 2) Biomedicine and Life Sciences; 3) Chemistry and New Materials; 4) Physics and Engineering.*

1. Project leader

Last name	
First name	
Initials	
Date of birth	
Title	
Function	
Organisation name (main*)	
Department	
Group*	
Address	
Post code and city	
Country	
Nationality	
Phone number (direct)	
E-mail	

More information

Site	Architecture Vendor/Technology	Contact
FZJ Germany	**MPP** IBM BlueGene/P	Michael Stephan m.stephan@fz-juelich.de
CSC-CSCS Finland+Switzerland	**MPP** Cray XT5/XTn - AMD Opteron	Janne Ignatius janne.ignatius@csc.fi Peter Kunszt peter.kunszt@cscs.ch
CEA-FZJ France+Germany	**SMP-TN** Bull et al. Intel Xeon Nehalem	Gilles Wiber gilles.wiber@cea.fr Norbert Eicker n.eicker@fz-juelich.de
NCF Netherlands	**SMP-FN** IBM Power 6	Axel Berg axel@sara.nl Peter Michielse michielse@nwo.nl
BSC Spain	**Hybrid – fine grain** IBM Cell + Power6	Sergi Girona sergi.girona@bsc.es
HLRS Germany	**Hybrid – coarse grain** NEC Vector SX/9 + x86	Stefan Wesner wesner@hlrs.de

- **WP5 leader: Axel Berg, axel@sara.nl**
- **WP7 leader/co-leader: F. Robin, francois.robin@genci.fr - J.P. Nominé, jean-philippe.nomine@cea.fr**
- **http://www.prace-project.eu/documents/first-scientific-workshop/nomine_ICT08-PRACE-Prototypes.pdf**

Addendum : Prototypes selected by WP8 (1)

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 selected by WP8 (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)	-
	March 10th: Selection of prototypes targeting energy efficiency	-

More information in general: PRACE Web site

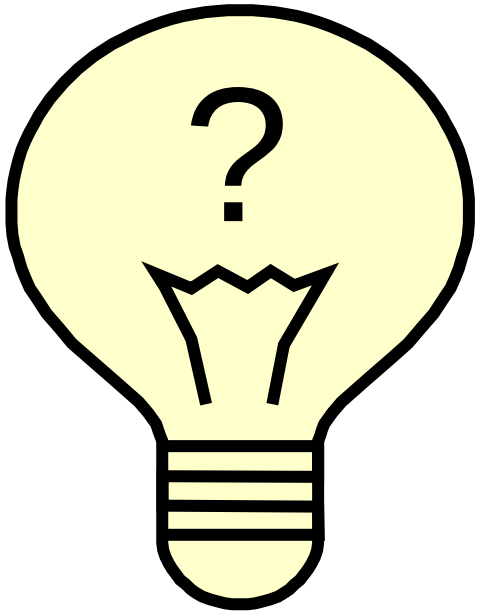
- The PRACE web presence with news, events, RSS feeds etc. http://www.prace-project.eu

The PRACE website, www.prace-project.eu

Outline

- What is PRACE
- Where we stand
- Questions

Thank you



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