

# NEWSletter

Volume 3 – September 2008

## Analysing Application Requirements for Petaflop/s Systems is a Major Challenge

Anni Jakobsson

**P**RACE's, Partnership for Advanced Computing in Europe, overall objective is to prepare for the creation of a persistent pan-European HPC service with three to five Petaflop/s scale systems. The project began on January 1st, 2008 and will last for two years.

The work of PRACE is organised into eight work packages which are designed to minimise interdependencies and to allow the preparation of the infrastructure to progress in parallel.

Dr. **Alan D. Simpson** from EPCC, in Edinburgh, UK, is co-leader of Work Package 6 (WP6) named "Software enabling for Petaflop/s systems".

WP6 is responsible for investigating software on future Petaflop/s systems. Its primary goal is to identify and understand the software libraries, tools, benchmarks and skills required by users to ensure that their application can use a Petaflop/s system productively and efficiently.

© EPCC



*Dr. Alan D. Simpson (in the picture) from EPCC, leads Work Package 6 named "Software enabling for Petaflop/s systems" together with Dr. Mark Parsons.*

"I am Project Manager for Work Package 6 which means that I am focused on delivery and management. I am responsible for coordinating all the activity (e.g. the deliverables and milestones) and for reporting to the Project Management Office – this coordination requires a significant amount of time", Simpson tells.

Alan leads WP6 together with Dr. **Mark Parsons**, from EPCC.

"Mark is the Task Leader and has a more outward-facing role, e.g. liaison with other Work Packages. Mark also coordinates the UK technical effort across the whole of PRACE. We could indeed reasonably be described as co-leaders, in that we jointly manage this Work Package along with **Peter Michielse** (NCF) and the other task leaders", he continues.

### The main objectives of Work Package 6

The initial work on porting and profiling the applications has resulted in a description of the likely applications requirements for European Petascale systems. WP6 conducted a major survey of the utilisation of HPC (High Performance Computing) resources across Europe. The objective of the survey was to understand the existing key applications in Europe and to identify likely applications that may be used on future Petaflop/s systems.

"Analysing the application requirements for these systems is a major challenge but is vital to ensure that the systems do meet the needs of European researchers", Simpson explains.

The result of this survey is a core list of nine applications which are representative of both the major scientific disciplines and the core algorithms using most of the compute cycles of the surveyed HPC systems. The list contains also a set of ten possible extensions.

Particle physics may well be the main scientific area on a future Petaflop/s system, alongside materials science and computational chemistry. Programming multi-core architectures is seen as a major challenge alongside parallel I/O and techniques for scaling applications to thousands of cores.

The results are being used to guide the se-

## Contents

### Page 1

Analysing Application Requirements for Petaflop/s Systems is a Major Challenge

### Page 2

The PRACE Award Winning Work Speeds Up Numerical Simulation Tools on Hybrid Hardware

### Page 3

PRACE a Highlight of ISC'08

### Page 4

PRACE Selects Promising Architectures for Prototypes

### Page 5

Featured Events

lection of applications for inclusion in the representative PRACE benchmark suite.

Next steps for the Work Package include defining the benchmark suite, which is a collection of benchmark codes.

"The major activity of WP6 will be focused on the optimisation and petascaling of the representative applications for the rest of the PRACE project", Simpson tells.

### Liaison with other Work Packages

WP6 works closely with WP7 and WP8 on architectural requirements for the first and the future generations of Petaflop/s systems.

"We provided the applications requirements that WP7 will translate into architectural requirements. This can help identify most relevant classes of architectures for the prototypes", Simpson explains.

WP6 also provides the benchmark suite used by WP5 for the assessment of the selected systems and works with WP3 on the dissemination.

"The benchmark suite will be an important tool in evaluating the prototypes. Subsets of the benchmark suite may be used for future Petascale procurements", he adds.

Work packages 1 and 2 are introduced in previous PRACE newsletters. They are available on [www.prace-project.eu](http://www.prace-project.eu)

# PRACE Award Winning Work Speeds Up Numerical Simulation Tools on Hybrid Hardware

Christopher Lazou

The applied mathematics team at Dortmund University of Technology, Professor **Stefan Turek**, **Dominik Göddeke**, **Christian Becker**, **Sven H.M. Buijssen** and **Hilmar Wobker**, were joint winners of the PRACE Award for the best paper by young scientists at ISC'08. The award was given for their work, UCHPC – UnConventional High Performance Computing for Finite Element Simulations. The prize is a sponsorship for Dominik Göddeke, a Ph.D. student to participate in a conference relevant to Petascale computing.

The world's fastest computer – first Petaflop/s system – formerly known as RoadRunner is the logical and latest culmination of a trend that has become apparent over the last couple of years: The number of installed supercomputing systems in the TOP500 list, that have been assembled from commodity hardware has been on the increase recently.

PRACE, Partnership for Advanced Computing in Europe, is preparing the creation of a persistent pan-European HPC (High Performance Computing) service, consisting of several tier-0 centres providing European researchers with access to capability computers and forming the top level of the European HPC ecosystem. The service will comprise three to five Petaflop/s-systems strengthened by regional and national supercomputing centres.

## Parallelism and chip specialisation

We do not know what the Petascale era will bring. What is clear however is that physical limitations (heat, leaking voltage, pin limits)



*Dominik Göddeke (left.), Sven Buijssen, Dr. Christian Becker, Prof. Stefan Turek and Hilmar Wobker are the winners of the PRACE Award. Behind the group is the mathematics department of the Technical University of Dortmund, Germany.*

have led to a paradigm change. Performance improvements are no longer driven by clock frequency scaling, but by parallelism and chip specialisation. Indeed the first Petascale system ever – the RoadRunner – assembled by researchers from IBM and Los Alamos National Laboratory, is a heterogeneous, hybrid, system. The RoadRunner comprises of 6,562 dual-core AMD Opteron chips and 12,240 Cell processors.

“Modern High Performance Computers are very complex systems and consist of many elements, processors, memory and an intricate network for interconnection so that all these elements can efficiently process data from an application and deliver the results to the user. Whilst the raw HPC processor performance is potentially outstanding, it does not necessarily translate to faster computation for all applications. The efficiency of many simulations is restricted by the cost of memory access, rather than by the available compute peak performance, and presently smaller advances are expected in memory and networking technology for moving data. The challenge imposed by this “memory wall problem” remains unaffected”, says Göddeke.

## Crash tests and aircraft design

This “memory wall” aspect is very important for Finite Element (FE) software. This software

technique is used for solving large systems of partial differential equations and has many applications in engineering and life sciences.

“From an engineering and mathematical point of view, FE approaches are very advantageous due to their flexibility and accuracy. Together with powerful and robust numerical solution schemes, they form the underlying fabric of many modern simulation tools. Important application domains that we pursue in this setting are computational solid mechanics (CSM) and computational fluid dynamics (CFD)”, says Göddeke.

Practical applications in CSM and CFD fields range from material failure tests (for instance crash tests in automotive industry) to fluid and gas flow of any kind, in aircraft design, weather prediction, nuclear reactors and in chemical or medical engineering, e.g. simulation of blood flow in the human body to predict aneurysms.

## Clever software methods required

On the one hand, technological trends lead to a heterogeneous, hybrid, hardware landscape, reflected in diversified HPC systems. On the other hand, mathematical theory continues to make great advances, and many arguments support the use of “high-end” numerical techniques for “high-end” (as in Petascale) computer systems. The challenge is to combine these often contradictory goals into codes that are at



the same time future-proof with respect to hardware evolution, obtain a reasonable amount of the available peak performance, scale well in parallel settings, and still allow extensibility and maintainability at reasonable costs. Experience shows that compiler technology is not able to efficiently parallelise serial codes automatically, despite promising advances of the latest compiler suites.

"The Finite Element Analysis & Simulation Tools (FEAST) project is our humble attempt to address these seemingly contradictory goals. Obviously, we cannot influence hardware design. Our aim is thus the continuous development and adaptation of numerical methodology respecting long-term hardware trends. In a nutshell, we believe that only clever software approaches will be able to harness the power of ever so heterogeneous Petascale systems", says Göddeke.

Implementation details aside, the general agreement in the community is that an optimal (in the mathematical sense) "textbook algorithm" does not automatically yield optimal per-

© Anni Jakobsson



*PRACE coordinator, Prof. Dr. Achim Bachem (right) presented the PRACE Award to Prof. Stefan Turek at the ISC'08 opening session.*

formance. Hardware characteristics do not play a role in numerical convergence proofs which is why different hardware usually requires modified algorithms. There is always a trade-off between numerical efficiency (in terms of number of operations per unknown to reach a reasonable solution) and parallel efficiency (in terms of strong and weak scalability). FEAST tries to optimise this trade-off and to hide it from the application level.

### Unconventional High Performance Computing

The success of the RoadRunner shows that

"unconventional" hardware is indeed becoming part of the hybrid Petascale landscape. In this particular case, the Cell processor drives performance.

The winning work concentrates on graphics processors (GPUs) due to their enormous bandwidth to off-chip memory. While GPUs certainly do not solve the "memory wall problem" (the codes are still limited in performance by the available bandwidth), they succeed in pushing the achievable performance by more than an order of magnitude.

This specialisation and the use of heterogeneous compute resources results in a major challenge for the programming model, as each co-processor is programmed in a different way and through different interfaces. Achieving maximum performance is hard, and can often only be realised by clever programming and novel algorithms. More importantly, programmers are usually not willing or able to adapt their applications to each new hardware generation that becomes available.

"We attempt to tackle this problem. Instead of accelerating individual operations, we take advantage of the block-regular data structures and accelerate entire solvers for sub-problems. We have demonstrated the success of this approach by achieving significantly faster execution of a large-scale solid mechanics code without any changes to the application; the user just "toggles a switch" and gets the answer to his problem much faster than before. Good scalability means that we have been able to solve a model problem with more than 1.2 billion unknowns in less than 50 seconds on a commodity based cluster with 160 GPUs, just two seconds slower than 300 million unknowns on 40 GPUs", says Göddeke.

### Free ride is over

Hardware aspects used to play a minor role in the development of numerical algorithms, as codes automatically ran faster with new hardware generations. Now this "free ride" is over!

Looking at current trends in numerical simulation methods and high performance computing techniques, dedicated hardware-oriented numerical methods developed in the FEAST toolbox, appear to be a key tool to significantly improve Finite Element (FE) software packages for highly challenging problems in engineering disciplines and life sciences.

FEAST not only uses "unconventional" hardware but also requires no reprogramming of user codes. Without FEAST like software, the general expectation that faster hardware automatically allows the simulation of more complex problems cannot be met. This is due

to the increasing gap between sustained and peak performance.

As Professor Turek says: "In fact, if we do not change our view now and respect technology trends in the design of numerical simulation tools, in a few years from now FE software will not be able to live up to engineering and industrial requirements".

This is the first time the prize has been awarded. The PRACE Award will be continued in the years to come to encourage young European scientists to work on innovative solutions for Petascale computing. The award winning paper can be found on: <http://www.mathematik.tu-dortmund.de/lsiii/static/showpdf/PraceGoeddekeBeckerBuijssenWobker2008.pdf>

## PRACE a Highlight of ISC'08

PRACE took an active role in the International Supercomputing Conference 2008. ISC'08 was the 23rd Annual International Supercomputing Conference, which was held at the Congress Center in Dresden on June 17-20.

PRACE was introduced in the ISC'08 opening ceremony by Prof. Dr. **Hans Meuer**, the ISC'08 general chair. Afterwards PRACE coordinator, Prof. Dr. **Achim Bachem**, Chairman of the Board Forschungszentrum Jülich presented the first PRACE Award.

PRACE was also among the 90 other exhibitors that attended the ISC'08 exhibition. As well as exhibiting a booth, PRACE arranged a Birds-of-a-Feather (BoF) session, entitled "Building the HPC Ecosystem". The session included presentations from Managing Director **Kimmo Koski** from CSC - IT Center for Science, Finland, the Director of RZG, **Stefan Heinzl**, from the Max Planck Institute, Garching Germany and Deputy Director of NCSA, **Rob Pennington**, from the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign.

Prof. Dr. Achim Bachem introduced PRACE in a conference session with a presentation: "PRACE: An Initiative for the Future European Supercomputing Infrastructure".

PRACE was mentioned as one of the highlights of ISC'08 in a HPCWire article on 21st of June.

PRACE will attend ISC'09 which will be held in Hamburg, Germany.

All BoF-session presentations can be found on [www.prace-project.eu](http://www.prace-project.eu).

# PRACE Selects Promising Architectures for Prototypes

PRACE, Partnership for Advanced Computing in Europe, has selected a broad coverage of promising architectures for Petaflop/s-class systems to be deployed in 2009/2010. Prototypes will be installed at six partner sites starting in 2008.

PRACE analysed key scientific applications and mapped them to suitable architectures. As a result six prototypes were selected including more advanced hybrid systems.

"Our objective is to build the best set of prototypes for preparing a timely and seamless deployment of production systems in 2009/2010 – not to attempt to select the best individual prototypes", says **François Robin**, from CEA/GENCI.

The prototypes will be installed at the following PRACE partner sites:

★ BSC (Barcelona Supercomputing Center) installs a hybrid prototypes combining IBM Cell and Power6 processors. The Cell processors are used for computation; the Power6 processors for system administration.

★ CEA (French Atomic Energy Commission) and FZJ (Forschungszentrum Jülich) jointly use Intel Nehalem/Xeon processors in their systems. Two shared-memory multiprocessors (thin node clusters) will be distributed over the two sites; a prototype produced by BULL at CEA and a larger system of the same architecture at FZJ.

★ CSC (IT Center for Science, Finland) and CSCS (Swiss National Supercomputing Centre) jointly evaluate the Cray XT5 architecture. This Massively Parallel Processing (MPP) prototype will be installed at CSC's facilities.

★ FZJ (Forschungszentrum Jülich) provides its already installed IBM BlueGene/P as a Massively Parallel Processing system.

★ HLRS (High Performance Computing Center Stuttgart) will evaluate a NEC SX-9 coupled to an x86 based cluster as a hybrid prototype.

★ SARA Computing and Networking

Services (the Netherlands) will evaluate the IBM Power6 architecture, a fat node cluster (shared-memory nodes) prototype. This prototype will be installed at SARA in Amsterdam.

## For many purposes

The prototypes will be used to evaluate the architectures in near-production situations with regard to application performance and scalability, as well as total cost of ownership and energy consumption.

They will also make possible the evaluation of software for managing the distributed infrastructure, the preparation of benchmarks for future Petascale systems allowing better understanding of user requirements, the scaling and optimisation of libraries and codes and the definition of technical requirements and procurement procedures for the PRACE Petaflop/s production systems for 2009/2010.

## Prototypes are of high importance

Supercomputing systems rely on new technologies and every system is individually planned and constructed. As this constitutes a high risk particularly considering the high costs and the requirements on the infrastructure (power supply, cooling etc.) an appropriate measure to reduce this risk is necessary. A proven way for reducing this risk is the deployment of one or more prototype systems preceding the actual production machine.

In this way all areas from operation, over application performance, infrastructure requirements up to usage strategies of a future system can be analysed in advance. This reduces the preparation time for the production system and supports the decision process. This is extremely important as the typical time during which a supercomputer remains a world-class system is only 2 or 3 years.

For PRACE the prototype phase is even more important since Petaflop/s scale computing imposes significantly new challenges for application software: utilizing the next generation supercomputers efficiently requires parallelization up to 100,000s of processor cores.

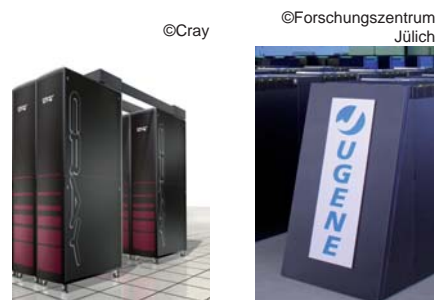
In addition, the current trend toward multi-/manycore processors and potentially hybrid systems built using different processor types within one single system make the situ-



BSC in Barcelona will evaluate a hybrid prototypes combining IBM Cell and Power6 (picture in left) processors. SARA Computing and Networking Services (the Netherlands) will evaluate the IBM Power6 architecture (right).



A BULL prototype will be distributed over CEA, France and FZJ in Jülich, Germany.



Cray XT5 hardware (left) will be installed at CSC's facilities in Espoo, Finland. IBM BlueGene/P (right) in FZJ, Jülich, Germany will be used as a MPP prototype.



NEC SX-9 (in picture) and x86 based cluster in Stuttgart, Germany will be used as a prototype.

ation even more complex. Accordingly, the codes and libraries must be tuned on the prototypes so that they are up to the task.

Also of major importance is the fact that the deployment of the prototypes will tighten relationships between the sites that will implement and operate a European infrastructure in a coordinated way, thus making the deployment and operation of the infrastructure more efficient.

"The prototypes have been selected after a careful analysis of the HPC market and

trends on the basis of proposals prepared by the PRACE partners. A key element in the success of the process is the outstanding expertise in HPC of the PRACE partners and their commitment to work together. The PRACE partners, from 16 European countries, will continue to work together and to share their expertise for the next phase: testing in near-production condition the prototypes", Robin concludes.

More information: PRACE work package 7 has been leading the evaluation work. Work package 7 leader François Robin from CEA/GENCI will give more information on request.

E-mail: [francois.robin@genci.fr](mailto:francois.robin@genci.fr)

[www.prace-project.eu](http://www.prace-project.eu)

## Featured events



### Cracow'08 Grid Workshop, CGW'08

13–15 October, Cracow, Poland

The main objective of the Cracow Grid Workshops, which were initiated in 2001, is to support the community of researchers, developers, and practitioners who work in the fascinating field of e-Science, grid systems, and their applications. PRACE dissemination material will be present at the event.

<http://www.cyfronet.krakow.pl/cgw08/>

### ICT-BIO 2008

23–24 October, Brussels, Belgium

The ICT BIO 2008 Conference 23–24 October 2008, will provide an insight into existing and future prospects on integrative research for complex human diseases at European and international levels. PRACE will be present at the event.

[http://ec.europa.eu/information\\_society/events/ict\\_bio/2008/index\\_en.htm](http://ec.europa.eu/information_society/events/ict_bio/2008/index_en.htm)



### SuperComputing 08, SC'08

15–21 November, Austin, USA

The SC Conference is the premier international conference for high performance computing (HPC), networking, storage and analysis. SC08 marks the 20th anniversary of the first SC Conference, then called Supercomputing, held in Orlando, Florida in 1988.

When SC08 opens November 15, 2008 in Austin, Texas, the conference will feature the latest scientific and technical innovations from around the world. Bringing together scientists, engineers, researchers, educators, programmers, system administrators and managers, SC08 will be the forum for demonstrating how these developments are driving new ideas, new discoveries and new industries.

PRACE will have an exhibition booth at the SC08 conference in the exhibition hall.

<http://sc08.supercomputing.org/>



### PRACE scientific workshop

26 November, 2008, Lyon, France

PRACE is arranging a scientific workshop on November 26th in Lyon to be held in parallel with ICT 2008.

The themes of the scientific workshop are Applications, Architectures and Training needs for the Petascale Regime.

Scientific workshop's evening session with presentations and discussion will take place on the evening of November 26th at Hotel Cite de la Concorde, just a stone's throw from the Lyon Congress Centre.

If you are interested in participating in the PRACE scientific workshop, please complete the registration form on the PRACE website or send an e-mail with your contact information to the following e-mail addresses:

[anni.jakobsson@csc.fi](mailto:anni.jakobsson@csc.fi), [tim.robinson@cscs.ch](mailto:tim.robinson@cscs.ch)

<http://www.prace-project.eu/events/>

[http://ec.europa.eu/information\\_society/events/ict/2008/index\\_en.htm](http://ec.europa.eu/information_society/events/ict/2008/index_en.htm)



#### Coordinator/Contact:

Prof. Dr. Achim Bachem:  
[prace-coordinator@fz-juelich.de](mailto:prace-coordinator@fz-juelich.de)

[www.prace-project.eu](http://www.prace-project.eu)



The PRACE project receives funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° RI-211528.