



PDC Center for
High Performance Computing

no:2²⁰¹³

PDC Newsletter

The Human Brain Project - Jeanette Hellgren Kotaleski, Anders Lansner & Erwin Laure, page 4

New HPC System for Neuroscience Research at PDC - Gert Svensson, PDC, page 7

CRESTA Year Two Report - Stefano Markidis, HPCViz, page 7

B2Share: A Simple Storage Service within EUDAT - Erwin Laure & Ilja Livenson, PDC, page 8

Introducing PDC Support - Genet Edmondson & Jonathan Vincent, PDC, page 10

EPIGRAM Emerges! - Stefano Markidis, HPCViz, page 12

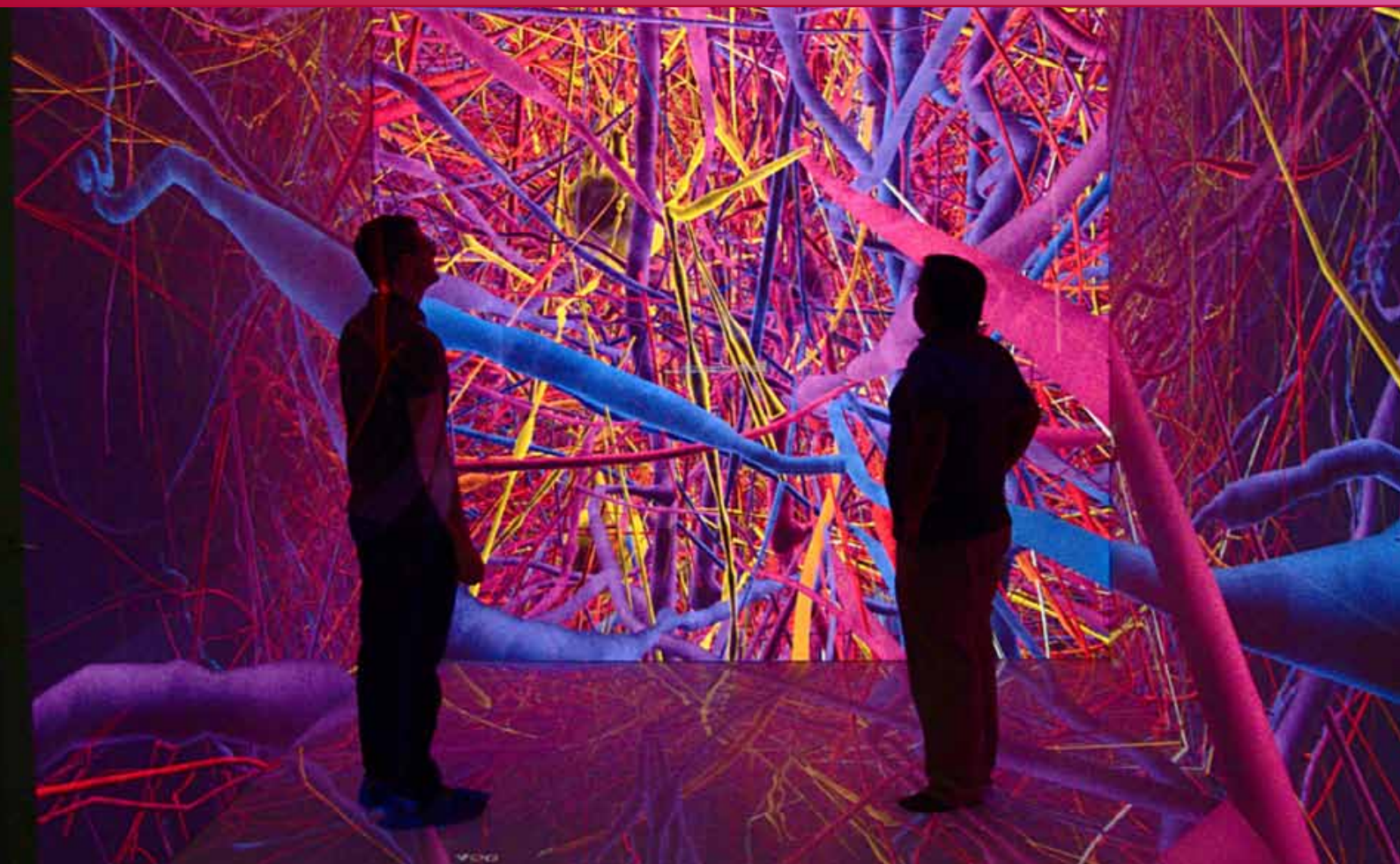
ScalaLife Legacy Lives On - Rossen Apostolov, PDC, page 13

PRACE-2IP Wraps Up - Michaela Barth, PDC, page 16

PDC Summer School - Stefano Markidis, HPCViz, pages 17

SNIC Galaxy Update - Åke Edlund, PDC and HPCViz, page 18

Inaugural NeIC Conference - Michaela Barth, PDC, page 19





Erwin Laure,
Director PDC & HPCViz

Published by PDC at CSC, KTH

PDC operates leading-edge, high-performance computers as easily-accessible national resources. These resources are primarily available for Swedish academic research and education. PDC, which is hosted by CSC, KTH, is one of the six centres in the Swedish National Infrastructure for Computing (SNIC).

Editor: Erwin Laure
Editorial staff: Genet Edmondson
Layout: Genet Edmondson
Graphic Design: Maria Malmqvist
E-mail: pdc-newsletter@pdc.kth.se
ISSN 1401-9671

Cover

The Human Brain Project

How to Contact PDC

Visiting address:
Teknikringen 14, "Plan 4",
KTH, Stockholm
Mail: PDC, KTH,
SE-10044 Stockholm, Sweden
E-mail: support@pdc.kth.se
www: <http://www.pdc.kth.se>
Phone: +46 8 790 7800
Fax: +46 8 247 784

On the 7th of October, one of the largest and most ambitious pan-European research projects was launched in Lausanne, Switzerland: The Human Brain Project (HBP). With a budget of more than a billion Euros for the coming decade, the project will be devoted to research aimed at giving us a much better understanding of the human brain. To tackle this far-ranging goal, a unique combination of researchers and resources from the fields of neuroscience, medicine, and computer science is being brought together. As Philippe Gillet (interim president of EPFL) pointed out during the HBP Summit, the project is comparable to the US Apollo program. PDC, together with neuroinformatics researchers from other departments at KTH and KI, is contributing to this exciting project. We will focus on developing what are known as "neuromorphic" computing techniques, namely novel computing solutions inspired by the way the human brain functions. To help with this work, we are also procuring a new HPC system for neuroinformatics research that will be coupled to neuromorphic hardware, namely the Spinnaker system developed at the University of Manchester. Consequently, the cover article of this edition is devoted to the Human Brain Project and we also give an update on the ongoing procurement process for KTH's new neuroinformatics computing platform.

Many of you will remember PDC's 20th anniversary celebration in autumn 2010 that also included the inauguration of our flagship computer, Lindgren. By January 2011 Lindgren had grown to reach its final size of over 305 teraflops, making it one of the largest super-computer systems in the Nordic countries at that time. Since then Lindgren has been a reliable and efficient partner in many scientific endeavours, but the sands of time have been running, and so Lindgren will be in its fourth year of operation next year, and will need to be replaced during the second half of 2014. The SNIC board has recently given the SNIC director a mandate to start negotiations with KTH regarding a potential follow-up system. We hope that these negotiations will be successful and that a new system with a capacity in the petascale range will be available to our users by late 2014.

As you well know, hardware is only one aspect of PDC's service offerings. Of vital importance is our support team who are dedicated to improving the experience our users have on the PDC systems, and who work towards a more efficient usage of our hardware. In this edition we begin a new mini-series looking into the internal workings of PDC, starting with our front-line lads in support. Improving the performance of applications is also the main goal of a number of related activities that PDC is pursuing, often in collaboration with

KTH's HPCViz department. In this issue we report particularly on the work performed by the ScalaLife project, our activities within PRACE-2IP, the CRESTA exascale flagship, and a new project looking at exascale programming models (EPiGRAM) that will start in November 2013.

While supercomputers like Lindgren are typically used in traditional batch mode, many researchers are not computer experts and therefore find it useful to have access to higher-level tools that provide easy-to-use interfaces for running their workflows on such supercomputers. PDC, together with Uppsala University, started a pilot to provide the Galaxy platform as a scientific workflow, data integration, and data and analysis persistence and publishing platform within SNIC, initially targeting specifically the bioinformatics user community.

Ease of use is an aspect we also work with in the EUDAT project, where we are developing B2Share – a simple storage service targeting the “long tail” of science, that is, data from users with pressing data management needs who are however not associated with larger organizations that take care of their data plans. Although there are currently no immediate plans to provide B2Share within SNIC, it could evolve into a useful complementary service within the SweStore portfolio.

Finally, NeIC, the Nordic e-Infrastructure Coordination, held its inaugural conference in May and we look forward to new NeIC-inspired initiatives in the near future.

May you all enjoy a rewarding autumn and winter season of research and benefit from our PDC resources!

Erwin Laure, Director PDC and HPCViz



PDC Summer School 2013 - Making new friends at the Summer School Picnic

In This Issue

Editorial

Erwin Laure 2

Staff Focus

Radovan Bast 4

Mohamad Rezaei 6

Peter Gille 10

Anders Järleberg 10

Magnus Helmersson 11

Jonathan Vincent 11

Åke Edlund 18

The Human Brain Project

Jeanette Hellgren Koteleski, Anders Lansner & Erwin Laure 4

New HPC System for Neuroscience Research at PDC

Gert Svensson 7

CRESTA Year Two Report

Stefano Markidis 7

B2Share: A Simple Storage Service within EUDAT

Erwin Laure & Ilja Levinson 8

Introducing PDC Support

Genet Edmondson & Jonathan Vincent 10

EPiGRAM Emerges!

Stefano Markidis 12

AVL Fire Licenses 12

ScalaLife Legacy Lives On

Rossen Apostolov 13

PRACE-2IP Wraps Up

Michaela Barth 16

PDC Summer School 2013

Stefano Markidis 17

SNIC Galaxy Update

Åke Edlund 18

Inaugural NeIC Conference

Michaela Barth 19

PDC-Related Events 20

HPC Sources 20

Staff Focus



Radovan Bast

Radovan Bast recently joined the PDC support group as an application expert for quantum- and computational-chemistry. Radovan grew up in the Slovak Republic, the Czech Republic and Germany. He studied chemistry in Marburg, Germany. After a research stay in Auckland, New Zealand, a Ph.D. in Strasbourg, France, a post-doc in Tromsø, Norway, and a research position at CNRS in Toulouse, France, Radovan finally came to Stockholm to join PDC and the Department of Theoretical Chemistry and Biology at KTH.

Radovan enjoys programming and believes in open source. He is contributing to the DIRAC, Dalton, and OpenRSP programs, as well as a number of side-projects. On rainy days, Radovan travels in time to admire the bitmap artwork and sound of the computer games from the late 80s and early 90s. On sunny days, he is attracted to skydiving drop zones!

The Human Brain Project

Compiled by Jeanette Hellgren Koteleski, Department of Computational Biology, CSC, KTH, Anders Lansner, Department of Numerical Analysis and Computer Science, SU, and Department of Computational Biology, CSC, KTH, and Erwin Laure, PDC and HPCViz

The Human Brain Project (HBP) is one of two EU Future and Emerging Technologies (FET) Flagship Initiatives. The project aims to improve our understanding of the human brain, so as to find better and more effective treatments for brain diseases, and, in the process of modelling the brain in more detail, we are also aiming to develop novel computing technologies inspired by the way our human brains work.

To speed up the process of understanding how our brains function, it is crucial that we have an appropriate strategy for integrating new experimental data with the knowledge we already have. Therefore, a primary goal for HBP is to develop an ICT platform called the Brain Simulation Platform – this system will include computer hardware and a software environment for building models of the human brain, and also running simulations of brain behaviour, at multiple levels of biological detail.

We need to be able to model the brain on different levels of biological detail as there are so many different levels of complexity to consider within the brain – for example, we can model the electro-chemical interactions between individual cells, or look at how different regions of the brain (such as the frontal lobe, visual cortex and so forth) interact with each other.

Our models of the different levels of the brain need to take into account what we already know about the way the brain functions on the relevant level, and should give results equivalent to what we find happening in real brains. Within HBP, we refer to such models as “unifying models”. Generally these unifying models start off representing one level of biological detail, but they can give rise to phenomena at a higher level.

For example, if we start with a model of the electrical behaviour of some individual neurons connected in a network (such as we have in our brains), and then consider the electrical effect that results from the combination of the electrical activities of all those neurons, the result could be an EEG signal. This result takes us to another level of brain description or model, as an EEG signal is not something that can be produced by an individual neuron.

The Brain Simulation Platform will provide novel services to the neuroscience community by providing an environment where neuroscientists can create and use software tools, algorithms and workflows in order to build these kinds of unifying quantitative

Staff Focus



Mohamad Rezaei

Mohamad Rezaei obtained his Masters degree in computer science (software systems) from the Iran University of Science and Technology. He is passionately interested in operating systems, virtual machines, resource management, and dependable computing. After his Masters studies, Mohamad worked in Iran's operating system research centre for four years as a senior researcher. During that time, he participated in projects related to development for the Linux kernel, Bitvisor and Xen hypervisor. That participation resulted in the implementation of several products that improved the trustability of Linux in virtualised environments. Mohamad also has more than three years of teaching experience in several universities in Tehran, and was also the leader of LPI's (Linux Professional Institute) academic relations in Iran.

Mohamad came to PDC to do a Ph.D. in data intensive computing for big data environments. Since he has been more focused on development and hacking, Mohamad felt the need to balance his experiences by putting more effort into mathematics and scientific computing to become a balanced "e-science" researcher. At the moment his focus will be dynamic resource management of applications in large-scale clusters.

The Brain Simulation Platform will not be the only platform in HBP – several other platforms will be needed to handle different tasks. We will develop another ICT platform, the HBP Neuroinformatics Platform, to supply data for building the quantitative brain models. There will also be a High-Performance Computing (HPC) platform that will provide supercomputing resources and support for simulating models of the brain. Furthermore, the models constructed using the Brain Simulation Platform will be used in closed loop experiments in a Neurorobotics platform or may be transferred for use in the Neuromorphic Computing Platform.

Where possible these platforms will be based on “neuromorphic” and brain-like hardware designs that will reduce the power envelop, footprint and price for chips and systems dedicated to these large-scale real-time brain simulations. This is expected to open up novel approaches to, amongst other things, artificial perception and motor control resulting in the potential to approach capabilities corresponding to those of living biological creatures. These approaches could be deployed in robotic or virtual agents of different forms for the purposes of assisting humans in many important tasks.

Several universities and research organizations are participating in the HBP work, with KTH involved in several different activities: Jeanette Hellgren Kotaleski is responsible for coordinating the construction of the initial quantitative brain models at various levels of biological detail. Anders Lansner and Erwin Laure will provide neural models and new software solutions for the Neuromorphic Computing Platform and will also be involved in transferring these simplified brain models to the neuromorphic hardware.

A consortium lead by Anders Lansner at KTH has recently been awarded a grant from the Swedish Research Council (VR) for a Neuroinformatics supercomputer which will provide an excellent test bed for these adventures. The new system (which is planned to be up and running at PDC during the first months of 2014) will consist of a standard HPC system coupled with a neuromorphic system based on the SpiNNaker technologies (<http://apt.cs.man.ac.uk/projects/SpiNNaker>). This system will also be used to explore the possibilities for connecting artificial “brains” (based in the HPC system) with robotic devices.

For more information about the Human Brain Project, see <https://www.humanbrainproject.eu>.

New HPC System for Neuroscience Research at PDC

by Gert Svensson, PDC

Early in 2013 the Swedish Research Council (VR) approved a proposal from a consortium – consisting of KTH, KI and the INCF (International Neuroinformatics Coordinating Facility) – to establish a computational resource for neuroinformatics applications, granting the consortium 10 million SEK for this purpose. The consortium will make the new system available as part of a computing platform they are developing for Swedish neuroscientists. This platform will form a basis for increased Swedish participation in large-scale international efforts such as the Human Brain Project and the BrainScaleS project (<http://brainscales.kip.uni-heidelberg.de>). The HPC system will be used with novel neuromorphic hardware (that is, hardware based on how the brain works) from SpiNNaker.

The new system will mainly be used for

- simulation of large-scale models of neural systems, especially parts of the human brain,
- interaction of such simulations (or neuromorphic hardware simulations) with simulated virtual environments,
- interactive visualization and exploration of large neuroscientific datasets and large-scale neural network models, and
- control of robots using large neural models.

The procurement of the new system is an open EU procurement process, which means that any interested vendors can place a bid. The final decision will largely depend on the performance of three benchmark codes using corresponding datasets from the neuroscience field, although we will also take into account the predicted total cost of ownership (TCO), which includes costs for maintenance, installation, power and cooling. Once the decision is made, the installation should go ahead swiftly and we expect the new system to be in place and operating early in 2014.

CRESTA Year Two Report: Designing the Development Environment for Exascale

by Stefano Markidis, HPCViz

During the second year of the CRESTA project, Work Package 3 (lead by PDC) has focused on designing the CRESTA development environment for exascale computing projects. The CRESTA development environment comprises a new auto-tuning framework, a new run-time system, Score-P and Vampir monitoring and analysis tools, an Allinea DDT debugger and the MUST correctness checker.

A Domain Specific Language (focusing on parallel auto-tuning aspects) and a new run-time system (that schedules and executes tasks depending on available resources and application performance) have been designed as part of this process. Extensions and modifications to existing Score-P, Vampir, Allinea DDT and MUST tools have also been prepared.

While designing the CRESTA development environment, we have been careful to take into account the experience gained from testing the CRESTA development environment against the CRESTA applications, in accordance with the main co-design philosophy of the CRESTA project.

For more information about CRESTA, see <http://cresta-project.eu>.





PDC Open House & Pub Afternoon

Back in the spring, PDC opened its doors one fine afternoon and invited everyone who is using the PDC computing resources (or interested in using them in the future) to come and meet the PDC staff and see the facilities. About forty people from various departments at KTH (such as HPCViz and Mechanics) enjoyed a quiz, along with refreshments (which included the onerous task of sampling a range of about a dozen beers). The PDC Pub also provided a good opportunity to play games (created by KTH students) on the Microsoft Surface 1 from HPCViz, as well as meeting other PDC users and visiting the PDC computer hall.

The PDC Pub is an annual event, so look out for the invitation in your email next spring!



B2Share: A Simple Storage Service within EUDAT

Compiled by Erwin Laure and Ilja Livenson, PDC

Many researchers are facing the problematic question of finding a simple, convenient and durable way of storing and sharing their data. They often have large numbers of small files, for example, files containing derived data in the form of spreadsheets or analysis results. Although the information in these small files is important, usually these files do not belong to large research collaborations with well-defined data storage plans. Consequently these types of files are known as “long tail data”, and are often stored on laptops and departmental servers. Hence there is a risk of losing such scientific data, either because other researchers do not have easy access to the data or because such storage is often relatively insecure. Within the EUDAT project, PDC is leading the efforts to provide a simple storage service (now called B2Share, but previously referred to as “Simple Store”) that will address this issue by

- allowing registered users to upload typical “long tail” data objects into the EUDAT store,
- enabling users to share such objects and collections with other researchers, and
- utilising other EUDAT services to provide reliability and data retention.

There are similar services offered by other initiatives (for example, the OpenAire FP7 project) as well as by industry (for example, Dropbox and Youtube) that are focusing on very specific areas. The EUDAT B2Share service will provide a simple submission process, proper management of rights, long-term persistence of data, and metadata enrichment for a wide variety of the data types. However, it should be noted that this service is not designed for scientific workflows that generate very large data sets; the purpose of this service is to take care of the smaller data sets that are in danger of not being taken care of otherwise.

Initially the B2Share service is aimed at researchers who

- do not have adequate facilities for storing and sharing data,
- cannot guarantee long-term persistence of their locally-stored data, and
- lack the means to easily share data, results or ideas with colleagues worldwide.

To understand how the B2Share simple store service is used, consider a researcher who has generated a number of small files to support a certain experiment. She wants to share those files with other researchers via the web, or to make them publicly available.

Using the EUDAT service, she will open a Simple Store submission page (or rely on upload tools offered by EUDAT), drag her collection of files to the upload area, enter information in a few meta-data fields, specify the access permissions and then submit the data.

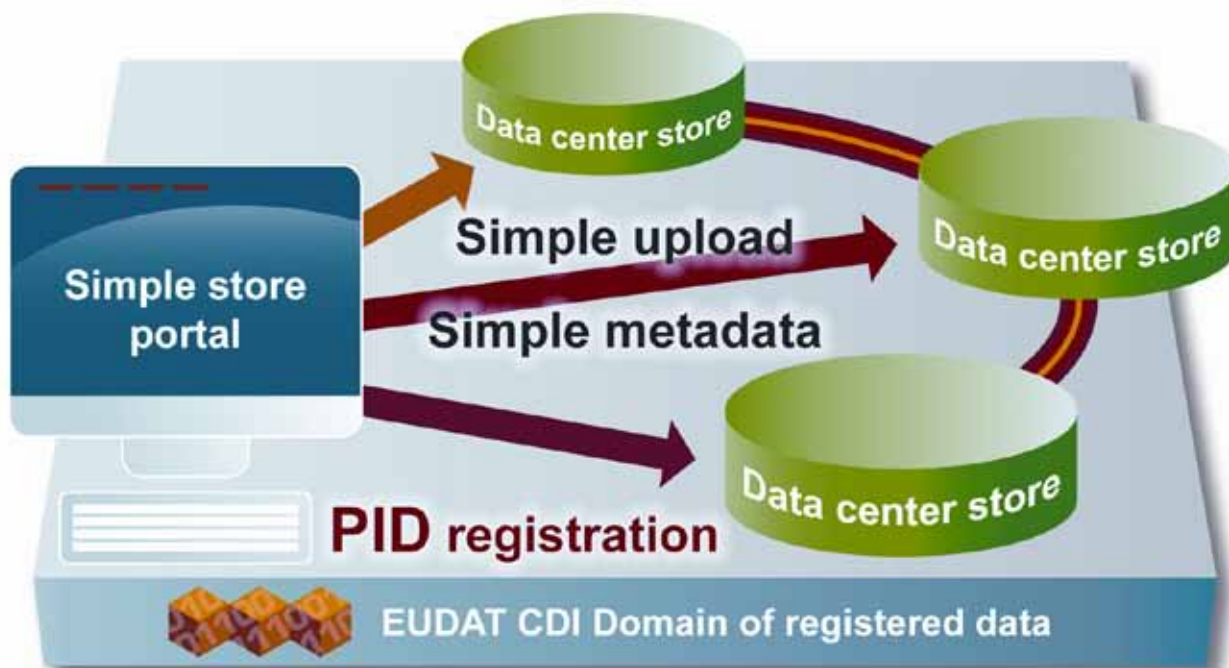
The data centre offering the service will respect the specified access restrictions, assign persistent identifiers (PIDs) to allow each object to be uniquely identified (for adding the objects to the EUDAT registered data domain) and store the data according to the service statements on the website. For the most common data types – such as text, audio and video files – custom players will be available so that authorized users will be able to directly view the contents of the data files. With a data service like this, it is important that all uploaded data relates to scientific research and is legally acceptable. Although it is hard to guarantee this, user registration and random content checks will be used to prevent the upload of non-scientific or illegal data, which will increase the level of confidence in the service.

Although the EUDAT project is initially funded for three years, we will ensure persistent access and services through mutually beneficial partnerships with data centres that have long-

term funding.

EUDAT will offer this service to all registered and reliable researchers throughout the world. To establish whether a researcher is reliable, EUDAT will require information such as the researcher's affiliation and contact details. Eventually the B2Share service will also support the Authentication and Authorization Infrastructure (AAI) mechanism offered by the AAI EUDAT service – this will simplify registration and login by offering trusted user credentials. To help researchers to use the B2Share simple store service, EUDAT will offer both educational material (such as screencasts) and a service helpdesk. To find out more details about this service, please contact the EUDAT B2Share simple store team: eadat-simplestore@postit.csc.fi.

EUDAT is happy to talk to any interested communities about all aspects of the B2Share service. While PDC is leading the development of the service, the initial service instance will most likely be offered by another EUDAT site. Discussions are also ongoing in the framework of NeIC as to whether a Nordic service should be offered. If you are interested in a local Swedish or Nordic instance of this service, please contact Erwin Laure (erwinl@pdc.kth.se).



Above: Overview of the B2Share simple store service

Staff Focus

Peter Gille is pursuing a Masters in computational chemistry at KTH, and has been working at PDC as a research engineer since last autumn. He provides user support and is also working on integration of the SNIC User and Project Repository (SUPR) to automate the handling of project members on PDC systems, which will give users a more consistent experience across the different supercomputer centres in Sweden.

In his spare time, Peter likes to read literature, watch movies and do some programming. He also volunteers at the Stockholm Early Music Festival every spring, although he does not play any instruments.

Anders Järleberg joined PDC's front line support in October 2012. In addition to answering support tickets, Anders also helps with documentation, changes to the website, maintaining the Lindgren CAC monitoring page, and installing software. Anders is in the last year of his Masters program in computer science at KTH, and is due to graduate in 2014, so he only works part-time at PDC in order to have time for his studies as well.

Anders is fond of music and has been playing the guitar for five years, and is also slowly learning to play the piano. In addition, he greatly enjoys reading and conversation, as well as movies and games for entertainment.



Peter Gille



Anders Järleberg

Introducing PDC Support

by Genet Edmondson and Jonathan Vincent, PDC

In this edition of the PDC newsletter, we start a gripping new series of articles unravelling the mysteries and workings of PDC. This time around, we tracked down Jonathan Vincent, Head of Support at PDC, to help us investigate PDC Support.

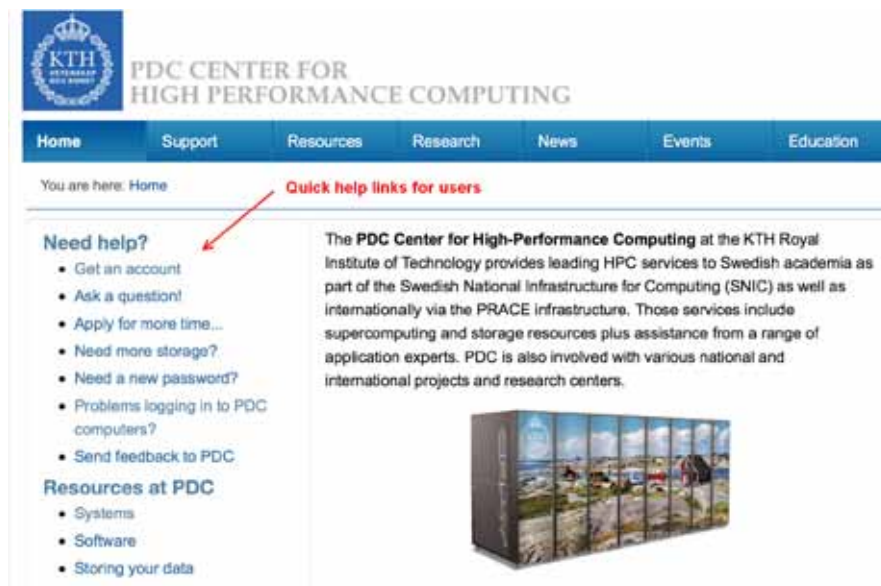
PDC Support is a group of about a dozen people, dedicated to helping researchers use the PDC supercomputing resources. There are three types of support provided by PDC – firstly, people need answers to basic questions such as how to get an account, how to log in, and what to do when a job is not executing or when more run-time is needed. Dealing with questions like these, installing software and taking care of other day-to-day tasks to do with user accounts all come under the umbrella of first-line support, which is mainly handled by Peter Gille, Magnus Helmersson, and Anders Järleberg. Peter, Magnus and Anders work part-time at PDC while completing their studies at KTH - you can read more about them in their profiles to the sides of this article.

PDC also offers second-line help in the form of Application experts – these are people who are experts in a particular area of research and who also have a lot of experience with parallel programming for supercomputers within their area of expertise. They deal with queries related to running programs in specific research areas. (In the next edition of the newsletter, we will introduce you to the Application experts and go into more detail about their research areas.)

Third-line support, which involves dealing with the more complicated queries, liaising with the system administrators, and taking care of other management tasks to ensure that everything is done correctly, is handled by Jonathan Vincent, who is the Head of Support at PDC and a full-time permanent member of the PDC staff.

Now that you know the basic structure of the support group, how do you get help from first-line support? As a first port of call, the PDC home page <http://www.pdc.kth.se> has a prominent section with links to many of the questions frequently asked by users.

Below: Screenshot of PDC home page showing quick help links



If you cannot find the answers you need on the PDC webpage, you can also send email directly to support@pdc.kth.se – in your email, please give details of your PDC username, the PDC system you are using and the relevant problem. If you need additional help and would like to meet someone from PDC support in person, you are welcome to book a time to come to PDC – please send an email to support@pdc.kth.se to request a meeting. Since many of the people who work in support work part-time, the best way to get an answer quickly is to send email, rather than trying to ring a specific individual.

PDC also invites anyone who is using our computing resources, or who has an interest in High-Performance Computing, to attend our ongoing seminars and best practise meetings, and also the annual PDC Pub (which is a good opportunity to meet with PDC staff and see the PDC computer facilities). Information about these events can be found on the PDC webpage <http://www.pdc.kth.se/events> or you can sign up for invitations from the pdc-announce mailing list - for details on subscribing, see <http://www.pdc.kth.se/about/mailling-lists>.



Magnus Helmersson



Jonathan Vincent

Staff Focus

Magnus Helmersson joined PDC in June 2013 and is currently working in the support group, where he helps users to debug problems and make the most of the resources. He is studying engineering physics at KTH and is currently completing an M.Sc. in scientific computing.

In his spare time, Magnus likes to produce Roots/Dub music and collect records in the same styles. He also enjoys spending time cooking, watching movies and playing games with friends and family.

Jonathan Vincent joined PDC as an application expert in 2009, and has been the Head of Support since May 2012. Prior to that, Jon worked with meteorological codes as a computational scientist in Reading, after post-doctoral studies in Uppsala, Santiago de Compostella, Leiden and Gothenburg. Jon's qualifications include a Ph.D. in semiconductor physics and a B.Sc. in theoretical physics, both from the University of Exeter, UK.

Outside of work, Jon does the usual boring things: watching sport (rugby and US football), playing computer games, and going to the pub. [Subeditor's note: We know for a fact that he sometimes does interesting things too like reading the latest Terry Pratchett.]

EPIGRAM

EPIGRAM Emerges!

by Stefano Markidis, HPCViz

PDC is leading a new EU-funded project, known as the Exascale ProGRAMming Models Project, or EPIGRAM for short. The project started on the 1st of November and will continue for three years. EPIGRAM is working on developing programming models for next generation supercomputers. The EPIGRAM project is being undertaken by a consortium consisting of five partners from four different European countries: KTH (Sweden), the Technical University of Vienna (Austria), The University of Edinburgh and Cray UK (United Kingdom), and the Fraunhofer society (Germany). The University of Illinois at Urbana-Champaign (USA) is also involved with the project as an associate partner. This consortium is being lead by Erwin Laure, PDC and HPCViz, (who is the project coordinator) and Stefano Markidis of HPCViz (who is the project manager).

The goal of the EPIGRAM project is to prepare Message Passing (MP) and Partitioned Global Address Space (PGAS) programming models for the coming exascale systems (that will be capable of delivering 10^{18} FLOPS). Although the Message Passing Interface (MPI) has emerged as the de-facto standard for parallel programming on current petascale machines, PGAS languages and libraries are increasingly being considered as alternatives or complements to MPI. However, there are severe problems with both of these approaches that will prevent them from attaining exascale performance. Therefore, the EPIGRAM project is addressing the main current limitations of the MP and PGAS programming models. The new concepts that EPIGRAM develops for MPI and PGAS (to enable them to reach exascale performance) will be developed and tested in relation to two particular applications in the domains of engineering (Nek5000 code) and space weather (iPIC3D code).

There will be a kick-off meeting for the EPIGRAM project on the 26th and 27th of November at KTH. If you would like more information about the EPIGRAM project, please contact Stefano Markidis (markidis@pdc.kth.se).

Sign Up Fast! Free AVL Fire Licenses at PDC

We have a limited number of free UPP licenses available for using the simulation software AVL Fire at PDC for research and educational purposes, as PDC has recently joined the Advanced Simulation Technologies (AST) University Partnership Program (UPP). For more information about this program see

<https://www.avl.com/web/ast/university-partnership-program>.

and for more information about AVL FIRE, see

<https://www.avl.com/web/ast/fire>.

If you are interested in running AVL FIRE at PDC, please contact support@pdc.kth.se to request a license and include details of the research project or course(s) where you would like to use AVL FIRE.



ScalaLife Legacy Lives On

Compiled by Rossen Aposotolov

Introduction

ScalaLife (Scalable Software Services for Life Science) is an EU-funded project which started in September 2010 as a collaboration between supercomputing centres in Stockholm (KTH), Barcelona (BSC), Munich (LRZ), and Oxford (OSC), together with research labs at KTH and IRB (Barcelona) and the Swedish SME Synectiv-eLabs. The main goals for the project were

- to develop new hierarchical parallelization approaches explicitly based on ensemble and high-throughput computing for new multi-core and streaming/GPU architectures,
- to establish open software standards for data storage and exchange, and
- to implement, document, and maintain such techniques in pilot European open-source codes (such as the widely-used Gromacs & Dalton codes, as well as a new application for discrete Molecular Dynamics simulations, Discrete).

ScalaLife also aimed to create a competence centre that would provide a training and support infrastructure for computational life science researchers, and establish a long-term framework for the maintenance and optimization of life science codes.

Boosting the performance and scalability of life science applications

The ScalaLife project has worked closely with the main developers of three pilot applications – Dalton (<http://www.daltonprogram.org>), Gromacs (<http://www.gromacs.org>) and Discrete (<http://www.scalalife.eu/discrete>).

Dalton is a powerful package for the study of electronic structure – development on this code started almost 30 years ago. The collaboration between Dalton developers and ScalaLife has

brought many improvements to the code including the following:

- multi-level parallelization through Echidna - a hybrid OpenMP+MPI implementation of Coulomb and Exchange Integrals,
- GPU accelerated kernels for hybrid-QM/MM (Quantum mechanics/Molecular Mechanics),
- I/O free integrator for DFT, including elimination of redundant I/O operations on slave nodes, and
- restructuring of the formation loop for response vectors.

The capabilities of Gromacs, a highly-efficient and fast code for classical molecular dynamics simulations, have been extended by the work with ScalaLife as follows:

- hybrid-OpenMP/MPI multi-level parallelization,
- new accelerated SIMD(for CPUs) and CUDA (for GPUs) kernels for non-bonded interactions – fastest GPU code for MD,
- much improved load-balancing, and
- accelerations for other architectures (namely Intel MIC, Sparc (K-computer), and BlueGene/Q).

Discrete is a newly-developed code that implements the Discrete Molecular Dynamics (DMD) simulation method. During the collaboration with ScalaLife, it has become a mature code with an increase of over an order of magnitude in the speed of the sequential simulations. Other developments from the collaboration work include the following:

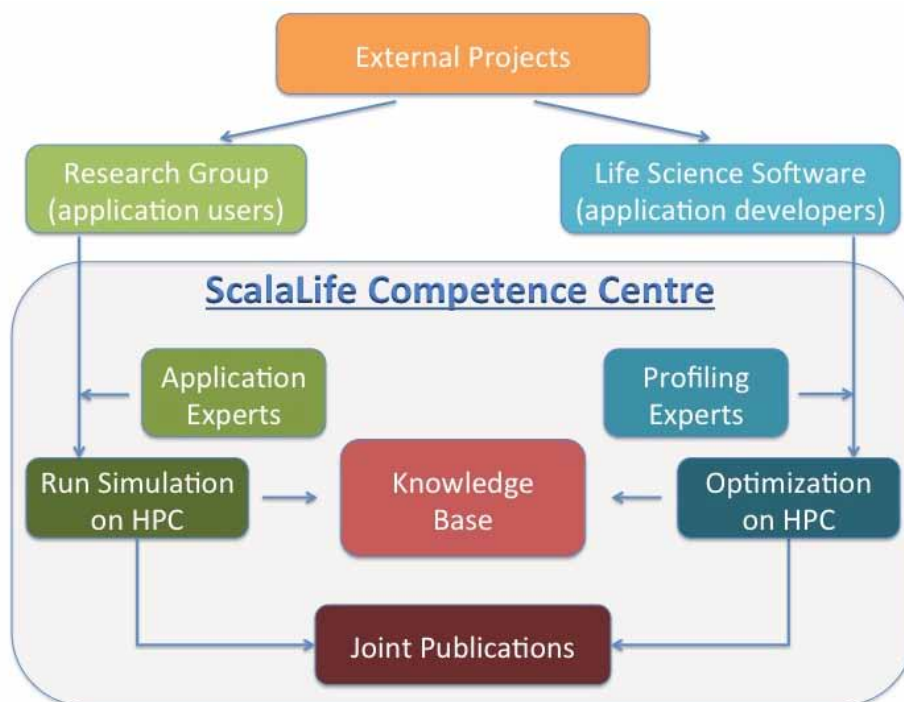
- parallelization using OmpSs (<http://pm.bsc.es/omps>),
- simulations of crowded systems, and
- re-mapping of the structures during setup from an atomistic description to a multi-scale one.

All the latest releases and patches of these codes can be found at <http://scalalife.eu/applications>.

Ensemble computing techniques

Ensemble computing is becoming more and more important when using molecular simula-

Below: Overview of the ScalaLife Competence Center



tions in life sciences. Such computations have traditionally been managed by hand (or using shell scripts), but this is a tedious and error-prone process.

Within ScalaLife, two different frameworks have been developed and adapted to simplify the process of running such simulations. The COMPSs framework, developed at BSC, is easy to set up and it takes little effort to integrate standard Gromacs simulation setups. COMPSs makes it easy to automate typical, standard biomolecular simulations, where multiple simulations are often required to achieve satisfactory sampling.

For more complex tasks, we developed the Copernicus distributed computing framework. Although this is a general framework, it has been tailored to the needs of some of the more advanced types of calculations that can be performed with the Gromacs package. Copernicus makes it possible to use ensemble simulation methods by integrating and combining current tools into a powerful workflow engine. It ensures high fault tolerance and recovery. Some ensemble simulation techniques, such as Markov state modelling, make use of adaptive sampling where one analyzes simulation data and decides if it is

required to start new simulations or extend current ones. In such cases, Copernicus is appropriate for automating the necessary workflows.

More information about different parallelization and ensemble computing techniques can be found at <http://scalalife.eu/content/parallelization-strategies>.

File format standards and APIs

In addition to code development work, the ScalaLife project designed new standard proposals for

- the storage of simulation data – e.g. TNG file format,
- data management systems – e.g. XML/UMM-MoDEL,
- automatic end-user interfaces – e.g. MDWeb, NAFLEX, and
- application programming interfaces - e.g. TNG and MOBY/R.

The API libraries, specifications and descriptions of the services are accessible through the ScalaLife Competence Center at <http://scalalife.eu/content/data>. These standards are intended to simplify collaboration and data sharing throughout the life science communities.

Competence Center for the Computational Life Science Communities

The ScalaLife Competence Center pilot was launched at the end of the first year of the project, and development continued throughout the remainder of the project. Although the ScalaLife project itself ended in August 2013, the fully-established competence centre still continues its operations, providing expertise in High-Performance Computing.

During the three years of the ScalaLife project, the Competence Center engaged in many collaborative projects with

- EU-funded projects (namely WeNMR, MAPPER, MMM, PRACE, EGI, and ERINA+),
- research communities (such as INCF, INSTRUCT, and HealthGrid),
- application developers (for codes such as MUSIC, ERGO, XMIPP, and SiMoNa),
- HPC centres (including NCSA/Bulgaria, EPCC/Scotland, CSC/Finland, and CINECA/Italy),
- local and international initiatives (such as, SeRC, SyBIT, RIKEN, and CTCC), and

- hardware vendors (for instance, NVIDIA, CRAY, Fujitsu, and Convey).

These projects have been very successful and have produced several high-impact publications, as well as establishing new collaborative initiatives.

Through training and the provision of support infrastructures, the ScalaLife Competence Center has established a long-term framework for the maintenance and optimization of life science codes. This is very important for the life science software ecosystem in Europe. The community will benefit greatly from efficient exchange of the knowledge acquired from the work in different projects (whether they are funded by the EU or other bodies), and especially from connecting software developers with experts in HPC software engineering, code optimization and method/algorithm development. It is also important to point out that improving the software efficiency directly translates to monetary savings for resource providers such as HPC centres. Keep in touch with us at <http://www.scalalife.eu>.



Above: The ScalaLife Competence Center remains available to help researchers with Life Sciences simulations.



PRACE-2IP Wraps Up

by Michaela Barth, PDC

The end of August 2013 marks a cornerstone date within the PRACE projects: PRACE-2IP, the second project phase of PRACE, was finished for all the work packages, apart from WP5 (Best Practices for HPC System Commissioning), WP8 (Community Code Scaling), WP11 (Prototyping) and WP12 (Novel Programming Techniques).

PDC will continue its efforts in the Best Practices for HPC System Commissioning work package and the Prototyping work package during the extension period of PRACE-2IP. For the Novel Programming Techniques work package, PDC's contributions will end so that just our partners at NSC will continue to be active in this work package.

At the same time, the overall PRACE project participation for the smaller SNIC centres (UPPMAX, C3SE and LUNARC) has come to an end, since they will not be directly involved in the third phase of the project.

Judging by the number of person month budgeted in the various project phases, the majority of the work within the different PRACE implementation phases had already been done during the part of PRACE-2IP before the start of the extension phase.

Sweden has its cake and eats it too!

When it comes to computer access awarded within PRACE-2IP, there were eight successful Tier-0 applications during Calls 1-7 that were at least partly Swedish, of which four had a Swedish principle investigator. (See <http://www.prace-ri.eu/statistics> for a comparison with other countries.) When looking at the Tier-1 applications, the numbers look even better: for the DECI7-DECI10 calls, Swedish projects made up 17% of all the DECI projects! Sweden was also assigned 22-50% more project time than our nominal share - the exact figure depends on whether one includes the 20% reserved for external projects. Practically speaking, Swedish projects were awarded 75 million DECI standard CPU hours (rather than 50 million), corresponding to 60 (rather than 40) million Cray XE6 CPU hours. And these figures are not even taking into account the projects that we moved away from Sweden so as to increase their chances of being approved.

In terms of the actual amount of computing time that was used, we can say that Swedish researchers used 117% of their time within the DECI7-DECI8 calls in comparison to the time that other European researchers really used on the Swedish PRACE Tier-1 resource, Lindgren.

During PRACE-2IP, quite a few projects received valuable help from application experts (including six Swedish projects within DECI) and a number of whitepapers were produced and reviewed as well. To follow-up on these successes, known Swedish "power users" have been approached directly and offered help to support them with their potential upcoming PRACE applications. On top of the continued efforts now going on in PRACE-3IP, SNIC has a further budget for PRACE: 6.5 million SEK has been ear-marked by SNIC which will be used for application enhancement in 2014.

Next page

Top: Students enjoying lunch and fine summer weather at the PDC Summer School Picnic 2013.

Lower left: Summer school students are introduced to Lindgren in the PDC Computer Hall

Lower right: During the visit to the PDC Computer Hall, the students were awed by the provisions in case of an outbreak of fire or a loss of power to the computer systems.



PDC Summer School 2013

by Stefano Markidis, HPCViz

This year's PDC Summer School "Introduction to High-Performance Computing" was held in glorious summer sunshine at the KTH main campus while Stockholm enjoyed one of the warmest and sunniest Augusts in many years. In the photos, you can see some of the 63 students who attended the summer school (and who were selected from nearly 100 applicants) enjoying the welcome picnic on the first day of the course, and visiting the PDC computer hall later on in the first week of the school.

The summer school (which is held annually) was organized by PDC, in collaboration with the KTH School of Computer Science and Communication (CSC), and is also endorsed by the KTH Computational Science and Engineering School (KCSE) with partial support being provided by Google.

The summer school students this year came from a wide range of different scientific backgrounds: some are studying computer science or computational biology, while others are focusing on mechanical engineering, material science, astrophysics, or numerical analysis. Most of the students are based at KTH or Stockholm University, although we were very pleased that other Swedish universities (such as Uppsala University, Chalmers and Linköping University) were also represented.

As usual, the summer school lasted two weeks with a series of lectures (many given by international speakers) in the morning, followed by lab work in the afternoons. The lectures covered the basics for High-Performance Computing (HPC) such as traditional programming techniques for parallel computers. Another important issue that was addressed was that of optimizing scientific program codes to achieve better performance. Moreover, students also learned about the architecture of Graphics Processing Units (GPU) when they are used for HPC. Finally, the novel programming techniques that are emerging were discussed, along with looking at how Google deals with vast amounts of data.

The labs gave students experience in parallel computing, optimization techniques and GPU programming as they worked on Lindgren, PDC's Cray supercomputer, and Zorn, the GPU cluster at PDC. During the lab sessions, experts from the PDC support group helped students to completing the given assignments and start work on their projects (which continue after the two weeks of the summer school finish). After two busy weeks, the school finally rounded off with a farewell dinner giving internationally-based students a chance to taste traditional Swedish crayfish. According to the feedback from the students, the school was very successful: 77% of the students rated this year's summer school as "very good", and the remainder rated it as "good".



Staff Focus



Åke Edlund

Åke Edlund joined PDC in April 2004 as a researcher and project leader – at that time he was appointed as the leader of the security activities within the Enabling Grids for E-Science project (EGEE). More projects followed (such as BalticGrid, NEON, and Venus-C) and lately Åke has been involved with the SNIC Cloud and SNIC Galaxy projects. He has also been a member of the EU Cloud Expert Group since 2010.

Prior to his appointment at PDC, Åke worked as Chief Architect for IS/IT Global Customer Services at Sony Ericsson Mobile Communications. Åke has been – and still is – engaged in a number of start-ups in the IT industry, for example, as the Solution Manager at Alzato (developing a novel real-time database, which is now part of MySQL) – an Ericsson Business Innovation venture, as the Product Manager for Cult3D (software for interactive 3D on the Internet) at Cycore, and as an advisor for start-ups like Severalnines.

Åke has a research background in applied mathematics, high-performance distributed computing and quantum chemistry and holds degrees from the Technion - Israel Institute of Technol-

...continuing on page 19

SNIC Galaxy Update

by Åke Edlund, PDC

Galaxy (<http://galaxyproject.org>) is a widely-used scientific workflow, data integration, and data and analysis persistence and publishing platform that makes computational biology accessible to research scientists who do not have much experience with computer programming.

The SNIC Galaxy project aims to add Galaxy onto the SNIC Cloud Infrastructure. The project, which started earlier this year at the end of March, has been moving on rapidly. PDC-HPC, together with UPPMAX, has put together an environment suitable for users from (mainly) the bioinformatics community. Although the initial plan was to develop the SNIC Galaxy for genomics research, as the project has progressed, Galaxy has turned out to be largely domain-independent. Since Galaxy can be used as a general bioinformatics workflow management system, we are also investigating other areas of use.

One of the main tasks for the project is to make the SNIC Galaxy an elastic resource, where the user will get the benefits of easy-to-use and automatic scaling properties. To make Galaxy suitably elastic, our developers have created Galaxy CloudMan. CloudMan had been used at PDC earlier, but only for Amazon cloud settings. Now we want the same functionality for our own cloud environments. To make it possible to add CloudMan to our PDC cloud (and the UPPMAX cloud), we are moving towards OpenStack (CloudMan has been ported to OpenStack), with experiments starting in September. In addition to user pilot studies on the new environments, we are preparing courses (including web-based courses) for our users, with a larger workshop planned for spring 2014 in Uppsala.



Above: Early version of the starting screen for SNIC Galaxy users

Inaugural NeIC Conference

by Michaela Barth, PDC

Around 160 scientists and technical experts from the Nordic countries and also from further afield attended the Nordic e-Infrastructure Conference 2013 held in Trondheim, Norway, on the 15th and 16th of May 2013, with additional workshops being held earlier on the 13th and 14th of May.

The “Nordic e-Infrastructure Collaboration: Lessons, Opportunities and Future Directions” conference was hosted by the Nordic e-Infrastructure Collaboration (NeIC), UNINETT Sigma AS and the Norwegian University of Science and Technology (NTNU) in collaboration with NORDUnet. The aim of the conference was to couple Nordic user communities with the skill base that exists within the national e-infrastructure organizations in order to develop ideas, and to elicit opportunities for cost-efficient common solutions and joint e-infrastructure services.

NeIC2013 was the first conference of its kind, and the event itself was an official starting point for ten further years of Nordic collaboration in the field of e-infrastructures and the handling of scientific data. The Nordic countries have all agreed on a Memorandum of Understanding for the years 2013–2023. This is the first time that the Nordic countries have made such a long-term commitment within the field of e-infrastructure cooperation.

The conference highlighted the fact that the Nordic countries are a natural playground for a regional collaboration, as the countries have similar goals and sufficiently complementary strengths for synergy, along with strong political support for such collaboration. For a Nordic collaboration to be successful, it must be flexible and be based on the fact that people really are the key part of any e-infrastructure. Ultimately, users are the driving force for any such endeavour, and any e-infrastructure solutions must serve their interests. So NeIC's goal is not to establish an e-infrastructure for its own sake, but to enhance the e-science resources available to all the researchers at Nordic institutions. The users can then choose the services they want, effectively forming and designing e-infrastructures themselves.

The next NeIC Conference is planned for 2015. Meanwhile, for more information on NeIC, see <http://neic.nordforsk.org> or contact Michaela Barth (caela@kth.se), the NeIC Coordinator for Generic Technologies.

Top right: Ebba Þóra Hvannberg, NHPC (Nordic High Performance Computing Project) Board chair and NeIC Board member

Middle right: Tom Langborg (NSC, Linköping), Mats Nylén (HPC2N, Umeå), and Thomas Svedberg (C3SE, Chalmers) enjoy mingling at NeIC 2013.

Bottom right: Pentti Pulkkinen, NeIC Board chair, during the opening session.

Staff Focus

...continuing from page 18...

ogy (Ph.D.) and Uppsala University (Tekn.Lic., M.Sc.). His “research stations” were at the Technion (Israel), Uppsala University (Sweden), the University of California at Berkeley (USA), and Rice University (USA). His main contribution was a generic parallel multidimensional solver for quantum physics applications. These days Åke is leading the HPCViz Data-Intensive Computing Group and the PDC-HPC Cloud Group.

NeIC Conference 2013



PDC-Related Events

ExaMPI13 - Workshop on Exascale MPI at SC13

22 November 2013, SuperComputing 2013, Denver, USA

<http://www.pdc.kth.se/exampi13>

PDC is organizing a workshop on the future of the Message Passing Interface (MPI) at the SuperComputing 2013 Conference (SC13) in Denver, USA, on the 22nd of November – see <http://sc13.supercomputing.org/>.

The goal of the workshop is to bring together developers and researchers to present and discuss innovative algorithms and concepts in Message Passing programming models, in particular related to MPI.

The workshop will feature several keynote speakers including Torsten Hoefler (from ETH, Zurich) talking about the future of MPI, and Pavan Balaji (from the Argonne National Laboratory, USA) presenting MPI from/to GPU and Intel MIC memories.

For further information about the workshop, see its webpage (details above), or contact Stefano Markidis (markidis@kth.se).

EASC2014: Solving Software Challenges for Exascale

2-4 April 2014, Stockholm, Sweden

<http://www.pdc.kth.se/easc2014>

EASC2014 aims to bring together all the developers and researchers involved in solving the software challenges of the exascale era. The conference focuses on issues of applications for exascale and the associated tools, software programming models and libraries. The invited speakers include

- Horst Simon, Lawrence Berkeley National Laboratory, USA,
- Misuhisa Sato, University of Tsukuba, Japan,
- Yutong Lu, National University of Defense Technology, China, and
- Erik Lindahl, KTH Royal Institute of Technology, Sweden.

EASC2014 is being organized at KTH in association with SeRC and the CRESTA and EPiGRAM projects.

The Call for papers for EASC2014 is now open, with a final submission deadline of 12 January 2014. For additional information, please contact Stefano Markidis (markidis@kth.se).

HPC Sources

We recommend the following sources for other interesting HPC opportunities and events.

CERN

<http://cerncourier.com/cws/events>

<http://ods.cern.ch/collection/Forthcoming%20Events?ln=en>

EGI

<http://www.egi.eu/about/events>

HPC University

<http://www.hpcuniv.org/events/current>

HPCwire

<http://www.hpcwire.com/events>

Linux Journal

<http://www.linuxjournal.com/events>

Netlib

<http://www.netlib.org/confdb>

PRACE

<http://www.prace-ri.eu/HPC-access>

<http://www.prace-ri.eu/PRACE-Training-Events>

<http://www.prace-ri.eu/news>

SNIC

<http://www.snic.vr.se/news-events>

XSEDE

<https://www.xsede.org/conferences-and-events>