High-end computing: modelling the world

In this issue...

2 ADMIRE: making data-mining easier
3 FPGA High Performance Computing Alliance
4 PRACE: towards a pan-European HPC service
5 EUFORIA: supporting Europe’s energy research
6 Simulating molecular dynamics
7 Parallel statistics with R
8 FilmGrid: Digital asset management
10 HECToR: the official launch
11 NEMO: improving ocean modelling
12 NextGRID: making the Grid fit for business
13 EGEE-II: Europe’s biggest Grid
14 DEISA2: a world-class Grid for Europe
15 OMII Europe: Enabling global research
16 DEISA Symposium: ‘Advancing Extreme Computing in Europe’
16 HPC-Europa++: the visitor programme continues
The first few months of 2008 see the start, and end, of a number of large European projects at EPCC. While major collaborations such as NextGRID, EGEE and OMII-EU are winding down, EPCC will continue to build on their successes, keeping a presence on the European stage with leading roles in DEISA2, PRACE and ADMIRE.

This issue of our newsletter will try to give you some insight into the work we deliver within the frame of these and other, smaller, projects.

Of course we can’t forget to mention an early highlight of the year: the launch of HECToR, the UK’s new national supercomputing facility, in January. This event marked the beginning of an exciting time for EPCC, as we will host the national HPC service for the coming six years.

Our involvement in HECToR and the major new European projects will give us the opportunity to continue showcasing our work at the forefront of European high-end computing.

ADMIRE: making data-mining easier

Radek Ostrowski

ADMIRE (Advanced Data Mining and Integration Research for Europe) aims to deliver a consistent and easy-to-use technology for extracting information and knowledge. The project is motivated by the difficulty of extracting meaningful information by data mining combinations of data from multiple heterogeneous and distributed resources. It will also provide an abstract view of data mining and integration, which will give users and developers the power to cope with complexity and heterogeneity of services, data and processes.

ADMIRE will take a radical approach to data mining and data integration, building on and integrating four key ingredients: data mining, data integration, semantic Grid technologies, and an advanced, industry-produced Infrastructure Service Bus (ISB). This integration allows the production of consistent and easy-to-use technology for extracting information and knowledge from distributed and heterogeneous sources of data.

As part of the project, two detailed data-intensive scenarios will be explored to test the technology: flood modelling and simulation, and customer relationship management.

The ADMIRE infrastructure will enable a set of gateways to be connected together over the Internet and Grid. The gateways will communicate with one another using ADMIRE-developed standard representations over an infrastructure service bus. Each gateway will provide a core set of data mining and integration services which can be driven using a high-level language as shown below.

ADMIRE’s high-level architecture

ADMIRE will develop a high-level language and tools that allow dynamic mapping of abstract, data-intensive processes onto a service-oriented architecture, thereby integrating the underlying components.

The proposed approach to integration will have a wide impact on enterprise-scale systems and will provide much commercial advantage as it allows a flexible and standardised way of process development and maintenance by service component composition. This will result in increased reliability of service-based process enactment and vastly improve the results from data-handling processes developed to run across distributed, dynamically assembled, evolving data services.

ADMIRE started in March and will run for three years. The University of Edinburgh is its main co-ordinator and the project has five international partners from both academia and industry. It is funded by the EU Seventh Framework Programme.

www.admire-project.eu
The FPGA High Performance Computing Alliance (FHPCA) brings together the key players in high-performance computing and field programmable gate arrays (FPGAs).

Led by EPCC, the FHPCA is developing high-performance computing solutions, using FPGAs to deliver new levels of performance into the technical computing market.

Hardware and software developed by the FHPCA have been used to build a large-scale demonstrator supercomputer called Maxwell. This is complemented by a campaign to raise industrial awareness and interest and to stimulate the market for the Alliance members’ commercial offerings.

Maxwell uses FPGAs and requires much less space and cooling than a conventional microprocessor system. It is also over 100 times more energy-efficient and up to 300 times faster.

Several Scottish companies have been using Maxwell since its launch in March last year. Impressive results have been achieved in the oil & gas, financial and medical imaging sectors.

Maxwell wins a medal at the British Computer Society IT Awards

Maureen Wilkinson

Maxwell, which was built in Scotland by the FHPCA with the support of Scottish Enterprise, was recognised at the prestigious British Computer Society IT Industry Awards in London.

Nominated in two categories, the FHPCA came home with a runner-up medal for the much coveted prize of the BT Flagship Award for Innovation. The BCS IT Industry Awards are the leading hallmark of success among practitioners in the IT industry today.

One of the first companies to use the supercomputer, Aberdeen-based Offshore Hydrocarbon Mapping plc (OHM), found that its application ran significantly faster on Maxwell. OHM is the world’s leading provider of controlled source electromagnetic imaging services to the offshore oil industry.

Dr Lucy MacGregor, Chief Scientific Officer of OHM, said: “Improving the performance of our data processing and visualisation services is key to our continued success and we are very excited about the code speed-ups we’ve achieved with Maxwell.”

The FHPCA was established in 2004 to promote the use of FPGAs as an alternative to microprocessors. With traditional microprocessor-based solutions hitting performance limits, there is a growing need for new technologies that address the need for ever greater processing capability without demanding large amounts of space and power. The Alliance is led by EPCC at the University of Edinburgh and comprises Alpha Data, Nallatech, Xilinx, Algotronix, Scottish Enterprise and the iSLI.

FHPCA Visiting Academic Programme

Rob Baxter

The Programme gives scholars the opportunity to work on their own research and to collaborate with staff at EPCC and FHPCA members. Scholars can apply for a visit of up to one month.

Visitors are given office space at EPCC and access to Maxwell and the people who built it. Funding is available to cover all travel and accommodation costs incurred during the visit.

This year so far, visitors from the University of Amsterdam, Oak Ridge National Laboratory, the National Centre for Supercomputing Applications and the University of Tennessee at Knoxville are either lined up or working on Maxwell already.

To apply, send us a CV and letter of interest (specifying the purpose and duration of visit) before 1 June 2008. Applications (in PDF or Word format) should be sent as an email attachment to: va@fhpca.org

FHPCA at MRSC’08

Rob Baxter

This year’s Multicore and Reconfigurable Supercomputing Conference will be hosted by Queen’s University, Belfast. Through EPCC, the FHPCA will run a tutorial at MRSC’08 entitled ‘HPRC Programming Methodologies’ or ‘How to make sure your first attempt at programming high-performance reconfigurable computers isn’t your last’.

This full-day, practical-based tutorial will introduce attendees to Maxwell and guide them through a software engineering methodology designed to increase the chances of the efficient and portable acceleration of HPC applications.

MRSC’08 runs from 1–3 April: www.mrsc2008.org

Maxwell wins a medal at the British Computer Society IT Awards

Maureen Wilkinson

Maxwell, which was built in Scotland by the FHPCA with the support of Scottish Enterprise, was recognised at the prestigious British Computer Society IT Industry Awards in London.

Nominated in two categories, the FHPCA came home with a runner-up medal for the much coveted prize of the BT Flagship Award for Innovation. The BCS IT Industry Awards are the leading hallmark of success among practitioners in the IT industry today.

One of the first companies to use the supercomputer, Aberdeen-based Offshore Hydrocarbon Mapping plc (OHM), found that its application ran significantly faster on Maxwell. OHM is the world’s leading provider of controlled source electromagnetic imaging services to the offshore oil industry.

Dr Lucy MacGregor, Chief Scientific Officer of OHM, said: “Improving the performance of our data processing and visualisation services is key to our continued success and we are very excited about the code speed-ups we’ve achieved with Maxwell.”

The FHPCA was established in 2004 to promote the use of FPGAs as an alternative to microprocessors. With traditional microprocessor-based solutions hitting performance limits, there is a growing need for new technologies that address the need for ever greater processing capability without demanding large amounts of space and power. The Alliance is led by EPCC at the University of Edinburgh and comprises Alpha Data, Nallatech, Xilinx, Algotronix, Scottish Enterprise and the iSLI.

FHPCA Visiting Academic Programme

Rob Baxter

The Programme gives scholars the opportunity to work on their own research and to collaborate with staff at EPCC and FHPCA members. Scholars can apply for a visit of up to one month.

Visitors are given office space at EPCC and access to Maxwell and the people who built it. Funding is available to cover all travel and accommodation costs incurred during the visit.

This year so far, visitors from the University of Amsterdam, Oak Ridge National Laboratory, the National Centre for Supercomputing Applications and the University of Tennessee at Knoxville are either lined up or working on Maxwell already.

To apply, send us a CV and letter of interest (specifying the purpose and duration of visit) before 1 June 2008. Applications (in PDF or Word format) should be sent as an email attachment to: va@fhpca.org

FHPCA at MRSC’08

Rob Baxter

This year’s Multicore and Reconfigurable Supercomputing Conference will be hosted by Queen’s University, Belfast. Through EPCC, the FHPCA will run a tutorial at MRSC’08 entitled ‘HPRC Programming Methodologies’ or ‘How to make sure your first attempt at programming high-performance reconfigurable computers isn’t your last’.

This full-day, practical-based tutorial will introduce attendees to Maxwell and guide them through a software engineering methodology designed to increase the chances of the efficient and portable acceleration of HPC applications.

MRSC’08 runs from 1–3 April: www.mrsc2008.org

Maxwell wins a medal at the British Computer Society IT Awards

Maureen Wilkinson

Maxwell, which was built in Scotland by the FHPCA with the support of Scottish Enterprise, was recognised at the prestigious British Computer Society IT Industry Awards in London.

Nominated in two categories, the FHPCA came home with a runner-up medal for the much coveted prize of the BT Flagship Award for Innovation. The BCS IT Industry Awards are the leading hallmark of success among practitioners in the IT industry today.

One of the first companies to use the supercomputer, Aberdeen-based Offshore Hydrocarbon Mapping plc (OHM), found that its application ran significantly faster on Maxwell. OHM is the world’s leading provider of controlled source electromagnetic imaging services to the offshore oil industry.

Dr Lucy MacGregor, Chief Scientific Officer of OHM, said: “Improving the performance of our data processing and visualisation services is key to our continued success and we are very excited about the code speed-ups we’ve achieved with Maxwell.”

The FHPCA was established in 2004 to promote the use of FPGAs as an alternative to microprocessors. With traditional microprocessor-based solutions hitting performance limits, there is a growing need for new technologies that address the need for ever greater processing capability without demanding large amounts of space and power. The Alliance is led by EPCC at the University of Edinburgh and comprises Alpha Data, Nallatech, Xilinx, Algotronix, Scottish Enterprise and the iSLI.

FHPCA Visiting Academic Programme

Rob Baxter

The Programme gives scholars the opportunity to work on their own research and to collaborate with staff at EPCC and FHPCA members. Scholars can apply for a visit of up to one month.

Visitors are given office space at EPCC and access to Maxwell and the people who built it. Funding is available to cover all travel and accommodation costs incurred during the visit.

This year so far, visitors from the University of Amsterdam, Oak Ridge National Laboratory, the National Centre for Supercomputing Applications and the University of Tennessee at Knoxville are either lined up or working on Maxwell already.

To apply, send us a CV and letter of interest (specifying the purpose and duration of visit) before 1 June 2008. Applications (in PDF or Word format) should be sent as an email attachment to: va@fhpca.org
PRACE, the Partnership for Advanced Computing in Europe, has been established to prepare the ground for the creation of a persistent pan-European HPC service. The PRACE consortium was formed through the signing of a memorandum of understanding in April 2007 and brings together 15 countries from across Europe: the UK, Germany, France, The Netherlands, Spain, Finland, Switzerland, Austria, Sweden, Italy, Poland, Norway, Greece, Portugal and Turkey.

One component of PRACE’s work is the leadership of a Framework 7 Integrated Infrastructure Initiative project – called the PRACE Project – which is funded by the European Commission. However, the overall ambition of PRACE extends well beyond the Project and is expected to lead to the establishment of an infrastructure of 4-5 petaflop-scale supercomputers sited at so-called Tier-0 centres. The University of Edinburgh, through EPCC, has ambitions to host one of the first petascale systems on behalf of the UK.

Before the infrastructure can be created, a lot of preparation needs to be done. This work has been ongoing for several years in Europe and has included a number of studies into the subject. The most comprehensive of these was published in 2006 by the European Strategy Forum on Research Infrastructures (ESFRI), which outlined a European Roadmap for Research Infrastructures drawing on the knowledge of a thousand experts from across all fields of scientific research. Research infrastructures are seen as one of the key pillars of the European Research Area (ERA).

The PRACE Project, which started in January 2008 and runs for two years, is designed to take forward the work of ESFRI and to put in place the necessary legal, managerial, financial and technical building blocks to enable the procurement and deployment of petaflop-scale systems by 2010. The PRACE Project partners include all the members of the PRACE consortium with the exception of Turkey. The UK is represented in the Project by EPSRC, which has subcontracted the technical work to EPCC, Daresbury Laboratory and Manchester University.

EPCC leads the work on Software for Petaflop Systems (Workpackage 6) and this nicely complements its involvement in other projects such as DEISA. We are also involved in the other technical workpackages, although to a lesser extent.

So what is the project trying to achieve over the next two years?

In order to procure petaflop systems, we need to understand what they will look like technically, based on existing and promised technologies. The Project is therefore engaged in the procurement and deployment of prototype systems to understand how the available technologies will meet user needs. User needs are a key driver of the project and a great deal of effort will be expended understanding the European HPC code base and ensuring the correct mix of technologies are considered in the procurement process. This process will be studied and set out in detail during the Project. It will rely on a number of benchmark codes that will be selected to be representative of the user code base. A key component of the Project will be the establishment of the necessary legal and financial models to ensure systems can be procured, paid for and deliver a service to widely-distributed users across Europe.

If successful, PRACE will be a leading example of how, by working together, European nations can develop and deliver bigger and better research infrastructures to benefit all the people of Europe.

Contact
Mark Parsons: m.parsons@epcc.ed.ac.uk
EUFORIA: supporting Europe’s energy research

Adrian Jackson

EPCC is at the start of a new three-year project, EUFORIA (EU fusion for ITER Applications). EUFORIA is a European funded project that aims to provide HPC and Grid resources to the European fusion research community. It will provide European fusion scientists with the simulation modelling tools and platforms that they will need to obtain experimental time on ITER.

ITER is the next generation of fusion devices designed to show that fusion can be used as a sustainable energy source for the future. ITER is an international collaboration between the European Union, Japan, the People’s Republic of China, India, the Republic of Korea, the Russian Federation and the USA. When the ITER machine is complete there will be a high level of competition for research access. Every application will have to be supported by full-scale predictive modelling of the event that is being investigated. This means that European fusion scientists have a need for high-quality simulation codes, and access to resources that can run such simulations, in order to be able to continue fusion research at a world-leading level.

Whilst we have a range of good fusion codes in Europe, most of these do not address the size or type of simulations required for a device as large as ITER, and when they do they require significant computational resources. Currently, access to computing resources for simulation and modelling relies on local or national resources, which does not support the pan-European modelling activity that will be required for ITER experiment applications and results analysis. EUFORIA will provide access to Grid and HPC resources for European fusion scientists, and software development effort to ensure that their simulation codes are able to run ITER-scale simulations. The goal is to bring together local code development and modelling activities, along with compute resources, to help refine the physics and modelling capabilities of European fusion physics.

Not only will EUFORIA focus on improving simulation and modelling of core and edge transport, and turbulence within the plasma, it also aims to provide the technology and development effort to couple modelling codes to allow more complex, and therefore more detailed, simulations to be performed. The ability to simultaneously run models, which work on different areas of the plasma physics and are coupled together (so they communicate data), enables scientists to perform much more realistic modelling, and therefore produce much better results. Ultimately EUFORIA aims to allow scientists to use both Grid and HPC resources within the same simulation, choosing the computational resources that match the computational modules required for a simulation, and orchestrating them into a single workflow.

As well as providing code development effort and access to computing resources, EUFORIA includes work on visualisation and data analysis tools to allow researchers to better interact with, and analyse, simulation results.

EPCC is involved in three areas of work in EUFORIA: optimisation and parallelisation of fusion codes, user support and training. We will be using HECToR, the new UK high-end computing resource, for code development and parallelisation. UKAEA Culham, the home of UK fusion research, is also involved in EUFORIA, providing fusion codes that will be optimised and developed during the project.

Contact
Adrian Jackson: adrianj@epcc.ed.ac.uk
The atoms of a hot bar. Initially in a regular austenite phase and clamped at both ends, it is cooled through the phase transition that sweeps across the material. The texture of martensite phase regions with different orientations (colours) and their boundaries are clearly seen.

Image courtesy of G. J. Ackland and O. Kastner.

Simulating molecular dynamics: investigating superelasticity

Kenton D’Mellow, Lorna Smith, Kevin Stratford, EPCC and Graeme Ackland, Institute for Physics

The martensite-austenite phase transformation is one of the most important in metallurgy and it exists in many materials. It has some interesting applications, for example in so-called superelastic shape memory alloys (materials that can regain their original shape after deformation). The martensite-austenite transformation occurs when moving from a high temperature, ductile (austenite) phase to a low temperature, hard (martensite) phase. In practical applications, such as shape memory alloys, both phases co-exist.

To study these materials, EPCC has teamed up with colleagues in the School of Physics to re-engineer and parallelise a simulation code – MOLDY – which can study martensites at the atomic level. The code can be used to understand the structure of the interfaces between martensite and austenite phases, and how this structure is affected by various material parameters.

Describing adequate surrounding material, to match both elasticity and atomistic detail, makes these molecular dynamics simulations very computationally intense. Using a traditional supercomputer is certainly an option, but a more cost-effective solution is becoming available through dedicated multi-core processors, the standard desktop hardware that will be prolific for the next few years.

Therefore, during the course of this project we will focus on exploiting shared-memory multi-core hardware in the parallelisation of MOLDY, and evaluate how multi-core systems can best be used in mid- to large-scale simulations in general.

Movies of the transition process can be seen at:
www.ph.ed.ac.uk/cmatter/gja/Mart

Spectacle frames are commonly made of shape memory alloys as they can withstand severe deformation and revert to their original shape.
Parallel statistics with R

Jon Hill

Statistical analyses as carried out in bioinformatics often utilise large data sets. For example, in genomic analysis it is possible to collect data of the order of 100000 x 100000 data points (sequences x samples). Such large datasets run the risk of outgrowing the memory of the systems used to perform the analysis. In addition they can also be complex to process, leading to excessive compute times. For example, small datasets require 8Mb of memory, but the analysis results in output of 800Mb. The largest datasets may require in excess of 10Gb to store in memory for an analysis of a small subset of data. A complete analysis requires terabytes of storage and the size of data sets will continue to increase.

To carry out the statistical analysis of genetic data, researchers at the Division of Pathway Medicine (DPM) in the University of Edinburgh use an open-source package called R; a statistical application that is commonly used by the bioinformatics community. While parallelisation techniques are available within R, they are generally focused on task-farm approaches. Unfortunately many statistical analyses, for example correlation, cannot be solved in parallel using such an approach.

The aim of this short project, funded by edikt2, is to design, build and test a prototype system that will allow statistical analyses to be done in parallel on standard HPC machines. This prototype will allow the analysis to be carried out with only minimal editing of existing R scripts and minimal knowledge of HPC systems. It is hoped the prototype will lead to the development of other techniques that will be added to R at a later date, as a preliminary grant application to the Wellcome Trust has been accepted for a full proposal to carry out this extension.

www.edikt.org/edikt2/ParallelRActivity
Film post-production requires collaboration between geographically distributed companies and individuals. The distribution and management of digital assets – the film footage and related data – in current post-production workflows is largely manual. These labour-intensive processes make it very difficult to maintain an up-to-date picture of the status of a production, leading to inefficiencies and complications.

Typical projects in the film industry involve the transportation via courier of terabytes of data on hard disk between collaborators. Security can be difficult to maintain and relies principally upon trusted couriers and collaborators. However, as faster networks become available and economically viable, network transfer becomes more attractive than courier. Even if the ‘effective bandwidth’ of a courier exceeds the network bandwidth between two sites, digital transmission can cut the lead time to near zero, so the receiver can start to work on the data almost immediately after the transfer is initiated. Security of assets transmitted digitally can be ensured through the use of strong encryption.

In all-digital production processes, such as those found in a computer-generated animation or within the Digital Intermediates (DI) phase of a filmed production, the data is stored in files. However, despite the move towards DI, tracking digital assets still proves very difficult. This can cause serious headaches for both technical and managerial staff, whose work depends upon knowledge of both the status and the location of footage.

An opportunity therefore arises to create a computerised system that can manage access to and transportation of these assets and so resolve these headaches.

Joe Dunton and Co. (JDC) and EPCC are producing a Grid-based software solution – FilmGrid – designed to address these issues by supporting effective management of the flow and availability of data.

**Significant features of FilmGrid include:**

- **Efficiency:** assets can be transferred quickly and easily, and so less time is wasted working on old versions of footage and ascertaining the status of shots.

- **Visibility:** management can see at a glance the current status of the production. Work which is running late can be quickly flagged up.
• Accountability: both scheduled work and completed work are clearly logged.

• Accessibility: assets can be accessed from anywhere, provided the user has access to the client.

FilmGrid is based on a distributed architecture where each site in the production hosts a server, which holds the assets required by that site. The FilmGrid client provides users with a view of the data across all sites in the production. The client software consists of the following tools:

• The Asset Manager, a data transfer tool for moving assets securely, reliably and efficiently.

• The Storyboard Manager, which provides a high-level visual overview of the production.

• The Scene Manager, a data management system designed to reduce the scope for confusion in post-production workflows. This provides detailed information on the scenes and shots in a film, including what work has been done to them and who has worked on them.

A usable and proof-of-concept implementation of FilmGrid was completed in February 2008. The technical aspects of FilmGrid were successfully disseminated in this year’s joint conference held by the Society of Motion Picture and Television Engineers (SMPTE) and the Video Services Forum (VSF) in Texas, USA. The theme was ‘Distributing Digital Content in a Networked World’. Many participants indicated their interest in FilmGrid and the application of Grid computing to this area. The conference also confirmed the importance of digital asset management to the film industry in general.

We are currently developing an exploitation strategy for FilmGrid. JDC plans to use FilmGrid as a marketing and dissemination tool to attract investment, develop it into a complete product and then sell the product and services based on it to post-production companies worldwide.

Look out for FilmGrid at the BEinGRID Industry Day, co-located with OGF-23 and OGF-Europe: June 3-5, 2008, in Barcelona.

FilmGrid is funded by the EC as part of the BEinGRID project, which aims to showcase the benefits of integrating established Grid technologies into small and medium-sized enterprises.
HECToR: the official launch
January 14th, 2008, Edinburgh

HECToR – the largest and most advanced supercomputing facility in the UK – was officially launched by The Rt. Hon. Alistair Darling MP, The Chancellor of The Exchequer on 14th January 2008. The launch was jointly hosted by the EPSRC and The University of Edinburgh and was attended by a hundred invited guests.

HECToR (High-End Computing Terascale Resources), will be used by researchers at the cutting edge of their fields. Four times faster than its predecessor and capable of 63 million million calculations a second, HECToR will facilitate innovative and world-leading research and represents the equivalent of approximately 12,000 desktop systems.

To get a sense of the capability of HECToR, imagine if it was possible for every person on Earth to carry out 10,000 thousand calculations, in a second, at the same time - this is how powerful HECToR is. As one of the largest and most advanced supercomputers in Europe, this UK Research Councils facility will play a key role in keeping scientists at the forefront of their fields of research.

Traditionally, scientific research has been based on the process of theory and experiment. Adding the power of supercomputing simulation to the route of exploration and discovery has moved science to another level. HECToR continues this process by taking high-performance computing up yet another gear.

HECToR will provide UK researchers with the means to undertake increasingly complex computer simulations across a range of scientific disciplines. This will include work in forecasting the impact of climate change, fluctuations in ocean currents, projecting the spread of epidemics, designing new materials and developing new medicinal drugs.

Based at the University of Edinburgh’s Advanced Computing Facility (ACF), the £113m service will run for six years and be operated by EPCC. The computer is a Cray XT4 system and support for applications software is provided by NAG Ltd. The procurement project for HECToR was managed by the Engineering and Physical Sciences Research Council (EPSRC) on behalf of the UK Research Councils.

www.hector.ac.uk
Over the past few decades, climate change has become an area of intense topical interest. The Earth’s oceans play a major role in understanding the current climate and the possible implications for the future.

EPCC is currently collaborating with researchers from the Ocean Modelling and Forecasting (OMF) group based at the National Oceanography Centre, Southampton (NOCS). The researchers at NOCS are part of the much larger Oceans 2025 project [1] funded by the Natural Environment Research Council (NERC). Oceans 2025 seeks to increase knowledge of the marine environment so that we are all better able to protect it for future generations. The project, a collaboration between seven UK marine research centres, will run from 2007-2012 with a budget totalling £120 million. Specifically, the NOCS group will work on Theme 9: Next Generation Ocean Prediction Systems, which aims to deliver the state-of-the-art ocean models needed for the next decade of UK marine science.

The French/European NEMO ocean model [2] will be used for the majority of the ocean modelling carried out by NOCS. The NEMO (Nucleus for European Modelling of the Ocean) model allows ocean-related components (eg sea-ice, biochemistry, ocean dynamics, tracers etc.) to work either together or separately. Unlike previous ocean models used by UK researchers (eg OCCAM), the NEMO model has not been specifically optimised for use on UK supercomputers. This is where EPCC fits in. Under a Distributed Computational Science and Engineering grant from NAG Ltd, we will work with the NOCS researchers to improve the performance of the NEMO ocean model on HECToR [3]. The main areas which we will concentrate on are:

- **I/O performance on intermediate and large numbers of processors**
  
  The manner in which data is input/output to the NEMO ocean model is not ideal for large numbers of processors. The OCEANS 2025 project involves the use of increasingly complex models at high spatial resolutions. With new machines, such as HECToR, the required throughput will only be achieved by increasing the processor count. We will work together with the researchers to investigate and subsequently optimise the I/O performance and to reduce the volume of I/O using lossless compression algorithms within the NEMO model.

- **Nested model performance**
  
  NEMO allows nested models to be used. These allow different parts of the ocean to be modelled with different resolutions but still within a global model context. This can help scientists to gain a better understanding of particular ocean features at a reasonable computational cost. Previously, setting up such nested models was extremely time-consuming and few attempts have been made to run such configurations in a high-performance computing environment. We will begin by investigating the performance of the current nested models and we will then attempt to improve the performance by optimising the model and code. A further aim of this work is to provide guidelines which make configuring and running nested regions within a global NEMO ocean model easier, such they can be more easily implemented by climate researchers.

We will be using the HECToR service (see p10) to carry out this work. The new HECToR system, A Research Councils UK High End Computing Service, is available for researchers at UK universities. The HECToR Cray XT4 system, which began user service in October 2007, comprises 1416 compute blades, each of which has 4 dual-core processor sockets amounting to a total of 11,328 cores which can be accessed independently. The service is currently the UK’s most powerful supercomputer and is ranked 17th in the world [4].

---

[3] www.hector.ac.uk
The NextGRID Integrated Project came to a close at the end of March 2008. NextGRID set out in September 2004 as a consortium of 22 partners led by EPCC and representing the elite of Europe’s Grid research community. Its goal was to define an architecture for Next Generation Grids that would overcome the barriers which thus far have prevented the widespread take-up of Grid technology in business. One of the major outputs of the project is a set of profiles and schemas that will allow the implementation of Next Generation Grids compliant with the principles of the NextGRID Architecture.

The NextGRID Architecture has, at its core, a set of architectural principles which were established after early analyses of business requirements. The requirements came from the analysis of diverse applications used by partners in the project, together with the views from service providers. The application areas included digital media production, financial risk analysis and product pricing, supply chain management, and Electronic Data Record processing. The service providers in the project were BT and T-Systems.

The primary architectural principles behind NextGRID are:

• Service Level Agreement Driven Dynamics: Service Level Agreements (SLAs) are critical building blocks in the NextGRID infrastructure and their dynamic behaviour is central to the operation of any Grid depending on that infrastructure.

• Dynamic Federation: The dynamic federation of resources is a key factor in establishing operational business Grids.

• Minimal Grid Infrastructure: Any Grid needs to be simple to ensure ease of maintenance and wideness of applicability. However it needs to have sufficient features to enable it to support viable business models.

Service Level Agreements (SLAs) are central to NextGRID. The use of a service in NextGRID is predicated by an SLA, dynamically created, and aimed at ensuring that the relationship between provider and consumer is well defined and understood. Follow-up capabilities allow for monitoring, violation management and audit. The SLA-based approach to all non-functional (as well as functional) aspects of NextGRID provides a uniform framework for the management and operation of all quality of service aspects of NextGRID, eg performance, security, provenance management and adherence to privacy regulations.

As a dynamic Grid infrastructure, NextGRID needs to provide extensive capabilities for service construction and composition, including traditional interface composition, various forms of workflow-enabled orchestration and support for dynamic extension of the capabilities of services.

All services operating in a NextGRID environment can expect to find, but are not required to exploit, a minimal level of capabilities either available in the environment or exhibited by peer services. These capabilities are further refined as communication protocols and languages, behavioural interfaces available on all services, support services from the environment, and a common infrastructure of underlying schemas.

From these architectural principles, the consortium has developed a component model comprising units which carry out the essential functions of NextGRID. The interactions between these units are expressed as the set of profiles and schemas mentioned earlier. The consortium has termed these ‘Generalised Specifications’ and they form the true definition of NextGRID. The Generalised Specifications have been published on the NextGRID website. The partners have also written a ‘cookbook’, which is a guide for developers showing how to combine the components of NextGRID to achieve useful functionality for Grid-based business.

The consortium is also working towards implementing parts of the functionality of NextGRID in some commonly used middleware releases.

www.nextgrid.org

NextGRID is supported by the EC’s Sixth Framework Programme as part of the Information Society Technologies thematic priority.

This article expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained in this article.
In September 2004, six months after the project had started, EPCC was invited to lead and develop software for the “Network Services Development” Activity of EGEE [1]. Tenacity and focus on delivery, together with an excellent team of collaborators from all over Europe, allowed us to produce on average two deliverables each quarter. Our work on Bandwidth Allocation and Reservation (BAR) led to an architecture and a prototype demonstrating the world’s first middleware component reserved a network service available across different domains [2]. The Network Performance Monitoring (NPM) team designed and provided implementations of an architecture that allows uniform access to the network performance information available from a heterogeneous set of monitoring frameworks [3].

In April 2006 when EGEE-II commenced, BAR gracefully gave way to financial pressures and the project’s focus on core activities. The NPM team was limited to EPCC only [4] and relocated to the Operations Activity, with more emphasis on the operational aspects.

Naturally building on the achievements of EGEE, all services have been overhauled and all documentation has been updated in the duration of EGEE-II, aiming to achieve production quality. The legacy of NPM is a set of software that now fulfils two objectives [5]:

- Our e2emonit NPM framework and the PCP generic scheduler are ready to be deployed across a large infrastructure to collect end-to-end monitoring data.
- The NPM services are able to interoperate with standards-compliant tools to provide access to and presentation of a variety of monitoring data from different sources [6]. NPM bowed out of EGEE-3, as further financial pressures and tighter focus on the EGEE core activities ensued. The legacy of the group lives on. Our webservice makes network status information available to EGEE’s site and system administrators [7]. Our contributions to the ETICS build system [8], and especially the OGF NM-WG group [9], have made a lasting impact. All our software is available from the EGEE software repository, complete with comprehensive documentation for all types of users and appropriate licence statements.

Funded by the UK Joint Information Systems Committee (JISC) [10] (who have co-funded the EGEE NPM work over the past 3.5 years together with the EC), EPCC continues its NPM work until 31 March 2009. The focus will be twofold: to disseminate the NPM results inside the JISC community and to embed NPM into DEISA2.


The 31st of March sees the end of EGEE-II. EPCC looks back with fondness on three-and-a-half years of successful involvement in the making of Europe’s biggest Grid.
Imagine being able to predict the flow of lava from an erupting volcano, or foresee the effects of a large-scale landslide. These are some of the aims of the GEOGrid project which is building a platform for global Earth observation – but it’s also an example of an international project that faces significant barriers because of the lack of interoperability between Grids.

Many research projects have already invested heavily in building Grid platforms, using middleware such as Globus, UNICORE and gLite. However as these Grids have developed independently from each other, using different configurations of software and hardware, it is difficult for researchers to properly utilise these resources, particularly when working together in international collaborations across multiple Grids.

The two year Open Middleware Infrastructure Institute for Europe (OMII-Europe) project, funded by the European Commission, involved 16 organisations from Europe, USA and China developing software designed to break down such barriers by adopting open standards from bodies like the Open Grid Forum to achieve interoperability. Software development teams focused on key areas where interoperability is a necessity for scalable and heterogeneous networks of resources: job execution, data integration, virtual organisations and accounting. This was backed by activities to ensure the development was robust and well engineered including an evaluation infrastructure and helpdesk run by EPCC, which enabled any researcher to trial the software developed by OMII-Europe in a supported environment.

Within OMII-Europe, EPCC has built on key strengths in data technologies to lead the porting of the popular OGSA-DAI data access and integration software to the gLite and UNICORE platforms. It also worked with the developers of the Virtual Organisation Membership Service (VOMS) to use this standard way of managing access to resources within a particular consortium. It is this work that has proved very useful for the GEOGrid project, which features data services running across platforms in Japan and Taiwan.

So, by implementing open standards to enable interoperability, it is possible at the most basic level to link Grids – and research can be shared more easily across the world.

DEISA has already proved its relevance for advancing computational sciences in leading scientific and industrial disciplines within Europe. In DEISA2, the consortium will continue to operate and develop the DEISA HPC infrastructure, incorporating new technologies and user services as these are positively evaluated. Support will also be provided to users for applications enabling, which is indispensable for the effective support of world-leading computational sciences in the area of supercomputing. The service provisioning model is extended from one that supports a single project (via DECI, the DEISA Extreme Computing Initiative) to one supporting virtual European communities. DEISA2 plans to support many projects and communities over the next three years, assisting them in scaling their codes to run efficiently on the largest computers in Europe and supplying them with cycles contributed by the partner HPC centres and by the national funding agencies.

Additional collaborative activities will be carried out with new European and other world-leading initiatives. Of strategic importance is the co-operation with the PRACE initiative (see page 4), which is preparing for the installation of several leadership-class supercomputers in Europe. DEISA2 and PRACE have complementary roles and, in the spirit of ESFRI, are co-operating with the goal of establishing a persistent European HPC eco-system.

DeIsA2: a world-class Grid for Europe

Alison Kennedy

The DEISA consortium has been funded for a further three years by the EC (via the DEISA2 project) from May 2008 to continue work on the deployment and operation of a distributed European high-performance computing environment.

Since 2005, the consortium has worked to integrate the national supercomputing systems of the partner nations using selected Grid technologies, both to add value to the existing national systems and provide an extra layer of European HPC services on top of them. The importance of building a European HPC ecosystem to support computational science at all levels was recognised in the ESFRI report of 2006 and the integration of leading national supercomputing resources on a continental scale is an important first step in this process.

DEISA2 continues to support and develop a distributed HPC infrastructure by consolidating the existing DEISA infrastructure and extending its collaborative environment for capability computing and data management. Further national computing centres will be integrated as associated partners. The resulting infrastructure is unmatched worldwide in its heterogeneity and complexity, enabling the operation of a powerful supercomputing Grid built on top of national services, facilitating Europe’s ability to undertake world-leading computational science research.

Enabling global research: interoperability through standards

Neil Chue Hong

Imagine being able to predict the flow of lava from an erupting volcano, or foresee the effects of a large-scale landslide. These are some of the aims of the GEOGrid project which is building a platform for global Earth observation – but it’s also an example of an international project that faces significant barriers because of the lack of interoperability between Grids.

Many research projects have already invested heavily in building Grid platforms, using middleware such as Globus, UNICORE and gLite. However as these Grids have developed independently from each other, using different configurations of software and hardware, it is difficult for researchers to properly utilise these resources, particularly when working together in international collaborations across multiple Grids.

The two year Open Middleware Infrastructure Institute for Europe (OMII-Europe) project, funded by the European Commission, involved 16 organisations from Europe, USA and China developing software designed to break down such barriers by adopting open standards from bodies like the Open Grid Forum to achieve interoperability. Software development teams focused on key areas where interoperability is a necessity for scalable and heterogeneous networks of resources: job execution, data integration, virtual organisations and accounting. This was backed by activities to ensure the development was robust and well engineered including an evaluation infrastructure and helpdesk run by EPCC, which enabled any researcher to trial the software developed by OMII-Europe in a supported environment.

Within OMII-Europe, EPCC has built on key strengths in data technologies to lead the porting of the popular OGSA-DAI data access and integration software to the gLite and UNICORE platforms. It also worked with the developers of the Virtual Organisation Membership Service (VOMS) to use this standard way of managing access to resources within a particular consortium. It is this work that has proved very useful for the GEOGrid project, which features data services running across platforms in Japan and Taiwan.

So, by implementing open standards to enable interoperability, it is possible at the most basic level to link Grids – and research can be shared more easily across the world.

www.omii-europe.org
DEISA Symposium:
‘Advancing Extreme Computing in Europe’
28–29 April, 2008, Edinburgh

Elena Breitmoser

DEISA, the Distributed European Infrastructure for Supercomputing Applications, is a unique European project that provides a structured layer for the coordination of national European supercomputing services. Since the start of the project over four years ago, it has been successfully used by scientists from a broad range of subjects (for example cosmology, material science, plasma physics, life sciences) who can perform simulations which without DEISA would not have been feasible.

This year’s annual DEISA Symposium, the fourth in a line of successful events, will take place at Our Dynamic Earth in Edinburgh. It will provide a forum where scientists from around the world can discuss HPC e-Infrastructures, and where DEISA users can share experiences and results. The Symposium also offers a great opportunity for new users to get involved.

The target audiences are computational scientists who provide and support Grid-enabled HPC-facilities and other scientists from a large cross-section of backgrounds, who are already using or considering the use of these facilities. But people working in a non-university setting, from industry and commerce, should also feel encouraged to attend and explore the possibilities offered by DEISA.

The DEISA Symposium will take place over two days. The first day will feature talks from key players in the field of HPC e-Infrastructures including the DEISA and PRACE (see p4) [1] projects. The second day will see prominent DEISA users reporting on their experience of the DEISA infrastructure and on their impressive results. As usual, the Symposium will start with lunch on the first day and end with lunch on the second day.

DEISA provides leading scientific users with transparent access to a Europe-wide pool of computing resources. With the UK service HPCx, EPCC is one of eleven principal partners. HECToR, the UK national supercomputing service, will become part of this HPC pool in May 2008.

EPCC also takes part in the DEISA Joint Research Activities, which are various computational science projects that explore this new infrastructure and provide useful software tools for users.

For a detailed agenda, travel information and on-line registration, see: www.deisa.eu/symposium

[1] www.prace-project.eu
Our EC-funded visitor programme HPC-Europa finished at the end of 2007. But for our visitors, hosts and applicants it is very much business as usual, as we have obtained funding for an additional year – the HPC Europ++ project extension.

The highly successful HPC-Europa visitor programme ran for four years and enabled 185 researchers to visit EPCC. We received applications from 23 of the 32 eligible countries (although 33 countries are eligible for HPC- Europa, UK-based researchers may not apply to EPCC). Our visitors came from a wide variety of research areas, including physics, chemistry, life sciences, earth sciences, engineering, maths and computer science. This diversity is clearly reflected in the fact that 35 different departments at 16 universities and research institutes around the UK were associated with HPC-Europa as host departments for our visitors.

HPC-Europa++ will continue to the end of 2008. As with HPC-Europa, it provides funding for collaborative visits to research institutes associated with one of six European supercomputing centres: BSC (Spain), CINECA (Italy), EPCC (UK), HLRS (Germany), IDRIS (France), SARA (The Netherlands).

One additional benefit of HPC-Europa++ is that we are able to offer visitors a significantly larger allocation of computational resources than was previously available – over the year of the project, the consortium as a whole expects to fund about 140 researchers across the six supercomputing centres, and will use a total of over 2 million CPU hours.

At EPCC, we can offer HPC-Europa++ visitors access to the national high performance computing services HECToR and HPCx, in addition to our Blue Gene and Ness HPC systems. We can also arrange access to a number of additional systems for visitors with specific requirements.

We have already welcomed our first visitors under HPC-Europa ++. The next closing dates are 15th May 2008 and 31st August 2008.

For more information and to apply, see: www.hpc-europa.org