In this issue...

2  Blue Gene tackles cancer genes
3  HEC training centre
4  Condition monitoring
5  FPGAs
6  QCD support
7  Research computing with eDIKT2
8  FLASH: modelling the Universe
9  DEISA
    EPCC at SC’06
10 HPC-Europa user group meeting
11 GlobusWORLD and GGF-18
    ISC 2006
12 EPCC at ScicomP 2006
The cow on the front of this, the autumn 2006 issue of EPCC News, brought to mind my childhood in rural Shropshire, where entertainment included wayward activities such as trying to detonate cowpats with dismantled fireworks, and then attempting to out-sprint a herd of cows as a result. (Such entirely trivial experiments presaged a scientific career, which now finds me as the editor of EPCC News.)

The local farmers took a dim view (to say the least) of such antics: cattle were important to the local economy and so should not be baited by young boys.

So it is with great interest that I read of EPCC’s work on using the latest technology to provide a system for the monitoring of farm cattle (see page 4). Processing the data involved (there are some 3 million cattle in the UK) makes use of the type of expertise that EPCC has gained in more exotic areas such as cosmology (page 8) and particle physics (page 6).

Masters of our field? A bold claim, perhaps. However, I hope that reading this edition of EPCC News will give you at least a feel for both the extent of that field, and how EPCC’s technical excellence makes us leaders in it.

Blue Gene tackles cancer genes
Fiona Reid

According to figures from Cancer Research UK [1], some 35,000 people each year are diagnosed with colorectal cancer (cancer of the bowel and rectum). Bar skin cancer, this makes it one of the most common forms of cancer in the country in both men (after prostate and lung cancer) and women (after breast cancer). While the development of effective treatments is clearly important, early identification of patients at risk would be extremely useful in prevention of the cancer. To this end, the Colon Cancer Genetics Group (CCGG) at the University of Edinburgh is involved in a project to investigate the relationship between genetic markers and colorectal cancer. Information about the different genes involved and their interactions should allow a better understanding of not only bowel cancer, but also related subjects such as natural history. Ultimately, the researchers hope to be able to identify individuals at risk of the disease by examining their genetic make-up.

The scale of the project is ambitious: it aims to use the largest genotypic data set that has been compiled anywhere in the world to date (with over 560,000 genetic markers being tested). The huge volume of data would be difficult to process on a regular desktop machine. For example, it is estimated that the analysis would take over 3 years to run on a standard PC. However, using the University’s Blue Gene/L service, the same analysis could take as little as 1 week.

Using its expertise, EPCC are therefore helping the CCGG to make use of the high performance computing facility to solve the problem. The work involves two main steps. First, the effects of each genetic marker individually will be investigated. This means moving the existing analysis code to the Blue Gene/L system so that the genotype data can be processed in parallel, releasing the CCGG to focus on the results. Second, together with EPCC, CCGG will develop a new code which allows the interactions between different genetic markers to be investigated. The added complexity from the interactions means that the problem can only be tackled on a large parallel computer such as Blue Gene/L. This would be the first study of its kind and represent ground-breaking work in understanding cancer biology.

This primary research work, including patient recruitment and genetic analysis, is funded by Cancer Research UK, the Medical Research Council (MRC), CORE (the Digestive Disorders Foundation) and the Scottish Executive. The CCGG is a University of Edinburgh research group based at the MRC Human Genetics Unit at the Western General Hospital in Edinburgh.

[1] www.cancerresearchuk.org
EPCC’s HEC Training Centre – now available to more students

Adam Carter

EPCC is entering its second year as one of two EPSRC-sponsored High End Computing (HEC) Graduate Training Centres, the other being at the Centre for Scientific Computing in Warwick. This new initiative provides a means for PhD students in a variety of scientific disciplines to add a component of HEC training to their PhD research. This is made possible by extending the length of the PhD by a year to four years during which time the student will also undertake EPCC’s MSc in High Performance Computing on a part-time basis. At the end of the four years, successful students will obtain a PhD from their own institution and an MSc in High Performance Computing from the University of Edinburgh.

To make this scheme as flexible as possible, EPCC re-runs many of its Semester 1 courses in April each year to allow HEC students to study for their MSc at a time which fits in with their PhD. Most students start their PhDs in September or October. They work on their PhD full-time until April, allowing them time to become settled in the research group in which they are working, and to make a start on their PhD research. In April, they attend EPCC for what is typically a seven-week block of lectures and practical sessions on HPC. The HEC programme allows us to reimburse the students’ travel costs to and from Edinburgh, and the cost of accommodation during their visit. The HEC centre also has a set of five scholarships to cover MSc fees.

Our first four students have already completed their first block of HEC training. Two more students are about to arrive to take their first set of courses beginning in September. This demonstrates the flexibility of the HEC training centre: for students who, for example, start their PhD later in the year, it is possible to postpone the first set of HEC courses until September. All the HEC students will then join our full-time MSc students next January for their second block of courses. The HEC students will be able to work part-time on their MSc dissertations throughout their third year with visits to meet with their MSc supervisor paid for by the HEC centre, and with remote supervision provided by email and videoconferencing. The MSc will be complete by the end of year three allowing the student to devote their final year to completing their PhD research.

If you are a PI at any UK university who would like to add a component of HEC training to your PhD students’ research programme there are now two routes to achieving this...

EPSRC HEC Studentships
EPSRC are offering 4-year HEC Studentships through their normal responsive mode. When applying for such a scholarship, you are advised to first contact the HEC centre to discuss your proposal, and the suitability of the HEC component. The HEC centre will then provide you with a proforma which can be attached to your application to EPSRC to demonstrate that we support the application. The applications are then subject to EPSRC’s usual peer review process. PIs are advised to start applying now for studentships beginning in September 2007.

HEC Centre Places for DTA-Funded Students
If you already have a place for a PhD student supported from a Doctoral Training Account (DTA), it is possible to add an HEC training component. The fourth year of living expenses for the student must be found from the DTA, but the HEC centre has a number of places which come with MSc fees paid, and all reasonable travel and accommodation costs for the student’s stay in Edinburgh reimbursed. If you are interested in this route, you should contact us directly to express an interest.

For more details about the HEC Training Centre, you can visit:
www.epcc.ed.ac.uk/hec. Or contact me for more details or to apply for the HEC Training Centre: A.Carter@epcc.ed.ac.uk

Note that applications should be made by PhD supervisors or PIs and not the students themselves.
EPCC has moved into an exciting new phase of the ITI Techmedia Condition-Based Monitoring (CBM) Programme.

The CBM Programme is the sixth research and development programme for ITI Techmedia, the organisation which develops market-driven intellectual property for the benefit of Scotland. CBM is a three-year, £4.75M programme which aims to apply sensors and networks technology to develop a condition-based monitoring platform with applications across a range of industries.

The Programme is initially focused on the commercial farming sector where the technology under development will enable the behavioural and physiological monitoring of beef and dairy cattle, allowing the animals’ conditions to be tracked continuously. The initial aims are to apply sensors and wireless communications technologies to provide health and fertility status information, to improve the monitoring and prediction of calving and post-partum conditions. Seven expert organisations are already contracted to undertake research on this multi-disciplinary programme. EPCC has been working closely with groups at the University of Strathclyde, Scottish Agricultural College and Royal (Dick) School of Veterinary Studies at the University of Edinburgh to prepare the ground for the interpretation of large volumes of novel sensor data. Ongoing farm studies will soon begin to generate the first of potentially petabytes of sensor data from trial cows, and the challenge for data analysis experts at EPCC and Strathclyde is to search this morass to find links from sensor readings to cows’ conditions. Future platforms will be tightly constrained by size, cost and a host of other restrictions, but right now it’s a case of generating as many data as possible for detailed analysis.

The project expects to generate hundreds of megabytes per cow per day, and EPCC will be deploying the resources of the major computer systems and 155 terabyte Storage Area Network at the University’s Advanced Computing Facility to manage and analyse this deluge. By applying a variety of novel analysis techniques, EPCC and Strathclyde hope to identify and capture key health and welfare indicators which can be encoded into future sensor platforms, providing an automated system of condition monitoring which will equal or exceed current farm best practice.

‘The application of sensors and networks technology presents numerous opportunities for ITI Techmedia,’ commented Techmedia CEO David Creed. ‘According to our market analysis, predictive intervention in animal health is an area of global market opportunity which Scottish skills and research expertise are well placed to address.’

The next twelve months will be an interesting time for the EPCC CBM team.
Field Programmable Gate Arrays (FPGAs) are a novel way of providing increased computing capacity using limited space and power. They can be configured, at a hardware level, to perform specific numerical algorithms and, as such, can be many times faster than conventional microprocessors.

Historically, programming FPGA hardware has required a high level of electrical engineering knowledge. A software developer was likely to regard the prospect of writing the Hardware Description Language (HDL) needed for such systems with trepidation. However, this may be changing thanks to a new generation of development tools that permit an FPGA application to be described as a collection of independent and concurrent software processes, prepared in a high-level language such as ‘C’ and following a design approach analogous to multi-threaded programming.

EPCC, with the support of the eDIKT2 programme (see page 7), has tested the latest FPGA development tools to investigate the feasibility of porting scientific computing kernels to an FPGA system by a High Performance Computing (HPC) software developer rather than an electrical engineer. The project has ported the well-known HPC benchmark, STREAM, across to an FPGA system. The experiences of the project team during the exercise have been documented and an evaluation of some of the high-level tools that are becoming available has been provided. The exercise has also created a useful tool for measuring the external bandwidth available on an FPGA device.

In parallel to this project, the FPGA HPC Alliance (FHPCA) — a consortium led by EPCC and involving Xilinx, Nallatech, Alpha Data, Algotronics, and the ISLI — are building a parallel FPGA supercomputer which, when complete, will be the largest such system in the world. This is part of the FHPCA’s commitment to developing real-world HPC solutions using FPGAs.

With the flexibility and economy that FPGAs can provide, is it time to consider another hardware breed when solving problems in HPC?

Further information on the FPGAs in HPC project led by EPCC can be found at www.fhpca.org/index.html while further information on the eDIKT2 Scientific Computing on FPGAs is available from: www.edikt.org/edikt2/ScientificComputingOnFPGAsActivity

The Scientific Computing on FPGAs project at EPCC, with the support of the eDIKT2 programme (see page 7), has tested the latest FPGA development tools to investigate the feasibility of porting scientific computing kernels to an FPGA system by a High Performance Computing (HPC) software developer rather than an electrical engineer. The project has ported the well-known HPC benchmark, STREAM, across to an FPGA system. The experiences of the project team during the exercise have been documented and an evaluation of some of the high-level tools that are becoming available has been provided. The exercise has also created a useful tool for measuring the external bandwidth available on an FPGA device.

In parallel to this project, the FPGA HPC Alliance (FHPCA) — a consortium led by EPCC and involving Xilinx, Nallatech, Alpha Data, Algotronics, and the ISLI — are building a parallel FPGA supercomputer which, when complete, will be the largest such system in the world. This is part of the FHPCA’s commitment to developing real-world HPC solutions using FPGAs.

With the flexibility and economy that FPGAs can provide, is it time to consider another hardware breed when solving problems in HPC?

Further information on the FPGAs in HPC project led by EPCC can be found at www.fhpca.org/index.html while further information on the eDIKT2 Scientific Computing on FPGAs is available from: www.edikt.org/edikt2/ScientificComputingOnFPGAsActivity

The Scientific Computing on FPGAs project at EPCC, with the support of the eDIKT2 programme (see page 7), has tested the latest FPGA development tools to investigate the feasibility of porting scientific computing kernels to an FPGA system by a High Performance Computing (HPC) software developer rather than an electrical engineer. The project has ported the well-known HPC benchmark, STREAM, across to an FPGA system. The experiences of the project team during the exercise have been documented and an evaluation of some of the high-level tools that are becoming available has been provided. The exercise has also created a useful tool for measuring the external bandwidth available on an FPGA device.

In parallel to this project, the FPGA HPC Alliance (FHPCA) — a consortium led by EPCC and involving Xilinx, Nallatech, Alpha Data, Algotronics, and the ISLI — are building a parallel FPGA supercomputer which, when complete, will be the largest such system in the world. This is part of the FHPCA’s commitment to developing real-world HPC solutions using FPGAs.

With the flexibility and economy that FPGAs can provide, is it time to consider another hardware breed when solving problems in HPC?

Further information on the FPGAs in HPC project led by EPCC can be found at www.fhpca.org/index.html while further information on the eDIKT2 Scientific Computing on FPGAs is available from: www.edikt.org/edikt2/ScientificComputingOnFPGAsActivity
EPCC has a long history of working with UKQCD [1], the UK collaboration of scientists working on Lattice Quantum Chromodynamics. Indeed, quite a few EPCC staff members are ex-QCDers! Recently, EPCC has renewed its relationship with UKQCD with a new project QCD Support.

QCD is the quantum theory of the strong interaction. In principle, QCD can predict the properties of hadrons, bound states of quarks and gluons, which can then be tested in experiments, both to verify QCD and to learn more about the properties of these fundamental particles. However, it is a non-linear, and therefore unsolvable, theory. In certain regimes of parameter space analytic solutions exist, but outwith these one can only attempt an approximate, numerical solution.

The calculation involves partial derivatives. These can be substituted with finite differences when continuous space-time is replaced with a discrete grid or lattice. This gives rise to a matrix equation, the size of which is proportional to the space-time volume used. As it is four- or sometimes five-dimensional, it quickly becomes a very big matrix. The condition number of this fermion matrix, which governs the number of iterations required to calculate its inverse, depends on the quark mass. The lightest quark masses, for the up and down quarks, is near zero, so the condition number gets very large.

In Quantum Field Theories, correlation functions are determined by a path integral, or ‘sum-over-histories’. That is, the probability that a particle at one point travelling to another receives contributions from every possible path, weighted by how likely that path is. In lattice QCD, this integral is replaced by a finite sum. Configurations of quark and gluon fields are generated at random, but weighted by the probability with which they contribute to the integral. This is known as importance sampling, or Monte Carlo method. The upshot of this is that the fermion matrix, its inverse and determinant have to be calculated many hundreds of times. At this stage, it’s a good idea to use the largest computer you can find!

In EPCC News 56, Peter Boyle reported on one such machine: the massively parallel QCDOC for which he is the co-designer. This dedicated machine can sustain around 2–5 Tflops for QCD application codes. Along with theoretical and algorithmic advances, this machine has enabled UKQCD in collaboration with another group, RBC [2], to do some genuinely ground-breaking research.

UKQCD codes are necessarily complex. There are two main code suites: the CPS, shared with the RBC collaboration, and qdlp++/chroma which is an open source code, supported and developed in part by the US SciDAC programme [3]. The QCD support project contributes to the support and development of these code suites, but there is more to support than technical code writing. Many UKQCD members contribute to these codes. Often this work is done by post-graduate students or post-doctoral researchers. This work is of great value to the collaboration, but as they are often only active in code-writing or data generation for around eighteen months. There is initially a steep learning curve for the researchers, and a danger for the collaboration that their expertise will be lost once they leave.

To help with these problems, we have hosted visitors from UKQCD and organised two workshops. The first workshop was for QCDgrid system administrators, and the second a three-day ‘teach ourselves how to do lattice QCD’ for UKQCD aimed at students and post-docs, called HackLatt. This had around 30 attendees and included international and UKQCD speakers.

Besides interacting with the systems staff who host the QCDOC systems at the ACF, the support project has significant overlap with another EPCC project, QCDgrid [4]. QCDgrid is a geographical data grid with replicas of files stored at several UKQCD sites. Access to this data is via the metadata, and the support project has been involved with developing data formats and a metadata language (QCDml) across the world-wide lattice QCD community known as the international lattice data grid (ILDG [5]). The QCDgrid project has been involved with...
The eDIKT2 project, funded by Scottish Funding Council and the University of Edinburgh, aims to support the research computing community around the University of Edinburgh. It spans a diverse range of organisations within the University from the School of Chemistry to the Brain Imaging Research Centre at the Western General Hospital.

eDIKT2 (e-Science Data, Information and Knowledge Transformation), is the follow-on project from eDIKT which emphasized managing and interconnecting diverse data sources (see EPCC News issues 46, 47 and 52). eDIKT2 has a broader remit than that of its predecessor and is intended to have a widespread and long-lasting effect on research computing around the University of Edinburgh by providing a grid portal, user support, software development and computing infrastructure. The project includes NeSC, the University’s Computing Service (EUCS), the School of Chemistry, the SFC Brain Imaging Research Centre for Scotland, the School of Biological Sciences, the Scottish Centre for Genomic Technology and Informatics (GTI) and, of course, EPCC, which is taking on the role of overall project manager.

So what activities are being carried out as part of the eDIKT2 project? It would be impossible in this short article to describe all the work carried out, so instead here are some of the highlights…

• The SFC Brain Imaging Research Centre for Scotland is creating sophisticated visualisations of medical images to help treat a wide range of neurological conditions and disorders such as strokes. A single Magnetic Resonance Imaging (MRI) scan of an individual can produce over 600 images. These can be combined to show information not obvious on the original images, such as blood flow, water mobility and metabolite concentration (see above). Taking scans several days apart can show changes in the brain structure but it is impossible to get the patient to lie in exactly the same position in the scanner for each scan. The images must therefore be processed to account for the differences in position. Currently all this analysis is carried out using command line programs that require a great deal of experience to operate. eDIKT2 will allow the Brain Imaging Centre to create a user-friendly front-end to the software and to develop more computational tools to help in patient diagnosis and treatment.

• GTI is embarking on an ambitious project to help researchers worldwide to store and share data from high throughput RNAi experiments. RNAi (RNA Interference) is a cutting-edge technique to silence the effects of genes in a cell. Currently researchers store data resulting from experiments and about the experiments themselves, but there are no guideline as to what data needs to be stored, no standardised formats to store the data and no tools to help with this storage. As part of the eDIKT2 project GTI plan to develop the data formats and tools necessary for the easy sharing of data. The project will culminate in a pilot study to test the system that has been built.

• The School of Biological Sciences generates vast amounts of imagery showing how RNA moves around a cell. As with the Brain Imaging Centre a vast number of images and video can be generated. It is estimated that a typical researcher working in the centre for 3 or 4 years will generate a massive 35,000 images. The eDIKT2 project will allow the development of an integrated and secure data storage system, which will allow the data generated from their high-powered microscopes to be stored and described in a systematic and rigorous way. In addition to the experimental data, the system will also store meta-data, including the microscope settings, dates, any image analysis carried out and other experimental conditions.

• EPCC is also carrying out a number of projects within eDIKT2 in addition to overall project management. Two of the projects (scientific programming of FPGAs and Cancer Genes on Blue Gene ) are described elsewhere in this issue. Another one of these projects is the development of the eDIKT2 portal in collaboration with NeSC. The portal currently consists of a collaborative wiki environment, which allows the editing of online material directly within a web browser, and lets the eDIKT2 participants to share knowledge as well as publicise their work online. In the near future the portal also aims to allow computing jobs to be submitted to the National Grid Service. To further facilitate discussion, there are also technical workshops held every three months, where the eDIKT2 participants meet to discuss their progress and exchange information. These meetings have already resulted in a potential collaboration between cell biologists and EPCC. This is hoped to be the first of many collaborations.

There is more information on the activities, including information on forthcoming workshops, on the eDIKT2 portal: www.edikt.org/edikt2
The Virgo Consortium [1], formed in 1995, is a large international collaboration of astrophysicists dedicated to carrying out the largest simulations of the evolution of the Universe. Indeed, Virgo has performed the largest simulations of the Universe undertaken to date, and is a significant user of high performance computing (HPC).

To provide Virgo with the required processing power EPCC has been developing Virgo’s cosmological simulations and associated tools to take advantage of the DEISA infrastructure. DEISA [2] (Distributed European Infrastructure for Supercomputing Applications) is a confederation of HPC centres providing an extremely large computational resource to scientists throughout the European Union.

Virgo currently employs so-called N-body simulation codes, such as GADGET [3], to model both gas and dark matter dynamics. The hydrodynamic and gravitational interactions between these components determine the evolution of structure in the Universe. However, particle-based N-body methods can have difficulty capturing shock fronts in the astrophysical flows. One solution to this shock-capturing problem is to use a mesh-based algorithm. Recently, Virgo have adapted FLASH [4], software developed by the United States Department of Energy and the Alliance Center for Astrophysical Thermonuclear Flashes at the University of Chicago. The code was initially developed for the US government to model thermonuclear flashes, but is now widely used by different scientific communities, including cosmologists who model gas flow where shocks are important.

Using a uniform grid, or mesh, with high enough resolution to capture shocks everywhere in a large system would be prohibitively expensive and wasteful. To get around this, FLASH uses a process known as adaptive mesh refinement (AMR). A simulation may start with a uniform distribution of gas and/or dark matter, the properties of which are represented on the cells of a uniform mesh. As the structure of the gas flow evolves, AMR can place more grid elements in regions of interest, while maintaining the original (low) resolution where gas and/or dark matter is more sparse. This allows highly inhomogeneous problems to be solved effectively. In addition, to improve the efficiency of the code in parallel, Tom Theuns of the Virgo Consortium has replaced the iterative Multi-Grid algorithm used for solving the Poisson equation in FLASH by a single-step fast Fourier transform method.

Virgo is keen to run very long simulations using FLASH which would otherwise fall foul of the time restrictions on individual jobs imposed by most HPC centres. To solve this problem we have been investigating code migration within the DEISA infrastructure. A job reaching the batch clock limit on one DEISA machine would then resubmit itself, migrate if necessary, and restart on the infrastructure at the earliest possible opportunity. Allowing the code to migrate in this way would ensure the fastest time to solution.

A simulation will run on a particular platform provided that the executable is installed. As DEISA is a heterogeneous cluster of resources, employing a staged executable, which requires binary compatibility, is not possible. EPCC is therefore involved in the process of porting and optimising both the standard release of FLASH and Virgo’s adapted version to the following DEISA platforms: IBM Regatta clusters in the UK (HPCx), France (IDRIS) and Germany (RZG); an IBM Beowulf in Spain (BSC); an SGI Altix in the Netherlands (SARA); and an NEC SX-8 in Germany (HLRS).

Having already ported Virgo’s GADGET code to DEISA, we are confident that FLASH will be made available on DEISA for Virgo to use in the near future. Furthermore, employing FLASH in addition to Virgo’s N-body codes should enable a more robust validation of these huge simulations.

---

**QCD support continued**

making the QCDgrid software interoperable with other ILDG partner software through defined common interfaces. This was successfully demonstrated at this year’s lattice conference in Tucson, Arizona, where we displayed a poster which demonstrated the aggregation of the metadata catalogues of German, Japanese, UK and US lattice groups. The poster won a prize for best poster [7].

---

[1] www.virgo.dur.ac.uk
[4] flash.uchicago.edu

Masters of our field

DEISA training
Special topic: performance and portability

The next DEISA training session will be organized at Forschungszentrum Jülich GmbH, in Jülich, Germany on October 23rd–25th, 2006. Scientists from all European countries and members of industrial organizations involved in high performance computing are invited to attend.

The purpose of the training is to enable fast development of user skills and know-how needed for the efficient utilisation of the DEISA infrastructure. The first part of the training will give a global description and introduction to the usage of the DEISA infrastructure. The second part of the training will be dedicated to the topic of performance and portability. The attendance is limited to 30 participants.

*The detailed agenda and the registration form are available at www.deisa.org/training. Registration will close on October 13th, 2006.*

---

EPCC at SC’06

EPCC has a strong presence at the SuperComputing 2006 conference (SC06) taking place in Tampa, Florida in November. The International Conference for High Performance Computing, Network, Storage and Analysis, to give it its full title, is the most important meeting of its kind in the calendar.

This, the 18th SuperComputing Conference, offers an extensive programme which includes the annual award of the Seymour Cray and Gordon Bell Prizes for personal and application achievement, respectively. There is a wide range of talks, exhibits, demonstrations and other related activities. There are also two days of tutorials, with 14 full-day and 12 half-day sessions on subjects ranging from introductions to high performance computing to programming on the latest hardware. EPCC contributes toward the tutorial programme with a full-day introduction to “Application Performance on the Blue Gene Architecture” (see below).

The EPCC booth is number 213 in the main Exhibit Hall next to our HPC-UK partners CCLRC Daresbury Laboratory and The University of Manchester. If you have any questions for EPCC staff, or about the activities of EPCC, please come along to the booth to say ‘Hello’. We hope to see you in Florida.

**EPCC tutorial at SC06: Application Performance on the Blue Gene Architecture**

We have extended the tutorial from that given at SC05 from a half-day session to a full day by adding hands-on sessions. Participants can gain direct experience of using the system and will have the opportunity to experiment with their own codes.

Blue Gene was designed as a special purpose system, offering unprecedented computing performance coupled with very low power consumption and cost for a limited set of applications. Adoption has, however, become more widespread and it is now being used in a relatively diverse range of scientific disciplines. Exploring the range and type of application that can make effective use of these systems, and the relevant architectural features influencing performance, is therefore highly topical.

The tutorial focuses on the performance of a series of applications and considers the techniques required to achieve optimal performance and scaling on the Blue Gene architecture. We discuss the factors and bottlenecks that influence performance, and cover techniques for compiler, memory and communication optimisation.

EPCC has a 1024-node Blue Gene system, the first to be installed in Europe in January 2005. The tutorial reflects our extensive experience of using the system, which is the flagship HPC resource for the University of Edinburgh.

*The tutorial is on Monday 13th November. It is presented by Lorna Smith, Mark Bull, Joachim Hein and Alan Gray of EPCC. Pre-registration and on-site registration are both possible (subject to availability): see http://sc06.supercomputing.org/techprogram/tutorials.php if you would like to attend.*
The HPC-Europa Transnational Access programme is now in its third year. More than 500 applications have already been approved for HPC-related visits to research groups associated with the 6 participating HPC centres. Applications have been received from 30 of the 33 eligible European countries so far – if anyone knows of any research groups using HPC in Iceland, Liechtenstein or Malta, please let us know!

The annual Transnational Access Meeting (TAM) gives HPC-Europa visitors an early opportunity to present the results of the research undertaken during their visit, and allows researchers who visited different centres at different times to come together and find out about each other’s work. It also provides an excellent opportunity for the visitors to meet up with the friends they made during their visits!

This year’s meeting was hosted by Barcelona Supercomputing Centre (BSC). Following keynote opening speeches by two eminent researchers from Barcelona (Prof. Mateo Valero, director of BSC, and Prof. Carme Rovira, of the Parc Científic de Barcelona), a total of 28 HPC-Europa visitors presented talks about their research, while 15 more presented posters.

HPC-Europa is an interdisciplinary programme, and the range of topics covered in the presentations highlighted the diversity of the visitors’ research backgrounds, from pure computer science through classical computational science fields - such as computational physics and chemistry - to emerging fields such as bioinformatics.

The high standard of the presentations demonstrates the value of the research projects supported by HPC-Europa. Copies of all of the posters and slides, as well as the visitors’ short papers presenting their results, can be downloaded from: www.bsc.es/TAM2006/.

In addition to the presentations, a round table session allowed visitors to give feedback on the programme and share their experiences, and the technology demonstration sessions provided a showcase for the other HPC-Europa activities. These include projects on data management and portability, performance analysis tools, and development of tools for collaboration using AccessGrid internet conferencing software.

BSC also offered all attendees the chance to visit MareNostrum, ranked the most powerful supercomputer in Europe in the Top 500 list at the time. It is not merely the scale of this system which is impressive, but also its stunning setting in an old chapel.

The TAM dinner was held on the terrace of a beachside restaurant, where a seemingly never-ending array of Spanish food just kept arriving on the tables throughout the evening. The beachside location also offered the opportunity for a post-dinner dip in the sea – an opportunity which was only taken up by a hardy few hailing from the North Sea and Baltic regions!

For more information about HPC-Europa, please see: www.hpc-europa.org

The next closing date for applications for HPC-related research visits in any discipline is 15th November 2006. There will be three more closing dates in 2007.
The second week in September saw Washington DC play host to both GlobusWORLD 2006 and GGF-18. For the first time these were run concurrently with GridWorld 2006. The Washington Convention Centre provided an air-conditioned oasis within the sticky humidity of DC.

A half a dozen members of the OGSA-DAI team attended the conference. There we hosted a developers session and a users forum and also contributed to a session on Globus data management tools. The new design of OGSA-DAI, planned for release in March 2007, was presented. This design includes a completely refactored activity framework, an expanded resource model, improved security and support for persistence, scalability and robustness. The proposals met with a positive reception. We also had a productive meeting with our Globus friends to discuss architectural issues.

For more information about OGSA-DAI see: www.ogsadai.org.uk

ISC 2006
Dresden, June 2006
Kenton D’Mellow

For most of its 20-year history, the International Supercomputing Conference (ISC, the main European supercomputing meeting) has taken place in the old University city of Heidelberg at the end of June [1]. This year, owing to the FIFA World Cup, the meeting had been moved to the less congested eastern city of Dresden.

The highlight of the meeting, the announcement of the 27th and latest Top500 [2] list of the world’s fastest supercomputers, took place on the opening morning. There were no particular surprises this time (the list is updated every six months), with the large IBM Blue Gene/L system at Laurence Livermore National Laboratory in the United States retaining top spot at 280 TeraFlops [LINPACK] performance. There was something of a stir when the fifth-ranked entry was announced: a Bull machine at the Commissariat à l’Energie Atomique in France timed at 42.9 TeraFlops. This is the fastest machine in Europe at present and the sole entry for Bull in the Top 500.

The remaining programme contained a number of traditional features such as the “hot-seat” session, in which a representative of different suppliers of HPC equipment is invited to give a short sales pitch, and is then grilled (lightly) by a panel of experts.

A range of technical and scientific sessions provided a great deal of interest. On the technical side, particularly eye-catching was the discussion of PetaFlop scale computing in Japan and the US. Kenichi Miura of the National Institute for Informatics described the Japanese plans to invest some $1 billion over seven years to achieve a significant PetaFlop scale facility to replace the now venerable Earth Simulator (which, remarkably, still appeared in the Top 10 list). Meanwhile, Steve Scott of Cray Inc. from the US talked about intentions for a 1 PetaFlop Cray machine to be in place by the end of 2008. To complement the hardware aims, the target is to have a real application perform at a sustained PetaFlop by 2010.

On the scientific front, some memorable work was presented by Theo Geisel of the Max Planck Institute for Dynamics and Self-Organisation in Göttingen, who discussed new models to describe the spread of epidemics such as “Bird Flu”. Geisel also provided an entertaining account of how to track human movement without breaching confidentiality legislation by taking the advice of Watergate source Deep Throat: “follow the money”. This used information from the dollar bill tracking wheresgeorge.com, which invites the public to log the location of US treasury notes via the internet [3]. Also of interest was the account by Galen Gisler, University of Oslo, of simulations of the Chicxulub asteroid impact around 65 millions years ago which is thought to have led to the demise of the dinosaurs.

ISC 2007 will also be held in Dresden to accommodate the growing interest in this meeting.

[1] www.supercomp.de
The headquarters of the US National Center for Atmospheric Research (NCAR) occupies an imposing position at the foothills of the Rocky Mountains near Boulder, Colorado. The Mesa Laboratory, as it is known, is itself a striking and unusual 1960s building which was designed by the Chinese-born architect I. M. Pei (probably best known for the Louvre’s controversial glass pyramid). His design was inspired by the monolithic structures of the local Anasazi Indian cave dwellings, and reflects the stark surrounding landscape. NCAR were the hosts of the 12th IBM System Scientific Computing User Group meeting (ScicomP), held over three days in July, attended by three EPCC staffers.

The meeting itself, held in the rather more down-to-earth surroundings of the UCAR Center Green building in downtown Boulder, is a forum which provides an opportunity for both operators and users of IBM supercomputers to meet with IBM staff and share experiences. The users can hear about not only the latest plans for hardware, but also learn about developments in compilers, libraries, and tools from the relevant IBM experts. These issues, which are important to high performance computer users, are of particular interest to EPCC, which runs a number of large IBM machines. These include the University of Edinburgh 2048 processor Blue Gene/L service and HPCx, the consortium which operates the UK national academic supercomputing service (a 1536 processor p575 Power5 machine).

NCAR is also a major user of IBM hardware, hosting a Blue Gene/L along with Power4 and Power5 systems, which they use in their research into the Earth’s changing climate and atmosphere. NCAR perform studies not just into the science behind such changes but also into the resulting environmental and social impacts. Local speakers contributed a number of talks to the programme, which included their most recent (dire) climate assessments and technical computing material. Contributors from other large US supercomputing facilities with IBM systems included the National Energy Research Scientific Computing Center and the Lawrence Livermore National Laboratory. European centres represented included Forschungszentrum Jülich and Rechenzentrum Garching from Germany, the European Centre for Medium Range Weather Forecasts, and EPCC.

EPCC contributed with talks by Kevin Stratford and myself. Kevin discussed Lattice Boltzmann calculations on Blue Gene/L machines. In addition to presenting results from running on EPCC’s system, he showed that his code scaled easily to 16,000 processors on the 20-rack system at IBM Thomas J. Watson Research Center. I described experiences from the recent upgrade to HPCx, and presented performance comparison results. I included results from utilising a newly available feature called simultaneous multithreading which allows multiple tasks to be performed simultaneously on one processor. This is a feature also becoming increasingly popular in consumer products.

This useful and interesting meeting was complemented by the fact that Boulder, with more than 300 days of sunshine per year, proved an excellent place to visit. The 2008 meeting will be in Garching, Germany.

www.spscicomp.org