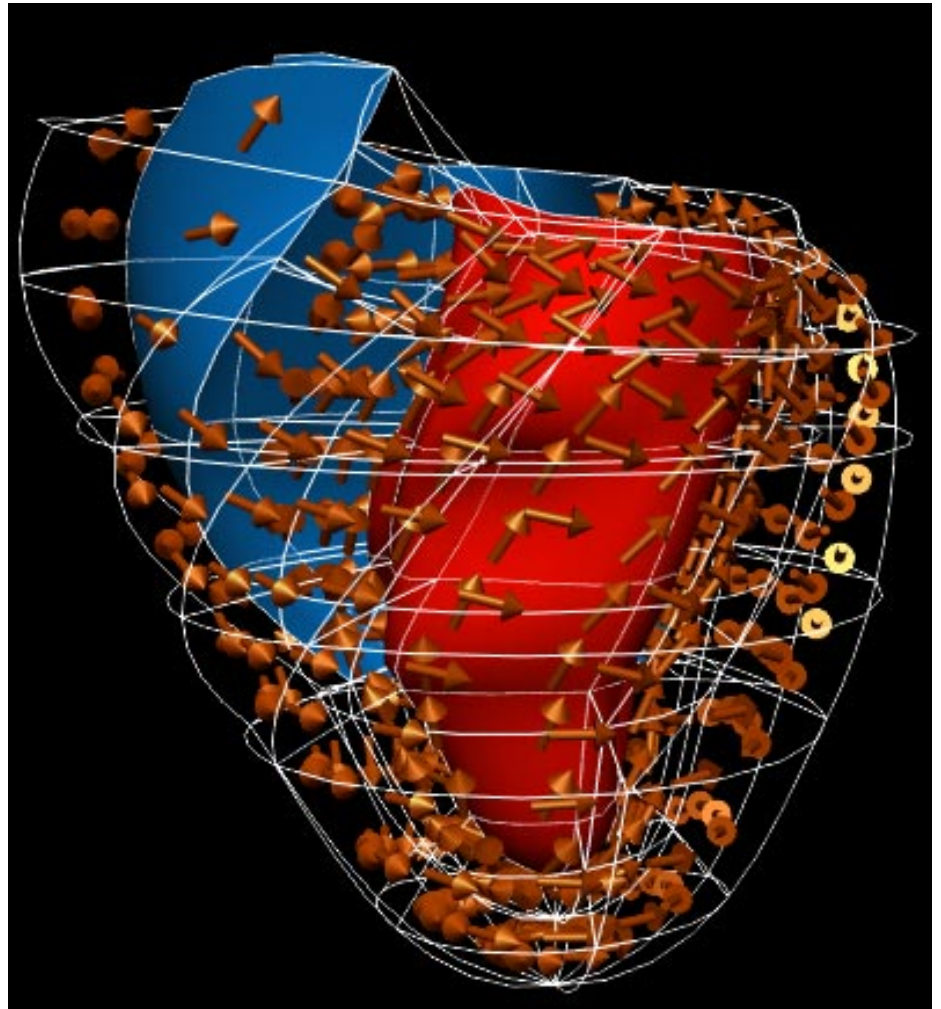


ISSUE 3, SUMMER 2001

# UKHEC

High-end computing



*Heart modelling image produced by the Bioengineering Research Group, The University of Auckland.*

# Contents

3	<i>News from EPSRC</i>	10–11	<i>The HPC Grid portal and related work</i>
4–5	<i>Courses, training and UKHEC seminar</i>	12–13	<i>Java for HEC</i>
6	<i>‘Visualization and Virtual Reality for Scientific Applications’ course</i>	14–15	<i>SCGlobal and the UK Constellation site</i>
7	<i>MSc in High Performance Computing</i>	16–17	<i>ePortal – a Grid computing solution for computational scientists</i>
8	<i>New UKHEC report on software engineering The CLRC e-Science Centre</i>	18–19	<i>Grid2000 Workshop and HiPC2000</i>
9	<i>Virtual Organizations and Mini-Grids</i>	20	<i>Contact information and feedback form</i>

---

## Editorial

*John Brooke, MRCCS*



Welcome to the first UKHEC newsletter of 2001.

Since this is a time of great change in High End Computing, UKHEC looks back on a hectic but exciting year 2000 and forward to a year 2001 that looks set to continue in the same vein.

Particular highlights of last year include the very successful and well-attended Java for HEC seminar at EPCC in November, the Grid seminar at Daresbury in June, (which included a live broadcast of the UK's first Globus tutorial) and the Visualization seminar at MRCCS whose key presentation featured the integration of visualization

and HPC (see front cover).

Last year was dominated by the word ‘Grid’ and this year seems set to be dominated by ‘eScience’. UKHEC has been active in tracking both, as can be seen from the contents of this issue.

However, amidst all this novelty, principles of good program design remain relatively constant and will become even more important in the heterogeneous environment of the Grid. Thus UKHEC is continuing to present best practice and to disseminate information about developing programming standards and development tools.

Supplementary details on many of these topics can be found at <http://www.ukhec.ac.uk>

## Technology Watch Panel visit to US

30 October– 2 November 2000

The High Performance Computing Strategy Committee's (HSC) Technology Watch Panel recently visited the USA. This visit followed on from the successful 1998 visit which concentrated on the developments at the Accelerated Strategic Computing Initiative (ASCI) sites.

The 2000 visit focused on the recent US Terascale computing procurements and Grid-based developments. The sites visited included the Argonne National Laboratory (ANL), the NSF Partnership for Academic Computational Infrastructure Centres (PACI), (namely, the National Centre for Supercomputing Applications (NCSA) at Urbana-Champaign and the San Diego Supercomputer Centre (SDSC)), the headquarters of the DOE, NSF and DARPA (Washington), and NASA Ames.

A detailed report can be found on the EPSRC website at <http://www.epsrc.ac.uk> where access to items is menu-driven. You should choose the link 'Support for Researchers'. Then select the item 'Centrally Supported Facilities', which describes the large-scale computing and experimental resources administered by EPSRC. From there you select the fourth submenu, 'Computing'. This describes the centrally-supported services available to UK researchers. Select the third item, 'High End Computing', which gives details of the Research Councils support for this area, including details of how to apply for resources, strategic issues for users to be aware of and information about forthcoming meetings. The Technology Watch Panel material is accessed via a link at the end of the text, the link is named 'Technology Watch Panel' (TWP). This has a section called 'Find out more information about' and the third item Visit to the USA 2000 is the link to the report.

The detailed report is at: <http://www.epsrc.ac.uk>

A summary of the report was given by Professor Ron Perrott in CSAR Focus Edition 6 (Winter 2000).

This is available from the CSAR WWW site at <http://www.csar.cfs.ac.uk/general/newsletter.shtml>

The report is in PDF format.

## HPC User meeting

26 September 2001

This meeting has been held annually for the last few years. It provides an opportunity for UK scientists involved in HPC to hear about the latest developments both in the UK and internationally and to then discuss these issues with representatives from the Research Councils. This year's meeting will feature presentations on the 40 Teraflop Earth Simulator currently being constructed in Japan and developments in High Performance and Grid computing for large meteorological simulations.

The venue is Whitehall Suite at 1 Whitehall Place, London SW1A 2HD. For further details contact Les Sims ([les.sims@epsrc.ac.uk](mailto:les.sims@epsrc.ac.uk)) or Jon Hunt ([jon.hunt@epsrc.ac.uk](mailto:jon.hunt@epsrc.ac.uk)).

## Royal Society discussion meeting on New Science from High Performance Computing

High Performance Computing (HPC) is one of the most rapidly developing areas of contemporary science and technology, and is one that is having an enormous impact on almost all areas of science and engineering. This discussion meeting will focus on the new scientific opportunities that are being opened up by this technology. The meeting will consider developments in hardware and hardware architectures, but will also include a survey of software and algorithm developments that are required if the capabilities of HPC are to be exploited for a wide and growing range of scientific applications.

There will be a wide range of UK and international speakers such as: Professor Thom Dunning, Pacific North West Laboratory, USA; Dr Jens Norskow, Danmarks Tekniske Universitet, Denmark; Professor Carlos Frenk, Department of Physics, University of Durham and Dr Valerie Daggett, Medicinal Chemistry, University of Washington, USA.

Further details can be found on the Royal Society website under [www.royalsoc.ac.uk/events](http://www.royalsoc.ac.uk/events)



# Globus course: Using the Computational Grid

*R.J. Allan, Daresbury Laboratory*

UKHEC will give a two-day Globus course on two dates during 2001. The course is intended for those who will be deploying, managing and developing applications for the emerging computational Grid which will connect academic sites with computers, databases and experimental facilities. Grid computing has been described as the Next Generation Internet. It provides remote access to a range of 'active services' and fosters the concept of 'virtual communities' working together on a common scientific problem.

The de facto software for Grid programming is currently Globus. In the first part of the course we

will focus on the steps necessary to install the Globus toolkit and the underlying security and data discovery software SSL and LDAP. Course attendees will be provided with a CD-ROM which has all the necessary software in a portable form for them to install at their own site.

The second part of the course will introduce concepts of managing Grid users and using Globus for data transfer and remote job submission. It is normally required for each user to have a private certificate and encryption key. This security mechanism, which is part of SSL and is an IETF Web standard, will be described. The same mechanism is used to authenticate resources as well as users. A grid-map file is used for managing access to local resources and will also be discussed.

Finally we will illustrate, through a set of example programs, how to integrate Globus with a scientific code. Typically this would involve a GUI front end which would have facilities for Grid login, data transfer and job submission to a remote machine with an appropriate application already installed. More advanced examples will show how web portals can be developed using the Globus software if time permits.

The course will be given in early September at the Daresbury Laboratory near to Manchester and in early December at the Rutherford Appleton Laboratory near to Oxford.

---

## UK workshops and conferences

The following events will present material of considerable relevance to UKHEC users. Both have workshop events at which access to advance techniques, hardware and software is available for evaluation and study.

### Euro-Par 2001

28–31 August 2001, Manchester.

Euro-Par is the annual European conference on parallel computing. It is dedicated to the promotion and advancement of all aspects of parallel computing. See <http://europar.man.ac.uk>

### 12th Daresbury Machine Evaluation Workshop

28-30 November 2001, Daresbury Laboratory.

Please see the DisCo website for information nearer the time. Proceedings of past events are available. See

<http://www.cse.dl.ac.uk/Activity/DisCo>

## Globus course dates:

- 1st week September, Daresbury Laboratory
- 1st week December, Rutherford Appleton Laboratory

Please see the UKHEC website for exact dates and registration procedure nearer the time.

# Keep up-to-date with HEC research at the UKHEC Annual Seminar

10 and 11 September

*Alan Simpson, EPCC*

EPCC will be hosting the second UKHEC annual seminar in Edinburgh on 10 and 11 September, 2001 – please put this date in your diaries.

The annual seminars are intended to provide opportunities for scientists and engineers to keep up to date with the latest high-end computing research in the UK and beyond. The themes for this year's seminar are particularly topical.

The first day will cover 'Programming next generation high-end systems'. With EPSRC's HPCx procurement due early next year, it is important for prospective users to understand the techniques which have been developed to program next generation architectures, particularly through ASCI-class SMP clusters.

This will be followed by a day on 'Exploiting the

Computational Grid'. While there is currently a lot of talk (and money!) linked with the Grid, it is still far from clear how users can benefit from it – we will attempt to shine some light on this.

The UKHEC annual seminar will address these issues through a series of talks from experienced researchers in the USA and Europe as well as dissemination of the results from each of the three UKHEC centres. The talks will focus on the practical issues that directly effect computational scientists and engineers in the UK.

Information on the speakers and how to register for the event will soon be available on:  
<http://www.ukhec.ac.uk/events/annual2001/>

---

## Events and training

UKHEC runs a series of seminars, workshops and training courses. The diary for the rest of 2001 is shown below – for full details and registration, see: [www.ukhec.ac.uk/events/](http://www.ukhec.ac.uk/events/) and [www.ukhec.ac.uk/training/](http://www.ukhec.ac.uk/training/).

### *Diary*

**Globus course: Using the Computational Grid**  
1st week September 2001, Daresbury

**UKHEC Annual Seminar**  
10-11 September 2001, Edinburgh

**Visualization and Virtual Reality for Scientific Applications**  
20-21 September 2001, Manchester

**Practical Software Development for Computational Scientists and Engineers**  
16-18 October 2001, Edinburgh

**Workshop: Numerical Algorithms for HPC**  
23 November 2001, Edinburgh

**Globus course: Using the Computational Grid**  
1st week December 2001, Rutherford Laboratory

The three UKHEC centres each organise their own seminars. Anyone is welcome to attend, but please let organisers know in advance.

For further information, see each centre's own website:

- Daresbury: [http://www.cse.clrc.ac.uk/Activity/Daresbury\\_CSE\\_Seminars](http://www.cse.clrc.ac.uk/Activity/Daresbury_CSE_Seminars)
- EPCC: <http://www.epcc.ed.ac.uk/epcc-tec>
- MRCCS: <http://www.man.ac.uk/mrccs/seminars>

# Visualization and Virtual Reality for Scientific Applications

*J. Leng, MRCCS*

This is a newly developed UKHEC two-day course to be held at MRCCS on 20 and 21 September 2001.

High-end computing facilities are producing huge volumes of data which need to be interpreted and assimilated. This data interpretation stage is becoming the bottleneck in some large-scale simulations and the aim of the course is to help provide visualization and virtual reality tools that can help with this. It is aimed at those who wish to use visualization rather than those studying visualization as a subject in itself and concepts are introduced in the light of scientific data from real simulations or experiments.

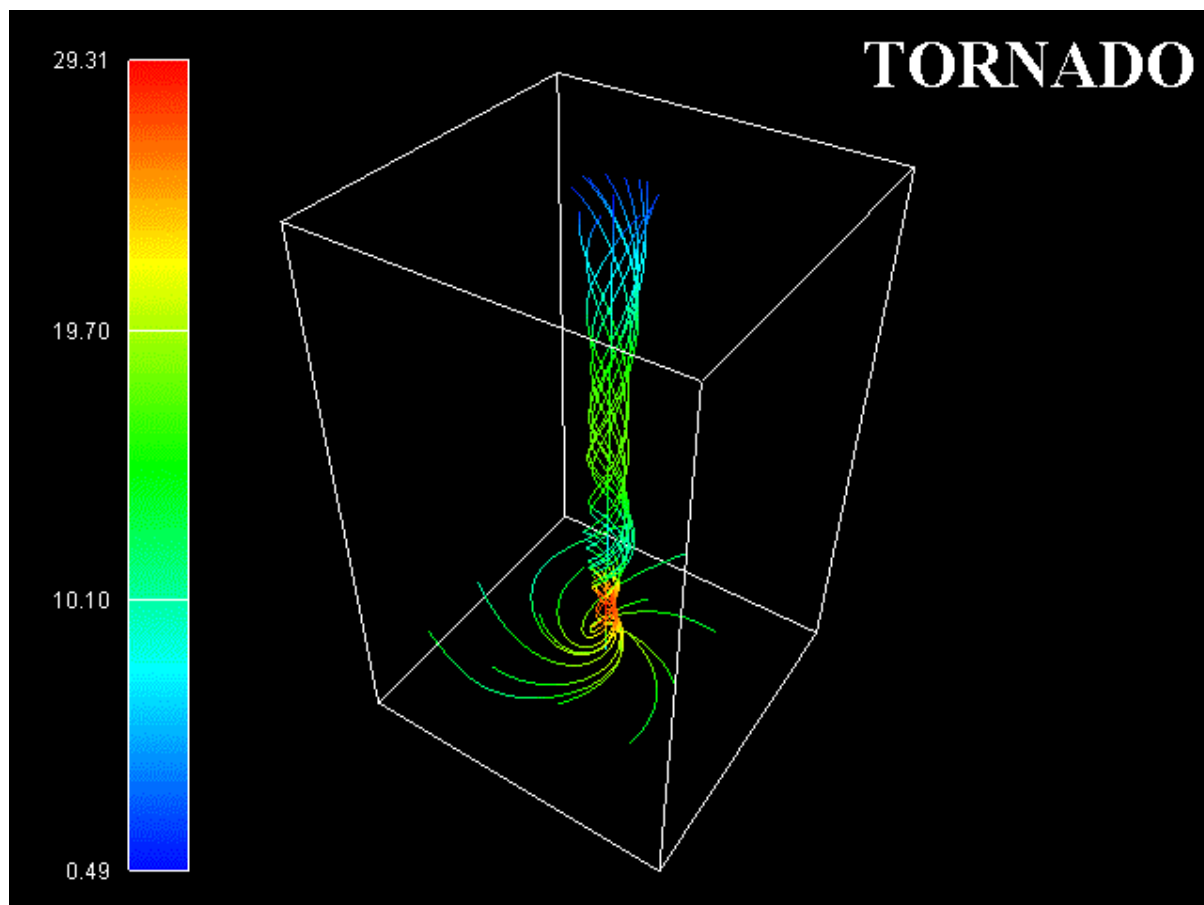
This two-day course will also provide an opportunity to utilise the Visual Immersive Projection Laboratory at Manchester. Such Reality Centres are becoming increasingly used at both academic and industrial sites, thus skills acquired will be of benefit for those wishing to pursue careers or

further study in scientific computing.

Further details and joining instructions will be available in the Training section of the UKHEC website.

The course is free to UK academics. We particularly welcome participation from postgraduate and postdoctoral assistants working on large computational codes on the UK high-end systems but attendance can be from a wider spectrum than this. The main requirement is the willingness to participate in an interactive course based on problem-solving and lab work.

*Visualization of a tornado produced by a student on a previous course. The students split into teams and each team used a visualization system to make images and animations suitable for a short presentation. This image was part of the winning presentation.*



# MSc in High Performance Computing

<http://www.epcc.ed.ac.uk/msc/>

EPCC, a technology transfer centre within the University of Edinburgh, will be offering an MSc in High Performance Computing (HPC) from October 2001.

Several EPSRC studentships which cover all fees and maintenance are available.

The course is taught as a series of intensive modules, each typically lasting around three days, with associated tutorials, course work and examinations.

## **Compulsory modules:**

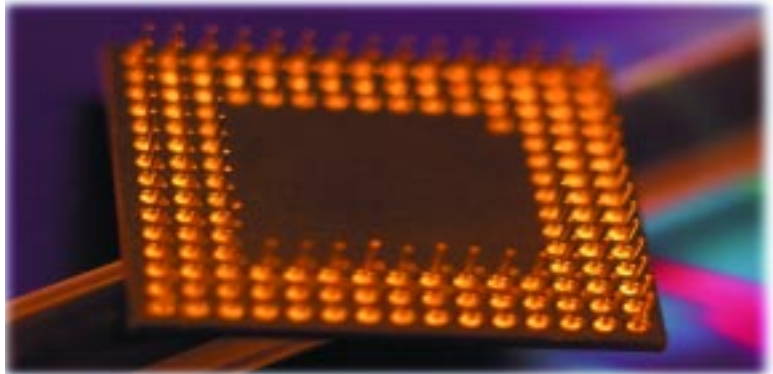
- Fundamental Concepts of High Performance Computing
- Principles of Parallel Programming: Theory and Practice
- Software Development: Methodologies, Tools and Techniques
- Object-Oriented Programming for HPC Applications
- Applied Numerical Algorithms
- Exploiting the Computational Grid
- Applied Computer Science: Architectures, Compilers and Algorithms

## **Optional modules:**

- Up to three modules in more specialized areas, according to each student's interest

## **Dissertation:**

- A mini-thesis based on a substantial practical project



EPCC has an international reputation in the application of novel computing solutions to real-life problems. This new postgraduate qualification extends EPCC's existing core HPC training courses, maintaining a strong practical focus and covering topics of key relevance to academia and industry.

MSc students will have access to an impressive range of leading-edge HPC platforms and technologies. They will acquire an extended skills base and expertise relevant to a wide spectrum of careers including computational science research and commercial software development. Graduates of this course will hold one of the few university-accredited post-graduate HPC qualifications in Europe.

Applications are encouraged from graduates of all areas of science, engineering, computer science and mathematics, and from those currently working in a relevant field. It is possible to study part-time, or accumulate credits towards a Diploma or Certificate in HPC. The entrance requirement for the MSc is a good honours degree, or equivalent work experience. Although no prior knowledge of HPC is assumed, candidates must be competent programmers in one of the following languages: Java, C++, C or Fortran.

*For more information and application details see <http://www.epcc.ed.ac.uk/msc/>.*

*Email: [msc@epcc.ed.ac.uk](mailto:msc@epcc.ed.ac.uk)*

# New UKHEC report on software engineering

Mario Antonioletti, EPCC

A new UKHEC report on 'Software Engineering and Code Development for HPC Applications' is now available from:

<http://www.ukhec.ac.uk/publications/>

The report, written by members from EPCC and Daresbury Laboratory, is targeted at anyone who will be developing codes within an HPC environment. Both existing and new tools are discussed in overview and, although not exhaustive in its coverage, it should serve as an excellent introduction for any researchers entering this fast moving field.

The report starts off by describing the different high level programming languages that are most often used for developing scientific and/or engineering applications. This is followed by a brief overview of how a compiler converts this high level source code to an executable. Knowledge of this process can help you get the best performance from a given platform. The report then goes on to discuss the different parallel programming approaches that can be used if performance is not adequate using a serial language. The historical background of each of these approaches is outlined together with the suitability for different classes of problems and/or

machines.

The next section covers software engineering practices and some of the principles behind writing good source code that is both easy to maintain and develop. This is an extremely important topic that is often not given enough emphasis in the scientific community. With a little common sense and good practice, good software engineering can make a huge difference to everyone who works on the code.

No matter how good your practices are, there will inevitably be programming errors that require backtracking to an earlier working version of the code. Source revision systems are therefore introduced, tools that facilitate the backtracking process and become essential if more than one person is simultaneously working on the same code. Finally, the hopefully familiar tools such as makefiles, debuggers and profilers are covered.

In short the report introduces and describes all the tools required for the complete code development cycle. It makes ideal reading for anyone who is about to embark on writing or maintaining new code, or would like a quick overview of software practices and tools of which they may not currently be aware.

---

## The CLRC e-Science Centre

*R.J. Allan and K. Kleese, Daresbury Laboratory*

Around six months ago, in response to pressure to make more efficient use of its facilities, CLRC set up its e-Science Centre. The Centre will act as a focus for e-Science activities in CLRC, and drive the development of e-Science as a key enabling technology for the next generation of major scientific facilities in CLRC. A number of pilot projects have been commissioned with internal funding. The HPCGrid Portal, which builds on the evaluations carried out by UKHEC, is described on page 10. Other projects include:

- Data Portal – developing a user access portal to a cross-facility range of scientific datasets using a common metadata framework
- StarGRID – Grid-enabling access to astronomical data and the Starlink software suite
- Earth Observation Grid – providing input to the Earth Observation aspects of the EU DataGrid project and to NERC's Grid requirements
- DataGrid development – establishing the first phase testbed for the EU DataGrid project
- SMART experiments – developing remote experimental control and data retrieval at the SRS

in partnership with a team from Birkbeck College, London

- Grid Reference – developing and supporting a reference implementation platform for Grid middleware
- Gigabit Networking – upgrading CLRC's internal networks to meet the future needs of facilities and the e-Science programme
- Petabyte Data Storage – upgrading the CLRC central data storage facility towards Petabyte capacity
- Industrial Collaboration – informing and involving industry in the development of the UK e-Science programme.

We have held a number of outreach workshops, eg the Industry Day on 20th February, and maintain a website with links to all the projects at: <http://www.e-science.clrc.ac.uk>.

We hope that you will find this of interest and value your comments on future directions and potential links to other projects.

# Virtual Organizations and Mini-Grids

*John Brooke, MRCCS*

At the recent IEEE ccGrid conference in Melbourne, Ian Foster of Argonne gave a keynote speech describing the development of the Computational Grid as a series of interlocking 'Virtual Organizations', sharing resources and utilizing common means of access to such resources, as provided by middleware such as Globus, Legion and Unicore for example.

The full argument can be found in a paper called 'The Anatomy of the Grid' by Foster, Kesselman and Tuecke available from the Argonne website. This paper is important as it moves the concept of the Grid from the analogy of an electric power grid where one can 'plug in' anywhere to access instantly available 'computing power'. This analogy, while useful in stressing the ubiquity and potential platform and location independence of Grid resources, is a very long way from practice in the near to medium future. More importantly, it does not emphasize the collaborative nature of the Grid as it is developing in the US, Europe and Japan.

The concept of a Virtual Organization, and the organic growth of overlapping virtual organizations seems to be a very effective way of what is happening to the Grid at the current time. The Global Grid Forum co-ordinates the developing activity of these Virtual Organizations. The Internet computing ventures *seti@home* and *climateprediction.com* are also Virtual Organizations and indeed the Internet has promoted the growth of such organizations.

This phenomenon has been noted in the UK and independently of the Foster, Kesselman and Tuecke thesis, an argument has evolved that the Grid is emerging around the interoperability of smaller units with structure on both sub and super scales.

In a paper delivered to the Grid2000 workshop in Bangalore MRCCS described an environment that we call a 'mini-Grid'. This is a collection of computational and data processing resources that has a heterogeneous structure (multi-architecture, multiple levels of data storage) but is organizationally simple, eg under the control of a

single organization.

We consider these mini-Grids to be of research interest in two ways. Firstly, a likely structure for the development of the worldwide Grid is the gradual connection and integration of such local Grids, each serving both its own hinterland and also acting as a node for wide-area applications. Secondly, because they are of sufficient complexity to test out key middleware components of the Grid, eg coupled applications, resource brokers, accounting and billing systems, without needing to address questions of local or national autonomy.

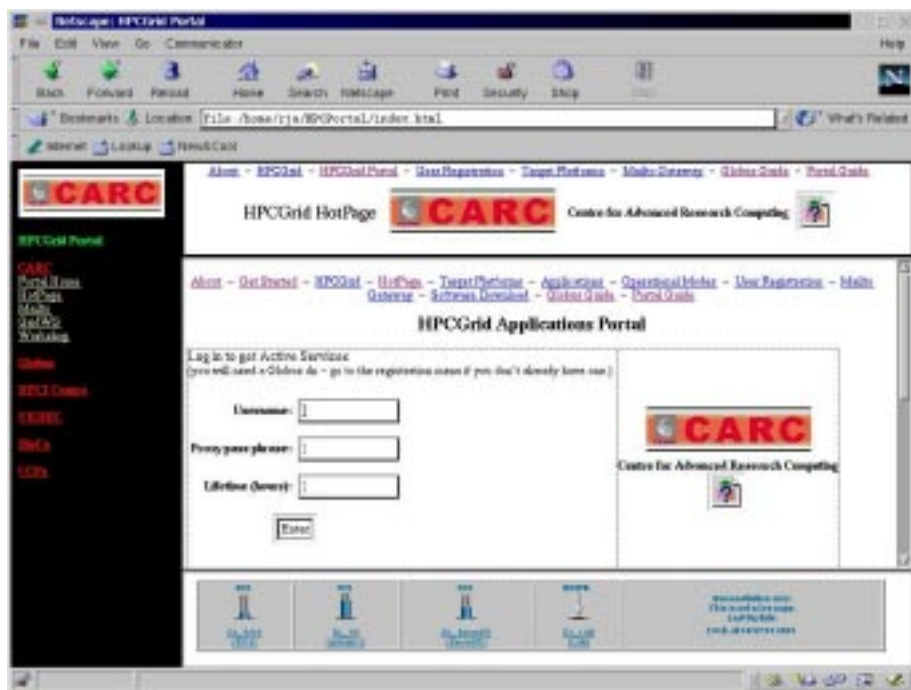
In this way the development of the Grid involves the simultaneous development of local and remote links – the development of important towns and cities around transport nodes is an example in another context.

The major UK centres of HPC computing have this dual structure and it can be accounted as a particular strength since it can potentially allow the access to the Grid of local organizations via their regional mini-Grid.

Such mini-Grids may evolve mechanisms for resource accounting and sharing both within their own Virtual Organization and between other mini-Grids and Virtual Organizations in the wider Grid structure.

The AccessGrid technology we describe in the article in this issue on SCGlobal (see page 14) is an important enabling technology, indeed in the U.S. it is seen as being a vital component in the promotion of healthy and innovative Virtual Organizations. One of the most important and often overlooked factors behind the growth of Grid computing is the enthusiasm of those involved and their desire to work together.

At all events, an important message is that the concept of the Grid is evolving as it emerges. There is no single blueprint for the Grid but there may be some key applications that shape it by being essential tools of the working scientist, as email and hypertext have already become.



# The HPCGrid Portal and related work

*R.J. Allan, Daresbury Laboratory*

Last year UKHEC ran a seminar and workshop on the Computational Grid. Proceedings, including a Globus Tutorial, are still on-line at <http://www.dl.ac.uk/TCSC/UKHEC/GridWorkshop> with a link from the main UKHEC website.

Following stimulating discussions at the workshop, we embarked on a fuller evaluation project, with additional funding from the CLRC e-Science Centre. This has resulted in a couple of draft evaluation reports, a collection of other material (online at <http://www.dl.ac.uk/TCSC/UKHEC/GridWG>) and a prototype HPCGrid Web Portal which interacts with the Globus middleware (see <http://www.globus.org>).

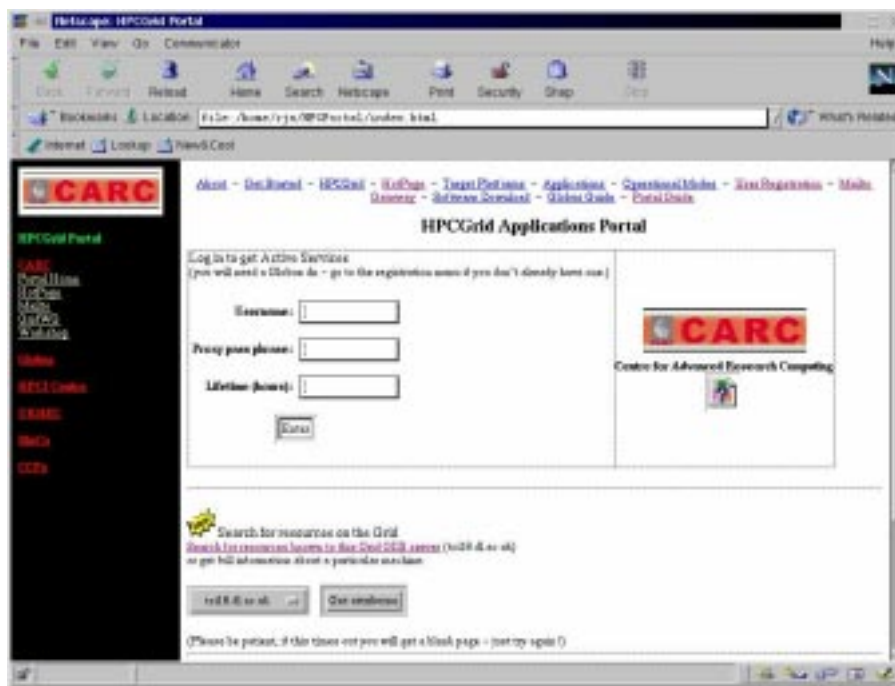
The aim of the HPCGrid Portal is to provide users of Grid-enabled HPC resource with a suite of tools to discover and select target systems, transfer files and run jobs on remote machines. It also provides a single login facility from a web browser. This is enabled using the MyProxy software from the National Center for Supercomputer Applications (NCSA) and National Laboratory for Network Research (NLNR) at Urbana-Champaign. MyProxy is a very secure way of

accessing Grid credentials, as it requires a secure server to be set up. Proxy credentials are issued which can then be accessed using a new password from a web browser. They have a limited lifetime so, even if the web server is hacked, will be of little use to anyone except the owner. We believe this is very important if the Grid is to be used in a routine way.

Once logged onto the Grid the user is presented with 'desktop' capability and the following functionality:

- resource discovery with specified attributes, eg machine type, memory size, etc.
- creation of a temporary file space on a remote machine
- file transfer
- running remote commands and submitting jobs
- automatic retrieval of 'stdout' and 'stderr' files onto the desktop
- automatically maintains a list of selected machines and remote environment variables on the desktop for subsequent use from the same browser IP address.

Currently this functionality is invoked using 'old-fashioned' HTML and CGI with C and Perl server-side programs and includes. It would however be relatively straightforward to re-produce this



functionality using Java if this proves to be the best way forward. Evaluation of the Java Commodity Globus Toolkit (CoG Kit) is currently under way.

Other work going on in the CLRC e-Science Centre may be of interest to HEC users. In particular the Data Portal will be integrated with the HPCGrid Portal. This will enable linking of computational resources with a range of experimental and modelling databases for research in chemistry and physics. It is currently linked to prototype facilities on SRS at Daresbury and ISIS at Rutherford Appleton Laboratory. We believe experimentalists may also wish to carry out simulations on HEC resources to support their scientific research. A Portal with an in-built expert system is a target for future work in this area.

For better resource monitoring and utilization we have added the HotPage software from the San Diego Supercomputer Centre (SDSC) to the HPCGrid Portal. With the use of JavaScript and remote cron scripts and http to collect resource information, HotPage provides a comprehensive view of the Grid. Icons which indicate availability and load of remote systems are dynamically updated. On-line registration and user manuals are provided for each system, with the system manager's co-

operation. HotPage is currently being re-written to fully utilise the capabilities of Globus and the underlying LDAP information register.

Our Portal also provides access to a wealth of on-line information and a user registration service. All the documents we have developed are available and a potential user may submit a request for an ID on any connected platform and automatically receive a Globus Certificate too. We hope to extend this in future to include the main national resources. Of course the local system managers have ultimate control and a peer review mechanism is included as on the current CSAR service.

Even without a Grid it is possible to examine what resources are available by accessing the search facility. The more selective search is however only possible once you are logged in.

The prototype is currently only visible from within the CLRC fire-wall, but a new server is being commissioned and it will be visible to everyone soon.

For more information please contact the author:  
r.j.allan@dl.ac.uk

# Java for HEC

Mark Bull and Lorna Smith, EPCC

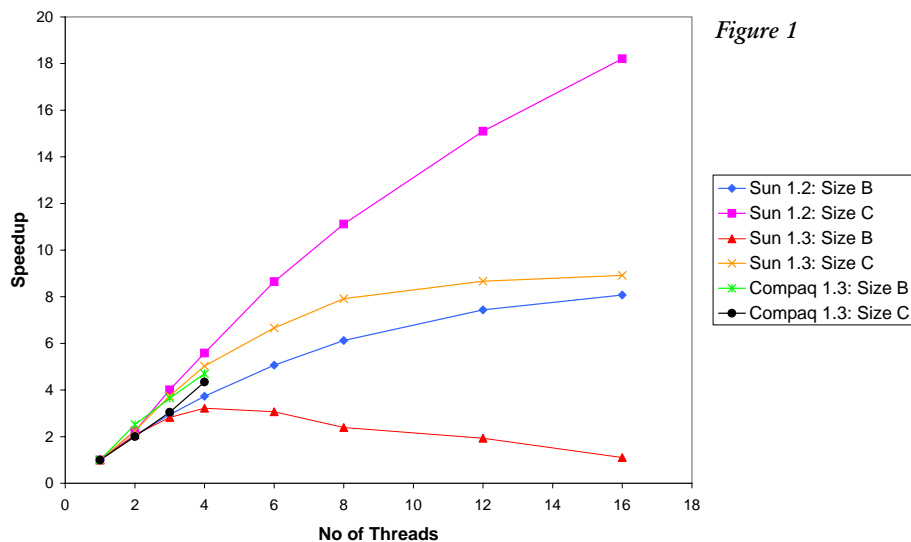


Figure 1

The aim of the serial Java Grande benchmark suite (see [www.epcc.ed.ac.uk/javagrande](http://www.epcc.ed.ac.uk/javagrande)) was to provide a standard benchmark suite for computationally intensive applications.

Java has a built-in thread class which can be used to write parallel programs for shared memory parallel architectures. By developing a multi-threaded version of the benchmark suite, the effectiveness of Java for high end computing applications on this type of system can be tested. The multi-threaded benchmarks consist principally of a subset of the serial benchmarks, together with some thread-specific low-level benchmarks.

## *The Multi-threaded Benchmark Suite*

In each benchmark, the code to be executed in parallel is separated into a new class which implements the `Runnable` interface.

A thread is spawned by creating an instance of this class; this object is then passed to one of the constructors of the `Thread` class. Each thread is given an individual identification number. This is an integer value and is passed to the constructor of the class implementing the `Runnable` interface.

Once the master thread has completed its parallel work, it will wait for the other threads to complete (to return from the `run` method) before continuing. This is achieved using the `join()` method, which causes the current thread to wait for the named thread to die.

Parameters passed to the thread are passed by reference and are shared between the threads. Class

variables are also shared. Instance variables in the new class, and local variables in methods of the class, are private to a thread.

Section I of the multi-threaded suite contains some thread-specific microbenchmarks, measuring the cost of two different barrier algorithms, thread fork/join, and synchronized methods and blocks.

Section II contains parallel versions of most of the Section II benchmarks in the sequential suite: Series, SOR, LUFact, Crypt and SparseMatMult. Of these, Series and Crypt consist of a single loop, in which each iteration is independent of every other, and the work may be distributed simply between the threads. The work of this loop is divided evenly between the threads in a block fashion, with each thread responsible for updating the elements of its own block.

In the LUFact benchmark, iterations of the double loop over the trailing block of the matrix are independent and the work is divided between the threads in a block fashion. Barrier synchronization is required before and after the parallel loop.

For the SparseMatMult benchmark, the loop over non-zero elements can be parallelized by sorting the non zero elements by row value, and adjusting the iteration block boundaries to ensure that no row is accessed by more than one thread. The original SOR benchmark is inherently sequential. To allow parallelization, the algorithm has been modified in the usual way to use a 'red-black' ordering mechanism.

Section III of the multi-threaded benchmark

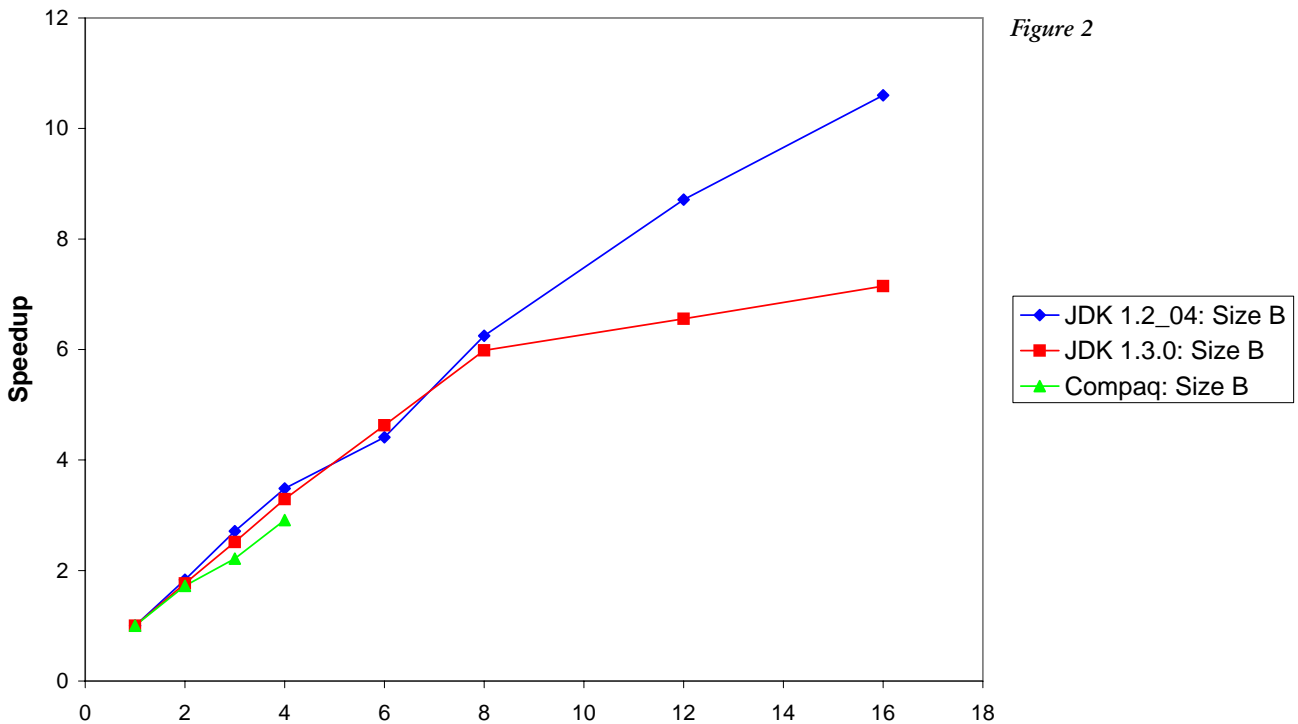


Figure 2

suite contains three of the larger benchmarks from the serial suite: MolDyn, MonteCarlo and RayTracer. The computationally intense component of the MolDyn benchmark is the force calculation, which calculates the force on a particle in a pair wise manner. This involves an outer loop over all particles in the system and an inner loop ranging from the current particle number to the total number of particles. The outer loop has been parallelized by dividing the range of the iterations of the outer loop between the threads, in a cyclic manner to avoid load imbalance. Local force accumulation is used in preference to atomically updating a shared data structure.

In the MonteCarlo Benchmark, the principal loop over number of Monte Carlo runs can be easily parallelized by dividing the work in a block-like fashion. For the RayTracer benchmark, each pixel in the image can be computed independently. The outermost loop (over rows of pixels) has been parallelized using a cyclic distribution for load balance.

### Sample Results

The threaded benchmark suite has been tested on a Sun HPC 6500 system (with 18 400MHz UltraSparc II processors and 18 Gbytes of memory running Solaris 2.7) and a Compaq ES40 (with 4500MHz EV6 processors and 4 Gbytes of memory running Digital UNIX V4.0F). The execution environments utilized were Sun\_JDK\_1.2.1\_04 and Sun\_JDK\_1.3.0 on the HPC 6500 and Compaq Java 1.3.0-alpha1 on

the ES40.

Figure 1 (see opposite page) shows the speedup of the LUFact benchmark from Section II of the suite. Here two data sizes, B and C are shown, corresponding to matrix dimensions of 1500x1500 and 2000x2000 respectively. On the Sun HPC 6500, the larger data size exhibits better scalability, whereas the opposite is true on the Compaq. This can be attributed to different Level 2 cache sizes on the two hardware platforms. Despite similar serial execution times, the Sun 1.3 JVM shows much poorer speedup than the 1.2 version. This may in part reflect the higher cost of barrier synchronization in the 1.3 JVM, but the phenomenon is also present in other benchmarks such as Crypt and Series (results not shown here), where no barrier synchronization is present. This suggests that the thread scheduling policy in the 1.3 JVM is non-optimal for parallel processing.

Figure 2 (see above) shows the speedup of the RayTracer benchmark from Section III, using data size B (500x500 pixels). Both Sun environments show slightly better scaling up to four threads than the Compaq, despite a much longer serial execution time on the ES40. The two Sun environments give similar performance up to eight threads, but for 12 and 16 threads the 1.2 JVM is significantly better than the 1.3 JVM.

The multi-threaded benchmark suite is available for public download. For details, please see: [www.epcc.ed.ac.uk/javagrande](http://www.epcc.ed.ac.uk/javagrande)

# SCGlobal and the UK Constellation site

John Brooke, MRCCS



## *What is SCSGlobal?*

SCGlobal is a very ambitious proposal to extend the SC2001 Supercomputing Conference (to be held at Denver from 11-16 November) to a truly globally distributed conference by using Access Grid Technology. This allows multiple sites to participate in shared meetings with

projections from each location visible on a large multi-window screen. Collaborative tools such as a distributed version of Powerpoint have enabled the technology to be used for distributed courses on MPI, Globus and Condor with excellent feedback from participants.

The UK TWP (see the *EPSRC News* item in this issue) recommended that the UK participate in SCSGlobal and set up at least one UK site. Accordingly UKHEC drafted a proposal which we are pleased to say has been accepted and is now part of the SCSGlobal schedule. The main access point for UK users will be an Access Grid node currently being built at the University of Manchester and funded as a joint initiative by EPSRC and the University.

We describe the UKHEC programme of events below. To find out about Access Grid and SCSGlobal see <http://www.accessgrid.org/> and <http://www.scsglobal.org/>

If you are interested in participating in or contributing to any of the four UK events described below please contact John Brooke ([j.m.brooke@man.ac.uk](mailto:j.m.brooke@man.ac.uk)).

## *Event 1 – Global computing via trans-Atlantic and trans-Pacific links*

This is a four-hour workshop. Partners include the University of Manchester (UK), HLRS, Albert Einstein Institute (Germany), PSC, Sandia (US),

AIST, JAERI (Japan), NCHC (Taiwan).

This meeting brings together major centres in Europe, USA and Asia who have been involved in experiments to exploit global metacomputing via trans-Atlantic links. The global metacomputing started with a trans-Atlantic coupled application demonstrated at SC97, was extended to include trans-Pacific links at SC99 (winning an award in the HPC Games competition) and was extended at SC2000 to include all the above sites. We also include a presentation from the eGrid testbed using the Cactus computational portal. The purpose of the meeting will be to discuss the results and experiences gained over the last few years, and to link this with current initiatives in Global Grid computing.

The scientific areas involved are: molecular dynamics, computational fluid mechanics, processing of radio-astronomy data and environmental prediction. This meeting is the culmination of a project to make the techniques of global metacomputing persistently available.

The meeting will be an opportunity to involve both computational and computer scientists in discussing leading edge problems that are potentially too large for any single machine. The method is via the coupling of large powerful supercomputers, thus the application fields are different to those exploited by the use of spare cycles on vast numbers of commodity PCs.

## *Event 2 – Collaborative use of STP Data*

This event is a BoF of two hours duration. Partners include University of Manchester, CLRC/RAL (UK), NOAA and JILA (Boulder) ARSC (U.S.) An important scientific component of the UK Global Metacomputing project is the work with the global solar-terrestrial physics (STP) community. They have requirements for the coordination of data from a variety of sources, satellite, ground-based radar, magnetometer readings. Their data processing needs



differ from those of the Virtual Sky Survey and other similar projects, in that STP data is observed continually and unique events need to be processed along with long term trends on a variety of time-scales.

It is particularly appropriate to host an SC Global event with a central node at Denver, since a major U.S. STP site is at Boulder, Colorado. Thus STP scientists from Boulder could hold a joint meeting with the WDC at RAL. We will also investigate the participation of other European sites and invite them to attend SC Global via the Access Grid Node at Manchester (if they do not have their own).

### *Event 3 – Grid Portals*

This event is a BoF of two hours duration. Partners include EPCC, CLRC/Daresbury (UK), FRZ/HLRS (Germany), SDSC, NCSA (US).

We wish to run a number of meetings to unite the major groups worldwide in Grid portal activities. These would include Daresbury Laboratory (HPCGrid Portal), San Diego Supercomputer Centre (NPACI GridPort), U. Illinois (NCSA), HLRS Stuttgart and FRZ (Unicore, EuroGrid) others. We note that HLRS and FRZ have already applied to be an SC Global site in Germany.

We believe that Grid portal technology is essential to widen the uptake of high-performance computer simulations and to combine access to computing facilities, databases and experimental instruments in a new way which will give scientists rapid access to information to test hypotheses and make discoveries.

The purpose of the meetings would be to collaborate on determining standards, security and appropriate middle-ware for these Internet-based projects.

### *Event 4 – Computation for Molecular Sciences and Materials Simulation*

This event is a BoF of two hours duration. Partners

include EPCC, CLRC/Daresbury (UK), SDSC, PNNL, Argonne, Oak Ridge, Utah (U.S).

There is interest from both communities in using the computational Grid to extend their scientific work through combining simulation and experiment, through collaboration and through inter-disciplinary and inter-scale working.

The purpose of the meeting is to give the UK community the opportunity to see how scientific researchers in the USA and elsewhere are using the Grid in this way. This includes a discussion of the UK codes and live demonstrations of US technology and new ways of working, such as computational steering, visualization and VR. Work in the UK is only just beginning in these areas and having the community understand the potential benefits would be very beneficial at this stage. Important potential contact sites would be SDSC, PNNL, Argonne, Oak Ridge, Utah.

The purpose of this meeting will be to plan the most appropriate way to Grid-enable the main UK codes DL\_POLY, GAMESS-UK, CASTEP and CRYSTAL, with a visual front-end and ability to combine remote resources. This will extend the work which is now starting.

### *The UK Access Grid Node at Manchester*

The room will be purpose equipped for the use of video-conferencing. The room can accommodate an audience of up to 50 people. Manchester is a major node on the UK SuperJanet4 backbone, with bandwidth of 2.5 Gigabits/second. SuperJanet4 accesses to Abilene via New York with a bank of four ATM connections at 155 Megabits/second and is looking to upgrade this connectivity. Manchester hosts one of the eight projects linking SuperJanet4 to Internet2, thus it has access to the latest connectivity between the UK and the US. The Internet2 partner in this project is Pittsburgh Supercomputing Centre who host one of the GigaPops on the US network.

## ePortal

WELCOME to the EPCC ePortal

To use the ePortal please complete the following tasks:

- First Select the relevant scientific consortium
- Second Select the scientific code you wish to run
- Third Select the HPC environment on which you wish to run the code

Press Submit

Scientific Consortium: MHD Consortium

Scientific Code: Larc2d

HPC Environment: St Andrews

ST Andrews Information

Hardware: Six Compaq E340's (each with 4 Digital v60 CPU's and 4Gbytes of RAM)

Software Installed: f77, f90, HPE, MP1, OpenMP, c, c++

ePortal home Submit

*A screenshot of the prototype MHD portal.*

# ePortal – a Grid computing solution for computational scientists

*Lindsay Pottage and Rob Baxter, EPCC*

High-performance computational science and engineering are now mature disciplines. There are a large number of stable, widely-used scientific software packages that exist on a range of computing platforms. Most computational scientists want to make use of these packages in as straightforward a way as possible, without having to worry about the vagaries of different environments on different platforms, data-file formats etc. EPCC believes that the use of the Grid for HPC packages has been somewhat neglected, and with this in mind we are developing a Grid computing solution for a specific yet broad class of users: computational scientists who want to do science rather than computing.

EPCC is working with the UK Magneto-Hydrodynamics (MHD) consortium to develop the prototype for a software toolkit to build a single Universal Portal for eScientists. The Portal will offer a single view of the range of available software packages for scientific simulations, in particular for some of the MHD consortium's solar physics codes. The MHD scientists will be able to select a computing service, including a 'don't care where it runs' service, to run their particular jobs using their favourite package, but in a transparent way that frees them from the details of the computing and allows them to concentrate on the science. No matter

whether their job runs on a Compaq Alpha cluster, a Cray T3E or a home-grown Beowulf system, their program interface will be the same, the details of job submission will be hidden, data format issues will be dealt with behind the scenes and the science will be brought to the fore.

The ePortal project will build on true heterogeneous network computing standards such as Java and XML. Using a combination of front-end applets, back-end interfaces to existing HPC job submission software and an intermediate portal services manager, the driving aim behind the project will be to keep the user interface as simple as possible, hiding the unnecessary details of job submission. A key feature of the design is extensibility, separating general portal functionality from application-specific details. This will allow us easily to extend the work beyond the initial area of solar physics by developing plug-ins for each new package.

EPCC and MHD hope that this pilot project will prove the benefits of freeing scientists from the details of their computers and pave the way for a wider development of portal services for science in general. We will be disseminating some of the key aspects of this work via a forthcoming UKHEC report on the practical use of XML in scientific applications.

# Review: 'Visualization Seminars For HPC Users' event

Manchester, December 2000

*J. Leng, MRCCS*

This event consisted of six seminars on a broad range of topics, from computational steering to visualization methods and case studies to VR. The speakers included Nick Avis, Helen Wright, Min Chen, Dan Kidger and Joanna Leng.

The key speaker was David Bullivant who came all the way from the Bioengineering Group in the University of Auckland, New Zealand. The Bioengineering Group is internationally respected, their aim is to understand the electro-mechanical behaviour of the heart and other organs from both mathematical models and experimental measurements.

They use HPC and they see particular importance in the visualization of the results fundamental to their work and will redesign computations to improve visual analysis. Their visualization tool also acts as a computational steering engine and is called CMISS. CMISS is really a mathematical modelling environment that allows the application of finite element analysis, boundary element and collocation techniques to a variety of complex bioengineering problems. It has a number of modules with a graphical front end for advanced 3D display while its

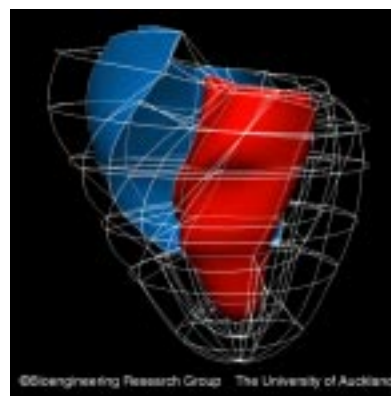
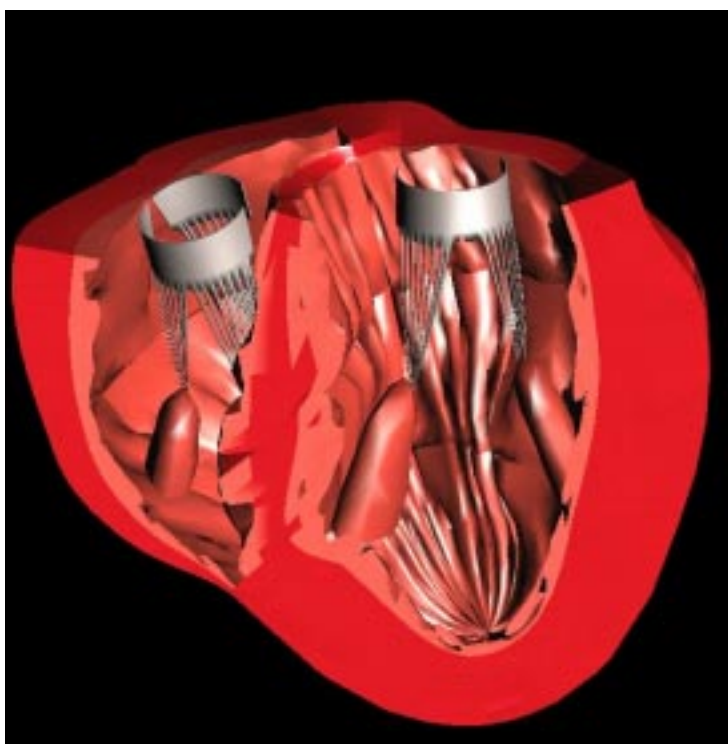
modelling capabilities are in a computational backend that can be run remotely on powerful workstations or supercomputers. His talk demonstrated how their methodology has solved particular problems that would have been impossible to answer otherwise.

David Bullivant stayed in the UK for a week-long study break in which time he worked with a collaborative group in the University of Oxford, demonstrated his software in Manchester and attended a local North West Visualization Community Club.

The event was well attended with about twenty people attending from across the UK. They came from a broad range of subject areas including some from virtual tissue engineering who are new to HPC.

Abstracts for all the talks can be seen at the URL  
[http://www.man.ac.uk/mrccs/viz\\_seminars.shtml](http://www.man.ac.uk/mrccs/viz_seminars.shtml).

The URL for the Bioengineering Group in Auckland is:  
<http://www.esc.auckland.ac.nz/Groups/Bioengineering/>



*Heart modelling images produced by the Bioengineering Research Group, The University of Auckland.*

*Left: details of endocardial trabeculations, heart valves and chordae tendinae.*

*Above: finite element model, based on prolate spheroidal coordinates and using linear Lagrange and cubic Hermite basis functions.*

Report

# Grid2000 Workshop and HiPC2000

Bangalore, 17-20 December 2000

*John Brooke, MRCCS*



HiPC, the High Performance Computing Conference of India, is not one of the best known conferences. However, this year the conference in Bangalore attracted a high-quality programme of invited speakers and contributed papers. It was also held in conjunction with the Grid2000 International Workshop on Grid Computing, which was sponsored by IEEE, ACM and the European Grid Forum (eGrid, now merged with the Global Grid Forum). The standard of speakers and papers reflects the growing recognition that India is becoming a major force in software development at many levels. Indian Universities and research institutes are developing strong links with US universities and national laboratories.

MRCCS attended the Grid workshop to give a paper outlining moves towards developing methods of accounting for resources on the Grid and outlining how these could be tested and developed in environments called Mini-Grids. We also wanted to talk informally about the work of UKHEC and brought along copies of the UKHEC newsletter, which we are happy to say were quickly taken up. Finally, it seems important to keep in touch with the very exciting work being done in India and in collaborations between Indian institutions and groups in the US, Europe and Japan.

The Grid2000 workshop attracted participants from 15 countries, presenting 20 papers. It was organised by Rajkumar Buyya of Monash University and Mark Baker of the University of Portsmouth. The keynote speaker was Wolfgang Gentzsch, now of Sun, called 'DOT-COMing the Grid'. A speaker from Sun had given a talk with a similar title at the eGrid meeting in Poznan last April and met with a mixed reception. This talk was much better received, perhaps because Wolfgang has a lot of experience of what has come to be known as Grid technology, through a series of EU projects and via companies such as Genias and Gridware Inc.

An interesting focus for the workshop was provided by the presence of Heinz Stockinger of the DataGrid project. This provided a focus that some of the more theoretical papers could organise around. It was also an opportunity to bring together work from the two major EU Grid projects (see last UKHEC newsletter). It was encouraging to see how strong Europe is becoming in Grid research with over half the papers coming from European research.

The proceedings are available as a volume in the Springer LNCS series, LNCS 1971 'Grid Computing – Grid 2000', R. Buyya and M. Baker (Eds.), Springer 2000. Grid 2001 will be in November 2001 alongside the SC2001 Conference. A call for papers has been issued with a final submission date of May 18th.

The main conference began on Monday 18 December. The programme was interesting and varied. For example, a very strong promotion of commodity computing as the way forward was given by Dan Reed of NCSA. This was counterbalanced two days later by Hans Zima of the University of Vienna who pointed out the problems of ever increasing demands for space and power by commodity clusters as they scale towards the petaflop level. Hans Zima described work towards the development of PiMA (Processors in Memory Architecture). This allows for an object-oriented approach to hierarchical memory by harnessing local microtasking of memory references with a global

layer of objects that can be used as the basis for higher level programming. There was also a very impressive talk on the management of distributed data from Robert Hollebeek of the University of Pennsylvania. This featured the use of HPC to organize hospital records in the US and to allow patient records to be retrievable across distributed locations as patients moved to different regions. Loss of paper-based records is a serious and sometimes life-threatening problem, so digitisation and data-base technology would seem to have a lot to offer here.

We found the liveliest parts of the conference to be those concerned with architecture, middleware and new technologies using wireless devices. This perhaps reflects the fact that these issues are of great concern to the emerging Indian computing infrastructure. India does not have very developed fibre-based networks and so the country has an incentive and an opportunity to explore wireless networks. It is striking that travelling through rural India, one of the distinguishing features of each village was its satellite dish.

The conference has a very strong emphasis on involving students, with an imaginative tutorial programme and a lively poster session, which gave a prominent place to student presentations. John Brooke of MRCCS spent a considerable amount of time talking to the students and researchers in the developing HPC centres. Their enthusiasm and willingness to learn was evident. They are focused much more closely than a comparable European group on actually physically building clusters and HPC systems, since there are problems with export restrictions on US equipment (due to political factors) and lack of hard currency. We explained how HPC is organized in the UK and made some contacts which we hope to follow up.

The conference organization was excellent, with social events featuring Indian classical music and the dance traditions of southern India. A big problem for future international participation is the timing; close to Christmas is not a good time for people to

undertake major travel and the flights are very heavily booked at this time. Against this is the fact that participants from outside India were curious about the country and wanted to visit. Our impression was that there is a groundswell of activity and research in India that will lead to this conference becoming more important in future.

This was summarized by an excellent presentation by N.R. Naranya Murphy of Infosys Ltd, an Indian software house competing successfully in India against the large multinational companies in the field of commercial data-bases for transaction handling. His talk was addressed to the need for India to develop its own strong independent base for software development.

We only took exception to one part of the talk: when discussing operating systems he implied that all the major ones had been developed in the US. We were able to point out from the floor that Linux was developed at the University of Helsinki and with strong European input throughout. Clearly most of the other Europeans felt similarly; you could feel the buzz in the hall as the statement was made. Almost for this reason alone we felt that it was worthwhile that the UK had a representation at the Conference and that the Indian participants became more aware of other contacts and collaborations than solely with the US. Otherwise there is a danger that a purely US-centric view will develop in the emerging software nations.

If you are interested in the HiPC series of conferences and might think about attending or contributing to next year's HiPC2001, information can be found at <http://www.hipc.org>.

The Conference Proceedings are available as LNCS 1970, 'High Performance Computing – HiPC2000', M. Valero, V.K. Prasanna, S. Vajapeyam (Eds.), Springer 2000.

The programme of this and previous conferences can be seen at <http://www.hipc.org>

# Feedback form

We'd like to know what you think of this newsletter. You can pass on comments by faxing back this form to: 0131 650 6555. This form is also online at: <http://www.ukhec.ac.uk/feedback.html>

1. Did you find any articles particularly interesting or useful? If so, which? \_\_\_\_\_

2. Why? \_\_\_\_\_

3. What topics would you like to see covered in future issues? \_\_\_\_\_

4. Please feel free to add any further comments \_\_\_\_\_

Your name \_\_\_\_\_ Job title \_\_\_\_\_

Company/institution \_\_\_\_\_

Address \_\_\_\_\_

Email \_\_\_\_\_

Telephone \_\_\_\_\_ Fax \_\_\_\_\_

*Was your copy of this newsletter properly addressed? Yes/No*

## Contact information

### MARTYN GUEST

Computational Science and Engineering  
Department  
Daresbury Laboratory  
Daresbury  
Warrington WA4 4AD  
m.f.guest@dl.ac.uk  
01925 603240  
<http://www.cse.clrc.ac.uk/Activity/HPCI>

### ALAN SIMPSON

EPCC  
JCMB, Kings Buildings  
University of Edinburgh  
Mayfield Road  
Edinburgh EH9 3JZ  
a.simpson@epcc.ed.ac.uk  
0131 650 5120  
<http://www.epcc.ed.ac.uk>

### TERRY HEWITT

Manchester Computing  
University of Manchester  
Oxford Road  
Manchester M60 9PL  
w.t.hewitt@mcc.ac.uk  
0161 275 6095  
<http://www.man.ac.uk/mrccs>  
<http://www.csar.cfs.ac.uk>

### ROBERT ALLAN

Computational Science and Engineering  
Department  
Daresbury Laboratory  
Daresbury  
Warrington WA4 4AD  
r.j.allan@dl.ac.uk  
01925 603207  
<http://www.cse.clrc.ac.uk/Activity/HPCI>

### DAVID HENTY

EPCC  
JCMB, Kings Buildings  
University of Edinburgh  
Mayfield Road  
Edinburgh EH9 3JZ  
d.henty@epcc.ed.ac.uk  
0131 650 5960  
<http://www.epcc.ed.ac.uk>

### JON HUNT

High Performance Computing  
EPSRC  
Polaris House  
North Star Avenue  
Swindon SN2 1ET  
jon.hunt@epsrc.ac.uk  
01793 444263  
<http://www.epsrc.ac.uk/hpc>

Main UKHEC web page: <http://www.ukhec.ac.uk>

This site is under construction and will be used to provide our main source of information.