CSAR on the Grid

John Brooke
Manchester Computing, University of Manchester

Realism and Opportunities

It is interesting for me to reflect on two very hectic and exciting years since I became the leader of the Special Projects research team at Manchester and worked in conjunction with my colleagues in the CSAR service and MVC to establish eScience and Grid Computing as a part of the service that we offer to UK High End Computing. From the beginning of the CSAR service we were interested in the possibilities offered by metacomputing, the linking together of computational facilities to provide more than could be gained from a single machine alone. The CSAR service right from its inception had a heterogeneous and differentiated architecture, with an Origin2000 coupled to a T3E. We were therefore interested in the work of other major European and US sites that were involved in metacomputing, such as HLRS Stuttgart, Pittsburgh Supercomputing Centre (PSC) and the Centre for Scientific Computing in Finland.

These contacts proved very fruitful and within two years of the start of the service we had been involved in the prize-winning collaboration in the HPC Games competition at SC99 and were members of the first large scale European Grid project, EUROGRID [1], which specifically concerned itself with linking large centres of computational excellence providing environments for complex problem solving in a variety of areas, molecular biology, meteorology, coupled models in engineering and also provided enabling technology for a European HPC Grid.

Through these contacts we became aware of the intense interest world-wide around the concept of Grid computing. Unfortunately this was a hazy concept, sometime advertised (at that time) as providing cheap, ubiquitous "computing power" on demand and without the user needing to concern him/herself with the origin of such power. The analogy was with the electrical Grid, since then events in the electricity supply industry itself have indicated that even with electrical power things are not that simple, witness the problems faced by the State of California and the Enron crisis. However it was clear that, despite the vagueness of this conception, integrative work of immense value was taking place in large-scale computing. It is greatly to the credit of the Globus project that in the midst of this they produced two very far-reaching research papers, "The Anatomy of the Grid" and "The Physiology of the Grid" [2] that clarified the basis for the developing architecture of persistent distributed collaborations. We were also aware, through EUROGRID, of the work that had gone into the development of the concept of seamless access to complex resources that forms the heart of various projects based around UNICORE [3].

At this time also, the UK eScience programme was initiated by John Taylor the Director-General of the Research Councils and Tony Hey was seconded from the University of Southampton to direct the eScience core programme. We found ourselves working with colleagues at EPCC and CLRC (Daresbury and RAL) to construct a fledgling UK Grid based using Globus. This programme also looked beyond the easy publicity associated with all things Grid, to identify what was the nature of the challenges posed by the integration of computational resources, experimental apparatus and data management and curation. A particular issue identified by the UK was the focus on the data and knowledge levels of the Grid. How does one provide infrastructure for large scale scientific collaboration, how does one ensure that differing data storage formats are made interoperable and what is the computational infrastructure required to support such large scale and persistent collaborations of research groups? These were severe and challenging problems, but because they were identified clearly they offered the prospects of real scientific progress that the vague concepts of a free ubiquitous "power grid" did not. There are now 10 active eScience Centres building the UK Grid, with others constructing regional Grids and a number of projects based on scientific applications throughout the range of disciplines covered by the UK Research Councils.

Current Opportunities for CSAR Users

In the previous section I have tried to show how groups in the UK are combining to form and define what a Grid infrastructure could be and how it could benefit UK science. At CSAR we have a particular focus on the very high-end of the Grid and we regard complexity and problems of scale as part of our particular problem domain. There is a large amount of Grid infrastructure work that is of great importance but does not touch these issues. Thus one can build Grids of clusters of workstations or PCs with a uniform architecture and operating system. This is done in the EU DataGrid project and is clearly the best choice for the problem of handling huge amounts of experimental data. High performance computing centres, however, are not dominated by clusters, they have highly differentiated and specialised architectures and for very good reasons. This has shaped our strategy in terms of project involvement and planning for services.

We are happy to announce that after several years of intensive effort we can now offer CSAR users what we consider to be production level Grid services on our major facilities. These are:

1. The UK eScience Grid links together the 10 eScience Centres all of whom are donating resource to initiate a persistent UK Grid. CSAR is pleased to report that our major resources Turing, Green, Fermat and Wren are enabled to participate in this Grid and have up-to-date and working deployment of Globus as recommended by the UK Grid Support Centre [6]. We have created an experimental project for those wishing to test their application in a Grid context, hopefully involving the participation of other resources at different sites. We are also part of the UK Grid Support Centre along with colleagues at CLRC and EPCC and provide input to the UK Engineering Task Force.

2. The UNICORE service based on the development of middleware to make submission of complex jobs easy for users and independent as much as possible of details of the underlying architectures utilised to run the jobs. It can be thought of as a means of composing complex workflows for solving problems on HPC systems. This allows users the possibilities of conducting experiments across a European High in Performance Grid. A demonstration of the potential scientific benefits of such an approach was given at the recent UK All-Hands Meeting Sheffield (September 2-4 2002) through a collaboration between members of the UK RealityGrid project and the EU EUROGRID project. The application demonstrated was a Lattice-Boltzmann simulation of mixing in two-phase liquids, where the simulation was run on one supercomputer, the visualization modules ran on another and the whole calculation was steered from a laptop on the conference floor with the user having the ability to change parameters and steer the calculation dynamically whilst directly visualizing the changes in the mixing properties of the two-phase fluid and seeing how this responded to the parameter changes. The whole steering was controlled within the framework of the proposed Open Grid Services Architecture [4] and is based on the first working demonstration of a scientific service based on OGSA devloped by Dave Snelling of Fujitsu Laboratories Europe as part of the EU-funded Grid Interoperability Project (GRIP [5]).

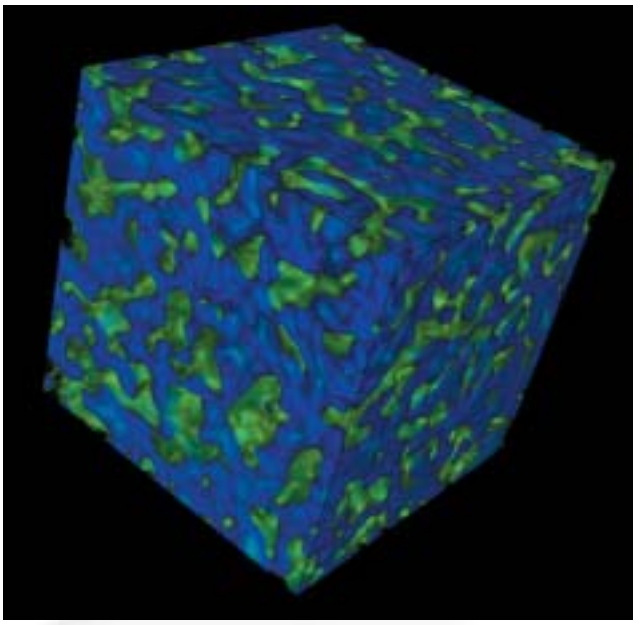

Figure 1: A visualization of data from the early stages of a Lattice-Boltzmann simulation of the phase-separation of an oil/water emulsion on a 256-cubed lattice.

To find out more about these services please contact me via the CSAR Helpdesk. Recently, we demonstrated our Grid technology at Supercomputing 2002, with an enhanced RealityGrid demonstration and demonstrations of metacomputing between world-wide resources to solve problems beyond the limits of a single machine. Some of these demonstrations were shown over the AccessGrid [6].

This is another area where we have successfully pioneered a new service for collaboration, details of this were given in the CSAR Focus December 2001 edition where we described how we organised and presented the UK SCGlobal Constellation Site programme. Our commercial partners in CfS, CSC and SGI very generously donated resources on the CSAR service and facilities at SC2002 which enabled us to provide demonstations comparable in quality and ambition with those being provided by the best centres of High Performance computation in the US, Europe and Asia-Pacific. This work was shown at many locations at SC2002 [7], on our own stand, on the UK eScience stand, the SGI stand, the EUROGRID/GRIP stand and on the stands of our metacomputing partners, such as HLRS Stuttgart, PSC and JAERI and AIST (Japan).

Future Challenges, the Economy of the Grid

As several people (including us) have pointed out, putting resources onto the Grid does not thereby make them free. At present the Grids with which we are working, the UK eScience Grid and EUROGRID have resources donated to them by participants, this generosity cannot survive large scale production use. It is here that the CSAR tokens economy and our experience of it have proved very valuable in helping us to develop proposals for how resources might be traded and exchanged on the Grid. We have been active with our partners in CfS and other UK sites which manage resources for HPC to create a proposal for a Grid Economy that can facilitate the sharing of resources across organisations (the remit is actually wider than this). This project is led by Professor John Darlington of the London eScience Centre and the eScience programme has recently indicated that it will fund the project and it will hold a workshop to publicise the idea and possibility of such exchange. We will work to make this project a success. It will give the UK a lead in this field and will lay the basis for large scale co-operative computing to maximise the UK resources and to provide UK researchers to resources on a global scale.

We are also working to enable significant codes to utilise such a metacomputing network and will report on this in future issues of CSAR Focus. Our experience confirms that the technical problems of metacomputing can be solved even in demanding environments with machines protected via firewall policies. In particular, PACX-MPI [8] developed by HLRS Stuttgart (partners in our global metacomputing work) has proved itself to be very able to operate in such conditions and we hope to be able to exploit the improved network connectivity of SuperJanet and GEANT. We welcome enquiries from CSAR users who wish to try metacomputing and especially who are interested in collaboration with colleagues who have resources on machines in the UK or outside.

Thus we are able to evaluate Grid middleware from a variety of projects and we are making it available to CSAR users. This will allow our users the ability to make informed judgements about the suitablility of such middleware and will enable them to be able to discriminate among the many claims being made for Grid computing. For our part, we are convinced that the Grid can bring many advantages to users of HPC facilities and look forward to working with you to demonstrate such benefits. The Grid is not the magic solution some people have seemed to claim, more it is a powerful tool which in the right circumstances can bring great benefits to computationally based scientific research.

References

[1] EUROGRID http://www.eurogrid.org
[2] Globus Project http://www.globus.org
[3] UNICORE http://www.unicore.org
[4] Open Grid Services Architecture (OGSA)
http://www.globus.org/ogsa
[5] Grid Interoperability Project (GRIP)
http://www.grid-interoperability.org
[6] UK Grid Support Centre (GSC)
http://www.grid-support.ac.uk
[7] SC2002, Supercomputing Conference
http://www.sc2002.org
[8] PACX-MPI, mpi library for metacomputing
http://www.hlrs.de/organization/pds/projects/pacx-mpi/

Globus and Unicore are available on the CSAR systems. If you would like to use either of these please do not hesitate to contact CSAR.