

# NeSC News

Issue 50, May 2007

www.nesc.ac.uk



## E-Science – enabled by grids by Malcolm Atkinson, e-Science Envoy

I write this editorial rather late, having just returned from a highly successful OGF20 in Manchester held jointly with the Second EGEE Users' Forum. With 925 delegates, it exceeded the previous record attendance of GGF5, held in Edinburgh in 2002. The change in maturity over that time is very significant.

Many readers may not appreciate the extent of the change achieved in five years so the planned editorial has been replaced with this personal view of grid progress that was impressed on me by the Manchester event. In 2002 the UK e-Science programme was full of excitement and promise, but the projects had only just started – there were no results to present and we were struggling to make any grids deliver sustainable operation. The GGF had only recently “gone global” and the excitement was launching groups and setting out to establish the foundations for architectures and data – there were no recommendations ready to use. Today there are plenty of examples of productive grids and these were in evidence in Manchester. There are now a bundle of standards recommendations being actively adopted, not only by academia but also by the software and service vendors. The standards for foundations for computational grids are well established, particularly through Job Submission Description

Language (JSDL), the OGSA Basic Execution Service and the OGSA-HPC profile. The JSDL working group received an award that was accepted by Stephen McGough – one of the many cases where UK e-Science has made a significant contribution. The Data Access and Integration Services working group, which started at GGF5 now has three standards recommendations and another underway.

The list could go on but standards alone are not enough. People need to implement systems and others need to use those systems to achieve business or research goals. This is where the sharpest difference between 2002 and 2007 is apparent. A small sample from the week's busy programme will illustrate this progress.

The packed Grids Means Business programme, organised by Grid Computing Now and led off by Dave Pearson of Oracle and Paul Strong of e-Bay, showed how vigorous and complex the commercial use of grids is becoming. This was reflected later in the Reference Model working sessions led by Dave Snelling of Fujitsu Labs Europe and Paul Strong. This is an effort to identify the pressing requirements that should drive the standards. A compelling case is the challenge of managing continuous services from data centres. Today's data centres

grow in scale and complexity. The cost of a mistake in their operation causing a loss of service can be massive – with literally millions of businesses being put in jeopardy. Yet new hardware and software must be rolled out without loss of service. Paul Strong likened it to changing the engines on a 747 while flying a full load of passengers at 30,000ft.

EGEE, presented at the joint opening of the two Forums on Wednesday by Ian Bird, is a remarkable example of an operational grid that is being used by an extensive community every day. The continuous operation of a global infrastructure of 200 sites running averages of 50,000 jobs per day last year is a great feat of engineering. All the more so, as it has been built by many independent organisations at a time when the software was changing rapidly. Nevertheless, Platform Computing remarked when we were discussing the need to train more people as professional grid managers, that among their several hundred customers there were some with grids larger than EGEE. Of course, those may be simpler to create and run because they are within one enterprise.

A new buzz of energetic R&D was brought to OGF by the exhibition floor. Many working grids were in evidence and being actively used. This floor was full of young software engineers and users at all times, with the UK much in evidence through the e-Science community, European projects and Grid Computing Now. If OGF can continue to co-locate with such activists and engage them in standards it will have a healthy future.

Peter Coveney gave an inspiring keynote. He demonstrated



## Registration Now Open

Registration is now open for The Sixth UK e-Science All Hands Meeting (AHM 2007) which will be held from 10-13th September 2007 at the East Midlands Conference Centre in Nottingham.

<http://www.allhands.org.uk/>



conclusively that he was able to do new science and identify new phenomena in complex materials by using grids. Most people find it hard to push the limits using one grid. But Peter is far more adventurous. He uses at least three different grids simultaneously (UK National Grid Service, the European DEISA grid, TeraGrid in the USA).

Peter made it clear that by combining such resources, unprecedented scales of simulation could reveal important properties of the many complex systems in our world. But he also identified criteria and software that would make such discoveries far more accessible. One aspect of these was consistent policies controlling their use – a job for the anticipated European Grid Initiative as well as the global community.

Such attempts to use multiple grids drive home the need for standards. The Grid Interoperability Now (GIN) working group at OGF directly addresses this goal. This builds on the rapid progress in agreeing how to describe and execute jobs that has leapt forward in the last year. Perhaps more important for users like Peter and his team, is the work of the Simple API for Grid Applications working group (SAGA), which has well advanced recommendations that should make the work of making applications work across multiple grids very much easier. OMII-UK are supporting implementation of this standard and the e-Science Institute theme on Distributed Programming Abstractions, led by Shantenu Jha, is mapping the future of efforts to make grid applications programming much easier.

The workshop on campus and community grids, led by Wolfgang Gentsch, director of D-Grid, and David Wallom from Oxford e-Research Centre presented many examples of grids that were delivering substantial increases in computing resources to their communities. Once again, the value of UK e-Science efforts was well-illustrated and current work

supported by OMII-UK and JISC will make such grids easier to run and use. Other workshops addressed some of the challenges that remain in grid technology, while another explored the potential use of e-science methods in the arts and humanities.

Tony Hey, corporate vice-president at Microsoft, in his keynote, issued a rallying cry to recognise a new category of grids – Social Grids. These build on the community action that is evident in Web 2.0 phenomena and seek to support data-centric science by using the established web mechanisms. Once again the difference between 2002 and now is dramatic. Today Yahoo offers information integration through pipes, Amazon offers a simple storage service (S3) and an elastic compute cloud (ec2). Tony's vision is that scientists will access the resources in the "cloud" of computing facilities provided by Amazon, Google, Microsoft and others using the convenient composition tools those providers are developing. It would have been hard to imagine such accessible services in 2002 – now, such services can be used by students, by researchers and by business. Of course, the cloud will depend on grids in the providers' data centres. Normally I carefully avoid a grid focus in these editorials and drive for recognition that any advances in methods, software and services can enable new science. But the powerful impact of OGF20 has brought the practical current contributions of grids to the fore.

The local hosts at Manchester University should be thanked for the provision of such a good environment for the event – the wireless network worked flawlessly and we could always find power. The OGF team had met the challenge of exchange rates and got the show on the road. That show was such a fine blend of interesting and productive meetings because of the outstanding efforts of the Programme Committee – Dave Berry and his team certainly deserve hearty thanks and congratulations.

## Full speed ahead for the Grid

Europe's biggest Grid event ever with more than 900 participants gathering for the 20th Open Grid Forum (OGF20) and the 2nd Enabling Grids for E-science (EGEE) User Forum. End-users, Grid experts and business representatives shared their experiences and thought about the future of the Grid, impressively demonstrating how far Grid technology has come.

"We are very proud that this event has brought together such a large and diverse part of the Grid community," said Massimo Lamanna from CERN, chair of the User Forum programme. "This really shows what a vibrant community we have and how important it is to organise events where scientists and grid experts can meet and discuss face-to-face."

Attendees of the joint event came from a wide range of technological and scientific areas already making use of the Grid, showcasing that the Grid has become a useful tool for many sciences. Organised in conjunction with OGF20, the User Forum not only helped to improve communication between users and between users and Grid experts, but also ensured that the needs of real Grid users feed into the development of key standards.

EGEE will hold its next conference, EGEE'07, in Budapest, Hungary, 1-5 October 2007 ([www.eu-egee.org/egee07](http://www.eu-egee.org/egee07)). Under the theme "building bridges" this conference will provide a platform to bring together users from different communities, Grid experts, projects, countries, and business to drive forward world-class Grid technologies.

Source: EGEE Press Release. For full release: [http://www.eu-egee.org/press\\_releases](http://www.eu-egee.org/press_releases)

# OGF 20: A Groundling's Perspective

by Iain Coleman

We live in a rapidly-developing electronic world, where people increasingly expect to simply reach out and grab the tools they need, when they need them. In this context, how can the grid community best meet the needs of an ever-greater diversity of users? This question came up in various forms, from high-level visions to presentations of real-world results, at the 20th Open Grid Forum (OGF), held jointly with Enabling Grids for e-Science (EGEE) second User Forum in Manchester on 7th – 11th May.

In a provocative keynote address, Tony Hey argued that the current standards-generating activities of OGF were being overtaken by other means of creating collaborative electronic infrastructures. He was not alone in this. David De Roure made similar points in an earlier presentation, drawing a close analogy between grids and Web 2.0 – although this argument was absent from the lively Grids Mean Business sessions, where presentations highlighted many successful deployments of grids in industry.

Web 2.0 is a generic term for technologies based on web standards that allow people to share, mix and annotate distributed information. This is much faster-moving than the structured development of scientific grids, and addresses users' needs almost by definition. Does this make grid standards obsolete? The counterargument was put by Alex Voss: Web 2.0 only allows you to do simple things simply. While the notion of simplicity is ill-defined, one can certainly identify areas where the grab-and-go approach is unlikely to provide a solution. Medical researchers, for example, need to show that their systems for data sharing are secure in some quite specific and stringent ways: a mashup of third-party software isn't likely to reassure an ethics board.

Tony Hey also warned that the work of OGF may be sidelined by some

commercial providers, despite the current industrial involvement with OGF. The likes of Google, Amazon, Yahoo and Microsoft are keenly aware of the competitive advantage of being first on the market with a product that meets users' needs, and they have the resources to develop impressive systems in-house. The major limitation of commercial systems is usually that interoperability ends at the company boundary: there are examples, such as MySpace blocking YouTube videos, of this placing a barrier in the way of the Web 2.0 ideal. There are, however, places where even this barrier is breaking down. Hey spoke of customer demand forcing authorisation systems like MS Passport and Google ID to work together, and it may be that where users demand interoperability the big commercial players will work to give it to them.

Nevertheless, the development of grids for science and business has produced some significant success stories. One scientific grid was particularly visible at the conference, thanks to the decision to combine OGF with the second EGEE User Forum. The Europe-based EGEE grid involves 91 partners in 32 countries, in Europe and beyond, and runs a range of scientific research applications. Its bread and butter is particle physics, with calculations for the Large Hadron Collider at CERN accounting for the bulk of its capacity, but it was clear from the posters and presentations around the conference centre that its capabilities are now being exploited for applications from radio astronomy to microbiology, with nuclear fusion research becoming increasingly prominent. Ian Bird, in his keynote speech, stressed the pragmatic engineering approach that underpins this scientific success. He advocated workable and appropriate standards based on established best practices in security, management and monitoring, warning against being too eager to fix standards in a technology that is not yet mature.

Bird also discussed the increasing importance of interoperability, not as an abstract activity but as a response to the demand from users for collaboration across different grids. Many of the gritty details of this work were later presented by EGEE's Erwin Laure, when he discussed the activities of the OGF-coordinated Grid Interoperability Now group in supporting astronomical research.

While production grids like EGEE are enabling a growing range of applications in the hard sciences, there is still little sign of their being exploited by archaeologists, musicians or mediaevalists. Will further developments such as OGF 20's Arts and Humanities workshop help to bring a more diverse user base on board? It is becoming clear that there is a demand for grid-based technologies in the arts and humanities, but those communities cannot simply piggy-back on the work that has already been done to support research in the natural sciences. This is not due to a lack of computer literacy – computers have been used in the humanities for nearly half a century – but because the needs of the arts and humanities are qualitatively different from most scientific requirements. As Mark Hedges pointed out, the hard sciences require fast access to large data sets, but things are different in the humanities. There, the data sets themselves may not be so large, but the data is complex, contextual, and often fuzzy and scattered. Preserving this complexity and context creates a requirement for seamless integration between highly structured repositories. This is an example of the tension that Sally Jane Norman described, between engineering that provides generic capabilities and the development of tools for specific purposes.

This user-oriented development approach was advocated by Peter Coveney in his keynote speech. He outlined examples of large-scale simulations in physics and biochemistry running across

multiple grids, and argued that these successes were due to lightweight, easy-to-use middleware and a focus on application-specific services. This illustrates that making things simple for users needn't mean a Web 2.0 approach. As Malcolm Atkinson pointed out, a lot of heavy engineering goes into making Microsoft Office sufficiently easy to use that 450 million people can work with it. The right design can make intrinsically complex computing accessible to almost any user.

Many of these points were summed up by Mark Linesch, who spoke in his keynote address of the gap between the vision of the grid as an invisible infrastructure and the over-complex reality. His message was that the mantra of "build it and they will come" doesn't work in reality. Successful grid technology needs a proven use case, whether commercial or scientific, and must be both less brittle and more pervasive than it is at present. Users shouldn't need to care that they're on a grid, they should simply receive the services and information they require. OGF's role, according to Linesch, is to understand and alleviate users' pain, to engage the community in an open forum and to find out what keeps them up at night.

Can the grid community deliver easy, reliable, near-invisible infrastructure that provides the distributed computing applications that users want? That's the challenge that emerged from this OGF meeting. Increased engagement between users and developers, and maintaining a pragmatic focus on delivering the goods, could see grid computing sitting alongside Web 2.0 and proprietary infrastructures as a system that just works, performing a valuable role in underpinning the foundations of intellectual and commercial life.

## The ES-LoA Project: Levels of authentication Assurance Survey

(The E-infrastructure Security: authentication Levels of Assurance)

The UK JISC would like to consult with service and identity providers and Grid/e-Science community for their views on using a risk-based authentication and authorisation scheme. Such a scheme, built on top of open standards based middleware such as Shibboleth, would take into account the level of confidence in users' identification through their identity and attribute assertions when deciding whether or not to permit access to a particular resource.

The principle objectives of the ES-LoA project are as follows:

- to raise the community's awareness with regard to authentication Levels of Assurance (LoA) as one of the factors that can be used to quantify the degree of protection for resources with varying levels of sensitivity in federated environments,
- to investigate potential applications of LoA to various types of resources, including grid/e-Science resources, library resources and e-learning resources,
- to build community consensus in using the appropriate definition of LoA reflecting the trustworthiness of a resource's authentication mechanism, and
- to raise service providers' interests in deploying a technology that can help them to achieve LoA linked fine-grained access control.

Please take the time to answer the ES-LoA Survey questionnaire, available from <http://www.es-loa.org/output>, by Friday 8th June 2007, and email it to [es-loa@manchester.ac.uk](mailto:es-loa@manchester.ac.uk).

## JISC sponsors 'Outstanding ICT Initiative of the Year' Award

Call to institutions to 'bring to wider attention examples of innovation in the use of ICT'

JISC is sponsoring an award this year which will showcase the most innovative and potentially far-reaching ICT initiatives across the UK. The award, one of the Times Higher's 2007 Awards, will 'recognise and reward an institutional ICT initiative which has demonstrated an innovative and strategic use of ICT in support of the goals of that institution.'

The award, for which all higher education institutions, teams or departments in the UK are eligible, is now open for entries until June 29th. The award will be presented at an event on the November 29th.

Speaking to the Times Higher this week, Professor David Baker, Principal of the College of St Mark and St John and Chair of JISC's Content Services committee (JCS) and one of the judges of the award said: 'There is a need for good practice and examples of innovation to be more widely shared. With the THES awards having quickly become a showcase for some of the best and most exciting work being done in higher education, we hope this ICT award will likewise bring to wider attention examples of innovative and far-reaching uses of ICT.'

For further information on the Outstanding ICT Initiative of the Year Award, and to enter, please go to: [www.thes.co.uk/Awards/2007/](http://www.thes.co.uk/Awards/2007/)

See also: [www.jisc.ac.uk/THESAwards](http://www.jisc.ac.uk/THESAwards)

# Database Preservation Workshop

by Iain Coleman

How do databases evolve in time, and how can we preserve data such that we can make use of it afterwards? The Database Preservation Workshop, held at the e-Science Institute on 23rd March, brought together experts from academia and industry to address these questions.

A database is a continually evolving environment, but the evolution owes more to the tinkering of an intelligent designer than to natural selection. An example is the common case of underused fields being recycled for other data, as the user's needs diverge from the original designer's expectations. George Papastefanatos (National Technical University of Athens) spoke of the need to understand the laws that govern the evolution of a database, as well as the importance of designing databases to allow smooth evolution. Along with several speakers at this workshop, he emphasised planning for change: when designing a database, support for data preservation must be built in right from the start.

There is increasing pressure to archive databases in the public sector, as Stefan Brandl (CSP GmbH & Co) pointed out in his presentation of the Chronos tool which enables the archiving of time-slices as the database evolves. One of the key points made by Kevin Ashley (ULCC) is that, when you want to preserve a database as a record of past activity, it is important not to change the past. This includes preserving errors, as these are informative in themselves. It also means finding ways to deal with the different varieties of missing entries that are a persistent feature of social survey databases. Another issue of particular importance to government work is the need to be able to restrict or redact databases, or entries within a database. These problems are not unusual, as Ashley pointed out, yet everyone seems to use homegrown software. Here there may be lessons to be learned from manufacturing

industry. An aircraft, for example, may have a working life of thirty years, and so the digital information about that aircraft has to last for at least as long. The methods that industry has developed to preserve information were discussed by Norman Swindells (Ferroday Ltd). Rather than have different programs representing the same data in different ways – akin to everyone making their own individual nuts and bolts – international standards for information models, such as the ISO standards for information engineering, can provide a formal specification for data with explicit rules for its interpretation. This allows information to be recorded and preserved in a consistent, unambiguous way that greatly assists in its long-term preservation and usability.

If such standards had been in place for radiobiology experiments in the forties, fifties and sixties, Jonathan Bard (Edinburgh) might have been saved a good deal of work. At that time, a large amount of research was performed on the biological effects of radiation, involving experiments on a range of animals from mice to monkeys. The data recording methods were inconsistent, and most of the scientists involved are now dead, creating a real challenge for scientists who are trying to turn these collections of results into a usable database. Redoing the experiments is not a realistic possibility: even disregarding the cost, many of these studies would now be considered wholly unethical. Making use of the legacy data is the only option.

And that's something that doesn't happen as often as it should, according to Michael Lesk (Rutgers). It all comes down to the incentive structures which channel people's work. The rewards go to the person who writes a new program, not to the person who finds out that it has already been written. It's cheaper and easier to use material already in the archives,

but those responsible for hiring and firing all too often regard this as a second-rate effort. If we want people to archive data properly, they must be trained, funded and rewarded for doing so. Just as importantly, people need to be given the incentive to reuse archived material. As well as rewarding the collection and mining of old data, we need to start regarding data preservation as being an essential part of data collection, just as we view publication and citation as an essential part of science. This means establishing a new public ethic for responsible scientists. If data preservation is ultimately a people problem, then the solutions will be at least as much social as technological.

Slides from this workshop can be downloaded from <http://www.nesc.ac.uk/esi/events/763/>

# Lighting the Blue Touchpaper- ESLEA Closing Conference

by Iain Coleman

## Open call for NCeSS research Nodes

The ESRC has issued a call for NCeSS research Nodes, which is open to existing Nodes and new applicants.

It is anticipated that up to eight Nodes will be funded. The average budget per Node is expected to be of the order of £1,060k over three years, of which the ESRC will meet £850k, representing 80% of the full economic cost, though there is scope for variation around this figure.

There will be a Briefing Meeting about the call from 2pm to 4.30pm on Friday 25 May at the Kohn Centre, Royal Society, 6-9 Carlton House Terrace, London, SW1 5AG. The meeting is free but registration beforehand is essential. Please complete the registration form found at <http://www.ncess.ac.uk/events/> to confirm your attendance

Deadline: full proposals must be submitted by 4pm on Tuesday 3 July 2007

Please see NCeSS (<http://www.ncess.ac.uk>) and ESRC (<http://www.esrc.ac.uk>) websites.

The key to progress in e-Science isn't just increasing bandwidth: it's improving the quality of the data transfer across networks. That was the recurring theme at "Lighting the Blue Touchpaper", the closing conference for the ESLEA project, held at Edinburgh's George Hotel on 26th – 28th March.

ESLEA (Exploitation of Switched Lightpaths for e-Science Applications) was a two-year project to demonstrate the benefits of optical networks, or lightpaths, to the UK e-Science community. In science fields from astronomy to particle physics to e-Health, the high-speed UKLight optical network has expanded the boundaries of the possible.

Increasing bandwidth allows a host of new scientific applications, from numerical solutions of Einstein's equations to modelling fluid flow in nappies. Even just being able to transmit large data files can open up new ways of working. There is, however, a bandwidth catch-22. Current bandwidth is often underused because people won't try things that they know won't work – this leads engineers to conclude that there's no need to increase the bandwidth, as there is still spare capacity. It's only when a substantial increase occurs, such as that offered by lightpath technology, that the latent demand is revealed.

One such advance, made possible by UKLight, is in very large-scale simulations of physical systems. For example, large-scale simulations of clay composites eliminate boundary effects, and reveal interesting properties such as thermal undulations to emerge that are not seen in small-scale simulations. The real system is so small in size that it's hard to experiment on it in the laboratory: simulations built up from individual atoms are now approaching size of real clay platelet. That's just one of many

simulations that achieve significant performance improvements by moving to an optical network. Indeed, grand challenge problems such as LB3D, a three-dimensional fluid mechanics simulation using lattice Boltzmann techniques, would be impossible without the combination of high performance computing and optical networks.

It's not just the hard sciences that need to handle a rapidly increasing volume of data. Archaeology, sitting at the interface between natural science and the humanities, generates a tremendous amount of complex spatial and numerical data. This comes from such diverse sources as three-dimensional scans of fragile ancient artefacts and detailed aerial measurements of landscape features. A striking example of what can be achieved is the visualisation of North Sea bed, based on data from 3-D seismic surveys carried out by the oil industry. This region used to be inhabited until around eight thousand years ago, before the sea levels rose and Britain was cut off from Europe. Archaeologists can identify likely areas of human habitation, and it is likely that the undersea environment has preserved an abundance of archaeological material in these ancient sites.

Transmitting and analysing large amounts of data is valuable, but it isn't always enough. For many applications, the high quality of service that lightpaths can provide is at least as important as the high bandwidth. In practice, high quality means minimising latency, the delay between transmission and reception, and reducing or eliminating packet loss, when chunks of data go AWOL on the network. Controlling these problems allows new kinds of scientific exploration. One example is the study of RNA translocation through protein pores, a fundamental biological process. This can be

modelled with grid infrastructure, allowing interactive computational steering to determine optimal values of the key parameters. Such interactive simulations are very sensitive to latency and packet loss, and require the quality of service that lightpaths can provide.

High quality data transfer can also generate new possibilities from established techniques. On such case is Very Long Baseline Interferometry, or VLBI. This is a method for vastly increasing the ability of astronomical telescopes to resolve small details by combining simultaneous observations from two or more instruments thousands of kilometres apart. This first became possible with the advent of atomic clocks and electronic data recording: observations from radio telescopes and their associated timing information would be recorded onto tapes and shipped to a central analysis centre where the data would be correlated, and highly detailed images produced. These days, of course, the data is transmitted over the network, but lightpaths can do more than save on air travel costs. If the data can be integrated and analysed quickly enough, the telescopes can go on to do follow-up studies of unpredictable, transient events like microquasar flares that have shown up in the first set of observations, in the brief window of opportunity before the activity ceases. Packet loss isn't such an important consideration here – a correlator originally designed to work on electronic tapes has to be able to cope with some missing data – but point-to-point lightpath connections with no problems of network congestion allow the telescope network to operate in new ways, obtaining new scientific results.

The most demanding application yet, as far as quality of service is concerned, is not in science but in music. Musicians are increasingly interested in using e-Science technologies for remote collaboration, but enabling professional musicians to work

together remotely in real time is a major challenge for networks. Musicians are sensitive to temporal shifts of the order of milliseconds, and this is the level of accuracy that they require from networks, far beyond the capabilities of the current internet. Ideally, musicians want near-zero latency and near-zero jitter. The latter can be achieved, but the laws of physics are such that some degree of latency will always exist: perhaps the best that can be done is to make it predictable. Remote teaching of music is even more demanding. This requires surround sound, which adds the need for a few hundred Mbps bandwidth to these already stringent requirements.

Adding video will be even more demanding. Here, however, help may be at hand from Hollywood. The commercial movie industry is now driving the development of ultra-high definition 4000 x 2000 pixel video, with latency of less than 0.1 seconds now possible thanks to optical networks. When this technology becomes readily available, it may prove to be a valuable tool for collaboration, replacing the grainy, jittery images of AccessGrid with startlingly realistic images.

While commercial entertainment pushes these new technologies, artists are starting to adapt the existing e-Science infrastructure for their own creative purposes. Unfortunately, their art is often constrained because they don't realise there are possibilities beyond readily available commercial technologies like Flash. Some artists have begun using AccessGrid in installations or performance, but this is often beyond the limited resources of a new media centre. The barriers that prevent artists from being exposed to the full range of technology deny them the opportunity to play with it and generate ideas. Play may sound frivolous, but it is to an artist what experiment is to a scientist. Artists need time and space to interact with technology, to find out how to use it, and to develop their

own conventions which might differ radically from what technology experts might imagine. As lightpath technologies become more widely used, it will be increasingly important to ensure that they are available to everyone who can make use of the opportunities they provide.

More information about this workshop is available at <http://www.nesc.ac.uk/esi/events/748/>

## John Wood Interview

John Wood, Chief Executive of the former Council for the Central Laboratory for the Research Councils, has recently taken on two important roles for JISC as chair of the JISC Support of Research committee and chair of the JISC Scholarly Communications Group.

In an interview in the latest issue of JISC Inform, Facing the data deluge, he outlines his vision for the work of the two groups, in particular the important issues around data management that JISC, the research community and institutions urgently need to tackle together.

The interview is available at [http://www.jisc.ac.uk/publications/publications/pub\\_inform17.aspx](http://www.jisc.ac.uk/publications/publications/pub_inform17.aspx)

## Digital repositories conference: Dealing with the digital deluge

Repositories are important for universities and colleges in helping to capture, manage, and share institutional assets as a part of their information strategy. A digital repository can hold a wide range of materials for a variety of purposes and users. It can support learning, research and administrative processes.

This event will showcase a range of research and development work in this area, and demonstrate how it can be practically applied by those working within institutions. It will put this work in the context of a vision for a UK repository network, whereby managing and sharing resources is straightforward and effective, and will put UK work in a global perspective.

Professor James Drummond Bone Vice Chancellor of the University of Liverpool and President of Universities UK, will be giving a conference keynote speech and there will be a number of strands of activity running throughout the conference covering every aspect of the repositories landscape along with opportunities to talk to experts involved with JISC work, demonstration and posters. Strands within the programme will include:

- Research data
- Sharing digital material for e-learning
- e-Theses
- Research papers
- Images
- Repositories and Preservation
- Repositories in a Legal Context

This two day event will be of interest to researchers, teachers, advisors and managers within universities and colleges. To find out more and register please go to: [http://www.jisc.ac.uk/events/2007/06/repositories\\_conference.aspx](http://www.jisc.ac.uk/events/2007/06/repositories_conference.aspx)

## Joint EGEE and SEE-GRID Summer School on Grid Application Support Budapest, Hungary, 25-30 June, 2007.

[www.egee.hu/grid07](http://www.egee.hu/grid07)

The Joint EGEE and SEE-GRID Summer School on Grid Application Support aims at introducing EGEE – SEE-GRID grid technologies to potential user communities and studying and practicing application development methods on the EGEE grid. Existing EGEE and SEE-GRID users can advance their knowledge on recent tools available for application developers and end users. Experienced lecturers from the EGEE and SEE-GRID projects will deliver talks and hands on exercises on grid concepts, services, application development methods and tools. During dedicated sessions the attendees – with the help of our grid experts – can port their own applications onto production grid infrastructures.

The summer school is open for anyone interested in the use of EGEE and SEE GRID tools, technologies and infrastructures. The primary focus of the school is on how end-users can apply the EGEE middleware and related software services to define and operate distributed applications on top of inter-organizational IT infrastructures. The school provides a perfect occasion to learn the usage of grid systems without the irrelevant low-level technical details. During dedicated sessions the attendees' own legacy applications can be ported onto EGEE and SEE-GRID.

Representatives of the EGEE and SEE-GRID application developer communities will

assist participants to understand the process of application gridification, to identify concepts and services that are relevant to their specific problem and to implement and install gridified applications on a production infrastructure. Besides getting first-hand grid user experience the school will be an excellent forum to meet the representatives of grid user communities, grid infrastructure providers and grid tool developers. The summer school will be conducted in English, no simultaneous translation will be available.

Budapest, one of the most beautiful historical capitals in Europe, situated on the banks of the Danube, will host the summer school. The architectural beauty with its monuments from neo-classic through baroque to eclectic art nouveau is prevalent in the city. Budapest is also world famous for its artistic abundance of concerts, operas, recitals, galas and exhibitions. The warm hospitality of people, excellent food and wine reliable and frequent public transportation, vivid cultural life, rich museums attract millions of visitors every year. Visitor friendly visa policy, value-for-the-price services and goods, and pleasant climate make Hungary one of the most popular meeting venues worldwide.

For additional information on grid computing, school program, registration and hotel reservation please visit [www.egee.hu/grid07](http://www.egee.hu/grid07)

## Grid Computing Now! Web seminar: The Semantic Web in Industry

The next GCN! Web seminar will be held on May 24, 2007, at 2.30pm.

The World Wide Web began as a tool for humans to access documents at known locations. Search engines have eased the task of finding relevant information but they are still based on searching for matching words, with no knowledge of what the search terms actually mean. The result can be pages of results and a tedious process of finding the one that matters to you.

Web services provide means for programs and agents to call remote services over the web. This raises more complex questions of service discovery and integration.

The Semantic Web addresses these problems. At its heart is a language for adding semantic annotations to web pages and web services. Search services can use these to find the resource the client wants. Integration services aid the task of linking separate services to address the problem at hand.

This webinar will describe the principles of the semantic web and show how it can already be applied to real industry use cases. Speakers will include John Davies, Head of Next Generation Web Research, BT Group Chief Technology Office and Paul Walsh, Segala CEO, BIMA Chair.

You will be able to take part online from your desk by watching live video, viewing slides and asking questions through an interactive webpage.

Registration details will be available soon on:  
[www.gridcomputingnow.org](http://www.gridcomputingnow.org).

## Call for Participation – High Performance Distributed Computing 2007 (HPDC2007)

Registration for the IEEE/ACM Conference on High Performance Distributed Computing 2007 and associated workshops is now open at <http://www.isi.edu/hpdc2007/>.

The meeting this year will be held in Monterey Bay California, at the Hyatt Regency, Monterey Bay.

As in previous years, HPDC will feature a high-quality single-track program, a hot topics session and an industry panel. This year, we are privileged to have keynotes be presented by:

- Prof. Dan Atkins, Director, Office of Cyberinfrastructure, National Science Foundation
- Dr. Urs Hölzle, Senior Vice President, Google Fellow, Google Corp.
- Prof. Satoshi Matsuoka, Professor, Tokyo Institute of Technology

This years workshops include:

- \* Grid Monitoring Workshop
- \* Workflows in Support of Large-Scale Science (WORKS07)
- \* Joint EGEE and OSG Workshop on Data Handling in Production Grids
- \* Challenges of Large Applications in Distributed Environment (CLADE 2007)
- \* Second Provenance Challenge Workshop
- \* Globus Scientific Users Workshop
- \* 2nd Workshop on the Use of P2P, GRID and Agents for the Development of Content Networks (UPGRADE-CN)
- \* Service-Oriented Computing Performance: Aspects, Issues, and Approaches

## 5th International GridKa School” - Grid Computing and e-Science”

The school will be held at Forschungszentrum Karlsruhe, Germany from September 10-14, 2007

GridKa School is a one-week summer school covering various topics of grid computing and e-science, this year focusing on Applied Grid Computing. The topics span the European EGEE project, the German D-Grid initiative, sciences such as High Energy Physics and Industry. About 50% of the school consist of hands-on tutorials, giving participants the possibility to gain actual expertise on different kinds of middleware, procedures and applications. Because these tutorials are offered in parallel tracks, GridKa school practicals cover topics well suited for beginners, experienced users and grid administrators.

The high grade talks and presentations will give an up to date overview on various grid subjects. Applicants from all scientific and industry disciplines are welcome to apply online at <http://www.fzk.de/gks07>.

Registration is now open!

Registrations before July 31 can take advantage of an early bird discount. The organising team reserves the right to limit the number of participants, if necessary.

For further details and registration please visit:  
<http://www.fzk.de/gks07>

## CALL FOR PAPERS

### Fourth International Conference on Life Science Grids (LSGrid2007)

6-7th September 2007,  
National e-Science Centre,  
University of Glasgow, Scotland

#### Paper Submission

Authors are invited to submit original and unpublished work. Papers should not exceed 10 single-spaced pages on A4 paper size, using at least 1 inch margins and 12-point font. Authors should submit a PDF or PostScript file that will print on a PostScript printer. Electronic submission through the symposium website ([www.lsgrid.org/2007](http://www.lsgrid.org/2007)) is strongly encouraged. Submission implies the willingness of at least one of the authors to register and present the paper.

#### Important Dates

Submission deadline of Abstracts for posters/demonstration (15th June 2007) (800 words)

Submission deadline of Papers (29th June 2007) (up to 10 pages).  
Notification of Acceptance (20th July 2007)

Submission of Camera Ready Version of Paper (3rd August 2007)

Note that this conference will take place the week after the Braemar Highland Games (<http://www.braemargathering.org/>) in Scotland and the week before the UK e-Science All Hands Meeting (<http://www.allhands.org.uk/>) in Nottingham, for those wishing to have more than one reason to be in the UK at this time!

<http://www.nesc.ac.uk/events/lsgrid2007/>

## International Conference on Information Society (i-Society 2007)

October 7–11, 2007, Merrillville, Indiana, USA

The International Conference on Information Society (i-Society 2007) is a global knowledge-enriched collaborative effort that has its roots from both academia and industry. The conference covers a wide spectrum of topics that relate to information society, which includes technical and non-technical research areas.

The mission of i-Society 2007 conference is to provide opportunities for collaboration of professionals and researchers to share existing and generate new knowledge in the field of information society.

The conference encapsulates the concept of interdisciplinary science that studies the societal and technological dimensions of knowledge evolution in digital society. The i-Society bridges the gap between academia and industry with regards to research collaboration and awareness of current development in secure information management in the digital society.

All papers accepted for the conference, including special sessions papers, will be published in the proceedings.

#### Important dates

Paper submissions  
31 May 2007 – paper submission  
15 June 2007 – notification of acceptance  
1 July 2007 – final paper submission  
Poster submissions  
15 May 2007 – poster submission  
31 May 2007 – notification of acceptance  
10 June 2007 – final poster submission

## 3rd International Digital Curation Conference

The UK Digital Curation Centre (DCC), the US National Science Foundation (NSF) and the Coalition for Networked Information (CNI) are pleased to jointly announce the 3rd International Digital Curation Conference to be held on Wednesday 12th – Thursday 13th December 2007 at the Renaissance Washington Hotel in Washington DC, USA.

Entitled “Curating our Digital Scientific Heritage: a Global Collaborative Challenge” the conference will focus on emerging strategy, policy implementation, leading-edge research and practitioner experience, and will comprise a mix of peer-reviewed papers, invited presentations and keynote international speakers.

Further details and a Call for Papers will be published shortly at

<http://www.dcc.ac.uk/events/dcc-2007/>

The event will follow on from the Fall 2007 CNI Task Force meeting which will be held on Monday 10th – Tuesday 11th December, also at the Renaissance Washington Hotel, Washington DC.

More information about the DCC can be found at: <http://www.dcc.ac.uk>

## Registration for the 3rd NGS User Forum Now Open

NGS 3rd User Forum and Training Event – 19th and 20th June 2007

The National Grid Service has now been in full production since September 2004 and in October 2006 entered phase 2 of its operation.

The third NGS user forum is meeting to bring together existing and potential NGS users and providers, to share experiences, understand future plans and provide an opportunity to influence the development of the NGS provision through the next few years. With users from a diverse range of academic disciplines and a growing number of resource providers, the NGS is constantly looking towards the future to better understand the service requirements of users and the applications required.

The event to be held in Oxford will also mark the official launch of the new NGS-2 Compute and Storage Clusters. An open drop-in session on the afternoon of the 19th June will enable users to informally speak to technical representatives from the NGS

Collocated with this user forum will be a NGS training day on the 20th June. This training day gives a practically-oriented introduction to the services that have recently been deployed by the NGS.

For further information, the detailed agenda and to register please visit <http://www.ngs.ac.uk/event.html>

## BELIEF-EELA e-Infrastructure Conference

Connecting the Knowledge of Today for the Value of Tomorrow  
25-28 June 2007, Rio de Janeiro, Brazil

The BELIEF-EELA e-Infrastructure Conference sees the teaming up of two EC-funded projects to platform what e-infrastructures can do in diverse fields, driving forward a global vision for adoption. Participation at this conference will ensure this vision can move forward and pave the way for the future sustainability of e-Infrastructures. With special focus on Latin America-EU collaboration, the event will bring together experts to exchange knowledge, experiences and best practices.

The programme features a broad spectrum of themes of mutual interest.

Collaboration between industry and research on e-Infrastructures will help accelerate the development of innovative technologies, facilitate technology transfer and lead to solutions that solve real business problems. The dedicated Business Session explores new opportunities for businesses, presenting case studies promoting the commercial benefits of e-Infrastructures.

To find out more about how your company, initiative and project can be part this conference, please visit and register at:  
<http://www.belief-eela.org/>

## Director of OMII-UK

The Open Middleware Infrastructure Institute UK (OMII-UK) is a unique opportunity to sustain the top technical outputs of the UK's e-Science programme.

We are a multi-site team with members at Southampton, Manchester and Edinburgh universities. We're dedicated to enabling scientists and researchers to use digital resources through open source software. Over the past 5 years we have built up a significant portfolio of software products and support services. You will represent OMII-UK nationally and internationally, promoting the use of OMII-UK products and developing future funding.

You will be able to devise and implement a strategic plan for the sustainability of the Institute, will successfully communicate with a broad range of stakeholders and users, and will be able to coordinate a highly skilled distributed team. You will be based at one of the three sites.

Information about the Institute can be found at <http://www.omii.ac.uk/>

The post is for 2 years in the first instance and would be suitable for a secondment arrangement.

For more information contact Prof David De Roure ([dder@ecs.soton.ac.uk](mailto:dder@ecs.soton.ac.uk)), Prof Malcolm Atkinson ([mpa@nesc.ac.uk](mailto:mpa@nesc.ac.uk)) or Prof Carole Goble ([carole.goble@manchester.ac.uk](mailto:carole.goble@manchester.ac.uk)).

To apply on-line for this position, please see <http://www.jobs.ac.uk/jobfiles/RH638.html>

The closing date for this position is 11 June 2007 at 12 p.m.

Please quote reference number 1086-07-E on all correspondence.

## Forthcoming Events Timetable

### May

30 - 1 June	Distributed Programming Abstractions, Models and Infrastructure	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/esi.html">http://www.nesc.ac.uk/esi/esi.html</a>
-------------	---	---------------------	---

### June

14-15 June	Support for e-Research: Filling the Library Skills Gap	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/770/">http://www.nesc.ac.uk/esi/events/770/</a>
18 June	Methods and Technologies for Enabling Virtual Research Communities	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/785/">http://www.nesc.ac.uk/esi/events/785/</a>
19 June	Ontologies and Semantic Interoperability for Humanities Data	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/773/">http://www.nesc.ac.uk/esi/events/773/</a>
20 June	Collaborative Text Editing	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/774/">http://www.nesc.ac.uk/esi/events/774/</a>

### July

2 July	Grid Enabling Humanities Datasets	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/786/">http://www.nesc.ac.uk/esi/events/786/</a>
6 July	e-Science and Performance	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/787/">http://www.nesc.ac.uk/esi/events/787/</a>
12 July	Semantic Integration Workshop	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/756/">http://www.nesc.ac.uk/esi/events/756/</a>
23-24 July	AHRC ICT Methods Network Workshop on Space and Time: Methods of Geospatial Computing for Mapping the Past	e-Science Institute	<a href="http://www.nesc.ac.uk/esi/events/772/">http://www.nesc.ac.uk/esi/events/772/</a>

### September

10 - 13	<b>UK e-Science All Hands Meeting - Registration Now Open!</b>	East Midlands Conference Centre, Nottingham	<a href="http://www.allhands.org.uk/">http://www.allhands.org.uk/</a>
---------	--	---	---



## Registration Now Open

Registration is now open for The Sixth UK e-Science All Hands Meeting (AHM 2007) which will be held from 10-13th September 2007 at the East Midlands Conference Centre in Nottingham.

<http://www.allhands.org.uk/>

**Welcome to.....**John Lyons who joined the Middleware Team this month as a Research Systems Consultant.

If you would like to hold an e-Science event at the e-Science Institute, please contact:  
Conference Administrator, National e-Science Centre, 15 South College Street, Edinburgh, EH8 9AA  
Tel: 0131 650 9833 / Fax: 0131 650 9819 / Email:

The NeSC Newsletter produced by:  
Alison McCall and Jennifer Hurst, email [alison@nesc.ac.uk](mailto:alison@nesc.ac.uk),  
Telephone 0131 651 4783  
The deadline for the May Newsletter is: 29 May 2007