NW-GRID Project Number N0000298
Evaluation Report - Revised
August 2009

Executive Summary

The aim of NW-GRID was to establish, for the region, a world-class activity in the deployment and demonstration of Grid middleware technologies (the software that glues together shared data and computing resources) and to exploit the capabilities of the Grid in leading edge computational science and engineering applications.

The objectives of the project included:

1. The establishment of a physical Grid infrastructure with large-scale computing systems connected by a high-speed network located at the four lead sites enabling access to computing resources for throughput and capability computations and visualisation.
2. The growth of a portfolio of projects aimed at developing and maintaining a competitive edge in research into the fundamental software (middleware) technologies that will enable efficient and effective exploitation of the Grid.
3. The development of a portfolio of academic and industrial applications that exploit the infrastructure in particular in simulation, data-intensive, collaborative and real-time monitoring and control applications of importance to the clusters supported by the NWDA.
4. Distributed collaborative working in specific projects and ready access to the best researchers in the North West England;
5. Demonstration of the middleware and applications developed at annual workshops and active promotion and marketing of the NW-GRID projects and infrastructure.
6. Development of a range of flexible training material to enhance the North West’s skills in Grid technologies and their exploitation.
7. Collaboration with overseas Grid activities in Europe, USA and Japan to ensure that the North West becomes a key partner in the emerging global Grid.
8. Development of a self-funding activity supporting the refresh and growth of the hardware infrastructure and of order 100 permanent high-technology jobs.
9. Additional leveraged income to the total of £15 Million to come from leveraged jobs, additional leveraged capital expenditure and outside contracts.

To achieve its aims and objectives the project was organised into 6 work packages, with associated milestones and deliverables to which each of the partners contributed.

WP 1 - Hardware Infrastructure (Objective 1, Deliverables: Procurement, installation and integration of Grid systems)
WP 2 - Operational Issues and Science Testbeds (Objective 1, Deliverables: Testing and demonstrations of systems capabilities)
WP 3 - Grid Middleware Evaluation and Extension (Objective 2, 7 and 8, Deliverables: Evaluation and development of middleware, generation of leveraged income and jobs)
WP 4 - Applications (Objectives 3, 4 and 8, Deliverables: Portfolio of application projects and through generation of leveraged income and jobs)
WP 5 - Knowledge Exchange (Objectives 5, 6 and 8, Deliverables: Workshops, courses)
WP 6 - Project Management

The outcomes expected from the project can be summarised as "the collaborative exploitation of a sustainable computational infrastructure that supports and enables future collaboration, scientific and industrial research to the benefit of the North West England economy and thereby supporting jobs in the region".

Project achievements include:

a) The establishment of a world-class physical Grid infrastructure at the four partner sites with a core aggregate computing capability making it arguably the largest regional Grid in the UK. We have augmented this with an additional 3 fold capacity from local computing resource at the partner sites. A private network connects the core sites for dedicated high-speed access.
b) The establishment of a long-lasting collaboration on among the partner sites plus other regional institutions such as the Proudman Oceanographic Laboratory becoming new partners.
c) The integration of NW-GRID facilities into local "campus" grids and the development of virtual research environments to make it easier for users to access NW-GRID facilities.
d) Enabling research of an international standing in areas as diverse as bipedal gait analysis (see the BBC News Article 21 August 2007, *T. rex 'would out run footballer'*), nano-materials modelling (see New Scientist News article 11 June 2008, *'Electron turbine' could print designer molecules)*, reducing pollution from organic contaminants on mineral surfaces or carbohydrate-ligand binding studies of interest to the pharmaceutical industry.
e) Support for commercially funded research projects (e.g. ship-airwake studies funded by SEA Ltd).
f) The generation of circa £15M in leveraged income by 30 September 2008.
g) The generation of c.100 jobs by 30 September 2008.

These achievements are largely in line with expectations at the start of the project. One unforeseen benefit was that the size of the core facilities allowed us to provide users with exclusive access to considerable compute resources for several days and this concentrated access allowed users to produce results in a more timely fashion than anticipated. This novel operational mode was named "science testbed".

There were several other, equally important, achievements that were not envisaged at the start. These include: the addition of new partner sites (Proudman Oceanographic Laboratory and University of Central Lancashire, with discussions going on with University of Keele), partnership in the National Grid Service and the Hartree Centre proposal. Additional resources from the various partners have been added to NW-GRID and campus grids, particularly at Liverpool and Manchester, have made access to NW-GRID facilities easier. Recently NW-GRID has spearheaded the formation of the Streamline Computing User Group (SCUG) to facilitate discussion between Streamline Computing and its larger customer sites, amongst which is NW-GRID.

The core NW-GRID hardware, funded by NWDA, continues to be highly relevant to the needs of regional researchers. There is considerable interest from new and existing projects to exploit these resources, particularly the new systems installed at Lancaster and Liverpool. The project Technical Board and Operations Board continue to meet on a bi-weekly basis via Access Grid. Project members are confident that we will continue to demonstrate innovative and exciting projects. NW-GRID has been accepted as a partner in the National Grid Service (NGS) with a number of early adopters and the majority of their users having access in early 2009. There is also a growing realization at the partner Universities that NW-GRID forms an excellent infrastructure to underpin the recently announced Hartree Centre (funding earmarked by DIUS) and the N8 Molecular Engineering Translational Research Centre (METRC) both of which will benefit from the infrastructure and skills available.

Project Overview

The NW-GRID project was funded by the NWDA, over the period April 2004 to March 2008, in order to provide North West England with a world-class activity in the deployment and demonstration of Grid middleware technologies and to exploit the capabilities of the Grid in leading edge computational science and engineering applications. The project had a budget of just under £5M, with approximately £1.6M ear-marked for large equipment purchases. At a peak aggregated performance of 16.5 Tflop/s this places the NW-GRID among the world's top 400 most powerful computer systems and represents excellent value for money.

The project directly funded nearly 40 staff-years of effort, with approximately 16 staff-years relating to applications development, 13 staff-years for system maintenance and the establishment of science testbeds and 6.5 staff-years for Grid middleware development and deployment. The remaining funded effort went towards procurement, and knowledge exchange as well as project management. In addition to direct NWDA support, the project leveraged an additional income of approximately £15M (from all sources) and created in the region of 100 full time jobs in North West England.

When discussions about NW-GRID first started in 2003 with NWDA, grids were perceived as the best way in which to share resources across geographically disparate sites so that users could have seamless access to world-class facilities while system support and run time costs could be shared easily among partners. In addition, grid computing was relatively new and it was felt that NW-GRID would provide a good vehicle to explore and enhance our knowledge of the area. This was particularly relevant for commercial users who might be looking to establish their own grid infrastructures.

Whilst much has changed over the intervening 5 years, it is still the case that providing regional users with seamless access to world-class computing facilities remains an important objective of NW-GRID. It is also the case that locating computing resource at each partner site is an effective way to share system administrative and running costs. Moreover, this approach has allowed complementary expertise to be developed at each site so that the collective NW-GRID system support is considerably more than the sum of its parts.

The project remains in a strong position for current and future work. The Lancaster upgrade of 536 Barcelona cores is just coming on stream to complement its existing 192-core system, both funded fully by NWDA. The Liverpool offering consists of three systems – one with 656 cores for serial and coarse-grain parallel jobs that was 100% NWDA funded, one system with 632 cores for large and fine-grain parallel jobs that was funded by NWDA, Liverpool University and research grants and the third system with 164 cores that was donated by the Proudman Oceanographic Laboratory with NWDA funding to re-commission the system. During summer 2008, a new version of the Liverpool campus grid software was deployed to make it easier for users to access any of the NW-GRID systems as well as systems of the NGS. A campus grid was recently established at Manchester to enable access to its 100-core system as well as "spare" Windows PC capacity on campus. The STFC Daresbury system with 384 cores was supplemented in 2007 by a BlueGene-P system with 4096 cores that was partly funded through NW-GRID. Daresbury Laboratory and the Cockcroft Institute are working together to establish a Daresbury Science and Innovation Campus Grid similar to the Liverpool campus grid, which will further widen the NW-GRID user base. The basic hardware infrastructure is described in more detail in the WP 1 section, the operational environment that evolved to support this infrastructure is described in the WP 2 section whilst the campus grid middleware is described further in the WP 3 section.

To illustrate how NW-GRID can be exploited and to explore important issues of relevance to any grid environment, NW-GRID ran a series of themed science testbeds – concentrated periods of use, with a focus on particular aspects of Grid usage. The first testbed, conducted after the deployment of the hardware purchased in 2006, concentrated on high-throughput, largely serial, jobs associated with conducting sweeps through parameter spaces (parametric search). All sites were involved, with particularly useful results being obtained by the Molecular Electronics group at Lancaster and by the e-Minerals project with users at Daresbury.

The second testbed involved fewer projects but explored more complex issues such as advanced reservation and co-allocation (booking multiple resources for simultaneous use), security and accounting which would later be relevant for commercial users. In addition, project requirements for checkpoint and restart mechanisms for longer-running jobs were refined and fed back to Streamline Computing, who had provided the initial NW-GRID systems. In the third testbed all of these features were put into practice via two ambitious projects that required simultaneous access to compute resources at more than one site. One of these also required computational steering. The execution of this testbed was delayed, firstly by an upheaval in staff at both STFC Daresbury and Manchester in early 2008 and secondly by delays in upgrading the operating system at core NW-GRID sites. The operating systems originally supplied by Streamline Computing reached the end of their support life in July 2007 and it was felt that any large scale changes in infrastructure to support project-wide advanced reservation and multi-site parallel jobs were best deferred until a new operating system was in place. This was to have happened early in 2008, but the upgrade has been delayed until the fourth quarter of 2008. Further details of the operational aspects of the project are described in the WP 2 section.

NW-GRID's substantial hardware resources would be less useful without middleware to make it easy for users to exploit its systems in the form of a Grid (or Cloud to use a more fashionable term). It had been decided at the outset to take advantage of the basic Globus infrastructure, which provided secure authentication via X-509 certificates and command level mechanisms to dispatch and monitor jobs launched remotely from a user’s system to one or more of the NW-GRID systems. This meant that NW-GRID would be compatible with the NGS and other activities of the UK e-Science Programme. However, Globus has a cumbersome interface and its client-side deployment is too complicated for most users, requiring privileged access both to the system and to site firewalls. The GROWL software developed at Daresbury addressed some of these issues by making the client-side deployment easier and by providing a set of lightweight scripts to submit, monitor and retrieve results from jobs run on NW-GRID from a user’s system. These scripts were useful both from the command line and from within application programs such as MATLAB or MATHEMATICA. Several Social Science applications involve large number of Monte-Carlo style jobs run with different pseudo-random number seeds. The R programming environment is a popular setting for such jobs and Lancaster developed a mechanism called multiR, based on GROWL, which allows large numbers of R jobs to be submitted and retrieved on NW-GRID systems.

The e-Minerals project required that NW-GRID jobs access centrally stored data files, both for input and output, that thousands of jobs could be submitted and that the key results from these large scale runs were extracted automatically and stored in a way to allow automated post-processing afterwards. The Storage Resource Broker from San Diego Supercomputer Centre was deployed to facilitate this and Condor was used as a scripting interface, a simple workflow management system and to submit jobs to the Globus gatekeeper. The overall infrastructure to support this was called “Remote My Condor Submit” or RMCS which is also used on campus grids, e.g. at Cambridge, UCL and Daresbury. Condor also played a role at Liverpool, where its basic ULGRID campus infrastructure was enhanced to make it easier to run jobs on the Liverpool NW-GRID systems. This was further enhanced in August to include all NW-GRID as well as NGS systems.

GROWL and RMCS are currently being integrated into the G-R-Toolkit at Daresbury Laboratory.

Whilst command-line interfaces are suitable for certain application groups that have experienced staff, simpler access mechanisms are required for other groups, particularly those who are relatively new to high-performance computing. The BioPortal at Manchester makes it easier for bio-sciences users to perform a variety of DNA search and investigation tasks. Certain computational chemistry applications only required a specified input file and returned a small number of output files. Web-based portals to specify the input files and to then submit NW-GRID jobs were developed at Daresbury and Liverpool.

As important as the submission of jobs is, of equal importance is the ability of scientists at geographically separated sites to work easily with one another, possibly to explore the results of large-scale testbed runs. The project itself needed such an infrastructure and made good use of the Access Grid combined

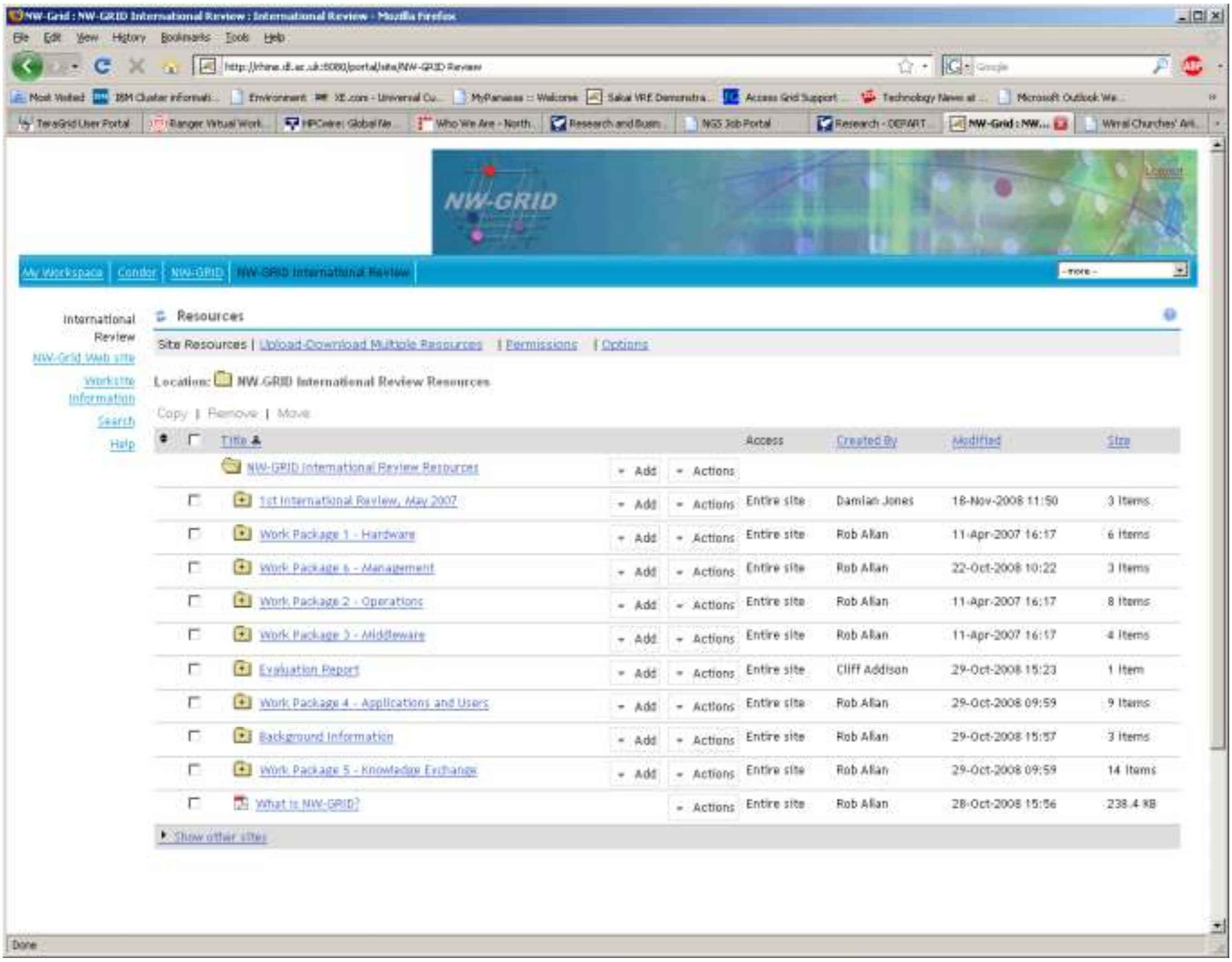


Figure 1: Sakai Portal used in NW-GRID project

with the Virtual Research Environment (VRE) provided by the Sakai Portal hosted at Daresbury. The Sakai portal technology has been particularly successful in supporting work on the Social Sciences e-Infrastructure, a collaboration among Daresbury, Liverpool and Manchester. Work is continuing to integrate the other project portal work into the VRE, thereby enhancing its utility further. Future effort will ensure that the tools from the NGS portal also inter-operate with NW-GRID via Sakai. Further detail on middleware developments is presented in the WP 3 section.

A demonstrable benefit of NW-GRID is evidenced in the results obtained from projects that have used it. Relatively soon after the first NW-GRID systems were operational, the Molecular Electronics group at Lancaster were faced with a challenge. A recent publication had spurred a series of investigations across the world. Being the first group to report on the results of these investigations would lead to a publication in a major journal. Unfortunately, these investigations required substantial amounts of computation and normal access methods on most computer systems meant that months of elapsed time would be required before they could be completed. The testbed model on NW-GRID offered an alternative. By providing near-dedicated access to the NW-GRID resources, the computations were performed in a small number of weeks, even allowing for the time required to set-up effective job submissions (GROWL was deployed to good effect). The group succeeded in their objective of being the first to complete the investigation with a resulting article in the **Journal of Physics: Condensed Matter**.

The group has been active throughout the lifetime of NW-GRID and recently received international attention for their work on nano-motors, which could lead to the "printing" of designer molecules for medicines and other areas (c.f. NewScientist.com 11 June 2008, , *'Electron turbine' could print designer molecules*).

Materials and Surface Science were also major areas of investigation in the e-Minerals project and in several projects at Liverpool. The e-Minerals project made substantial use of NW-GRID to investigate the removal of contaminants from a polluted environment and other highly practical topics. Work at Liverpool covered topics such as constructing electronic devices from single molecular building blocks to investigations into scanning tunnelling microscopes. In addition, a separate group was investigating different ways of storing hydrogen safely and compactly to make it more attractive as an everyday fuel.

Not all NW-GRID applications related to Physics or Chemistry. Researchers at Manchester and Liverpool used over 170,000 core hours on research into bipedal movement during which they produced a virtual race illustrating several species, some living and some extinct. As an effective dissemination vehicle, the "race" received international attention (c.f. BBC Science News article T.rex would outrun footballer (12/8/07)). Work continues on this topic, partly because the movement models employed were relatively simple and partly to expand the information source on gait – for instance, information that can be obtained from fossilised footprints.

Engineers at Liverpool required a similar amount of computing effort to study air turbulence around a ship's superstructure as a helicopter is landing. The simulation needed to study the effects of many different wind angles in the same model so that several levels of parallelism were exploited to produce results for a commercial customer in a timely fashion.

Researchers in the Computer Science and Environmental Science Departments at Lancaster University have been investigating grid-related technology to improve flood monitoring and prediction. For example, they investigated selectively shifting the execution of prediction models to the wireless sensor network used for monitoring floods. Computations organised in this way not only provide more timely flood warnings, but also help to adapt the wireless sensor network dynamically and thus optimise it for current or predicted environmental conditions. For example, the network can employ a low power, low throughput organisation in quiescent conditions, and switch to a high power, high throughput organisation when flooding is imminent. Future plans are to integrate this localised "mini-grid" with larger scale NW-GRID computation resources in order to perform more detailed calculations in the presence of a flood threat and to then ensure that the relevant local and rescue services are apprised of the pending problems.

Further information about applications run on NW-GRID is provided in the WP 4 section.

One of the major goals for NW-GRID was to create a self-sustaining infrastructure for the region that would last long after direct NWDA funding had ended. Commercial involvement with NW-GRID was considered an important route to sustainability. However, the pace and direction of commercial interests in grid computing have been considerably different from project expectations at the start in April 2004. For a variety of practical reasons, commercial users have remained more interested in the specific solution to their particular problems than the use of grid computing to achieve this solution. Similarly, large commercial concerns, if they did adopt grid-computing technologies, chose to do so entirely in-house. Smaller commercial concerns are often not familiar with using the kind of computational resources and expertise offered by NW-GRID. Thus there was a shortfall in the knowledge exchange that took place from the project directly to industry in the North West. However, NW-GRID has facilitated commercially relevant work by its academic users by enabling them to perform computational modelling on a larger scale and in a shorter time frame than groups elsewhere in the UK. One such new research direction relates to the work on nano-scale motors mentioned earlier.

Regional academic groups will always be essential conduits to certain commercial users because their work relates directly to problems these users need solving. However more direct routes are available to other commercial partners. Independent software vendors, such as Fluent, have difficulties tuning and testing their codes on systems that approach the power and size of compute clusters run by their customers. They are willing to pay for access to such systems and discussions have taken place on the possibilities of NW-GRID providing some of this access.

The number of industrial customers buying clusters is increasing. However, even modest clusters require specialised machine rooms, with considerable cooling and electricity demands. In addition, small clusters often represent relatively poor value for money because a high proportion of their cost is taken up in basic network connectivity, the main server node and file store. On paper, it is frequently more cost effective to simply buy time on third-party clusters or storage servers, but there are offsetting concerns with security and connectivity. Therefore cluster vendors have been exploring ways to assist potential customers in their decision making process, whilst also improving their competitive advantage. Discussions have been held about the possibility of these vendors buying time of NW-GRID to provide access to prospective customers on a “try-before-you-buy” basis. Similarly, rather than academic groups buying small clusters, it is more cost effective for all involved if NW-GRID could lease access to a certain number of nodes over a period of time.

A fourth way to increase NW-GRID commercial contact is by identifying potential partners, both academic and industrial, to target strategic applications as identified by the Technology Strategy Board. Pursuit of this particular funding avenue has been delayed because of work relating to the Hartree Centre and because of lingering issues with the 2007 procurement – the selected vendor made promises to all of NW-GRID that were crucial for forward planning but which still have not been effected some 9 months after the promises were to have been carried out.

WP 5 was also responsible for broader dissemination. This covered several activities. One important way to get information out to interested parties was via the project web page (www.nw-grid.ac.uk). Presentations on NW-GRID specific activities have been given at conferences and workshops (notably the All Hands Meeting (AHM) in 2007) and NW-GRID had booth displays at the Open Grid Forum (OGF20) and at the 2007 AHM. A set of flyers on NW-GRID applications and key features was developed and a leaflet describing the NW-GRID Service went to press in October 2008.

In addition, NW-GRID ran several themed workshops covering topics from “Middleware and clusters for grid computing” through to “Campus-Grids” and “Computational Steering”. These workshops were held at different partner sites and were generally well received and well attended. Further details about dissemination activities are provided in the WP 5 section.

WP 1 – Hardware infrastructure

The core of NW-GRID centres around four sites at STFC Daresbury Laboratory and the Universities at Lancaster, Liverpool and Manchester. The clusters are based on multi core, multi processor AMD Opteron nodes with at least 8GB memory per node and larger nodes of 16 and 32GB. The nodes at each site are:

- Daresbury: 96 nodes 2.4 GHz twin dual-core CPU
- Lancaster: 48 nodes 2.6 GHz twin dual-core CPU
- Lancaster: 67 nodes 2.6 GHz twin quad-core CPU
- Liverpool: 104 nodes, 2.2 GHz twin dual-core and 2.3 GHz twin quad-core CPU
- Liverpool: 108 nodes, 2.4 GHz twin dual-core and 2.3 GHz twin quad-core CPU and InfiniPath network
- Manchester: 25 nodes 2.4GHz twin dual-core CPU

Daresbury, Lancaster and Liverpool have 8 TB of storage accessed by Panasas file servers. In addition to this, there are RAID arrays of 2.8 TB at Manchester and 24 TB at each of Lancaster and Liverpool. Nodes are connected by separate data and communications interconnect using Gigabit Ethernet and Liverpool's second 108-node cluster is connected with InfiniPath.

Around this core are other computer systems that are connected to the NW-GRID.

- Daresbury: IBM BlueGene-L (2048 cores)
- Daresbury: IBM BlueGene-P (4096 cores)
- Daresbury: 2560-node IBM 1.5 GHz Power5 (HPCx - subject to approval)
- Daresbury: 32-node Harpertown cluster
- Daresbury 32-node Woodcrest cluster with ClearSpeed accelerators
- Lancaster: 124-node Streamline/Sun cluster 2.6 GHz twin dual-core (HPCF users)
- Liverpool: 96 node, 196 core Xeon x86 cluster, contributed by Proudman Oceanographic Laboratory, with 5.7 TB of GPFS storage;
- Liverpool: 960-node Dell cluster, Pentium IV processors (Physics)
- Manchester: 44 node dual-processor Opteron cluster, 2.5 TB RAID storage based on 2 GHz Opterons with 2 GB RAM.
- Manchester: SGI Prism with 8 Itanium2 processors and 4 ATI FireGL X3 graphics pipes.

The details of the procurement for the first set of NW-GRID compute resources (with NWDA funding totalling £760,000 inc. VAT) are available on the International Review Portal, along with details of the second procurement, see International Review Resources/ Work Package 1 - Hardware/ D1.2 - Phase 1 Systems. The second procurement (with NWDA funding totalling £794,000 inc. VAT) is described on the portal at NW-GRID International Review Resources/ Work Package 1 - Hardware/ D1.3 - Phase 2 Systems but details are reproduced below.

Introduction

In Phase 2 of this workpackage funds were available for the partner sites Daresbury, Lancaster and Liverpool to refresh equipment. To meet the requirements of Lancaster and Liverpool, an EU procurement process would be required, whilst Daresbury requirements would be best met via participation in an ongoing project. On behalf of the North West Grid, and more particularly Lancaster and Liverpool, the procurement was lead by Prof. Terry Hewitt from the University of Manchester.

Preparation of the Invitation to Tender

It was decided to model this purchase upon the successful Phase 1 procurement, including lessons learned. The tender process used for this procurement is outlined below:

- One EU procurement lead by the University of Manchester;
- Procurement to be organised into 2 lots (Lancaster and Liverpool);
- Vendors would be invited to a “Vendors’ Day” meeting where a presentation about the procurement and an opportunity to ask questions was offered;
- To ensure vendors are aware and able to meet the site-specific requirements, such as a water-cooled solution in Lancaster, vendors were encouraged to visit the sites by appointment;
- Lancaster and Liverpool would each undertake the evaluation of their lot;
- There was no preference placed on procurement of identical equipment, a more heterogeneous approach would be acceptable;
- The tender document would be structured as one series of requirements, with a small final section with site specific requirements;
- The UK NGS minimum software stack to be used as a minimum specification for the required grid software.

In addition to preparing the Invitation to Tender [2] a document describing the evaluation methodology was prepared [5].

Due to the “lessons learned”, the tender documentation was modified by:

- Using only electronic submission;
- Reducing the number of minimum requirements;
- Reducing the number of benchmarks to be supplied;
- Increasing the number of highly desirables and information requests;
- Introduction of the option of liquidated damages to cover costs of delays in delivery.

Vendors requesting the tender documentation were invited to attend the vendor’s day, which took place in Lancaster on the 4th of October 2007. A presentation was made [3] and questions from vendors were answered. The majority of vendors made site visits; to comply with regulations, queries that arose during these visits were clarified to all vendors [4].

Evaluation of the Tenders

3.1 Evaluation panels

The evaluation panel was chaired by Terry Hewitt. Local panels in Lancaster and Liverpool completed the evaluation of their respective lots, with input from the panel members in Daresbury and Manchester. The local panels recommended a solution to the evaluation panel, using the agreed and published procedure and evaluation forms [5]. The evaluation panel then endorsed these decisions after which Prof. Hewitt announced the winner. This was also the moment at which the 10-day cooling off period would start.

The members of the evaluation panel were:

- Terry Hewitt (Chair)
- Rob Allan (Daresbury)
- Cliff Addison (Chair local panel Liverpool)
- Ties van Ark (Lancaster)
- Mike Pacey (Chair local panel Lancaster)
- Mike Houlden (Liverpool)
- Andrew Jones (Manchester)

Bids were received from ClusterVision, Dell, Eurotech, OCF, SGI, Streamline and Viglen. ClusterVision only bid for the Lancaster lot, SGI and Viglen each posted a choice of two bids.

3.2 Evaluation criteria

The bids were evaluated on 5 criteria agreed by the Evaluation Panel, with criteria E1, E2 and E3 further subdivided. Each criterion had an attached weighting so that for example the score on criteria E1 would result in 35% of the total points available.

	Evaluation Criteria	**Weighting**
E1	Fitness for purpose, including evaluation of benchmarks	35
E2	Value for money, cost effectiveness (including total cost of ownership aspects)	30
E3	Technical merit, including aesthetic and functional characteristics	15
E4	After-sales service and technical support	10
E5	Support towards achieving the aims of the North-West Grid	10

Table 1: Evaluation Criteria and Weighting

Each criterion was awarded a score from 0 to 4 on the following basis:

0	poor/ unsatisfactory	Solution falls below reasonable expectation for the criteria. This is the highest score possible if some of the relevant IRs or MRs are not fulfilled.
1	minimum standard	Solution satisfies a reasonable expectation for the criteria.
2	Good	Solution exceeds reasonable expectations for the criteria, for example fulfils several relevant HRs.
3	exceeds expectations	Solution exceeds expectations and exceeds the average offering of the submitted tenders in several important categories.
4	Exceptional	Solution exceptionally exceeds reasonable expectation for the criteria, for example fulfils almost all of the relevant HRs. This score should be used very sparingly.

Table 2: Scoring principles

The scores assigned to each of the evaluation criteria E1-5 (e.g. Table 1) were then multiplied by one quarter of the weighting factors in column 3 of this table and summed to give the overall score for each solution, so the maximum score possible was 100. The highest score would be the recommended solution.

3.3 Evaluation results

The local teams at Lancaster [8] and Liverpool [9] identified Streamline Computing as their preferred bidder. The Evaluation panel endorsed these conclusions and letters of congratulations to Streamline and commiserations to the other vendors were issued on the by Prof. Hewitt on the 26th of November 2007.

During the cooling-off period Manchester received a complaint from one of the vendors. This complaint was robustly rejected by a coordinated response from Lancaster and Liverpool [6].

Agreement on additional Terms and Conditions

Both Lancaster and Liverpool exercised their right to make variations to the lots offered. A significant variation was that Lancaster decided to supply its own preferred water cooling solution [12]. This resulted in a significant saving on the cost of cooling and enabled Lancaster to order additional nodes and memory from Streamline Computing.

The systems tendered by Streamline were based on twin AMD Opteron "Barcelona" quad-core processors. After the evaluation process was completed, new information about the availability of the selected Barcelona processors owing to manufacturing difficulties arose. In negotiations with Streamline, Sun and AMD it was agreed that the initial system would be delivered with twin dual-core Opteron processors with these to be replaced by quad-core Barcelona processors within a hundred days of general availability. Lancaster and Liverpool sought additional guarantees from Streamline and Sun to ensure that financial risks were minimised [11]. Additional guarantees, final configurations, timetables, acceptance tests and liquidated damages were agreed in additional terms and conditions [10]. Purchase orders were issued late December/ early January 2008. The initial systems were delivered in late February/ early March 2008 - these were upgraded by Sun to Barcelona processors in July/ August 2008.

Daresbury

The requirements of STFC Daresbury were best satisfied with a NW-Grid share in the purchase of an IBM BlueGene-P frame. This was affected via the re-negotiation of a contract between STFC and IBM.

Connectivity throughout the North West

The compute clusters at the partner sites are complemented by a high-speed private network that can be enhanced and configured to meet user requirements for secure access and data transfer between clusters and storage systems. The costs of this network infrastructure were largely met by other sources. For instance, the initial one-gigabit connection between Lancaster and Daresbury made use of fibre installed as part of the UK Light project. The one-gigabit connections among Daresbury, Manchester and Liverpool were funded by the Net North West consortium. The intention is to update all of these connections to 10 Gbp/s and part of the infrastructure costs to do this between Liverpool and Daresbury was funded from the NW-GRID project.

Documentation on the NW-GRID portal

The following documentation is available on the portal under NW-GRID International Review Resources / Work Package 1 - Hardware / D1.3 - Phase 2 Systems:

[1] Summary of 2nd procurement stored as file `081123, Proc 2 –Summary, Summary of 2nd procurement`.

[2] Tender document for 2nd procurement stored as file `070904, Proc 2 - Tender Sept 2007`.

[3] Presentation to vendors as made on vendors' day stored as file `071004, Proc 2 - Vendors Day`.

[4] Clarifications of tender as issued to vendors stored as file `010116, Proc 2 - Tender Clarifications`.

[5] Evaluation procedure for tender responses stored as file `071120, Proc 2 – Evaluation Criteria`.

[6] Response from Lancaster/Liverpool to SGI stored as file `071217, Proc 2 - SGI reply`.

[7] Winning response from Streamline Computing to the invitation to tender stored as file 071126, Proc 2 – Streamline Response.

[8] Matrix with evaluation results as comprised by the local evaluation team in Lancaster stored as file 071123, Proc 2 – Lancaster Evaluation Matrix.

[9] Matrix with evaluation results as comprised by the local evaluation team in Liverpool stored as file 071123, Proc 2 – Liverpool Evaluation Matrix.

[10] Agreement on details, delivery and schedule with Streamline Computing stored as file 071210, Proc 2 - Terms and Conditions.

[11] Letter from Sun guaranteeing delivery of Barcelona stored as file 081218, Proc 2 - fr Sun, guarantee.

[12] Quotation for Rittal water-cooling solution as selected and ordered by Lancaster University stored as file 071221, Proc 2 – Lancaster Cooling Quotation.

WP 2 – Operational Issues and Science Testbeds

WP 2 was designed to ensure a fully-tested production environment for end users of the NW-GRID infrastructure. The middleware available for building wide-area Grid infrastructures changed radically over the period March 2004 to date. We undertook a number of middleware evaluations and were involved in several related projects to assess the impact of changes on Grid-based applications, but remained focussed on providing a reliable service. Functionality of the middleware deployed had to meet stated application requirements and ensure inter-operability with other UK and overseas Grid systems, e.g. National Grid Service (NGS) which included the HPCx and CSAR facilities at the time http://www.ngs.ac.uk , White Rose Grid http://www.wrgrid.org.uk, NorthGrid (GridPP production facility) http://www.gridpp.ac.uk , the European EGEE project http://public.eu-egee.org and the US extended TeraGrid Facility http://www.teragrid.org. The more recent emergence of Campus Grids and Cloud Computing has also been addressed.

We note that staff at Manchester and Daresbury are involved in management and operational tasks around the NGS clusters in addition to HPCx, CSAR and now HECToR. HPCx is part of the European DEISA Grid in addition to NGS.

This WP was implemented in three phases, in line with the rest of the NW-GRID project. Operational and related issues (excluding policy) were overseen by the NW-GRID Operations Board chaired by Dr. Rob Allan (Daresbury). Its first meeting was 9/6/2005 and it has met regularly since that date using Access Grid.

Part 1 - Operational Areas

Prior to the first major tender for equipment in late 2005 (see WP 1) NW-GRID partners had access to test systems and knowledge from previous work. This enabled a Rollout Plan [1] and Technology Roadmap [2] to be written in which we recommended system software, configuration and a middleware stack based on Linux and Globus Toolkit 2 (pre Web services). This, together with conditions of access [3] and user-management procedures [4] acceptable to JISC, and UKERNA ultimately led to our partnership with the NGS for which a Service Level Description was written [5] in late 2007.

The original WP 2 plan identified a number of areas of work which at the time were believed to be relevant to a future world-class Grid service. We briefly address how we have evaluated these areas (now referred to as “Special Interest Areas”) and how priorities have changed over a 3-year period which has seen a huge flux of new software technologies coming into Grid computing. Information has been shared using a Wiki in the Operations Board portal work site.

System Software

Systems were maintained with a base level of software consistent with science and engineering uses of high-performance compute clusters. This software was based on that recommended by the systems vendors as far as possible in compliance with the original contracts. Some additional software was deployed as requested by users.

To address issues of maintaining consistent and up-to-date system software across multiple sites we formed the Streamline Computing User Group in September 2008. The first meeting at Daresbury Laboratory on 7/10/2008 was attended by systems staff from Daresbury, Durham, Lancaster, Liverpool, Manchester, Nottingham and Oxford. Expressions of interest were received from other sites including the White Rose Grid. Feedback on issues raised have been received from Streamline Computing and a second meeting was held on 1/12/2008 before the 19th Daresbury Machine Evaluation Workshop.

Storage and Data Integration

All NW-GRID clusters have substantial storage capacity, either RAID file store or proprietary Panasas servers. Data was preserved as required by users for the duration of their projects. Panasas had originally been considered in order to provide a wide-area file store with share file name space. The Storage Resource Broker (SRB) was implemented as required by some projects, e.g. e-Minerals. An Oracle DB infrastructure was available for some time at Daresbury and was, for instance used by the NW-GRID Portal. Later this data was migrated to the Oracle DB running on the NGS cluster at Manchester.

- Panasas, RAID arrays, Sun x4500 'Thumpers' – used for primary data storage on each cluster;
- Backup – carried out on a site-by-site basis;
- SRB: Storage Resource Broker – used as a distributed logical file system;
- Oracle service at Daresbury and later Manchester interfaced to Sakai and SRB and RMCS meta-data services;
- OGSA-DAI – was investigated in the UK OGSA testbed project, but there have been no requests from users so a full services was not provided.

Network and Communications

A unique feature of NW-GRID is the provision of a private fibre network linking the core clusters, as described above. Network communication for data transfer and parallel applications was also subjected to testing and maintenance.

- Private Fibre Installation and configuration – functionality and bandwidth testing;
- SCORE – high performance MPI over Gb/s Ethernet used for parallel applications;
- Infiniband – low latency, high bandwidth for capability computing;
- Testing – parallel libraries and communication interfaces such as OpenMPI, MPICH and MPICH2.

Batch Scheduling and Advanced Reservation

Batch scheduling was chosen as the normal operational mode for the core clusters, but other modes were possible, e.g. interactive use during science testbeds or co-allocation. Advanced reservation was required to implement the latter on a routine basis.

- SGE v6 – used for routine job submission and monitoring purposes;
- MOAB – was evaluated for some purposes such as advanced reservation. SGE v6.1 now supports this and there is full support in v6.2.

Infrastructure Middleware

This special interest area was a continuous evaluation of available middleware for building Grid infrastructures (as differentiated from the user-facing and novel middleware addressed in WP 3). This area focussed initially on Globus, VDT and Grid Services as implemented at the time in OGSA and OGSI. Grid services were later radically changed when the WSRF standard emerged and we were unable to track subsequent development due to lack of resource.

- Globus GT2.4.3 (or a VDT supported compatible release) – used for authentication (via GSI), resource discovery (MDS), file transfer (GridFTP) and job submission (GRAM). In fact MDS was not used but in future a BDII service will provide information via the NGS;
- Globus GT3 (Grid Services) – was evaluated along with OGSA-DAI during the EPSRC-funded UK OGSA Testbed project. A report is available http://dsg.port.ac.uk/projects/ogsa-testbed;
- Globus GT4 pre-WS – used in place of GT2 as soon as available – currently GT4.0.7 is deployed on most core clusters and is a dependency for other middleware such as MPIg;

- Review WSRF specification against requirements – this was done in the EPSRC funded WOSE project (Workflow Optimisation Services for e-Science) resulting in a set of comprehensive tutorials: http://www.grids.ac.uk/WOSE/tutorials/index.html

Whilst OGSI is no longer relevant, we note that Daresbury and Manchester were involved in the OGSA Testbed project and Manchester has worked on OGSI::Lite which later became WSRF::Lite. Nevertheless the pre-web services Globus has remained the most important middleware for us to support a production Grid environment as on the NGS. We also have consolidated a lot of experience of Web services (WS-I) in GROWL and RMCS.

Accounting and Monitoring

Accounting and monitoring are important requirements, particularly for commercial users.

- RUS – as used for the NGS usage/ accounting system;
- GOLD, part of MOAB suite – was evaluated, see ref. [10];
- Evaluation of NGS user management and VOMS systems as also used on GridPP.

Authentication, Authorisation and Virtual Organisation Management

In common with other Grid infrastructures, NW-GRID uses authentication and security based on GSI, the Grid Security Interface from Globus. This uses X.509 certificates for users, hosts and services (such as portal, GROWL and RMCS) issued by the UK e-Science Certification Authority http://ca.grid-support.ac.uk . Other certificates have been used for training purposes issued by NGS-TOE.

Establishment of a scalable set of authentication and authorisation services based on Shibboleth and PERMIS for UK academics is a longer-term objective being pursued by JISC. Staff from Daresbury and Manchester have participated in the ShibGrid, SHEBANGS and more recent SARoNGS projects to investigate inter-operability with Grids and portals.

Co-allocation and Meta-computing

- HARC – evaluated and deployed, see ref. [8]
- MPIg, PAC-X – ongoing evaluation, see ref. [9] for preliminary results

NW-GRID Web and Portal

We developed a project Web site to publicise the work of NW-GRID and its users, its content is further discussed under the section on WP 5. We also established a NW-GRID portal to complement the Virtual Research Environment and Virtual Learning Environment demonstrators being investigated by JISC. We were funded to develop the Sakai VRE Demonstrator in the VRE-1 programme, so this was extended for the NW-GRID work and later to support NCeSS and other e-Research projects.

- Portal collaborative working interface – the Sakai VRE Demonstrator was extended;
- Portal job submission interface - a clone of the NGS portal has been tested https://portal.ngs.ac.uk. More recent NGS portlets funded under OMII remain to be ported into the Sakai environment owing to lack of resource, although this was demonstrated with earlier versions;
- Portal data management interface – a portlet version of DataPortal was written and demonstrated in Sakai. This was designed specifically for use on the Diamond Synchrotron in conjunction with data analysis services running on NW-GRID (e.g. e-HTPX for protein structure determination);
- Other portal developments are described in the section on WP 3.

Campus Clouds

Recent work has pursued the concept of a "Campus Cloud" based on earlier work on Condor-based Campus Grids. Condor and other services based on SRB, Oracle or MySQL and Globus are now deployed at Liverpool and Manchester and on the Daresbury Science and Innovation Campus. For information about the latter and relevant documentation see http://tardis.dl.ac.uk/Condor. The NW-GRID clusters are utilised in such clouds either through a Condor-G interface or RMCS or by back-filling so that vacant nodes become part of the campus resource pool.

Part 2 - Science Testbeds

The NW-GRID adopted a novel policy toward Grid usage. In addition to regular remote access using Globus middleware it can be booked as a "facility" for "Science Testbeds". These have proven very effective, enabling access to all core compute clusters for up to one week for intensive scientific work by a group of scientists collaborating on a specific project. Results of some of the testbeds and lessons learned are published on the Web site at http://www.nw-grid.ac.uk and two papers illustrating the scientific use of NW-GRID resources in the first year of operation were presented at the 2007 UK e-Science All Hands Conference [6, 7]. Some highlights are given in the section on WP 4 in this report.

Up to 20% of the NW-GRID resource is available for application and middleware development which may be novel or ground-breaking and may also be used to evaluate Grid solutions appropriate for other sites (e.g. via ETF, the Grid Engineering Task Force). This work is described in other sections (WP 3, WP 4) of this report.

The NW-GRID Project was expected to develop software and run three science testbeds. We here briefly describe the findings from this experience in terms of deployed operational software, user experiences and further requirements. Details of the access middleware developed and applications deployed are given in sections on WP 3 and WP 4.

T1 (2006-7): High-performance and parametric computing

This testbed was led by groups from Daresbury and Liverpool. Parameteric computing (searching parameter space or doing sensitivity analysis) requires many repeated runs of an application with slightly different input conditions. These are sometimes referred to as "ensemble" runs and can be used for subsequent statisitical analysis or graphical interpretation. They can also be used as a starting point for more sophisticated modelling once regions of interest are identified. This methodology is ideal for Grid computing using loosely-coupled resources, indeed it has been referred to as a "killer app" and a number of projects made use of it on NW-GRID.

Issues with parametric computing arise from the related "bookkeeping" activities. It is important to record the parameters used for each run and to correlate the results. We extended existing middleware to do this by building on the work of the NERC-funded e-Minerals project. This resulted in RMCS, SRB and RCommands being deployed on NW-GRID and tools such as RGem used to extract and present results from the data collected. Most significantly, RCommands enabled automatic creation of job-related meta-data which solving the bookkeeping problem. An example of this methodology is shown in the section on WP 4 in the studies glasses, minerals and $CaCuO_2$ antiferromagnetism. These issues are discussed in ref. [6].

To enable this testbed series to be completed it was necessary to adapt the Globus jobmanager scripts so that it interfaces to Condor-G could be used to submit the many jobs across the Grid in a consistent manner.

T2 (2007-8): Safe, bookable and accountable Grid

This tesbed was led by Manchester and addressed some of the future needs of commercial Grid users and also underlying requirements for T3. In order to provide commercial access we have to provide assurances that data will be kept confidential, applications only provided to users with the necessary licenses and in some cases even information about projects kept secure. In order for commercial users to pay for the services they receive, an appropriate and accurate accounting system must be in place mapping resource usage to individuals and projects, but keeping this information confidential. Finally it is necessary to book access to resources for particular high-priority tasks, which could relate to business-critical phases such as validation, urgent social situations such as disaster avoidance and recovery or support for medical procedures. This could involve co-allocation, for instance of computing, experimental collaboration (Access Grid) or visualisation resources.

Staff at Manchester evaluated the RUS, GOLD and HARC middleware to satisfy the above requirements. Work in Lancaster continued to develop the coupled monitoring and environmental prediction framework. Work at Daresbury involved dialogue with ISVs to identify new business models for licensing applications.

T3 (2008-9): Collaborative real-time Computing

This testbed series was led by groups at Lancaster and Daresbury. Part of this built upon work in projects funded by JISC, around collaborative Web portals using Sakai and the Grid portal as deployed for the NGS. Manchester and Liverpool contributed to evaluating portal interfaces for bio-informatics and chemistry. The portal interfaces are available for use as a Virtual Research Environment encapsulating research, administrative and collaboration tools for sharing information. In the future it is anticipated that bespoke portals or "science gateways" can be constructed enabling turnkey access for both academic and commercial use of deployed applications. The science gateway will manage the access, accounting and licensing software required.

Campus Grids (or Cloud to use a more modern term) are being deployed at Liverpool, Manchester and Daresbury linking local resources to NW-GRID resources to enable easier access and matching resource to application requirements. This has been particularly successful in extending the parameter sweep range of applications as started in T1. Access to the Cloud is via a portal as described above.

Several projects expressed a requirement to do meta-computing and / or computational steering. In GENIUS, the Grid Enabled Neurosurgical Imaging Using Simulation project, various facets of high-performance computing are used which go well beyond the typical batch job submission including: (a) *in situ* visualization; (b) real-time remote visualization and steering using RealityGrid steering; (c) distributed computing using MPIg; (d) advance reservations using HARC and urgent computing capabilities using SPRUCE; and (e) automated job launching using the Application Hosting Environment (AHE). GENIUS achieved the capability to steer and visualize in real-time blood flow through patient-specific vasculatures within reservations on remote resources throughout the UK (NGS and NW-GRID) and USA (LONI and TeraGrid).

GCOMS, the Global Coastal Ocean Modelling project led by the Proudman Oceanographic Laboratory also planned to use MPIg. This work has progressed more slowly owing to a lack of staff effort.

Documentation on the NW-GRID portal

[1] R.J. Allan "Rollout Plan" (NW-GRID, 22/2/2006) 6pp

[2] R.J. Allan, T.J. Franks, J.M. Brooke, S. Nadeem, M. Pacey and C.A. Addison "NW-GRID: A Technology Roadmap for the North West Grid" (NW-GRID, 22/2/2006) 42pp

[3] R.J. Allan "Regulations for Use of the North West Grid" (NW-GRID, 18/5/2005, updated 17/12/2007) 13pp

[4] R.J. Allan, T.J. Franks, S. Nadeem, A.J. Lee and C.A. Addison "Joining Procedures for the North West Grid" (NW-GRID, 16/3/2007, updated 24/2/2008) 18pp

[5] R.J. Allan, J.M. Brooke, M. Pacey and C.A. Addison "Service Level Description for the United Kingdom National Grid Service Resources from the North West Grid" (NW-GRID, 22/11/2007) 11pp

[6] J.M.H. Thomas, R.P. Tyer, R.J. Allan, J.M. Rintelman, P. Sherwood, M.T. Dove, K.F. Austen, A.M. Walker, R.P. Bruin, L. Petit and M.C. Durrant "Science carried out as part of the NW-GRID project using the e-Minerals Infrastructure" (UK e-Science All Hands Conference, Sept'2007) 8pp

[7] J.M.H. Thomas, J. Kewley, R.J. Allan, J.M. Rintelman, P. Sherwood, C.L. Bailey, A. Wander, B.G. Searle, N.M. Harrison, S. Mukhopadyhay, A. Trewin, G.R. Darling and A.I. Cooper "Experiences with different middleware solutions on the NW-GRID" (UK e-Science All Hands Conference, Sept'2007) 8pp

[8] J. MacLaren "HARC Deployment" (NW-GRID, Mar'2007)

[9] T.J. Franks "MPIg Deployment" (NW-GRID, Mar'2007)

[10] T. Robinson "Accounting and Allocation Management on Man2" (NW-Grid, 2008)

WP 3 – Grid Middleware Evaluation and Extension

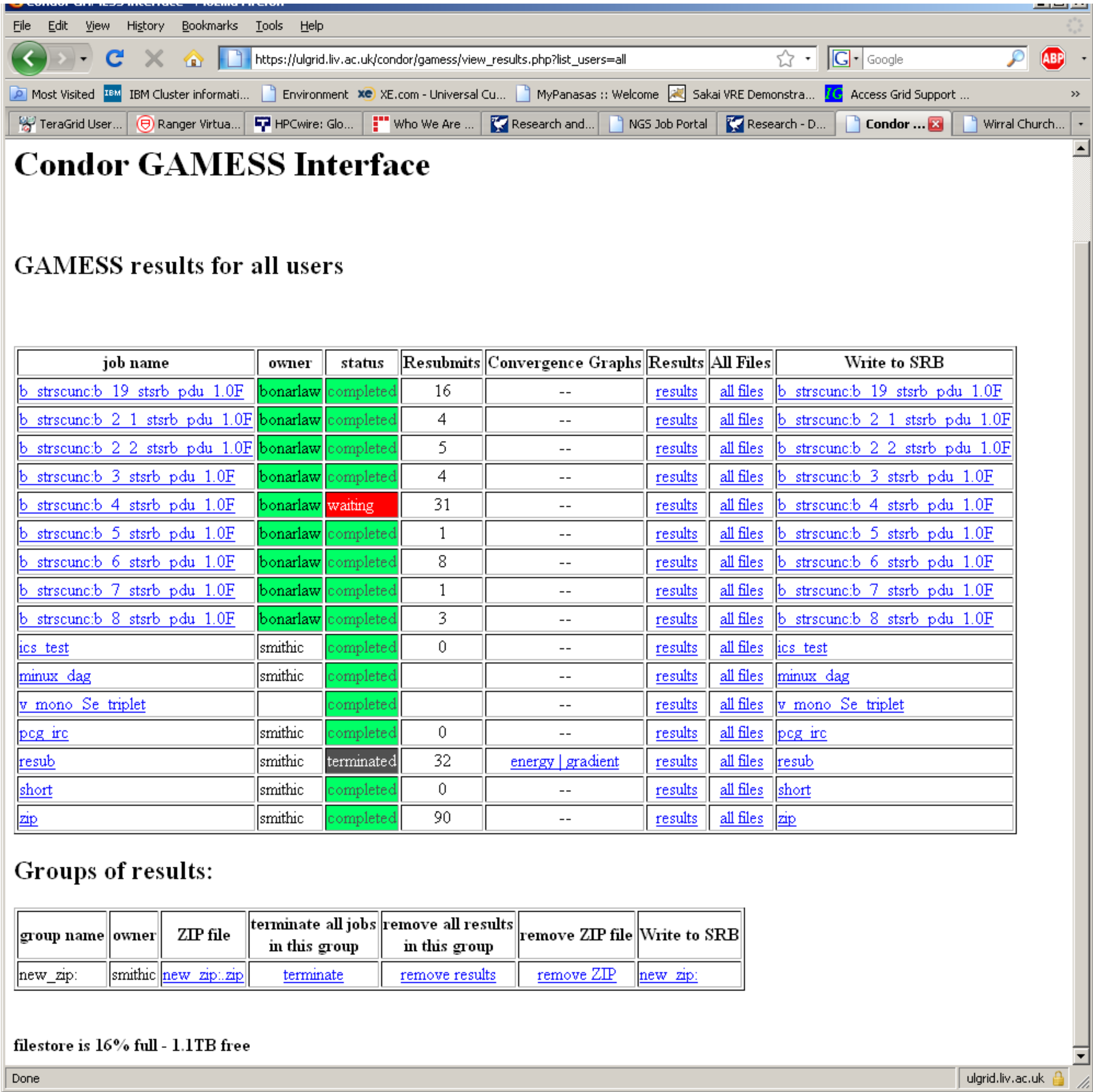
This WP was implemented in three phases, in line with the rest of the NW-GRID project. The initial phase consisted of evaluation and choice of middleware for deployment. Work had been done at all four sites as part of the wider UK e-Science Programme, so we were already familiar with available middleware and requirements of HPC users, having participated in the establishment of the National Grid Service and associated Grid Support Centre. In the Rollout Plan [1] and Technology Roadmap [2] (part of WP 2) we therefore recommended a middleware stack based on Globus Toolkit 2 and user-management procedures compatible with the NGS. An evaluation of the status of Grid middleware available in 2004-5 was carried out [3] and this identified a number of gaps. A technical seminar entitled "Middleware and Clusters for Grid Computing" was held on 13/12/2005 at which our findings were presented for the first time.

WP 3 in its entirety addressed two key dimensions. The first dimension investigated short to medium term extensions to the infrastructure with a view to increasing usability of the software. The second dimension, which we called middleware development, has addressed longer-term issues relating to the very structure of Grid middleware to enable richer sets of applications to be supported.

Short-Medium Term Grid Extensions

We were fortunate in the NW-GRID in having the flexibility to deploy and evaluate, with local users of demanding HPC applications, user interfaces and middleware which we were developing via funding from a variety of additional sources (JISC, EPSRC, NERC, BBSRC, ESRC, CCLRC and DTI). A review was carried out resulting in a plan to use this software to fill the gaps and meet the user requirements [4]. In Phase II therefore, work at Daresbury Laboratory and Lancaster focussed on the "Grid Client Problem", namely how to provide lightweight access from a researcher's desktop PC to large-scale Grid compute and data resources thereby encouraging uptake. This work has two strands: (1) lightweight Grid toolkits, resulting in GROWL, MultiR, AgentX and RMCS [10, 11, 14-20]; (2) Web-based access and collaboration portals as deployed for NW-GRID, NCeSS, N8 Materials Consortium, NGS and other e-Research projects [6-9]. Related work at Manchester resulted in the BioPortal. Work at Lancaster and Manchester has investigated portal access to advanced audio-visual collaboration technologies, Lancaster delivered the Sakai-based Agora software and Manchester the Portal Access Grid, both of which have been evaluated in a range of research areas. More novel work at Lancaster which focussed on future middleware to address a range of novel requirements, for instance linking flood monitoring wireless sensor networks to large-scale environmental modelling for rapid response to potential natural disasters is described below. This resulted in the prototype GridKit middleware [5, 22]. A workshop entitled "Lightweight Grid Computing" was held in Castleton 2-3/5/2006. An international conference entitled "Portals and Portlets 2006" was hosted at the National e-Science Centre, Edinburgh 17-19/7/2006.

In Phase III, work at Liverpool and Manchester (with input from Daresbury) focussed on deploying Campus Grid infrastructures utilising NW-GRID resources plus those from departments across the universities. Figure WP 3-1 is a snapshot of one of the portal interfaces to the Liverpool Campus Grid. GROWL and RMCS are now being combined in the new G-R-Toolkit, which will be deployed for the Daresbury Science and Innovation Campus Grid. Work on portals has continued with extensions to become full Virtual Research Environments combining on-line project management, collaboration, virtual meeting and resource access tools. These are all components required to transition the NW-GRID to a Cloud for researchers with a rich set of deployed applications and user interfaces. A technical seminar was organised entitled "Campus Grids" on 31/10/2007 and one on "Computational Steering" 18-19/3/2008. A "Portal Usability Workshop" was held 12/10/2007. NW-GRID members contributed to "OGF-20" on 7-11/5/2007 (the international conference of the Open Grid Forum) and are contributing to the "NGS Innovation Forum" and "JISC Tools and Technologies Roadshow" 4-6/11/2008.



Condor GAMESS Interface

GAMESS results for all users

job name	owner	status	Resubmits	Convergence Graphs	Results	All Files	Write to SRB
b_strscunc:b_19_stsrb_pdu_1.0F	bonarlaw	completed	16	--	results	all files	b_strscunc:b_19_stsrb_pdu_1.0F
b_strscunc:b_2_1_stsrb_pdu_1.0F	bonarlaw	completed	4	--	results	all files	b_strscunc:b_2_1_stsrb_pdu_1.0F
b_strscunc:b_2_2_stsrb_pdu_1.0F	bonarlaw	completed	5	--	results	all files	b_strscunc:b_2_2_stsrb_pdu_1.0F
b_strscunc:b_3_stsrb_pdu_1.0F	bonarlaw	completed	4	--	results	all files	b_strscunc:b_3_stsrb_pdu_1.0F
b_strscunc:b_4_stsrb_pdu_1.0F	bonarlaw	waiting	31	--	results	all files	b_strscunc:b_4_stsrb_pdu_1.0F
b_strscunc:b_5_stsrb_pdu_1.0F	bonarlaw	completed	1	--	results	all files	b_strscunc:b_5_stsrb_pdu_1.0F
b_strscunc:b_6_stsrb_pdu_1.0F	bonarlaw	completed	8	--	results	all files	b_strscunc:b_6_stsrb_pdu_1.0F
b_strscunc:b_7_stsrb_pdu_1.0F	bonarlaw	completed	1	--	results	all files	b_strscunc:b_7_stsrb_pdu_1.0F
b_strscunc:b_8_stsrb_pdu_1.0F	bonarlaw	completed	3	--	results	all files	b_strscunc:b_8_stsrb_pdu_1.0F
ics_test	smithic	completed	0	--	results	all files	ics_test
minux_dag	smithic	completed		--	results	all files	minux_dag
v_mono_Se_triplet		completed		--	results	all files	v_mono_Se_triplet
pcg_irc	smithic	completed	0	--	results	all files	pcg_irc
resub	smithic	terminated	32	energy | gradient	results	all files	resub
short	smithic	completed	0	--	results	all files	short
zip	smithic	completed	90	--	results	all files	zip

Groups of results:

group name	owner	ZIP file	terminate all jobs in this group	remove all results in this group	remove ZIP file	Write to SRB
new_zip:	smithic	new_zip:.zip	terminate	remove results	remove ZIP	new_zip:

filestore is 16% full - 1.1TB free

Done ulgrid.liv.ac.uk

Figure WP3-1: ULGRID portal from Liverpool

Longer-Term Middleware Development

The section relates to Lancaster University Computing Department's Activities in WP 3 of the North West Grid Project.

In more detail, the work reported here focused on providing more flexible and extensible next generation Grid middleware software with a view to supporting richer classes of application and also richer deployment environments. To this end, we have developed a new Grid middleware platform called GridKit. This provides a more radical communications architecture based on the use of *network overlays*. The GridKit approach supports the natural integration of (for example) peer-to-peer protocols and also support for alternative interaction styles such as publish-subscribe. The approach also supports configuration of the software for a variety of non-conventional deployment environments including sensors interconnected by *ad hoc* wireless networks, mobile devices using a variety of interconnection technologies depending on location, and more traditional large-scale cluster computers connected by high speed networks (this our vision of what we call the "Divergent Grid"). As part of this, we have also investigated the use of Model-Driven Architecture tools to generate appropriate families of GridKit-based middleware from higher-level specifications.

We have especially focused in the NW Grid project on the application of GridKit to *water management*. This built on the experience of the consortium in Environmental Science but also provided an excellent example of a Divergent Grid, featuring both large-scale simulation running on clusters and also a state of the art *wireless sensor network* embedded in the natural environment. More specifically, we have designed a Divergent Grid for predicting flooding events in river valleys in the North West area. This system has been actively deployed at two sites (on the River Ribble and the River Dee; the latter work has been funded by the NERC FREE Project: see later) and is producing real data that is of direct use and relevance to Environmental Scientists.

Since the project ended, the River Dee deployment is continuing. This site is posing a number of additional challenges not experienced on the Ribble site. From a Grid middleware/ networking perspective, this includes the requirement to cover a larger geographical area, while from the flood modelling perspective it includes complex rainfall, tidal and drainage interactions. The supporting infrastructure of all but one of the River Dee nodes has now been deployed and we anticipate that the system will go live before Christmas 2009. Additional work has focused on providing more sophistic deployment support, including the development of a genetic algorithm capable of calculating the most efficient positioning for bridging nodes based on the wireless network characteristics and the topology of the deployment site.

We believe that the GridKit work has been a highly successful in meeting the agreed outcomes of WP 3. We have designed, built, deployed and evaluated a large and complex distributed system, produced a significant number of quality publications, and sparked significant future research work, which has brought in further funds, employed more people, and enhanced the visibility of the NW-GRID in general and in the use of advanced Divergent Grids in flood monitoring and management.

The GridKit software is available under open source licence from Sourceforge (http://sourceforge.net/projects/gridkit; over 750 downloads to date) and has already been adopted by a number of systems-oriented research groups internationally (including LAAS, France; U. Tromsø, Norway, LNCC Brazil). In addition, we have attracted a number of industry contacts in recent months: e.g.: a US-based company is looking at applying GridKit in a flood-warning system in Mumbai, India; Anglian Water are looking at our approach in a sensor-network-based project of theirs called Sewernet; and BAE Systems are considering the use of GridKit in wireless shield surveying systems.

Major Scientific Highlights are as follows:

- Development of a 'Divergent Grid' scenario in which a sensor network delivers real time sensor data in a tightly-coupled manner to a computational model running on a back-end computational resource (i.e. the NW Grid cluster resources). This scenario involves advanced sensor network concepts such as dynamically reconfiguration of sensor networks, self-repair of sensor networks, and dynamic trade-offs of computation and communication across both the sensor network and the back-end computational resource.

- Development of a new and powerful sensing platform with both low-power and high-performance personalities – the GridStix platform.
- Refinement of Lancaster's next-generation middleware research to support a wireless sensor network (WSN) environment and the GridStix platform.
- First deployment of Grid middleware in a WSN environment and demonstration of the advantages of this approach in maintaining network resilience and high performance.
- Investigation of the use of overlay network technologies in a WSN setting as supported by GridKit's "open overlays" abstraction.
- Pioneering use of novel sensor types (e.g. image-based flow measurement).

The Divergent Grid vision that came out of the project is set out in Coulson, G., "A Possible Future for the Grid", Editorial article, Concurrency and Computation: Practice and Experience 2007; **19**: pp 1879-1884, available at: http://www.comp.lancs.ac.uk/computing/users/geoff/Publications/CCPE06.pdf .

Documentation on the NW-GRID portal

[1] R.J. Allan "Rollout Plan" (NW-GRID, 22/2/2006) 6pp

[2] R.J. Allan, T.J. Franks, J.M. Brooke, S. Nadeem, M. Pacey and C.A. Addison "NW-GRID: A Technology Roadmap for the North West Grid" (NW-GRID, 22/2/2006) 42pp

[3] R.J. Allan, R.P. Tyer, P.A. Couch, J. Kewley, G. Coulsen, D. Hughes, R. Crouchley, D. Grose, J.M. Brooke, S. Nadeem, J. MacLaren, C.A. Addison and I. Smith "Status of Grid Middleware and Choices for NW-GRID" (NW-GRID, 2004-5) 51pp

[4] R.J. Allan, J.M. Brooke, G. Blair, G. Coulsen and C.A. Addison "NW-GRID WP3: Extensions to the Grid Middleware" (NW-GRID, 2005 updated 12/12/2005) 14pp

[5] G. Blair and G. Coulson "Future Grid Middleware" (NW-GRID, 8/3/2007) 5pp

[6] R.J. Allan, X. Yang, A.L. Braimah, J. Kewley, R. Crouchley, A. Fish, M. Gonzalez, D. Hughes, C.A. Addison, A. Sykes, J.M. Brooke and B. Kattamuri "NW-GRID Developing a Virtual Research Environment for NW Researchers" (NW-GRID, Mar'2007) 17pp

[7] C. Severance, J. Hardin, G. Golden, R. Crouchley, A. Fish, T. Finholt, B. Kirshner, J. Eng and R.J. Allan "Using the Sakai collaborative toolkit in e-Research applications" Concurrency, Practice and Experience 19 (2007) 1643-52

[8] R.J. Allan, X. Yang and R. Crouchley "ESRC e-Infrastructure: Portal Requirements" (NCeSS, Dec'2007) 29pp

[9] R. Crouchley, R.J. Allan, M. Fraser, M. Baker and T. van Ark "Sakai VRE Final Report" (JISC, 10/12/2007) 57pp

[10] M. Hayes, J. Kewley et al. "GROWL VRE Final Report" (JISC, 31/1/2007) 11pp

[11] D. Grose "High Throughput Distributed Computing using R: the multiR Package" Journal of Statistical Software (2008) submitted

[12] NGS Portal https://portal.ngs.ac.uk/

[13] Sakai Portal for NW-GRID http://rhine.dl.ac.uk:8080/portal

[14] RMCS User Manual

[15] RMCS Guide on NGS Wiki http://wiki.ngs.ac.uk/index.php?title=Category:Community_Software

[16] P.A. Couch, P. Sherwood, S. Sufi, I.T. Todorov, R.J. Allan, P.J. Knowles, R.P. Bruin, M.T. Dove and P. Murray-Rust "Towards Data Integration for Computational Chemistry" (UK e-Science All Hands Conference, Sept'2005) 7pp

[17] J.M.H. Thomas, R.P. Tyer, R.J. Allan, J.M. Rintelman, P. Sherwood, M.T. Dove, K.F. Austen, A.M. Walker, R.P. Bruin, L. Petit and M.C. Durrant "Science carried out as part of the NW-GRID project using the e-Minerals Infrastructure" (UK e-Science All Hands Conference, Sept'2007) 8pp

[18] M. Hayes, L. Morris, R. Crouchley, D. Grose, T. van Ark, R.J. Allan and J. Kewley "GROWL: a Lightweight Grid Services Toolkit and Applications" (UK e-Science All Hands Conference, Sept'2005) 8pp

[19] J.M.H. Thomas, J. Kewley, R.J. Allan, J.M. Rintelman, P. Sherwood, C.L. Bailey, A. Wander, B.G. Searle, N.M. Harrison, S. Mukhopadyhay, A. Trewin, G.R. Darling and A.I. Cooper "Experiences with different middleware solutions on the NW-GRID" (UK e-Science All Hands Conference, Sept'2007) 8pp

[20] J. Kewley, A.L. Braimah, R.J. Allan, M Hayes, P. Borsson, R. Crouchley, D. Grose, T. van Ark "GROWL Scripts: Lightweight Access to Grid Resources" (UK e-Science All Hands Conference, Sept'2007) 8pp

[21] D. Hughes, *AdaPtP - a Framework for Building Adaptable Peer-to-Peer Systems*, Ph.D. Thesis (University of Lancaster, August 2007).

[22] D. Hughes, P. Greenwood, G. Blair, G. Coulson, F. Pappenberger, P. Smith and K. Beven. "An Intelligent and Adaptable Grid-based Flood Monitoring and Warning System", Proceedings of the All Hands Meeting (2006). (Recipient of the best conference paper award)

WP 4 – Applications

A number of different applications have successfully used NW-GRID resources. A list of NW-GRID projects and users is included in Appendix 2. Flyers (double-sided A4 sheets) with case studies illustrating several of these applications are available from the portal. Rather than attempt an enumeration of these projects, this section provides some highlights of a selection of indicative projects. This information draws heavily upon that provided in several of the project flyers.

Molecular Electronics

Molecular electronics work at the University of Lancaster on NW-GRID is partly aimed at developing a state-of-the-art suite of simulation tools, capable of predicting material-specific properties of hybrid nano-structures. This project is producing a major leap forward, by forming a bridge between *ab initio* electronic structure calculations, nano-scale transport and molecular electronics. Devices and materials relevant to current and future industrial needs will be modelled and results will be used to initiate pre-competitive industry-based research projects.

This new capability will lead to a database of tight binding parameters for materials and their interfaces, which can be used to predict and optimise the properties of nano-structures of arbitrary complexity, e.g. Figure WP4-1.

In parallel with this over-arching activity, a number of new projects are being tackled, examples of which include:

- Developing a theory of superconducting and ferromagnetic proximity effects in carbon nano-tubes and hybrid metallic nano-structures;
- Developing a theory of how the proximity to step edges on a semiconductor surface can be used to organize nano-tubes;
- Understanding the effects of docking molecules on transport through nano-tubes;
- Modelling transport through single molecules of immediate interest to experimental programmes in QinetiQ and elsewhere, including conductance through mono-molecular arrays;
- Investigating quantum properties of polyacene crystals, which are currently the subject of intense worldwide investigation;

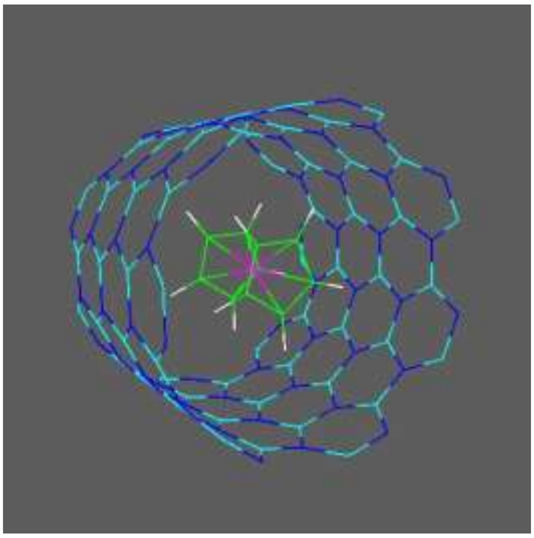


Figure WP4-1: The atomic positions of Cobaltecene inside an “armchair” carbon nano-tube.

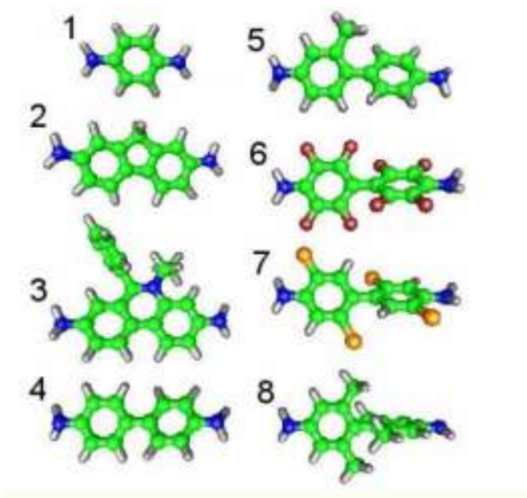


Figure WP4-2: Possible constituents of molecular wires reveal a dependency of conductance on angle of rotation.

In a first series of runs, the DFT code SMEAGOL (Spin and Molecular Electronics in an Atomically Generated Orbital Landscape) has been used to calculate the conductance through twisted single molecule junctions based upon bi- or tri-phenyl rings with amine link groups, see Figure WP4-2. The calculations to find the lowest energy configurations of various molecules takes some time to complete and the unique opportunity to use the full computing power of the NW-GRID to complete this initial stage was fortuitous as it gave the Lancaster group the edge (in terms of time) over most other research groups in the field. After initial tests to confirm that results were numerically equivalent across the Grid, we proceeded with the calculations, which were successful and now form a core of data which has been used to complete the research publication *Conformation dependence of molecular conductance: chemistry versus geometry* in the **Journal of Physics: Condensed Matter** (2008).

This work was carried out by Colin Lambert, Iain Grace, Victor Garcia-Suarez, Skon Sirichantaropass and Chris Finch from University of Lancaster.

Four researchers were principally involved in the job submissions, and they ran a variety of jobs of different sizes. Once a NW-GRID client gateway server with the necessary Globus components was written, the users found the Grid accessible and efficient, apart from the time taken to transfer files from the remote machine to the working site. Naturally once the initial files are transferred the process is relatively easy. This is improved even more by using a client toolkit such as GROWL. The main advantage as far as the NW-GRID is concerned has to be that there is a core of users who are familiar with job submissions and from this critical number it is likely that the entire group will use the Grid in the future.

One example of the impact of our work is the world-wide adoption of these non-equilibrium Green’s function methodologies, covering electron transport through nanoscale superconductors, magnetic multi-layers, carbon nanotubes, single molecules, atomic wires, nano-electromechanical systems, quantum dots and semi-conducting 2-DEGS. In part, this has involved the development of leading-edge techniques for *ab initio*, material-specific modelling of molecular-scale electronics, encapsulated in the electron-transport code SMEAGOL.

The highly-cited paper announcing this breakthrough in capability and establishing the new field of molecular spintronics appeared on the front cover of Nature Materials (see Figure WP4-4).

Figure WP4-3. Map showing a sample of the distribution of SMEAGOL users. The number of users is currently over 100.

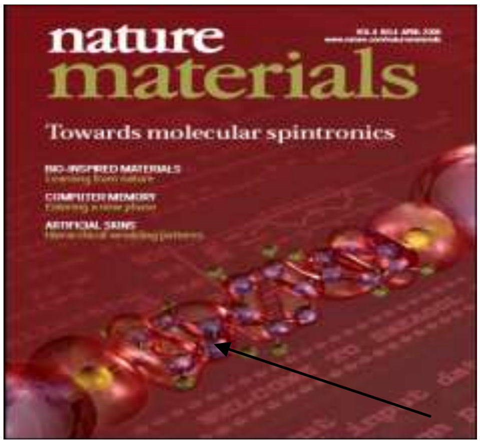


Figure WP4-4. The release of SMEAGOL in 2005, showing its welcome page appearing on the front cover of a journal.

Since 2005 the above paper and a sister paper giving technical details have collected over 170 ISI citations.

SMEAGOL represents a culmination of 20-years technical development at Lancaster. SMEAGOL is jointly developed by Universities of Lancaster, Oviedo (Spain) and Trinity College Dublin (Ireland). Further information about SMEAGOL and its academic license can be found at http://www.smeagol.tcd.ie. Since its release in 2005, SMEAGOL has been used by over 100 research groups spread over 29 countries; the 16 groups located in the USA include Harvard, Stanford, Cornell, Illinois, Minnesota, Northwestern, Ohio State and North Carolina (see Figure WP4-3).

Another example which highlights the continued use of the Grid to produce internationally acclaimed papers can be seen in a more resent publication. Here reported in PHYSORG.com in July 2008.

Carbon Nanotube Windmills Powered by "Electron Wind"

Theoretical physicists from Lancaster University in the UK have designed a nano-motor that operates by a novel mechanism: an electron wind.

As Steven Bailey, Ilias Amanatidis, and Colin Lambert explained in a recent issue of Physical Review Letters [3], the new drive mechanism could be useful for future NEMS (nano-electromechanical structures) technologies.

The researchers describe the nano-motor as a carbon nanotube (CNT) windmill, although the device looks more like a telescope than the conventional spinning blades of a windmill. It consists of a double-walled CNT, where the outer tube is clamped to two external electrodes, and the shorter inner tube is free to move and rotate. In a version called the CNT drill, the outer tube is clamped to just one electrode, while one end of the free inner tube is in contact with a mercury electrode, but is still free to rotate. The devices are called windmills because they're powered by an applied dc voltage between the electrodes, which produces a wind of electrons.

As it moves through the CNT, the flow of electrons acquires angular momentum, producing a tangential force that causes the inner nanotube to rotate. The researchers calculated that the electron wind can produce a force that significantly exceeds the inter wall friction – sometimes by as much as three orders of magnitude. Depending on the applied voltage, the rotating inner tube can reach CNT breakdown velocities of up to 8000 meters per second.

The electron wind powered nano-motor could have a range of applications. For instance, by using a voltage pulse to make the inner tube rotate at a specific angle, it could be used as a switch or memory element in nanoscale magnetic memory devices, or, by putting the CNT in contact with a reservoir of atoms or molecules, the nano-motor could act as a nano-fluidic pump.

The physicists also suggest that the motor could be powered in different ways. In one possibility, the electrodes could be replaced with reservoirs of atoms or molecules. Then, an applied pressure difference could drive the atoms or molecules across the CNT, and their angular momentum could cause the inner tube to rotate. Similarly, a temperature difference between the ends of the CNT could create a flux of phonons that could also drive the motor.

Other CNT-based nanomotors have been developed, including multi-walled CNTs with a similar structure to the CNT windmill. However, these previous designs have been powered by electrostatic forces that require metallic plates and gates, which aren't needed in the new design. The researchers hope that the efficiency and simplicity of the CNT windmills will provide advantages over electrostatic and other nanomotors in the future.

Funding for this work is from a variety of sources including, the EU Commission, EPSRC, DTI, Royal Society, NWDA, QinetiQ.

eMinerals

Results from eMinerals, a major project funded under the Natural Environment Research Council's (NERC) e-Science programme, are helping to solve two pressing environmental problems [1]. One finding could help to avoid arsenic contamination of drinking water extracted from man-made wells. Another could lead to improved methods of removing the now-banned industrial chemical, dioxin, from soil. The results were obtained using e-Science techniques and grid computing to simulate all the possible interactions between these contaminants and rock or soil.

Arsenic often appears in minerals rich in iron and sulphur, such as pyrite (fools' gold). Scientists working as part of eMinerals have found out precisely how arsenic is taken up and held in the pyrite structure and the factors likely to lead to its release. "We now know that arsenic replaces the sulphur in pyrite rather than the iron, and that pyrite is likely to dissolve more easily when arsenic is present," says Dr Kate Wright, who worked on the project. Further work could identify ways of stabilising arsenic-containing iron sulphide rock by introducing additives that slow the rate at which it dissolves.

The eMinerals project found that a dioxin molecule will bind more strongly to clay surfaces the more chlorine atoms it contains, irrespective of the position of the chlorine atoms in the dioxin molecule. It also found that binding is stronger the greater the electrical charge on the surface. However, water competes with dioxin to bind to surfaces and, in practice, a dioxin molecule's ability to bind to a surface is a balance between the binding strength of the dioxin to the surface, the water to the surface, and the dioxin to the water.

Both examples involved performing numerous simulations of the interactions between the different minerals in soil and rock with all the known variants of the contaminants. For example, there are 76 different variants of the dioxin molecule and numerous mineral surfaces in the environment to which they can attach, so hundreds of calculations are necessary, each of significant size.

The project has developed a grid infrastructure consisting of clusters and condor pools (including campus grids) at the collaborating institutions and resources held on the National Grid Service and the North West Grid. High performance computing resources can also be accessed for particularly large simulations if necessary.

Without access to such grid resources, researchers would have to perform all of the simulations sequentially, taking too much time to be practicable. Using the eMinerals infrastructure, they can submit all these jobs at once and see the results within a few hours. Results are automatically returned to a distributed data store with an interface that shows the files as if they are part of a single system. The data can be accessed remotely by collaborating scientists, as well as by those who originally submitted the job.

Example: Antiferromagnetism in $CaCuO_2$

$CaCuO_2$ is a prototype of the high temperature superconductors. In its pure form, it is an anti-ferromagnetic insulator. Upon doping with Sr, which introduces holes in the CuO_2 layers, it undergoes a Mott metal-insulator transition to a non-magnetic metallic phase, which occurs at a hole concentration of about 0.15.

A complete study of the properties of materials of this kind requires a very large number of separate calculations, including:

- scans over different configurations of localized and delocalized states;
- scans over the lattice constant;
- scans over different magnetic structures;
- scans over the temperature.

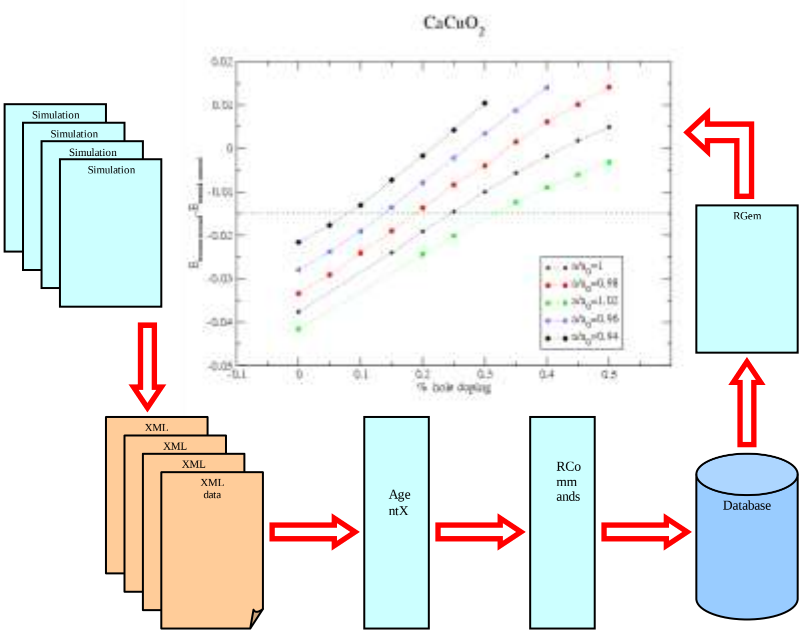


Figure WP4-5: Workflow for parametric studies of antiferromagnetism using RMCS

All results must then be analysed in order to extract the relevant information.
This sort of problem is an ideal example for the use of Grid technologies:

- (Semi-) automatic generation of input data;
- Meta-scheduling across suitable Grid resources;
- Matching requirements to available resources to minimise queuing time;
- Simulations can be run across multiple heterogeneous compute resources, using the abstraction provided by the Grid middleware;
- Both the input and output files are automatically archived within the data Grid, creating a complete audit trail for the calculation;

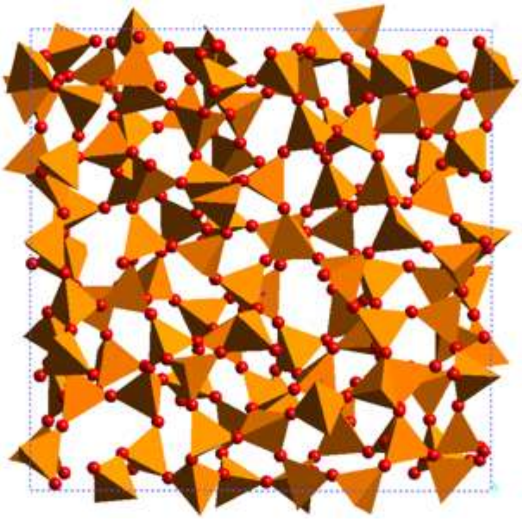


Figure WP4-6: Structure of silica glass

Key data values from the calculation are harvested from the output files and stored within a metadata database, which significantly facilitates the processing of data from a large number of related calculations.

Example: Explaining the Equation of State of Silica Glass

Experiments on silica glass (amorphous SiO_2) have revealed a number of unusual properties including negative thermal expansion and an increase in compressibility with increasing pressure - most materials increase in volume on increasing temperature and become harder to compress on increasing pressure. The key aim of the current work is to aim to explain these properties by performing atomic-scale simulations. We are probing the effect of temperature (T) on the behaviour of the model, to explore the change in position of the compressibility maximum on increasing T, and to see how the mechanisms leading to the compressibility maximum and negative thermal expansion interact with each other.

Key to this study is the ability to run many simulations under slightly differing P (pressure) and T conditions. We need many state points because we are interested in the derivatives of volume with respect to P and T, and for materials that are behaving in unusual ways we do not have the possibility of fitting a well behaved function to a small number of well spaced points. Along with the need to run many individual calculations, a task ideally suited to Grid resources, comes the problem of managing the torrent of output produced. Fortunately we have developed methods to do just this with the eMinerals Grid software deployed on NW-GRID was compatible with our existing tools. In effect we were able to plug the NW-GRID machines into our existing compute and data Grid and massively boost the compute power of our resources without significant modification to the job submission and data management strategy.

Bipedal Gait Analysis

A consortium of academics named APEMEN (Agent-based Predictive Environment for Modelling Expansion in the Neogene) based around NW-GRID is building a collaboratory to develop evidence-based models for research into primate and human evolution. They are firstly simulating locomotion and carrying out gait analysis to understand the energy costs and locomotor capabilities of extinct and extant species using metric evidence from fossil remains and observational field work. A high resolution movie of a virtual race among several bipeds is available for download from http://www.animalsimulation.org/bipedal. Figure WP4-7 is a screen shot from this movie. The consortium is also beginning to develop agent-based models of the Pleistocene Out-Of-Africa migration event that led the African hominins to spread to other parts of the world. Project investigators for applications of this work on NW-GRID are Professors Robin Crompton (Liverpool) and William Sellers (Manchester).

The scientific and technical objectives of work on NW-GRID are to obtain a better understanding of the mechanics of bipedal motion in both living and extinct species. For instance, given the fossil information about *Australopithecus afarensis*, can researchers determine the most efficient, and therefore most likely, mode of bipedal motion of this hominid? Similarly, what are the likely locomotor capabilities of bipedal dinosaurs?

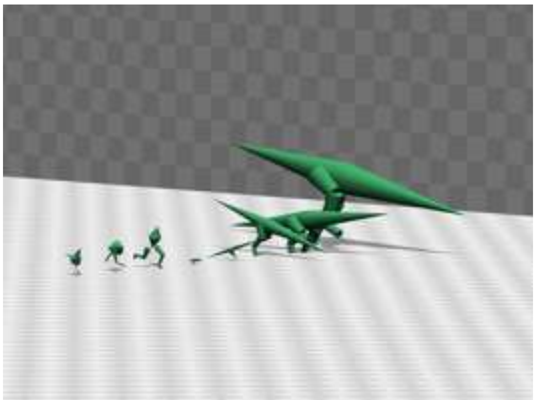


Figure WP4-7: A virtual race with (from right) a *Tyrannosaurus rex*, an *Allosaurus*, a *Dilophosaurus*, a *Velociraptor*, a *Compsognathus*, a human, an ostrich and an emu.

The modelling of bipedal motion relies heavily on genetic algorithms to perform necessary optimisation. This is computationally intensive. Furthermore, the computational requirements increase as better physics is added to the model. Application software is mostly in-house codes, although engineering packages such as Abaqus and ADAMS on HPC systems are also used.

Approximately 170,000 core hours of computation were performed in the latest computational study, which was reported on at a Society for Experimental Biologists conference in Glasgow at the end of March 2007 and in a paper published in the Proceedings of the Royal Society B in August 2007.

One of the co-authors of the Royal Society paper presented information about dinosaur gait on the Fossil Detectives television series (http://www.open2.net/fossildetectives/westwales.html).

Ship Airwake Modelling

One of the challenges that helicopter pilots face when landing on a ship is the air turbulence around the ship's superstructure. Prof. Gareth Padfield at the University of Liverpool is leading a team of researchers to model this turbulence using the Fluent computational package in order to understand the phenomenon better and to provide a better representation of the turbulence in helicopter training simulators.

The detailed modelling of airflows around a ship is computationally demanding and produces large volumes of data. To make this modelling feasible for Prof. Padfield's team, parallelism at several levels must be exploited. Firstly, the Fluent package itself is parallelised, so that parallelism within a model can be exploited. This is essential as each simulation performs 1500 time steps, representing about 15 seconds of simulation time and takes roughly 4500 minutes of real time (just over 3 days) on 32 cores - 8 nodes to complete. The second level of parallelism arises because for a given model, the simulation is repeated to cover 28 different wind angles. Given sufficient computing capacity, multiple wind angles can be studied simultaneously. Approximately 300 GBytes of data are generated for each simulation, so naively storing the raw output requires nearly 9 TBytes of storage.

The NW-GRID, with over 1500 cores on 4 sites, has been approached to provide an effective and efficient solution to resolve the computational challenges of Prof. Padfield's research team. The challenge for the NW-GRID is to deal with the data volume effectively and to reduce the throughput time enabling the scientists to work more efficiently.

Data requirements

While the data requirements for a set of simulations are large, the data are not of uniform importance.

About 1 GBytes is taken up with input files - these are copied over from an Engineering system and are static during a run. About 6 GBytes of data are generated every few hours to provide a restart basis for the code in the event of a system crash. Copies of this data are useful while the job is running but can be purged upon job completion.

As mentioned earlier, 15 seconds of simulated time generates about 300 GBytes of output, at the rate of 200 MBytes of data every time step. These files are simple text files. Once timestep files are created, they are compressed (and then uncompressed as required in the post-processing stage). The compressed raw data needs to be archived in case the post-processed data needs to be recomputed for a different structured mesh (see below) and also for audit purposes. Compression reduces the data volume by a factor of about 2.5.

About 50 GBytes of data is generated at the end of each run by interpolating the raw data from a simulation onto a coarse, structured mesh. This data is downloaded onto data storage drives within Engineering and is the essential simulation output. Maintaining a backup copy of this data is highly desirable (so about 1.4 TBytes in total, but built up over time - at most 100 GBytes every 3 days assuming two simulations are run at once).

For the restart data, a separate background job runs every 12 hours to copy the latest restart data, so the restart data files can be overwritten.

Another background script is run hourly to copy over the raw files and to compress them. Once the compressed files have been backed up, they too are deleted, although a formal archive process would be desirable. If there is sufficient disk space available, the 28×(120+50) = 4.8 TBytes of results data can simply be kept on disk, protected by a backup.

Improving throughput time
With its novel approach of science testbeds, NW-GRID is able to allocate sufficient resources to Prof. Padfield's team to complete 28 runs in 6 days, while still meeting NW-GRID commitments to other projects.

Future Plans
A simulated 15 seconds is really the minimum suitable time – it just covers the period as the helicopter is coming down to land. A simulated two minutes (i.e. 8 times longer) allows the full final approach sequence to be modelled. However, this has important implications. Firstly, there is an 8-fold increase in storage requirements – from 5 TBytes storage to just under 40 TBytes of storage. A refinement in what raw data is kept and for how long is necessary. The essential data output would be about 11 TBytes in size, which provides a minimum storage requirement.

Another desire is to reduce the time for a simulation to one day. This requires a 24-fold increase in parallel performance, so that something like 800 cores would be involved in each simulation. This might be possible running Fluent on the Daresbury Blue Gene system (where 4096 cores are available). Feasibility tests using Blue Gene and the 400 cores of Liverpool's NW-GRID High Capability cluster will be conducted soon.

Flood Monitoring

Flooding is a growing problem in the UK that has a significant effect on residents, businesses and commuters in flood-prone areas. The cost of damage correlates closely with the warning time given before a flood event, which makes flood monitoring and prediction critical to minimizing the cost. Grid computing is being used to reduce these problems.

Current approaches to flood monitoring (e.g. in river valleys) involve statically deploying depth and ultrasound-based flow sensors across flood-prone areas, and feeding the collected data to hydrologists off-site as input to Grid-based computational models which predict flooding trends. At Lancaster University, we believe that there is considerable scope for improvement in such scenarios. In particular, we are investigating selectively shifting the execution of prediction models to the sensor network itself, which thus acts as a 'mini-grid'. Computations organised in this way not only provide timelier flood warnings, but also help to adapt the wireless sensor network (WSN) dynamically and thus optimise it for current or predicted environmental conditions. For example, the network can employ a low power, low throughput organisation in quiescent conditions, and switch to a high power, high throughput organisation when flooding is imminent.

To achieve this vision we have developed a sensor network framework based on an appropriate combination of software and hardware. The software consists of our component-based run- time reconfigurable GridKit middleware.

Figure WP4-8: model used to demonstrate GridStix functionality

This provides flexible networking support, service binding, resource discovery, resource management, and security. Our hardware platform uses the Linux-based Gumstix embedded computing platform [2]. We call the combined hardware and software framework “GridStix”. This is capable of integrating with remote fixed-network Grids such as NW-GRID for computationally intensive flood modelling purposes, and also capable of performing on-site flood modelling. This combination yields significant benefits, for example, local computation can be used to provide timely warnings to local stakeholders, and a combination of local and remote computation can inform adaptation of the sensor network to maintain optimal performance in changing environmental conditions. Fully predictive models can them be run on the NW-GRID using input which has been collected and uploaded in real time.

Figure WP4-8 shows a model used to demonstrate GridStix technology. This has been shown at a number of workshops and demonstrates, in real time, the overlay network topology being maintained by the GridStix modes.

Figure WP4-9: GridStix deployment at Cow Bridge

The first deployment site is at Cow Bridge, on the River Ribble in the Yorkshire Dales. This site is prone to flooding for much of the year and thus offers good potential for evaluating the system under real-world conditions. Flooding affects the nearby village of Long Preston, which thus additionally presents us with a motivation for evaluating warning systems for local stakeholders.

Traditional flood monitoring approaches impose a rigid separation between the on-site WSN used to collect data, and the off-site computational Grid that is used to analyze this data. Our GridStix approach provides more on-site “intelligence” with next generation Grid middleware used to realize our goals.

This work is by Danny Hughes, Phil Greenwood, Barry Porter, Paul Grace, Geoff Coulson, Gordon Blair and Francois Taiani of the Computer Science Department and Florian Pappenberger, Paul Smith and Keith Beven of the Environmental Science Department, both at University of Lancaster [2].

Documentation on the NW-GRID Portal

[1]	Jens M. H. Thomas, Rik Tyer, Rob J. Allan, Jamie M. Rintelman and Paul Sherwood, Martin T. Dove, Kat F. Austen, Andrew M. Walker, Richard P. Bruin, Leon Petit, Marcus C. Durrant Sc*ience carried out as part of the NW-GRID project using the eMinerals infrastructure*

http://rhine.dl.ac.uk:8080/access/content/group/NW-GRID%20Review/WP4%20-%20Applications/Selected%20Publications/rmcs_and_science.pdf

[2] Danny Hughes, Phil Greenwood, Gordon Blair, Geoff Coulson, Florian Pappenberger, Paul Smith and Keith Beven *An Intelligent and Adaptable Grid-based Flood Monitoring and Warning System* http://rhine.dl.ac.uk:8080/access/content/group/NW-GRID%20Review/WP4%20-%20Applications/Selected%20Publications/639.pdf

[3] Bailey, S.W.D; Amanatidis, I.; and Lambert, C.J. *Carbon Nanotube Electron Windmills: A Novel Design for Nanomotors*. Physical Review Letters 100, 256802 (2008)

WP 5 – Knowledge Exchange

WP 5 covered a range of important activities:

- the maintenance and support of the NW-GRID web site;
- the organisation of technical, industrial and scientific seminars/ workshops across the region;
- the provision of training courses for users as well as input into formal post-graduate course material – this was done in association with the NGS Training, Outreach and Education group;
- project marketing and business outreach, including working with the NGS publicity officer;
- sustainability activities, with a particular emphasis on obtaining industrial collaboration.

Further details about these activities can be found on the NW-GRID portal, under NW-GRID International Review Resources/ Work Package 5 - Knowledge Exchange

The list of seminars and training courses that NW-GRID hosted or was actively involved with included:

- ReDReSS (Resource Discovery for Researchers in e-Social Science) – Quantitative Methods in e-Social Science – April 2005
- Middleware and Clusters for Grid Computing – Dec 2005
- NCeSS (National Centre for e-Social Science) Training School – Feb 2006
- Bioinformatics and the Grid – March 2006
- Industrial requirements of the grid – May 2006
- Lightweight Middleware – May 2006
- NW-GRID Training course – Jan 2007
- Materials Modelling on NW-GRID – Mar 2007
- Campus Grids – Oct 2007
- NCeSS Portability Usability Workshop – Nov 2007
- MEW Industrial Day – Nov 2007 and Dec 2008
- Computational Steering – March 2008
- NGS Innovation Forum 2008 – November 2008
- JISC Advanced tools and technologies for collaborative research – November 2008
- MEW19 Industry Day –Dec 2008

In addition, project members gave presentations at several workshops and conferences, including two NW-GRID specific presentations at the All Hands Meeting in 2007. Other presentations from individual projects using NW-GRID, such as e-Minerals, Flood Monitoring, VRE and GROWL where work was being undertaken to develop middleware and services, were presented at this and previous All Hands events.

To improve the efficiency of keeping the NW-GRID systems updated working with the project's main provider, NW-GRID spearheaded the setting up of the Streamline Computer User's Group (SCUG) in 2008. All Streamline Computing user sites with "large" clusters were invited to join. There have been three meetings so far, with another scheduled for September 2009. All the main UK academic sites have joined or expressed an interest to join SCUG. It is hoped that the group will provide an ongoing opportunity for users and Streamline Computing to work together to address difficulties in support and procurement, to provide a forum where Streamline can present their future plans, a means to enable collaboration between Streamline and its customers and for all sites to share information, hints and tips for efficient system management.

We note in a similar vein that all the NW-GRID sites actively contribute to the UK HPC Special Interest Group which comprises all academic groups managing large computer resources and the newly formed Campus Grid SIG which is promoting better utilisation of the whole spectrum of compute resources in HEIs, including addressing issues of green computing.

Marketing efforts include the development of A4 flyers that described different NW-GRID projects (available from NW-GRID International Review Resources / Work Package 5 - Knowledge Exchange / Academic Case Studies) and basic project information. Support materials for exhibition materials included rolling project presentations used at OGF 20 and AHM 2007 as well as A0 posters and an extremely popular hands-on demonstration of the flood monitoring work. In late 2008, an A5 brochure on the project was completed and will be used to provide background information to potential commercial collaborators.

Within the institutions involved, internal marketing is slowly leading to a culture of joint working across departments with emerging Campus Grids. For instance at Daresbury, the HPCx service can be accessed (subject to peer review) and a range of novel architecture clusters (Intel Nehalem and Harpertown, ClearSpeed, IBM Cell and nVidia GPU) have been added. Other clusters at Lancaster, Liverpool and Manchester plus Condor workstation pools are also available. There are ongoing discussions with NorthGrid and White Rose Grid about federating or sharing resources.

Marketing has been successful in recruiting a number of high-profile academic users to the system including: VEC, the Northwest Virtual Engineering Centre; the Cockroft Centre; N8 Materials Consortium; the Centre for Materials Discovery; the Knowledge Centre for Materials Chemistry; and the SuperSTEM microscope facility. The Proudman Oceanographic Laboratory became an early partner site co-locating one of its clusters (Intel Xeon system) with other NW-GRID resources at Liverpool. Most recently, and following extended discussions, University of Central Lancaster (UCLAN) agreed to become a partner site bringing an SGI Altix system as a new component to NW-GRID.

As suggested in the Overview section, sustainability was and remains an important project objective. It is a key criterion in NWDA's assessment of the project's success. Industrial collaboration is seen as central to this sustainability in addition to growing and maintaining the Grid infrastructure itself. Whilst there have been several successes in this area, the road to overall NW-GRID sustainability has proven more difficult to negotiate than was envisioned at the start of the project. The salient reasons for this difficulty, together with a strategy towards industry that should address some of the problems encountered are presented in the latest version of the NW-GRID Business and Development Plan. Information available to us suggests that similar difficulties have been met by other Grid initiatives, but that our mode of using the Grid resources as a facility for provision of skilled staff, applications and data in joint research projects will be successful in the longer term.

NW-GRID – Continuing Commitments and follow-on Funding

All four sites have continuing commitments to see NW-GRID grow and be exploited. This includes institutional commitment to fund staff working on the systems, cover necessary room and electricity charges, to add more resources and to bid for funding for project work.

Daresbury:
Will keep the core service running, including paying for some percentage of operational support (Tim Franks) to run it. Will keep e-Science software (GROWL and RMCS) running and pay for some percentage of a developer's time (currently Dr. R.J. Allan). Will keep Sakai portals running for NW-GRID projects, pay for the electricity and room charges, contribute other resources to NW-GRID, i.e. BlueGene-P and several newer clusters as listed above. We have an agreement with our knowledge transfer organization CLIK that they will match funding for small commercial projects using NW-GRID as a pump-priming exercise. There are also opportunities to obtain STFC funding to set up KTPs. We aim to grow and exploit a Daresbury Science and Innovation Campus Grid (including Cockroft Institute and SMEs, not just Laboratory). All staff involved in NW-GRID will contribute to Hartree Centre development and Cloud infrastructure.

Follow-on Funding:

- Hartree Centre for Computational Science, value: £50m from the DBIS Large Facilities Capital Fund (TBD early 2009, now postponed until late 2009)
- Knowledge Centre for Materials Chemistry with Liverpool and Manchester (see below)
- Centre for Earth Systems Intelligence (see below)

Liverpool:
Liverpool in January 2008 hired Dr. D. Love on a 2-year contract to support and help manage the clusters bought with NW-GRID money. Dr. C. Addison's job remit includes user support and expanding the use of this system across the University. That is unlikely to change. Dr. I. Smith has developed the University Campus Grid and has recently enhanced its facilities to make it easier for users to access all NW-GRID, as well as NGS, facilities through a single system interface. Both of the latter posts are permanent positions.

Generally Liverpool is committed to keeping the core NW-GRID service running until April 2011 (3 years from last procurement). This includes the room costs and cooling maintenance. The University centrally is paying electricity costs.

A researcher added 23 nodes to this core service in October 2007, mainly, but not exclusively, for his own group's use.

The University is leasing some nodes to research groups over 3 years starting October 2008 and part of that lease agreement covers system administration support costs, while most of that payment will contribute to a recharge pool for replacement systems in the future.

Follow-on Funding:
- Centre for Materials Discovery, Jan'2007, value £9.6M (see below)
- VEC: Virtual Engineering Centre, 1 July 2009, value: £3M (see below)

Lancaster:
Dr. Mike Pacey will continue to maintain and run the NW-GRID clusters. The University has agreed to contribute room and electricity charges. The University is also about to embark on a major hardware procurement, drawing on the lessons learned in NW-GRID and upon the expertise built in Lancaster during the course of this project.

Follow-on Funding:

- Divergent Grid Project, EPSRC funded, October 2005, value: £256,644
- Open Overlays Project, EPSRC funded, October 2005, value: £483,096
- FREE Project, NERC funded, October 2008, value: £87,140
- WISEBED Project, EU funded, June 2008, value: 305,415 Euros
- N8 Materials Consortium bid (funded , see below)
- CRIB, Collaborative Research in Business with Daresbury – JISC funded project from April 2009

Manchester:
Following the establishment of Research Computing Services (now part of ITS) Manchester are exploiting NW-GRID resources as part of their growing Campus Grid and are seeking to unite activities across many university departments under a common umbrella so that operational expertise can be shared and economies of scale exploited. The recent discussion on cloud/ grid etc. computing fits into this model. The University has agreed to contribute room and electricity charges.

Research Computing Services also hosts nodes of the National Grid Service, including an Oracle database node. This continues to be supported and hosts the NW-GRID VRE (Sakai portal) database and backups.

Follow-on Funding:

- EPSRC platform grant proposal with Daresbury and Lancaster (not funded).
- ESRC e-Infrastructure Project, coordinated by NCeSS, value: £544,450
- Centre for Innovation through Materials Science, Chemistry and Engineering. Manchester University in partnership with Lancaster University have received a £5m grant to research applications of graphene - a new form of 'super carbon', discovered in the UK in 2004.
- National e-Infrastructure for Social Simulation with Daresbury, £1.7M JISC funded (including other sites) from 1st April 2009 (see below)

Active NW-GRID-related Projects

Most of these projects have funded one Research Assistant, and a number of Ph.D. studentships have also been funded. Some funded considerably more.

The N8 Molecular Engineering Translational Research Centre (METRC). METRC, funded as a Northern Way initiative, is a virtual laboratory pooling expertise from leading research centres across Universities in the North of England. The focus is on soft nano-technology and its applications in UK industry. Together, the partners deliver a strategic R&D services needed to drive the research agenda in this area and stimulate economic growth.
METRC makes it possible for academic and industry researchers to work alongside each other to deliver research into innovative new projects using a combination of experimental and modelling methodologies. This combined strength means that METRC partners can tackle challenges and capitalise on opportunities beyond the scope of any single one.

The Centre for Materials Discovery (CMD) was opened in January 2007 and was funded through a combination of £9.6 million support from the University of Liverpool, the NWDA, EU Objective 1 and industry. Located on the University of Liverpool city campus and adjoining the Department of Chemistry, CMD acts as a focal point for the high throughput (HT) activities of a broad community of research scientists with wide ranging expertise and know how. A multidisciplinary team of professional scientists supports and delivers research in the Centre, facilitating the discovery and optimisation of new materials for future commercial exploitation. CMD is targeted towards: Enhancing competitiveness and innovation within business and academia through application of HT technologies; Widening access to specific HT expertise; Strengthening the cross sector relationships and interactions within the industrial and academic communities

Knowledge Centre for Materials' Chemistry (KCMC). A project involving the Universities of Bolton, Liverpool and Manchester and STFC Daresbury (CSED) that was launched officially in March 2009. With approximately £15M in funding from the NWDA (almost £500k to CSED) over the next 3 years, KCMC is a one-stop-shop for materials' R&D services for NW England-based companies. The NWDA funding is to upgrade the resources of the partners and to subsidise the cost of the projects. It will also fund a Knowledge Transfer Team, tasked with selling R&D projects to industry across the Region. CSED has committed some £200k in-kind contribution via free access to the NW-GRID node here at Daresbury (incl. the Blue Gene P).

The Centre for Earth Systems Intelligence. A project involving IBM, CSED and Lloyds Insurance, this was begun in November, 2008. This is intended to study the Earth's various systems that contribute to weather and climate events. This will utilise the Daresbury node of the NW-GRID and Lloyds has provided £300k to allow "free at the point of use" access to the node for projects in the Centre's portfolio.

The Virtual Engineering Centre (VEC). This is a project initiated by the Faculty of Engineering at the University of Liverpool over eighteen months ago. The prospect started 1 July 2009 and is currently recruiting. Based in the A Block at Daresbury, the ‘real’ Centre of the project will have up to 16 individuals working on systems, methods, tools and resources for virtual engineering ("virtual prototyping"). The NWDA is providing £3M over 3 years to help set up this Centre and to facilitate access for NW-based companies. The primary objective is to enable the Region to remain a top class supply base for the aerospace and other high-tech engineering sectors by providing a Centre of Excellence for accessing skills and resources and a steady stream of qualified graduate and postgraduate engineers skilled in the art of virtual engineering. CSED has committed in kind contribution to the project, again by providing free at the point of use access to the NW-GRID node here (incl. the Blue Gene P).

National e-Infrastructure for Social Simulation (NeISS). The NeISS project has been funded by the JISC as part of its Information Environment Programme. It will provide a platform to meet the demand for powerful simulation tools by social scientists, public and private sector policymakers. Social Simulation is an expanding field due to its forecasting applications for scenarios in transport, housing, education, healthcare etc. The NeISS project will build a generic production quality social simulation e-Infrastructure covering the social simulation lifecycle. It will introduce social scientists to new ways of thinking about social problems, and provide new services, tools and research communities to support them. It will provide an e-Infrastructure framework capable of being deployed for a diverse range of social research domains (i.e. not limited to simulation). The tools will enable researchers to create workflows to run their own simulations, visualise and analyse results, and publish them for future discovery, sharing and re-use. This will facilitate development and sharing of social simulation resources within the social science community, encourage cooperation between model developers and researchers, and help foster adoption of simulation as a research method in the social sciences, and as a decision support tool in the public and private sectors.

WP 6 – Management

NW-GRID has three levels of project management.

The top-level group, the Project Board, has a remit that includes:

- To set and agree the strategy for the project;
- To provide direction in issues relating to foreground and background IPR and exploitation of the technology developed in the project;
- To coordinate the testbeds, workshops and external project reviews;
- To coordinate information dissemination and interactions with related projects;
- To provide advice to the project manager relating to the partners' wishes on project management;
- To resolve matters that cannot be resolved at lower levels in the project.

The Project Board tries to meet every four months and is chaired by the Project Director, who is currently Prof. Colin Lambert, Lancaster University. It is important to notice that one Project Board representative from each partner was also a Technical Board member. The frequency of Technical Board meetings meant that Project Board activities could be sustained through the Technical Board with e-mail discussions among Project Board members.

The Technical Board is more involved in day-to-day project activity. Its original remit was to agree the NW-GRID software environment and to agree software quality assurance standards, applications and Grid resource requirements. In addition the Technical Board is the mechanism to ensure progress against the work plan on each work package. Work package leaders are responsible for the production of all deliverables within their associated list of tasks, and report progress against the plan, resource usage and problems to the Technical Board. The Technical Board's remit was expanded over the course of the project, to include policy-related activities such as approving project requests to use NW-GRID resources. Many of the details relating to testbeds, Grid resource and applications deployment were deferred to the Operations Board. Procurements were handled jointly between the Technical and Operations Board, with overall approval of recommendations being made by the Project Board.

The Technical Board originally planned to meet monthly. However, once the Access Grid facilities at each partner site became available in 2005, this board met fortnightly. Up until the end of February 2008, the Project Director was also chair of the Technical Board. Since then, this task has been delegated to the Project Manager, who is currently Dr. Cliff Addison from Liverpool University.

The Operations Board was established to deal with the operational details of firstly the prototype systems and then, from April 2006 onwards, the operational details of the production systems. This included things such as the allocation of project and user ids, the functioning of the Web site, security and systems updates as well as Grid and application deployment. Joining procedures and conditions of use were worked out in the Operations Board for final approval by the Technical Board. Throughout the project, the chairman of the Operations Board has been Dr. Rob Allan, STFC Daresbury. Dr. Addison is also an Operations Board member and the overlap of Board members has contributed positively to the running of the project. Agreement of a Service Level Description with the NGS was handled by members of both Technical and Project Boards.

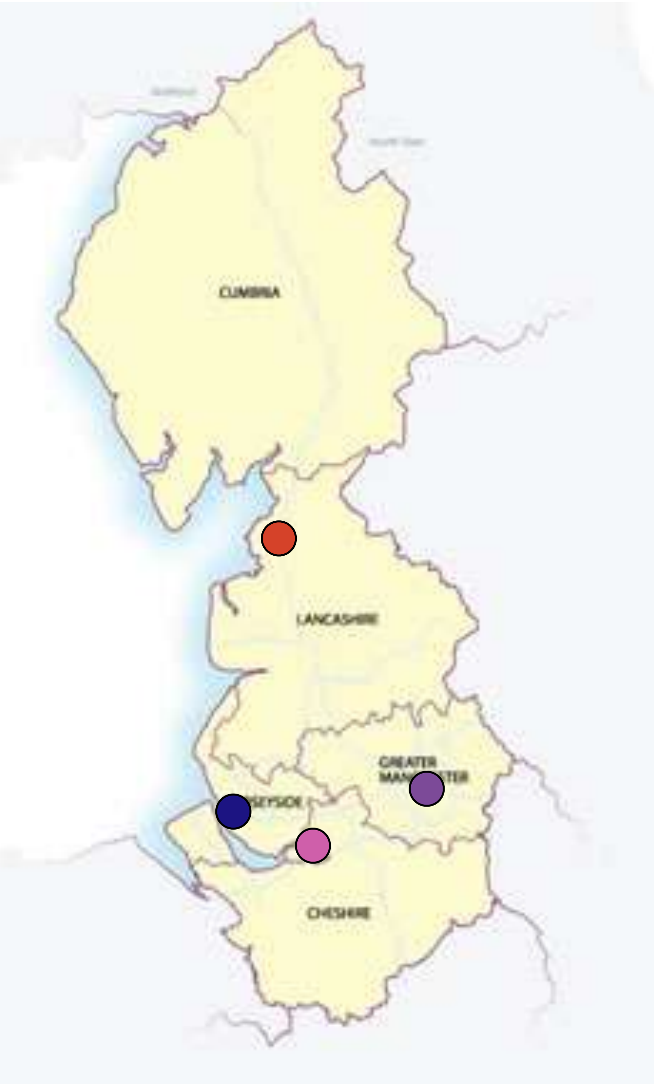
Initially, the Operations Board met every three weeks. However, once the procured systems arrived in 2006, meetings became fortnightly (once again via Access Grid) and have remained so until the present. The operational details for NW-GRID systems tend to be relatively complicated and all the partners cooperated to share best practices and to keep each other informed of important changes and developments. The Operations Board e-mail list is the most heavily used of all of the NW-GRID lists. It is fair to say that the Operations Board demonstrates the benefits of multi-site systems in a project in that the communal systems support is considerably more than just the sum of that support at each partner site.

Decisions in both of the lower tier boards were almost always arrived at by consensus. At no point were disputes contentious enough to warrant their promotion to the Project Board.

Appendix 1 - Background information

About England's Northwest – from NWDA website

England's Northwest is a fast growing, vibrant region, combining a dynamic business base, cosmopolitan urban centres, breathtaking landscapes and an internationally recognised creative and cultural scene.



It boasts Manchester, the largest media hub outside of London; Liverpool, one of the world's most famous waterfronts; Cheshire, home to AstraZeneca's largest global Research & Development premises; Lancashire a world class centre of excellence in advanced manufacturing and engineering; and one of Europe's leading national parks, the Lake District.

With almost seven million inhabitants and 230,000 companies, the Northwest is a thriving economy. It is worth a remarkable £106 billion - a tenth of the overall UK GDP. It is the UK's largest regional economy, larger than several European countries, including Denmark and Finland, and it is one of only three regions to contribute positively to the UK's balance of trade.

And it's not just excellent for business. A huge 29% of the region is designated as National Park or area of outstanding beauty. The Northwest boasts the largest lakes and mountains in England, not to mention the longest stretch of undeveloped coastline.

These rich natural assets are part of the reason why England's Northwest has been voted the top location for quality of life in the Reward Group's cost of living survey for two years on the run, while a poll by Ipsos MORI in 2006 rated the region highly on areas of outstanding beauty, culture and nightlife.

It is an unbeatable combination of commercial dynamism, cultural diversity and natural beauty that is helping to make England's Northwest an internationally-recognised success story.

Appendix 2 – Projects using NW-GRID

Biology and Bio-Informatics

Bioinformatics - Dr. Robert Stevens (Manchester)

Biological Modelling - Dr. Magnus Rattray (Manchester)

Bipedal Gait Analysis - Prof. Robin Crompton (Liverpool). Studies of homonin and dinosaur movement

CCP4 – Dr. Martyn Winn (Daresbury). Development and applications of high-performance software for protein crystallography.

e-HTPX: e-Science Resources for High-Throughput Protein Crystallography - Dr. Rob Allan (Daresbury). BBSRC/ DTI funded structural biology pipeline from gene expression to 3D protein structure analysis. Now finished, but some services have migrated on the the Diamond Synchrotron.

e-Fungi e-Science Infrastructure for Comparative Functional Genomics (Manchester) - database design and management

Chimatica Ltd. (Manchester) - Drug design lifecycle using myGrid and Taverna workflows.

Integrative Biology – Kerstin Kleese van Dam (Daresbury). BBSRC-funded project Investigating the causes of heart disease and cancer.

Computational Chemistry

Optical Properties of Materials - Dr. Paul Sherwood (Daresbury)

Modelling Polymers for Hydrogen Storage – Abbie Trewin (Liverpool)

Bio-molecular Chemistry - Dr. Richard Bonar-Law (Liverpool)

Quantum Directed Virtual Evolution – Dr. Marcus Durrant (Newcastle upon Tyne) hosted at Daresbury. Studies of catalyst behaviours using genetic algorithms.

Astra Zeneca and CCP5 – Dr. C. Yong (Daresbury). Complexation of cyclo-dextrins.

Stability properties of complex metal oxides – Dr. George Darling (Liverpool)

Computational Materials

Exhaustive Phase Space Search of Atomic Configurations - Dr. Gabor Csanyi (Cambridge) - host site Daresbury

Solid State Physics - Prof. Werner Hofer (Liverpool)

Ab initio Materials Modelling and CCP3 – Christine Bailey (DL)

Engineering

Ship Air Wake Studies - Prof. Gareth Padfield (Liverpool)

MACE - Dr. Juan Uribe (Manchester). Development and testing of Code Saturne for engineering applications

Medicine

GENIUS: Grid Enabled Neurosurgical Imaging using Simulation – Dr. Rob Haines (Manchester). Real-time simulation and visualisation using co-allocated Grid resources.

Natural Environment

e-Minerals: Environment from the Molecular Level - Dr. Rob Allan (Daresbury). NERC-funded project with University of Cambridge simulating mineral properties to better understand how to control pollution and encapsulate nuclear waste. Now finished, but the RMCS software which was developed is still being used on NW-GRID and the NGS.

Global Coastal Ocean Modelling - Dr. Jason Holt (POL) - host site Daresbury. Studies of shelf-sea environments around continents.

Unified Model Centre of Competency – Dr. Stephen Pickles (Daresbury), joint with UK Met Office.

Rapid Portals for Seismological Waveform Data – Dr. Jano van Hemert (Edinburgh) – host site Liverpool.

Physics

Hadronic Physics - Prof. Alan Irving (Liverpool)

GridPP - Dr. Mike Houlden (Liverpool). Analysis of particle physics data.

Nano-Technology - Prof. Colin Lambert (Lancaster). Design and simulation of hybrid nano-structures

NW Accelerator Grid - Dr. Hywel Owen (Cockcroft Institute) - host site Daresbury. Multi-particle physics simulation for accelerator design and beam steering.

Exploring the Multiverse: Producing a suite of Cosmological N-body Simulations - Prof. Bob Nichol (Portsmouth) - host site Daresbury.

Simulations of Antihydrogen Formation - Dr. Svante Jonsell (Swansea) - host site Liverpool

Monte Carlo Simulation for Radiotherapy applications – Dr. Colin Baker (Liverpool) – host site Liverpool. Joint with Clatterbridge Centre for Oncology.

Statistical Analysis

CQeSS: Collaboratory for Quantitative e-Social Science - Prof. Rob Crouchley (Lancaster).

Efficient Stochastic Process Models - Dr. Neil Lawrence (Manchester)

Synthetic Small Area Estimates - Dr. Paul Hewson (Plymouth) - host site Lancaster

Text Mining and Information Retrieval - Prof. Paul Watry (Liverpool)

Sabre/ SabreR - Prof. Rob Crouchley (Lancaster). Statistical analysis software for longitudinal data sets used in the ESRC-funded CQeSS node.

ESRC e-Infrastructure - Dr. Rob Allan (Daresbury). ESRC-funded project coordinated by NCeSS, the National Centre for e-Social Science, in Manchester.

National e-Infrastructure for Social Simulation (NeISS) – Prof. Mark Birkin (Leeds).

Software and HPC Developments

ACET Software Development - Prof. Mark Baker (Reading) - host site Daresbury

Daresbury Science and Innovation Campus Grid - Dr. Rob Allan (Daresbury)

Daresbury Local - Dr. Rob Allan (Daresbury). Installation and testing of software and applications

DisCo: Distributed Computing Support Programme - Jonathan Follows (Daresbury). EPSRC-funded project is making novel cluster-based resources available for evaluation via NW-GRID.

Diamond e-Infrastructure – Dr. Rob Allan (Daresbury). Development and testing of middleware and applications for analysis of data on the Diamond synchrotron.

Grid Markets – Dr. Steven Newhouse (Imperial), host site Manchester. Investigating accounting, negociation and payment for Grid services. Project finished.

GROWL: Grid Resources On Workstation Library - John Kewley (Daresbury). JISC-funded project has now finished, but software now forms part of G-R-Toolkit and the GROWL Scripts are in use on the National Grid Service.

UK OGSA Testbed – Prof. Mark Baker (Reading), host site Daresbury. EPSRC-funded middleware evaluation project now finished.

MultiR - Prof. Rob Crouchley (Lancaster). R scripting software based on GROWL and used in the ESRC-funded CQeSS node.

Application Benchmarking - Dr. David Baker (Southampton) - host site Daresbury

e-CCP - Dr. Phil Couch (Daresbury). STFC-funded project finished, developed the AgentX toolkit for meta-data and semantic interoperability of computational science applications.

Sakai VRE - Dr. Rob Allan (Daresbury). Developing portal services for e-Research, including the NGS, NCeSS, NW-GRID, N8 and Hartree Centre Portals. Original JISC-funded project has finished.

GridKit: Advanced Middleware Infrastructure Support for the Grid - Prof. Gordon Blair (Lancaster)

OpenORB: Next Generation Middleware - Prof. Gordon Blair (Lancaster)

Visualisation – Dr. Richard Blake (Daresbury)

Workflow Optimisation for e-Science – Prof. David Walker (Cardiff) host site Daresbury. EPSRC-funded project now finished.