White Rose
university consortium
Universities of Leeds, Sheffield & York

The White Rose Grid Experience

Report Commissioned by JSCR

September 2005

Dr Joanna Schmidt, White Rose Grid e-Science Coordinator

Dr Sarah Clark, White Rose Project Officer

Executive Summary

The White Rose University Consortium (White Rose) is a strategic partnership between Yorkshire's leading research Universities of Leeds, Sheffield and York. In 2000 the Vice-Chancellors of the three Universities decided to combine computing resources using the emerging grid technology. The White Rose Grid (WRG) was launched in 2002 and White Rose considers it one of its most successful and important initiatives. This report on our unique experiences in creating the WRG was commissioned by the Joint Information System Committee (JISC) for the Support of Research (JSCR). The WRG was chosen because of its pioneering role in the establishment of a multi-campus regional grid. This report provides an overview of the aspects of the WRG that might be of value to other institutions that are considering a similar approach and wish to benefit from our experience.

The WRG builds on the international expertise of our Universities to deliver grid and High Performance Computing (HPC) services based on integrated computational and data resources of the three institutions. Our innovative infrastructure enables close research collaboration both locally and globally. The WRG operates as a virtual organisation with dedicated resources in each institution. It is supported by a combination of researchers and Computing Services support staff; personnel from different institutions undertake complementary roles. The organisational structure continually evolves to meet the developing needs of our users and the grid.

Our distinct approach to building the WRG has been to bring together the provision of HPC services with emerging grid technology. We decided to use the existing grid technologies to integrate resources newly acquired by the three Universities. WRG activity successfully supports e-Research and projects on grid nodes distributed across the consortium and integrated by grid middleware. We have cutting edge expertise in grid system engineering, grid-enabled decision support systems and visualisation but our development and use of the WRG has also given us the following strengths:

- **Proven ability to collaborate effectively at every level**
- **Professional user support and grid training**
- **Productive engagement with international grid activities**
- **Successful support of production grids** – we have built and supported the WRG and host and operate a core node of the National Grid Service and GridPP Tier 2 Node.
- **Successful engagement with a broad application user community within the White Rose Universities** - our users fully exploit state-of-the-art WRG facilities and our current nodes support a huge amount of research across a wide range of disciplines (e.g. Earth Sciences, Engineering, Particle Physics, Life Sciences, Social Sciences, Medicine, Mathematics, Computing, as well as Art and Humanities). Much of this research is multi-disciplinary. We envisage that as the grid usage expands the new users will come increasingly from non-traditional disciplines.
- **Outreach and dissemination** - The WRG two-year Yorkshire Forward funded outreach project has enabled us to work with regional science-based companies. We have given training in the technology to companies and considered with them the

potential demand for grid facilities in the region and the issues pertinent to a grid-based business service. The White Rose Grid conference has now become an established event drawing in regional, national and international delegates. Five demonstrators have been developed to support dissemination and promote further collaborative projects. Further dissemination of information about the WRG has been achieved through our websites, workshops, regular User Group meetings at the three Universities, e-Science publications, collaboration with other e-Science Centres, working closely with IT partners, provision of grid training and other outreach activities.

The WRG has been established with extensive funding from the three Universities who recognised the importance of e-Science and the need for its support. Generous funding from the HEFCE Science Research Investment Fund phase 1 and phase 2 has been augmented with funds from the Yorkshire and Humber Development Agency, Yorkshire Forward, towards business outreach activities; the UK e-Science Core Programme contribution towards the establishment of the WRG e-Science Centre of Excellence, and support from Esteem Systems. Project funding has come from a variety of public and private sector sources.

The WRG has been developed independently and in parallel with the UK e-Science Grid. In 2003 the unique pioneering work of the WRG was formally recognised by the UK e-Science Core Programme through the award of the status of a UK e-Science Centre of Excellence in visualisation and distributed diagnostics. Since then, and in particular following our winning bid to host one of the core nodes of the dynamic UK National Grid Service, it has been much easier for the WRG to operate and share grid technologies experiences and knowledge with the global e-Science community.

We have found numerous benefits to operating a multi-campus grid including skills enhancement through sharing knowledge and experience, good system back up, a large resource pool, ease of collaboration, and flexibility to stagger usage and continuity of service. The whole process has been a steep learning curve for us as it has been a resource intensive activity in a non-trivial technology. Over the years a vast and unique experience has been built up with the deployment and operation of the White Rose Grid. Many complexities have had to be overcome in order to successfully complete a very large procurement process, establish WRG facilities, deploy and operate grid services, as well as to firmly establish and support a considerable White Rose e-Science community.

'**e-Science is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it**.' *John Taylor, 2001, Director General of Research Councils, Office of Science and Technology*

1 Background

This report was commissioned by JSCR to describe the approach of the White Rose University Consortium (White Rose) to grid computing. White Rose was chosen because of its pioneering role in the establishment of a multi-campus regional grid. Our unique grid, the White Rose Grid (WRG), comprises three partners, and is an excellent example of a regional production grid that is much more manageable than any other large grid yet reflects many of the issues related to the support of a distributed collaborative grid infrastructure.

This report, whilst not intending to offer a comprehensive review of all our activities, provides an overview of the aspects of the WRG that might be of value to other institutions which are considering a similar approach and wish to benefit from our experience. It aims to offer an insight into the WRG by describing our approach and highlighting some of the challenges we faced.

2 About the White Rose University Consortium

Established in 1997, the White Rose University Consortium is a strategic partnership between Yorkshire's leading research Universities of Leeds, Sheffield and York. By combining strengths, particularly in science and technology, its goal is to ensure that its partner institutions and the Yorkshire region can prosper through unity and benefit from a range of opportunities and initiatives. These include collaborative research, exploiting commercial and business opportunities, industrial partnerships and joint postgraduate scholarships. White Rose, which has been featured in the HEFCE White Paper *The Future of Higher Education* as an exemplar of university collaboration, has generated additional funding and research projects worth more than £55m for the White Rose Universities. Other successful initiatives, in addition to the White Rose Grid project, include the Faraday Packaging Partnership, the Centre for Excellence in Teaching and Learning of Enterprise and the National Science Learning Centre amongst others.

3 Introduction

In 2000 the Vice-Chancellors of the three Universities made a strategic decision to combine computing resources using the emerging grid technology. The term grid computing had been coined three years previously at a workshop by Foster and Kesselman [1] as a development from metacomputing. Grid technologies promise to solve large problems through delivering significant computational power of a virtual computer integrating extensive distributed computer systems. White Rose established the WRG to pioneer the use of these evolving grid technologies to harvest resources across the three campuses (see Figure 1). The decision was based on a number of perceived advantages.

The high level objectives behind the formation of the WRG were as follows:

(i) to create and operate an enabling e-Infrastructure, the Grid, that supports scientific collaboration across the three Universities;

(ii) to offer a much larger pool of resources which can be easily selected by individual researchers to best match their projects' needs;

(iii) to develop collaboration across the three Universities;

(iv) to deliver stable high performance computing (HPC) services at all three sites;

(v) to expand the amount of e-Science research and infrastructure and increase the e-Science research funding;

(vi) to work with regional partners and businesses to gain the experience of delivering a stable grid service.

The project brings together e-Science and grid researchers as well as technical teams from Computing Services and Computer Science departments in the three Universities. They are involved in building, enhancing and exploiting the innovative technology, which employs many components to create a collaborative environment for research computing in the region. By working together we can selectively apply to our joint projects our skills, knowledge and expertise which are drawn from a much wider pool than those available to any one University. The WRG enhances the competitive position of the three Universities in attracting e-Science funding.

Figure 1: UK e-Science Centres

The WRG underpins a large number of e-Research projects, grid middleware developments as well as traditional HPC projects. The WRG was built with the intention of achieving the above objectives by the three White Rose sites; the scalability of our approach was less important. The aim was to design it in such a way that it best serves the White Rose community and interoperates with other grids as an integrated unit rather than by expanding through the incorporation of additional sites.

4 About the WRG

The WRG operates as a Virtual Organisation with dedicated resources in each institution. The Grid Executive provides strategic direction for the project. A senior academic representative from each University, who sits on the WRG executive and coordinates local actions, heads each site. The White Rose Chief Executive Officer (CEO) chairs the Grid Executive Board

which includes representatives from each University, the e-Science Centre of Excellence, and from our strategic commercial partner Esteem Systems, as well as the WRG Manager. The Grid Executive reports through the CEO to the White Rose Executive Board comprised of the three Vice-Chancellors with Pro-Vice-Chancellor representation so there is high-level institutional commitment to maintaining the grid. Figure 2 illustrates the structure of the WRG consortium.

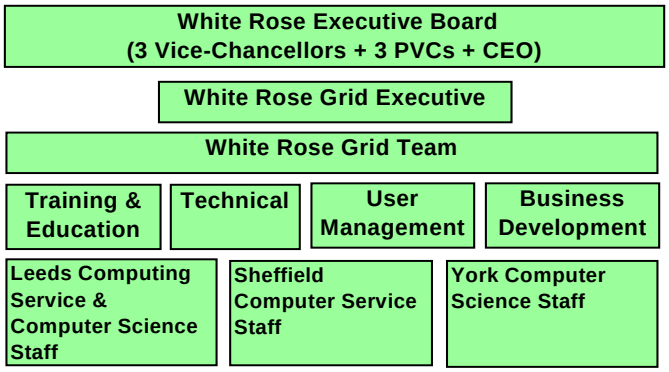

Figure 2: WRG consortium structure

The WRG is supported by a combination of researchers and Computing Services support staff; personnel from different institutions undertake complementary roles. At the White Rose level Leeds has a WRG e-Science Coordinator and Grid Middleware Engineer (part-time), Sheffield a Grid Training Manager and Grid Development Officer and York a WRG Grid Manager and a Senior Experimental Officer. At the institutional level the local computer officers support grid nodes and their users. Research staff identify and evaluate the potential of the latest grid technologies for Computing Service staff who deploy and maintain our grid services. The members' varied experiences of and approach to production services and research are invaluable in supporting the WRG. The WRG team is enlarged by staff supporting the National Grid Service (NGS) node hosted in the Computing Service (ISS) at the University of Leeds.

There are different staff management structures in each institution. This complicates resolution of issues but difficulties are overcome by working closely, good communication, and most importantly through guidance from the Grid Executive. This is a strategic project of the three Universities that is directed from the top.

Over the time four support teams have been instituted to undertake particular aspects of the WRG. The Training and Education team collaborates in the development and provision of training courses in e-Science and grid computing; the Technical team ensures the provision of a well-managed and stable service and makes decisions on the service development; the User Management team, when needed, resolves issues related to registration, authentication, authorisation and accounting; the Business Development team works with regional companies and universities. These teams report through the WRG Coordinator and through their academic representative to the Grid Executive. The WRG Coordinator liaises with the internal and external e-Science academic community and takes day-to-day responsibility for the White Rose e-Science Centre of Excellence and the Sun Centre of Excellence. In its first two years the WRG had a Business Manager who worked with the business community. This role has been extended and is now part of the remit of the WRG Manager who is to be appointed.

The organisational structure evolves to meet the developing needs of our users and the grid; in the early stages there was an Architecture team to oversee procurement and installation of equipment and to ensure compatibility between the three sites.

5 WRG Activities

Briefly our activities within White Rose, or more broadly in the region, nationally and internationally may be divided into four groups:

(i) **Production grids:** We have built the substantial experience in establishing and operating production grids. These include the WRG, the NGS node and a Tier-2 node of GridPP (Grid Particle Physics).

(ii) **Scientific research**: A huge number of e-Research and HPC projects are computed on the WRG. Examples of these include projects in grid-system engineering such as GridWeld, in grid-enabled decision support systems such as DAME, HYDRA or CLEF, in visualisation and programming environments – gViz, or e-Viz, and in application driven projects such as Integrative Biology or Epitheliome. These projects and many more others utilise the WRG e-Infrastructure and benefit from the WRG service.

(iii) **Education and training**: Substantial efforts have been invested into the development of a range of relevant training courses as well as into training the White Rose user community and our regional collaborators. This is crucial to encourage widespread utilisation of the grid.

(iv) **Collaboration and outreach**: Both these activities are an integral part of our strategy as we continually aim to enhance and expand the WRG. By working with Yorkshire Forward we intend to enlarge and consolidate our regional outreach. We will continue our close collaboration with the Chinese grid community, including CNGrid and ChinaGrid, and to support the Worldwide Universities Network (WUN) Data Grid.

6 Construction Process

The core of WRG resources was acquired in a joint procurement (of value £2.8m) that was agreed in February 2002. Prior to this, different models of computing infrastructure and its procurement were considered. The overall aim was to agree a model that offers each site autonomous control over its resources and services but as the whole delivers a single, easily-accessible infrastructure that supports research collaboration at every level between the three Universities.

The following models were considered:

(i) Each site acquires its own resources, which are to be integrated through grid technologies to deliver a virtual pool of resources for use by all WRG researchers. This approach was rejected as the pool of resources acquired independently may not integrate easily, and though it might meet diverse users' requirements at local sites, will not necessarily ensure that there is no resource duplication and may not deliver benefits of efficient utilisation of resources at the three Universities.

(ii) Each institution acquires its own computational resources and a common shared storage resource is acquired and located at one of the sites. Resources are to be integrated by grid technologies. This is a variant of approach (i) but it offers a tangible ground (common filestore) for developing collaboration of distributed teams but at the

time of procurement there were technological difficulties with its implementation because the bandwidth available between the sites (about 100Mb/s) was not sufficient to implement a single shared store.

(iii) A joint procurement is undertaken so as to ensure easy integration of resources delivered by one supplier as well as services optimisation and a good value for money. The key aim in selecting node specialisation was to ensure that each site specialises in services that best meet its local users' requirements. A virtual pool of resources was to be implemented through grid technologies. A single supplier would ensure that distributed grid nodes are easily integrated and maintained.

The most acceptable model, and therefore selected as the way forward, was the third one as it offered a mixture of centralised and distributed approaches to WRG distributed resources and services. Furthermore the idea of a single supplier responsible for procurement and later on for systems maintenance and grid components was a very persuasive one.

Small working groups drawn from the potential user community for each grid node liaised with the suppliers and local computing services to make decisions on defining requirements. Video and multi-site telephone conferencing facilities were used but did not always work well. Face-to-face meetings were found to be always more effective than virtual interactions despite the added travel times. Negotiations with the suppliers were challenging and protracted, especially as we had to ensure that we were getting the latest equipment meeting our users' requirements and offering good value for money. This large joint procurement was successful due to the determination and extensive effort of the three Universities' Purchasing Offices, the potential user community, the Executive and technical staff.

7 White Rose Grid Resources

Our distinct approach to building the White Rose Grid was to bring together the provision of HPC services and the emerging grid technology. In parallel with grid technologies the White Rose Grid offers high performance computing services for our researchers. All WRG computational resources are divided into two pools: 25% are allocated by the WRG Executive to projects that grid-enable applications, develop grid middleware, are of a collaborative nature, or require access to a resource at a remote site; whereas the remaining 75% are controlled by local sites and are used to support more traditional local high performance computing (see Figure 3). This split was implemented to balance local ownership issues with the view to gaining a grid sharing capability. Users needed to know that they could continue to utilise their own local resource while also having access to the external capability when needed. Grid technologies permit us to integrate the three Universities' computational and data resources which are under local site control. Our approach ensures the WRG is able to provide a larger pool of resources to researchers and at the same time offers access to grid technologies.

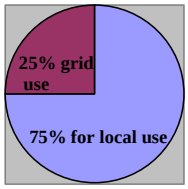

Figure 3: Allocation of WRG resources

7.1 WRG Resources 1: Nodes

As a result of the aforementioned procurement four grid nodes were formed (all named after varieties of white rose) to reflect different specialisation services which would best benefit

local user communities. Two of these nodes are located at the University of Leeds, one at Sheffield and one at York.

The Leeds Grid Node 1 (**Maxima)** offers general purpose high throughput computational facilities with modest parallel capabilities for both message-passing and shared-memory applications; in particular it was acquired for users developing and running shared-memory programs utilising OpenMP.

The Leeds Grid Node 2 (Snowdon) is a large distributed-memory system that supports highly parallel applications implementing message-passing based parallelism. It is mainly intended for users running Computational Fluid Dynamic (CFD) codes.

The Sheffield Node (**Titania/Iceberg)** has been dedicated to an application package service offering a large selection of software products. Additionally it provides modest parallel capabilities for users developing and running shared-memory and distributed-memory programs. Its recent upgrade/replacement offers the capability for running larger parallel applications.

The York Node (Pascali) offers a general computational node with a bias towards bioinformatics and biomedical applications. It also hosts pattern recognition web services as part of the e-Science projects DAME/BROADEN. It provides modest parallel capabilities for users developing and running shared-memory and distributed-memory applications via single user front-end machine.

All these four nodes provide significant computational power for the three Universities' researchers and their industrial partners; further details of their specification are in Appendix A. They are interconnected by the network managed by YHMAN and operate as the White Rose Grid. Subsequently to the joint procurement each site has enlarged its initial pool of resources by integrating other independent resources or by acquiring upgrades to the current nodes to meet the ever changing requirements of our users. A new node (Leeds Grid Node 3) is currently being established at Leeds.

The nodes (see Figure 4) are integrated through grid middleware that supports a set of services and provides a virtual computational infrastructure. Access to WRG resources will be simplified by the provision of the WRG portal.

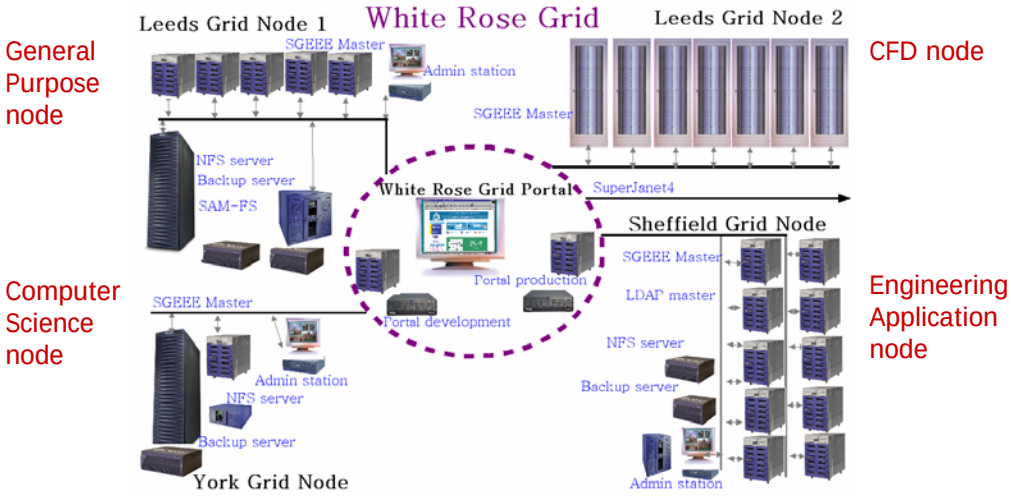

Figure 4: WRG resources in 2003, and their specialisation

7.2 WRG Resources 2: Software Stack

The WRG software stack is composed of a variety of components, many of which are open source software. Table 1 summarises our basic software stack, the selection of which was carried out in consultation with our users and our IT partners.

Component	Description	Access
WRG portal	Application specific portals: (i) the e-Social science portal; (ii) the DAME portal.	From a web browser.
Grid	The grid aspect of WRG is provided by Globus Toolkit (GT) 2.4.3; GRAM, MDS and GridFTP are supported. --- The Storage Resource Broker offers data management.	Deployed on all WRG nodes. --- Deployed on selected WRG nodes.
Queuing and resource scheduling system	The Sun Grid Engine Enterprise Edition (SGEEE) is an advanced batch processor that schedules jobs, submitted by users, to appropriate systems under its control according to the accepted resource policy.	All WRG nodes; some without the Enterprise Edition module.
Compilers	Sun, Portland Group and Intel compilers are available for Fortran95, Fortan77, HPF, C, C++, and Java.	Selected WRG nodes.
Libraries and tools	OpenMP, MPI (Message Passing Interface) Library, Sun Forte Developer, Sun Workshop, Sun One Studio, Sun Performance Library, Sun Scalable Scientific Subroutine Library, HPC Cluster Tools, NAG Fortran library.	Selected WRG nodes.
Operating System Environments	SunOS, Linux RedHat.	Selected WRG nodes.
Scientific Visualisation	IRIS Explorer, IBM Explorer, IGNUplot.	Available to selected groups under limited licences.
Application packages	Fluent, Ansys, Abaqus, Matlab, DIRAC, DL_POLY, NW Chem, Dalton.	Available to selected groups under limited licences.

Table 1: WRG resources: software components

8 White Rose Grid Services

Broadly, services offered by the WRG can be classified into the following groups:

- Computation – this includes high performance computing services delivered through traditional means as well as much more innovative grid computing (delivered through grid middleware);
- Data (& metadata) storage and management (delivered via Storage Resource Broker);
- Software environments & applications;
- User management and support – this includes user registration, authentication, authorisation and accounting as well as helping users to get started on WRG facilities and offer them support;
- Training - this is discussed in more detail below (see section 10).

These services are implemented across the three sites and are supported by virtual teams distributed across the three Universities.

8.1 Grid Middleware

The approach taken to building the WRG was to use the existing grid technologies to integrate resources newly acquired by the three Universities. The Globus Toolkit (GT) developed originally by the Argonne National Laboratory, and now the Globus Alliance, was selected for this purpose, though its limitations such as lack of a resource broker or a simple, easy to use API for job submission were well known. This choice was dictated mainly because of the requirement for compatibility with the UK e-Science Grid which used this grid middleware. Earlier installations of GT version 2 were causing problems and there were stability issues. These were resolved by the deployment of version 2.4.3 which is currently installed on all systems. This version has proved to be stable and reliable. GT version 3 and now version 4 have been deployed for experiments and research on local systems.

8.1.1 Grid Components on the White Rose Grid

Grid components implemented on the WRG can be divided into three categories: core services; monitoring, information publishing, and accounting services; and advanced services.

The WRG core services include:

(i) **Authentication:** the WRG has deployed the Public Key Infrastructure authentication system using digital certificates for authentication in identifying users and resources. The WRG operates its own Certification Authority (CA) which implements a very simple procedure for issuing a certificate. Initially the number of WRG users requiring digital certificates was small so this offered a practical and time-efficient solution. Our grid nodes have been set up to accept digital certificates issued by both the WRG CA and the UK e-Science CA. To support our users engaged in e-Research projects computing on the UK e-Science grid we have set up the UK CA Registration Authority at each University so local staff can vet our users' identities. This resulted in the need to increase support offered to our users as they often experienced problems with getting a UK e-Science digital certificate. These problems were related to the complexity of the processes required to get the UK e-Science CA certificate as well as to technical issues. For example a user whose surname comprised just two letters had his application for a digital certificate persistently rejected by the CA server until he added additional characters to his surname.

(ii) **Certificates management**: the MyProxy server has been installed on one of our grid systems. The software deployed provides a credential repository for the White Rose Grid users authenticating to the WRG with personal X.509 v3 digital certificates.

(iii) **The WRG implements authorisation (through grid-map file) based on Globus GSI: m**apping of grid-wide Distinguished Names (DNs) between digital certificates and Unix ids is performed manually in grid-map files on request from one of the authorised system administrators at each site.

Additionally access to compute resources is implemented through assigning users to the appropriate projects managed by Sun Grid Engine Enterprise Edition (SGEEE). This offers a coarse-grained authorisation. The fine-grained authorisation could not be implemented due to lack of appropriate tools (see also below section on *Authentication and Authorisation Issues*).

(iv) **Job submission/batch services:** WRG nodes use the Globus Grid Resource Allocation and Management (GRAM) protocol to support secure access to grid resources (via `globus-job-run, globus-job-submit and globusrun`). All jobs with the exception of those being computed on the development cluster are routed through SGEEE resource management software which is used for control of resource allocation within each node. SGEEE enforces the site-specific management policy which is implemented as an agreed SGEEE *share-based* policy. This policy assigns the level of service according to: the share owned by individual users; their past usage of this share; and their currently intended use of the systems. The integration between GT and SGEEE is achieved through the Globus SGE job manager and information provider packages from Imperial College, London. These are Perl modules and shell scripts that perform the necessary translation of generic Globus-level requests to the equivalent resource scheduler commands understood by SGE. The variation between the way each server is set up resulted in the need to tune each of the local SGE managers to match the local configuration, for example to use the correct values for the LD_LIBRARY_PATH variable.

(v) **Login:** To enable our grid users to carry out interactively some of the necessary developmental tasks, the Grid Security Infrastructure enabled version of the Secure SHell (`gsissh`) is supported for login to the nodes.

(vi) **Basic data management:** To transfer input and output files to and from nodes, the GridFTP has been deployed on all WRG systems. See also paragraph below on *Data access/integration services*.

The WRG monitoring, information publishing and accounting services include:

(i) The Globus component **Monitoring and Discovery Service** (MDS) is supported on WRG systems to provide information on the status of grid nodes. Individual grid nodes, which run GRIS (Grid Resource Information Service) collect and send resource information to a local site Grid Index Information Service (GIIS) which reports to a collective MDS service. The MDS is integrated with SGEEE which allows information such as queue length to be passed up through the Globus layer.

(ii) The **GITS** (Grid Integration Test Scripts) developed by the UK e-Science community are deployed for monitoring status of our grid.

(iii) Real-time system load monitoring is achieved using **Ganglia**, a distributed monitoring system for high-performance a computing system, which is deployed on the majority of our systems.

(iv) **Accounting:** Sun Grid Engine is used across WRG nodes for usage statistics collection and to produce reports. Monthly usage statistics are uplifted to the WRG site at York for federation into a single web site.

The WRG Advanced Services are enhancements that are useful to or required by our users:

(i) **Data access/integration services:** the San Diego Supercomputing Centre Storage Resource Broker, which provides a comprehensive distributed data management solution, has been deployed on some of our systems.

(ii) **WRG portals** [2]: To hide the complexity of heterogeneous computer systems and grid middleware technologies integrating our distributed systems, the WRG has undertaken the development of application specific portals; for example

(a) A portal based on the EASE web application technology has been developed for e-Social scientists.

(b) The DAME portal (see Figure 5), which utilises the Struts framework with a JSP Model 2 Architecture, supports the case study from the EPSRC pilot e-Science project Distributed Aircraft Maintenance Environment [3]. It provides a collaborative environment with the three main profiles that address specific needs of Maintenance Engineers, Maintenance Analysts and Domain Experts. The portal itself interacts with the workflow manager which orchestrates an array of grid services providing the core functionality of the DAME project.

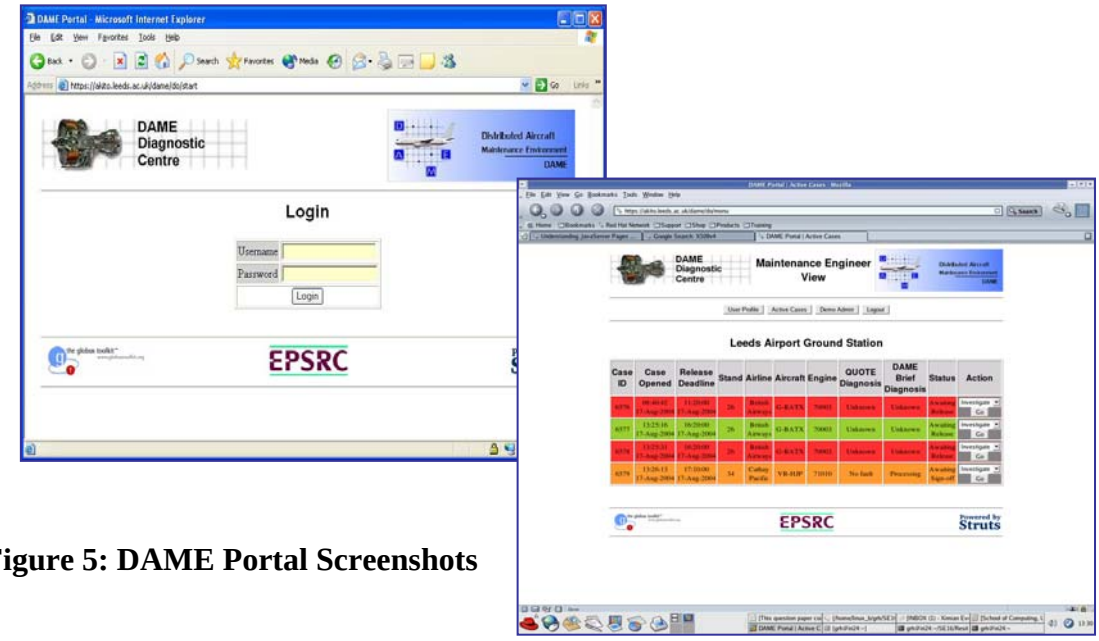

Figure 5: DAME Portal Screenshots

9 User Management and Support

Our novel multi-campus collaborative service required the establishment of a working model for user management and support in the White Rose distributed environment. The following factors needed to be taken into account when developing and implementing new procedures:

(i) **The differences in existing local registration procedures across the three sites**: Ultimately the three participating institutions had different procedures implemented for user support and management of their local users. These needed to be researched carefully and appropriately modified or extended to include users of the new inter-university service. Some historical dependencies of software products used for user management as well as the existence of and dependency on other user registration procedures needed to be considered.

(ii) **The different types of users i.e. academics, industrial collaborators and business users**: The username scheme had to be designed in such a way as to differentiate between academic users, potential industrial collaborators as well as business users so our software automatically could assign users to an appropriate category; for example this was important for authorisation to access software products.

(iii) **The transfer of users from local users to cross-site users:** Over time some of our local users might become cross-site users. Procedures needed to be defined for user administration staff which will allow them to update the user category (i.e. assign remote accounts, alter the project name and move filestore).

(iv) **The geographically distributed registration sites:** The information workflow had to be agreed so all sites had full information about new users and any other amendments to user registration details.

9.1 User Registration

A new common application procedure for WRG resources has been designed and implemented. The individual site's Executive representatives are responsible for approval of their users' applications, and their users' usage of the grid.

As our grid is clearly defined in terms of which sites are involved, our users accept regulations applicable to all three Universities' *Use of Computer Resources* when applying for grid resources.

A user registration workflow has been designed and is presented schematically in Figure 6.

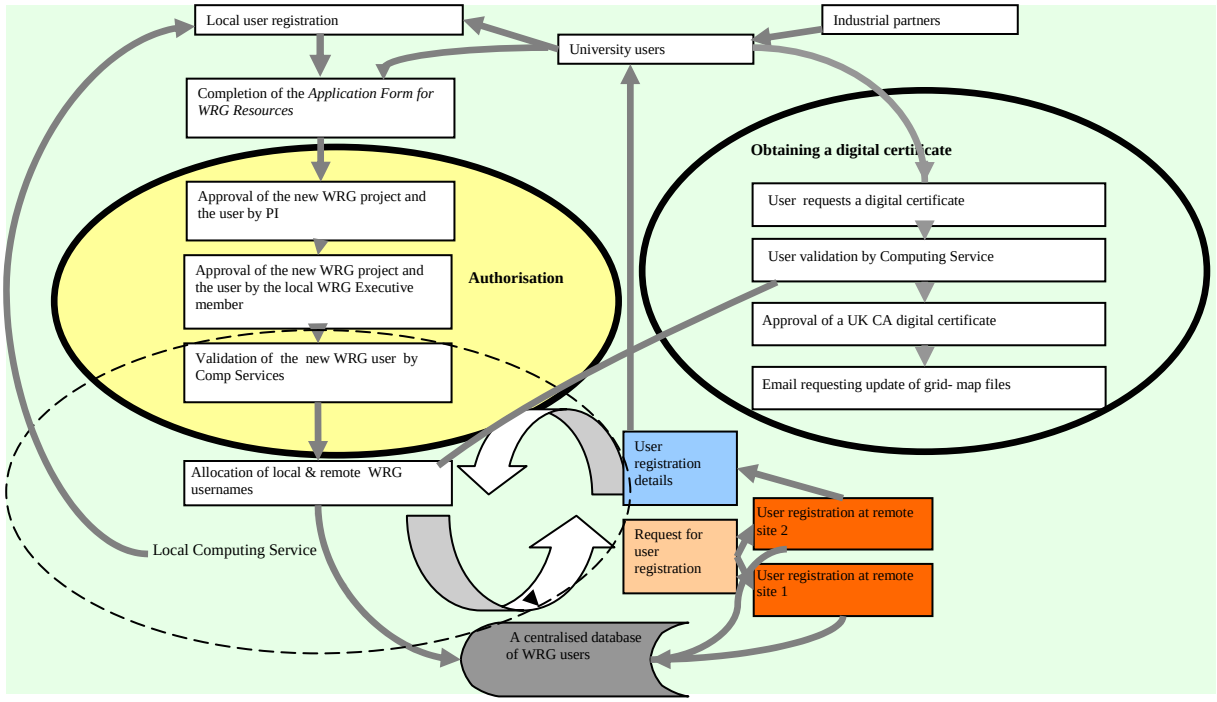

Figure 6: Workflow for WRG user registrations

9.2 Authentication and Authorisation Issues

The issues below relate to personal digital certificates issued by the UK e-Science Certification Authority:

- In our opinion, the procedure for obtaining a UK e-Science certificate is complicated and often confuses users. Each user has to make a personal visit with an identity document to a local Registration Authority in order for his identity to be vetted. Also every year users need to renew their digital certificate. This causes delays and additional effort is required to support users.
- Personal certificates are issued individually on a user by user basis, therefore some sites (in our case Sheffield) may have no link back to institutional identity management databases so it is difficult (impossible in practice) to ensure that rights are withdrawn when the status of an individual changes (for example leaves an institution). This means that users who left the institution still own valid personal certificates until their renewal date, though they have no authorisation to access resources at their previous institution.

Further issues are as follows:

- WRG user registration requires authorisation of projects and individuals; a slicker registration would have been a significant development effort not helped by differing cultures at the three universities; this proves to be a barrier to use at Sheffield.
- A number of UK institutions are currently adopting Shibboleth middleware (supported by the JISC) for authentication and authorisation for access to resources. There are also developments aiming to integrate fully this technology with Permis and grid-enable it that then might be adopted by the e-Science community. This technology is more scalable because the host institution can authorise and authenticate its own

members in an automated fashion. The National Grid Service plans to carry out the first integration of Shibboleth into their authentication system in 2006.

9.3 User Support

The support of grid users was a new element of our service and required a fresh approach. We have accepted a distributed approach to user support. This means that each local site is responsible for support of all its users accessing its systems: local and those computing across the three platforms. At the beginning of service this approach caused delays in responding to users' queries as local support staff did not know users from other Universities and there were communication problems. Also face-to-face consultations have not been easily achieved. With the introduction of better documentation and training courses these issues have virtually been resolved.

Currently perhaps the most common issue is applications from users who intend to use specific software products. They usually assume that the grid nodes' offering across sites is uniform and that all sites have installed all software products.

9.4 Collaborative Tools

Figure 7: Working with the e-Science community over the Access Grid

To enable White Rose collaboration four Access Grid nodes (at York and Sheffield from the funding allocated to the e-Science Centre) have been installed for use by our researchers. These nodes support group-to-group communication over the Internet, and offer interactive video-conferencing facilities for multiple sites participating in the session (see Figure 7).

Video conferencing was not well developed at the three sites at the start of the project. When it became available staff preferred to travel rather than use video conferencing, which is possible due to the relative closeness of Leeds, Sheffield and York). However once staff have got used to using video conferencing (access grid) it has proved to be effective though face-to-face meetings continue for critical issues (such as interviews).

The nodes are housed centrally at each of the WR Universities and there are booking systems in operation for all meetings and seminars. The White Rose-wide Access Grid service is available free of charge to all three Universities' researchers.

Access Grids offer an essential means for supporting collaboration regionally and nationally (e.g. used for WRG Executive meetings or weekly NGS meetings) and internationally (e.g. meetings with members of the University of Utah, WUN, and with Beihang University in China for effective collaboration).

10 Education and Training

The White Rose Grid Training Team coordinates provision of training for users within the White Rose community. Members of this team meet regularly and collaborate in the

development and provision of training courses, including e-Science and grid computing, as well as working closely with the National e-Science Centre training group. A variety of training courses has been developed, which can be divided into the following categories:

- *Getting Started* (e.g. introduction courses for new users)
- *Application Development* (e.g. using Fluent)
- *High Performance Computing* (e.g. use of MPI)
- *Grid Computing* (e.g. using Java COG kit, Globus Toolkit)

Furthermore the team provided courses as part of the ADVICT (Advanced Information and Computer Technology) project funded by the Learning and Skills Council (North Yorkshire) to:

- the Bio-Science Community
- Grid Special Interest Group of Digital South Yorkshire
- Information Studies Instructors from Sheffield Hallam University

The training group is currently establishing over Access Grid a short series of presentations on grid computing to be held jointly between White Rose Grid sites. At Sheffield a number of the courses are accredited as part of the research training programme undertaken by research students; these encourage use of the grid.

Training in high performance computing tools and techniques (for example MPI and Sun Grid Engine) and cluster computing is a prerequisite for grid computing. This is an important area for training as it benefits a wider community of application scientists than this involved with the Globus Toolkit or Storage Resource Broker.

11 WRG and Its Users

The WRG is engaged with a broad application user community within the White Rose Universities. Our users fully exploit state-of-the-art WRG facilities and our current nodes support a huge amount of research across broad range of disciplines at White Rose (e.g. Earth Sciences, Engineering, Particle Physics, Life Sciences, Social Sciences, Medicine, Mathematics, Computing, as well as Art and Humanities). Much of this research is multi-disciplinary. We envisage that as the grid usage expands the new users will come increasingly from non-traditional disciplines.

There are a significant number of projects in: (i) e-Science, (ii) a range of major application areas, (iii) core computer science. Grid usage has increased year on year and now most of the time our systems run at full capacity (see Figure 8). Cross-site usage has also significantly increased (see Figure 9) and has recently started to exceed the 25% resource allocation set for this purpose on a regular basis.

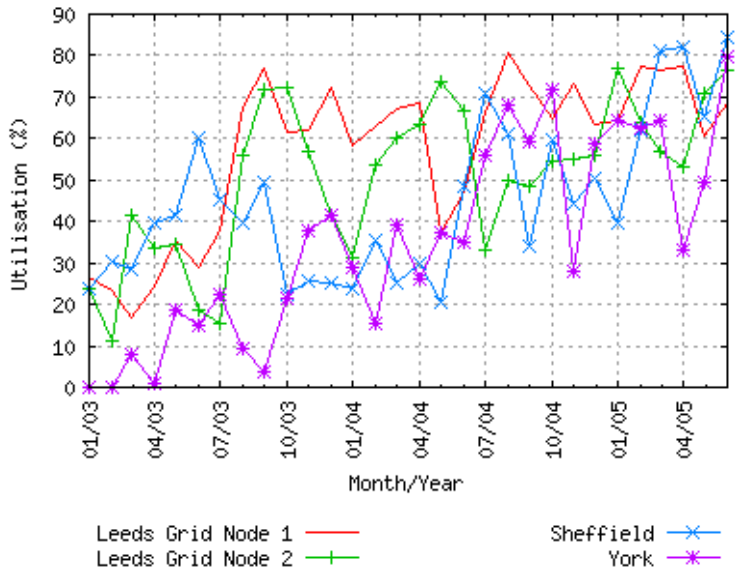

Figure 8: Usage of WRG systems

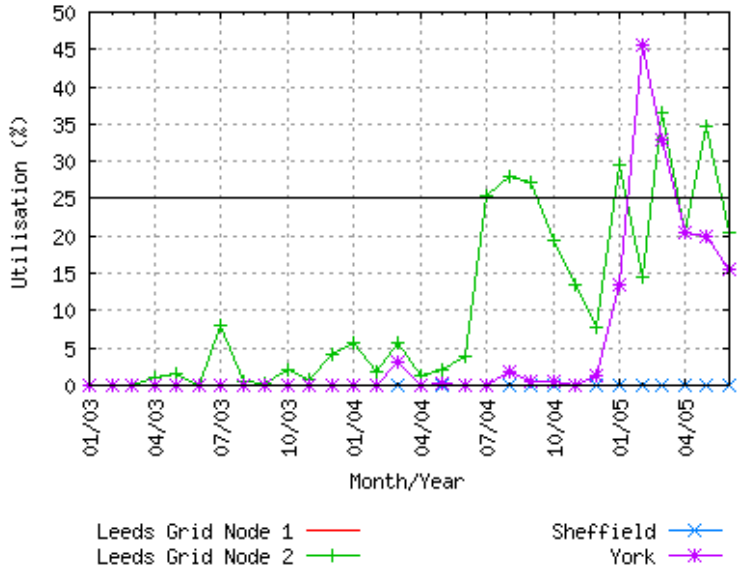

Figure 9: Cross-site usage

The White Rose Grid naturally brings together two communities: traditional HPC users and e-Science researchers, as their work concentrates on the relevant grid nodes forming the WRG**.** Regular meetings of the User Group as well as node shareholders take place at which HPC, WRG, and e-Science issues are considered and resolved. These meetings foster links and multi-disciplinary collaboration among our e-Research community.

Two large conferences have been organised for potential grid users in business. Each of the two events attracted over 120 participants mainly from industry. These brought together researchers and grid computing providers along with a large number of e-Science practitioners from diverse parts of the supply chain. The events sought to provide an insight into the commercial benefits of grid technology, what types of companies would benefit; and

what were the barriers to take-up; what types of applications and services might be offered on the grid; what were the different types of grids (enterprise optimisation grids, utility grids etc) and what were the benefits and merits with each type; how might the grid evolve; where were the likely markets?

In addition to these grid conferences, regular workshops are organised for WR users with the grid and e-Science agenda (e.g. *Open Meeting on e-Research and Parallel Computing* at Leeds with over 80 attendees, and *What is the Grid and How Can it Help Your Research?* at Sheffield with over 45 participants). They aimed to explain what grid technology is and describe its capabilities; to give examples of projects using this technology, to explore funding opportunities, and to introduce the WRG support team.

12 Outreach to the Region

The WRG two-year outreach project supported by Yorkshire Forward has enabled us to work with regional science-based companies. At the start of 2002 there was little awareness of grid technology, especially beyond academia although a few of the larger organisations, such as Rolls-Royce had begun to take an interest. WRG has sought to build relationships from across the whole supply chain to further understanding of commercial and organisational interest in grid technology and promote collaboration between White Rose, companies and organisations. Outreach to business has included giving presentations on the WRG at a number of events, and building a number of demonstrators (e.g. DAME, HYDRA, GOSPEL) to support dissemination and promote further collaborative projects. We have also given training in the technology to companies and considered with them the potential demand for grid facilities in the region and the issues pertinent to a grid-based business service.

Our outreach experience has been enlarged through working with Hull University Computer Science staff who jointly with their partners have developed software (BMUSim) to simulate bone remodelling caused by ageing and disease and prototyped a parallel version of this code on the WRG. Following our preliminary discussions we are actively seeking funding for collaborative grid projects with Hull, Sheffield Hallam and Huddersfield Universities.

Commercial interest in grid computing is growing and WRG has developed a business model following identified regional opportunities for a grid-based service. Our experience indicates that there are a number of issues that need to be resolved before full commercial adoption of grid is achievable. Some of these are:

- Security – commercial companies need to be confident that grids offer an acceptable level of security for their data and applications;
- Stability – business users expect 100% grid stability which may be difficult to achieve using evolving technologies;
- Grid-enabling applications – Applications Service Providers need to respond to new requirements for grid enabling software;
- Resistance – need to build trust so as to overcome natural reluctance of business users to new technologies;
- Company culture – there needs to be a behavioural change as traditionally companies have invested in their own hardware and software, and there is reluctance to introduce a different model of operation;
- Licensing – to introduce grid licences for software products used by businesses;
- Standards – to widen standardisation for grid technologies;

- Ease of use – grid access needs to be simplified. A steep learning curve to the Resource Specification Language offered by Globus Toolkit is a barrier to its usability. Commercial take up is likely to become more widespread when the grid can be accessed through a graphical interface from a web browser accessible from the users' desktop;
- Installation and maintenance – there is a need to make grid products deployment less complex and reduce the large amount of maintenance effort;
- Documentation - there is a definite need for more thorough and easier to read documentation for grid products.

13 Dissemination

The WRG disseminates information through:

(i) Web sites; each of the local sites supports its own pages (i.e. http://www.leeds.ac.uk/iss/wrgrid/ http://www.shef.ac.uk/wrgrid/index.html, https://www.wrg.york.ac.uk/WhiteRoseGridYork/) which are integrated through the WRG web site at http://www.wrgrid.org.uk/. Institutional web pages are used to cover differences in internal procedures for each institution and advertise internal support, whereas the main web site covers common aspects of the WRG project and provides a cohesive structure to all its varied activities;

(ii) WRG leaflets; a number of informative leaflets for circulation amongst current and potential users (http://www.wrgrid.org.uk/leaflets.html) has been produced. As well as covering information about the White Rose Grid portals, technologies, service level agreement management and outreach activities; the leaflets also include exemplars of grid projects.

(iii) Conferences (http://www.wrgrid.org.uk/conference2003_slides.html, http://www.wrgrid.org.uk/conference2004_slides.html and workshops (http://www.leeds.ac.uk/iss/wrgrid/OpenDay2005_proc.html, http://www.wrgrid.org.uk/workshop2005_slides.html);

(iv) User Group meetings at the three Universities;

(v) Academic routes via e-Science publications and collaboration with other e-Science Centres. The WRG team have also written a number of articles about their experiences in developing the grid for the annual All Hands Meetings; or describing the White Rose Grid [4];

(vi) Working closely with IT partners (Esteem, IBM, SUN, Streamline Computing);

(vii) Provision of grid training (see section 10 above);

(viii) Our outreach activities.

14 Funding

14.1 Infrastructure Funding

The WRG has been established with extensive funding from the three Universities which recognised the importance of e-Science and the need for its support. Over £5.5m has been invested into building and operating this e-Infrastructure that is believed to be a state-of-the-art production grid supporting new e-Research (e-Science) projects, and online scientific and engineering communities across the White Rose jointly with their industrial collaborators. Generous funding from the HEFCE Science Research Investment Fund phase 1 and phase 2, jointly with the Universities' strategic funds and other internal initiatives added up to approximately £4.9m. This has been augmented with funds from the Yorkshire and Humber Development Agency, Yorkshire Forward, towards business outreach activities; as well as the UK e-Science Core Programme contribution towards the establishment of the WRG e-Science Centre of Excellence; and support from Esteem Systems.

Furthermore the WRG has already secured £2.3m for its further development over the next two years. This includes funds from the Higher Education Innovation Fund round 2 (HEIF2) which is a partnership between the Department of Trade and Industry/Office of Science and Technology (DTI/OST), HEFCE, and the Department for Education and Skills (DfES). HEIF2 funds are designed to encourage the transfer of knowledge and expertise and enhance the relevance of programmes of teaching and research to the needs of employers and the economy.

14.2 Projects Funding

Project funding has come from a variety of public and private sector sources. EPSRC has funded many through their e-Science Core Programme, including e-Demand [5], GOSPEL, gViz (Visualisation Middleware for e-Science) [6,7], IBHIS (Integration Broker for Heterogeneous Information Sources), COVISA-G (Collaborative Visualisation), myGrid (developing middleware to tackle problems of bioinformaticians), Integrative Biology Project (with CCLRC) and DAME (Distributed Aircraft Maintenance Environment) one of six large projects launched in the first phase of e-Science funding. A number of these were in collaboration with other universities and with industry, for example Integrative Biology Project **(**joint project with many UK universities, led by Oxford University, and involving IBM), myGrid (Manchester, Newcastle, Nottingham, Southampton, nine industrial partners including GSK and IBM) and DAME (Oxford, Rolls-Royce, Data Systems & Solutions). Other research council funding has come from ESRC for HYDRA, a demonstrator which seeks to exploit grid technology in support of the decision making process in healthcare planning, its continuation through the HYDRA2 project and MoSeS (Modelling and Simulation for e-Social Science) as well as PPARC and MRC. Further project funding has been obtained from the European Union Framework 5 IST programme for GEMSS (Grid-Enabled Medical Simulation Services). The DAME project has been so successful that the DTI has partially funded a follow-on project, BROADEN, which includes the White Rose Universities, Oxford, Rolls-Royce, Data Systems & Solutions, Electronic Data Systems, Lost Wax Media Ltd, Streamline Computing Ltd, Oxford Biosignals Ltd and Cybula Ltd as industrial partners.

The above are the examples of e-Science projects in which the WRG members are actively involved; many others projects are also supported by the WRG.

15 Relationship with the UK e-Science Core Programme

The WRG has been developed independently and in parallel with the UK e-Science Grid. In 2000 at each of the three Universities there were already well established high performance computing communities exploiting local computational facilities and the potential benefits offered by the grid encouraged the Universities to support this very innovative initiative. As the competition for UK e-Science regional centres was not opened publicly, the WRG was not been given the opportunity to bid to become a regional e-Science centre. The need for the development of a collaboration platform for the three Universities and an enabling infrastructure for e-Science dictated the development of the WRG at the time. Therefore the WRG has been developed in parallel to the UK e-Science Grid. This meant that at that time we did not benefit from the experience and knowledge developed within the e-Science community formally supported by the UK e-Science Core Programme as there were no effective mechanisms for access to the UK e-Science resources and experience.

The lack of direct involvement in the UK e-Science Core Programme from the beginning meant that:

(i) The WRG had to independently research grid technologies issues that may have been already solved by the Core Programme efforts as we were not informed adequately about UK e-Science grid developments;

(ii) Substantial staff effort in training and gathering experience were required as there was no formal mechanism in place to share UK e-Science Core Programme outputs. The Grid Support Centre did not fulfil this function for us;

(iii) Additional efforts were required to build trust when outreaching into the region as often regional companies were enquiring about our relationship with the official UK e-Science grid.

In 2003 the unique pioneering work of the WRG was formally recognised by the UK e-Science Core Programme through the award of the status of an UK e-Science Centre of Excellence in visualisation and distributed diagnostics. Since then, and in particular following our winning bid to host one of the core nodes of the dynamic UK National Grid Service, it has been much easier for the WRG to operate and share grid technologies experiences and knowledge with the UK e-Science Core Programme community and the global e-Science community. Refer to Figure 1 for the distribution of UK e-Science Centres.

16 Relationship with the National Grid Service

The WRG consortium operates one of the NGS four core nodes and is a founding partner of the National Grid Service (NGS) and the Grid Operations Support Centre (GOSC). This relationship [8] is invaluable as it enables our staff to keep abreast of developments in grid technologies. It makes sharing of experiences easier and enables us better to support both WRG and NGS e-Science users.

Our bi-weekly NGS Operational and Technical teams' meetings over Access Grid with NGS partners - Oxford, Manchester, Rutherford, Bristol, Cardiff and Edinburgh - enable us to develop and grow the UK e-Science Grid and help to enhance the WRG. New NGS partners such as Lancaster have broadened experience of the grid community and allowed better understanding of grid scalability issues.

17 Benefits from a Multi-Campus Grid

Operation of our multi-campus grid demonstrates numerous benefits. Some of these are as follows:

- Computing services staff skills and knowledge are enriched through working jointly. Parallel programming skills previously mainly established in Leeds have been propagated to Sheffield, and *vice versa* knowledge about application packages from Sheffield has been passed onto Leeds support staff. It is easier to develop the web-service and portals by exchanging ideas.
- Disaster recovery equipment can be located at a remote site. Leeds has located their backup storage system at Sheffield and operates it over the secure network connection.
- Users have access to a larger resource pool than that offered by a single university.
- The existence of the grid infrastructure encourages collaboration between the three sites at every level: users, technical and research.
- Some of our users require a significant amount of computational power during certain periods. The multi-campus service offers the flexibility to stagger usage across campuses.
- Upgrades and phasing in and out of machines are made easier for users as systems at other campuses can be utilised to provide continuity of the service, e.g. Sheffield is shunting Sun jobs to York as they run down Titania and bring Iceberg on-line.

18 Challenges and Lessons Learned

Over the years extensive and unique experience was gained with the deployment and operation of the White Rose Grid. Many complexities needed to be overcome in order to successfully finalise a very large procurement, establish WRG facilities, deploy and operate grid services as well as firmly establish and support a considerable White Rose e-Science community.

The many challenges faced include the following:

- **Innovative technology:** at the beginning of this project there was very little knowledge of grid technologies existing within the consortium and in the UK e-Science community. The grid middleware was very immature as is even now evolving. Lack of comprehensive documentation and adequate guidance delayed grid deployment.
- **Geographically distributed teams:** frequent meetings demanded regular travelling within White Rose which placed additional demands on our staff's busy schedules. At Leeds there was an Access Grid node for enabling video-conferencing but until the other two sites established Access Grids all meetings were face-to-face.
- **Large team:** the very large number of technical staff involved lengthened the decision process in all aspects of the project.
- **Human interaction factor:** the project has crossed organisational boundaries and any interaction problems faced by the team comprising the three Universities' members are managed through strong leadership and by clearly defining members' responsibilities.
- **Trust and ownership:** these questions of ownership and trust were regularly posed, particularly in the procurement process and equipment location.

- **Reaching agreement**: there is a real difficulty in getting agreement over issues and priorities between institutions in a virtual organisation where there are differing requirements from the participants and a set of open goals for the project rather than specific deliverables. This is compounded by the difficulty of resource management in a virtual organisation.
- **Effective communication** is crucial for the continuing success of the WRG**.** Communication problems were often due to the organisational complexity of our support team, which comprises academic and technical staff governed by different management models.

Some of lessons learned from our procurement:

- **Beowulf type systems**: if these systems are part of grid facilities, consideration should be given to the substantial amount of space, air conditioning power and electrical power required by them so that they can be installed within the planned time-scale; at White Rose, new machine rooms had to be built to accommodate new computational facilities.
- **Joint procurement:** this offered value for money and helped to form a close working grid support team but at the same time additional significant resources were required to coordinate and agree this very large equipment procurement.
- **Separate procurements**: these are much easier to handle by individual sites though they must ensure that the procured systems (computer systems, storage systems) will couple using grid technologies.

Lessons learned from operating the WRG:

- Support of the grid requires a larger support team than the combined number of staff supporting services at local sites as additional issues related to the collaborative service need to be resolved; in our project there are additional staff at the White Rose level.
- Grid needs new operational procedures agreed between sites.
- Historical and local dependencies need to be considered when developing new procedures for grid.
- Currently the support team needs to include staff experienced in service provision as well as research staff. These two categories of staff with different skills and approaches are needed as grid technologies have not yet matured and are not ready for full production service off the shelf.
- Grid technologies must be further developed to offer a comprehensive package of services which can be easily deployed and supported by service staff.
- User trust to new technologies needs to be built through both training and development of use cases to show their functionality and benefits.
- Users' requirements are ever changing and the grid resources need to be constantly upgraded.
- Comprehensive and easy to read user and system administration documentation needs to be developed.

- At White Rose many problems were swiftly resolved due to the effective management of the project by the Executive and high level support from the White Rose University Consortium.
- The issue of software licences for grid users caused many problems. It is vital that software suppliers make available licences for software products use over the grid.
- It is essential to consider the authorisation issues related to software licensing which arise where commercial software is used and the licences commonly prevent software licensed to institution A being used by members of institution B.
- Simplified access to grid is essential for rapid uptake grid technologies. There is a need to provide the means to access all grid resources through a simple web-based interface.
- There is a need to establish the necessary trust relationship between sites over who (for example staff, research, masters and or undergraduate students) should be authorised to use which resources.
- In the grid environment there is a clear need to operate a market for computational services. Currently there are a number projects investigating the associated issues and developing software for this purpose (e.g. “A Market for Computational Services” - the project led by the London e-Science Centre at Imperial College, London).
- There is a significant cost in the development of local HPC facilities into a shared grid and it is important to set this against identifiable benefits.

19 Development Plans

The WRG serves its community by underpinning a large number of White Rose projects. Our development plans include upgrades to individual node facilities as well as adding additional resources. For example currently a new node at Leeds is being established that will support bioinformatics, geography and medical users as well as offer upgrade facilities to users of Leeds Grid Node 1 and 2. There are firm plans to add a visualisation node at Leeds. Each University is now working towards establishing a campus grid (of a different constitution) and linking them together within the WRG. This will ensure further expansion of computational resources available to the White Rose community. Experiments are underway to develop demonstrators for a WUN Grid, building on the experiences gained with the WRG. New projects are being developed utilising the UKLight network. New collaborations are being formed with China’s e-Science Programme and through the new WUN project with Penn State.

It is envisaged that the follow-on activities from our outreach project will enable us to further our collaboration with regional partners in the future.

20 Conclusions

The WRG initiative has been a great success; we have demonstrated that by working together we can selectively apply to our joint projects our skills, knowledge and expertise, which are drawn from a wider pool than that available in a single institution. We have also delivered stable HPC services at all three sites and expanded the amount of e-Science research and infrastructure within White Rose.

Having a single supplier responsible for procurement and for systems maintenance and grid elements has facilitated installation. We have found that senior level support is essential in overcoming difficulties caused by the differences in culture within the participating institutions and that organisational structures need to be fluid and responsive to changing needs. Experience has shown us that there is a need for face-to-face meetings to build trust and agreement between partners and that care must be taken to ensure registration procedures and conditions of use comply with those of all participating institutions. Good documentation and training courses are essential to ensure consistent, high quality user support across all sites. It is important to co-operate and exchange information with the wider e-Science community in order to avoid duplication of effort and ensure compatibility of systems. There are many challenges to be overcome in establishing a multi-campus grid but the benefits accrued make the effort worth it.

Our future priorities are to continue to expand the WRG e-Science community, the portfolio of grid-enabled applications and the number of e-Science projects. We also wish to encourage users to utilise all functions of the WRG and to further develop grid accessibility.

Acknowledgments

(i) The White Rose Grid Consortium members are grateful for funding received in support of this project from their Universities, EPSRC and the UK e-Science Core Programme, Yorkshire Forward, Sun Microsystems and Esteem Systems.

(ii) The authors of this report are grateful for comments and suggestions received from the WRG Consortium members.

(iii) The authors wish to thank presenters at the White Rose Grid Workshop 14 April 2005 for giving their permission to reference their PowerPoint presentations.

Selected Glossary

API	Application Programming Interface
CA	Certification Authority
CCLRC	Council for the Central Laboratories of the Research Councils
CFD	Computational Fluid Dynamics
CLEF	Clinical e-Science Framework
CNGrid	China National Grid
COVISA-G	Collaborative Visualisation - e-Science Demonstrator Project
DAME	Distributed Aircraft Maintenance Environment
DIRAC	FORTRAN (and a bit of C) code for relativistic molecular calculations based on the Dirac-Coulomb Hamiltonian
DL_POLY	General purpose serial and parallel molecular dynamics simulation
eViz	An advanced environment for enabling visual supercomputing
Fortran	FORmula TRANslation - high level language for solving problems with significant arithmetic content
GIIS	Grid Index Information Service
GITS	Grid Integration Test Scripts
GRAM	Globus Grid Resource Allocation and Management protocol
Grid FTP	Transport protocol designed by the Global Grid Forum for bulk data transport in a grid environment
GT	Globus Toolkit
gViz	Visualisation Middleware for Grid users
HPC	High Performance Computing
HPF	High Performance Fortran
HYDRA	A Grid-Based Spatial Decision Support System
IGNUplot	An interactive plotting program
IRIS Explorer	Tool for developing customised visualisation applications
ISS	Information Systems Services (Leeds Computing Service)
JISC	Joint Information System Committee
JSCR	JISC for the Support of Research
JSP	JavaServer Pages
Linux	A a free Unix-type operating system originally created by Linus Torvalds with the assistance of developers around the world
MPI	Message Passing Interface
NAG	Numerical Algorithms Group
NGS	National Grid Service
NWChem	Computational chemistry package designed to run on high-performance parallel supercomputers and conventional workstation clusters
OpenMP	A specification for a set of compiler directives, library routines, and environment variables that can be used to specify shared-memory parallelism
SGEEE	Sun Grid Engine Enterprise Edition
Sun OS	Sun Operating System
Unix	Operating system
WRG	White Rose Grid
WUN	Worldwide Universities Network
YHMAN	Yorkshire & Humberside Metropolitan Area Network

References

[1] Parallel Computing Research Newsletter, Volume 5 - Issue 4 (1997) Argonne Workshop Explores Construction of a National Computational Grid http://nhse.cs.rice.edu/CRPC/newsletters/fal97/news_grid.html

[2] M. Thompson, J. G. Schmidt, P. M. Dew, "White Rose Grid Portals: Practice and Experience", *Poster in Proceedings of the UK e-Science All-Hands Meeting 2004*, Nottingham, 31 August -3rd September 2004, ISBN 1-904425-21-6.

[3] J. Austin, T. Jackson, M. Fletcher, M. Jessop, P. Cowley and P. Lobner. "Predictive Maintenance: Distributed Aircraft Engine Diagnostics", in: I. Foster and C. Kesselman (Eds.). Grid2: *Blueprint for a New Computing Infrastructure. Chapter 5*, Morgan Kaufmann, 2nd edition, 2003.

[4] P. M. Dew, J. G. Schmidt, M. Thompson, and P. Morris, "The White Rose Grid: Practice and Experience," *in Proc. UK e-Science All Hands Meeting,* Simon J. Cox Eds, Nottingham Conference Center, U.K., Sept. 2nd-4th, 2003, ISBN 1-904425-11-9

[5] N. Looker, M. Munro, and J. Xu, "Simulating the Failure of Grid services," *Special Issue on Grid Performability*, *International Journal of Simulation Systems*, Jan. 2005.

[6] K.W. Brodlie, D.A. Duce, J.R. Gallop, J.P.R.B. Walton and J.D. Wood. "Distributed and Collaborative Visualization, Computer Graphics Forum", Volume 23, Number 2, pp 223-251, 2004.

[7] K.W. Brodlie, D. Duce, J. Gallop, M. Sagar, J. Walton, J.Wood "Visualization in Grid Computing Environments" in *Proceedings of IEEE Visualization 2004*, edited by Holly Rushmeier, Greg Turk and Jarke J. van Wijk, pp155-162. ISBN:0-7803-8788-0,2004.

[8] A. J. Richards, J. G. Schmidt, P. M. Dew, F. Youhanaie, M. Ford, N. Geddes, "Production Quality e-Science Grid", *in Proceedings of the UK e-Science All-Hands Meeting 2004*, Nottingham, 31 August -3rd September 2004, ISBN 1-904425-21-6.

Appendix A

Table containing hardware specifications of WRG nodes.

Site	Node Name - High-Level Description	Brief Description
Leeds Grid Node 1	Maxima - designed to offer a robust and resilient service to support a large number of researchers in the University of Leeds and the White Rose Consortium.	A constellation of shared-memory systems which include a Sun Fire 6800 with 20 UltraSPARC III Cu 900Mhz processors, and five Sun Fire V880 servers, each with 8 UltraSPARC III Cu 900Mhz processors.
Leeds Grid Node 2	Snowdon - two Beowulf type clusters: a 28 processor development cluster and a 256 processor production service cluster; each using a Myrinet 2000 switch for interconnecting nodes. The production cluster supports both capability and capacity jobs.	The main cluster comprises a head-node configured with a dual 2.2 GHz Intel Xeon and 256 processors configured out of dual processor nodes based on 2.2 and 2.4 GHz Intel Xeon processors. The development cluster is of similar configuration.
Sheffield Node	Titania – the originally established node comprising a set of shared-memory servers. Iceberg - is a 160 processor Beowulf type system, supplied by Streamline Computing and Sun Microsystem.	Titania was originally configured with 10 Sun Fire V880 servers, each with 8 UltraSPARC Cu processors. Iceberg is made up of a mixture of 2-way nodes each with 4GB memory and 4-way nodes; all nodes are based on AMD Opteron processors.
York Node	Pascali & Fermata - two shared-memory multiprocessor systems. Nevada - a medium size distributed memory system.	Fermata is a Sun Fire 6800 with 20 UltraSPARC III Cu 900Mhz processors, and Pascali is a Sun Fire V880 server with 8 Ultra SPARC III Cu 900Mhz processors. Nevada – a Beowulf type cluster based on Intel Pentium III processors.