NW-GRID Developing a Virtual Research Environment for NW Researchers

Rob Allan, Xiaobo Yang, Adam Braimah and John Kewley
Daresbury Laboratory

Rob Crouchley, Adrian Fish, Miguel Gonzalez and Danny Hughes
Lancaster University

Cliff Addison and Adam Sykes
University of Liverpool

John Brooke and Bharathi Kattamuri
University of Manchester

e-Mail: r.j.allan@dl.ac.uk

D R A F T
Commercial in Confidence

Abstract

This is a proposal to develop a Virtual Research Environment for users of the NW-GRID and other researchers in the North West of England. This NW-VRE will be developed during the final phase of the NW-GRID project – as part of the "Real Time and Collaborative Grid". It will then continue to be used to support a growing community of academic and commercial end users for a following minimum of three years.

NW-GRID WP3 is concentrating on developing, evaluating and deploying Grid middleware to meet the NW-GRID testbed phases. However with a small amount of additional funding a rich user interface can be provided to link users to applications via this middleware.

The VRE will comprise a suite of powerful Grid computing, data management and collaboration tools which will be provided for Web-based portal access and for linking into users desktop applications. It will be both a "one-stop shop" for academic users and a "turnkey solution" for commercial users.

Contents

1 Background: Virtual Research Environments

The concept of a Virtual Research Environment or Collaborative Research Environment has grown from the activities of the UK e-Science Programme, JISC Development Programme and work on the Sakai project in the USA. VRE was established as a distributed way of working using a Web-based portal to access a wide and growing range of on-line tools. These include access to Grid-based computing and data management sytems as well as collaboration tools, some based on Web 2.0. Alternative non-portal interfaces will be available where appropriate for linking the underlying services into other desktop environments and applications, such as bespoke GUIs for scientific research or more familier packages such as Stata, SPSS and MatLab.

A VRE will be deployed for use by the NW-GRID partner sites using technology and components which we have developed in other projects and which are already in use. It will be hosted on an IBM BladeCentre at Daresbury Laboratory and use a professionally maintained Oracle DB service. In fact the NW-GRID is already using the Sakai portal maintained in this way for its project administration. The developers will then work with NW-GRID users to create project-specific tools which will be embedded into the VRE framework or can be deployed as separate instances where required. This will effectively link the VRE to Grid-enabled applications using the middleware which is being developed in WP4.

During the course of this work we will thus deploy a flexible integration framework for multi-institutional research and research-related administrative services, validating this framework and its components against the requirements and practice of several discrete end-user research communities. This will act to both guarantee existing investment made in frameworks and components – particularly the significant range of available portlets from the project partners, and to reduce time to the implementation of potential solutions for new research communities.

This VRE deployment will not only retain and maintain the skills of software developers in the region, but has been designed to test and extend our understanding of the following:

- How can portal frameworks be configured to best suit the expectations and work practices of different research user communities and institutional or organisational contexts, including commercial partners?
- Can tools from multiple institutions and organisations be brought together coherently to enable sharing of information and processes to support collaboration?
- Can community-specific tools be integrated meaningfully alongside generic and remotely-hosted Web tools?
- Can a portal based approach provide the flexibility to enable effective use by both researchers and administrators?

It will be valuable to assess different requirements in different research groups, but also to gather proof or evidence of the applicability of both particular tool combinations and the general approach.

Analysis of the requirements and desires of the research community participants, a number of such statements already exist from our users and as outcomes of the seminars we have organised.

The first project phase will culminate in the deployment of "a first cut" integrated VRE for the research communities using NW-GRID. The subsequent project phases will focus on the iterated improvement and development of this environments, and addition of further validating communities requirements with tools appropriate for their work.

1.1 The Sakai VRE Demonstrator

Daresbury Laboratory and University of Lancaster with partners at Universities of Oxford and Reading were funded by JISC to develop a demonstrator VRE using the Sakai portal framework. This is available for trial use at `http://rhine.dl.ac.uk:8080/portal`. It is this portal which is already in use for NW-GRID project management purposes. To see the full demonstrator site please the select "Sakai Demonstrator" tab or log on as `id/passwd = guest/demo`. See also `http://www.grids.ac.uk/Sakai`.

Through the work of this project the Sakai framework contains both tightly-coupled collaboration and e-Learning tools and also JSR-168 support for more loosely-coupled portlets appropriate for e-Research.

The Sakai User Manual is available from `http://www.scrippscollege.edu/dept/it/guides/sakai/Sakai2.3Manual.pdf`.

2 A VRE Service for NW-GRID

Several salient characteristics of VREs have emerged from the early work of the JISC VRE Programme. The broad conception of a VRE begins with requirement for highly granular flexibility of both tools and presentation, driven by recognition of the diversity of user needs and requirements the environment is designed to support. It begins with a requirement to integrate not only with systems within a single institution, but also, driven by the dispersed nature of research communities, with a variety of systems and resources outwith the institution. These may range from different-but-similar systems at a collaborating HEI or other organisation (for example, a research council or commercial partner), through those developed specifically to meet the computational or collaboration needs of research, such as Grid services, to those provided by the broader Internet environment, in order that researchers might obtain maximum advantage from the investment being made in this area. VRE integration with the emerging landscape of institutional and other repositories, including (but not limited to) those supporting open access publication of research output including data is also important.

These emerging characteristics of a VRE are increasingly overlaid with a requirement to provide support for the creation, further development, or enhancement of a research community in virtual space – a "Virtual Research Community". The OST report of March 2006 indicated that VRCs have *the potential to open exciting new opportunities to collaborate in research and thus realise significant gains at institutional, national and international levels.* Support for a rich variety of these interlocking communities is likely to remain a significant national and international objective for VRE development for the foreseeable future, whilst noting that this will most likely be influenced by both planned and emergent developments. The increased use by researchers of social networking applications, for example, adds urgency to the requirement for VREs to adhere to open and published standards and

specifications.

Given this rich, and potentially confusing, contextual landscape, how might a rational and coherent course for the development of the VRE to be charted? This proposal outlines one approach, using the Sakai portal framework, which has obtained a measure of consensus in the virtual learning and virtual research communities. Technically, we will build on (fuse) a range of existing components fro previous development projects. The use of Sakai in the existing programme provides a solid basis for our VRE with the flexibility to include tightl-coupled collaboration tools alongside a loosely-coupled portlet standards based framework. The development of service registries within an architecture based around service-oriented design is being integrated within this effort, in order to facilitate ease of deployment of particular services at partner institutions. The considerable experience within the NW-GRID partners of Grid, informational and institutional adoption and deployment of portals and portlets to support personalisation and customisation will provide strong support for Grid-based application developers and help to ensure a coherent end user experience.

By using, adapting and integrating largely pre-developed components into a flexible VRE we can minimise time to implementation for a great many communities and projects, and act to ease problems associated with integration with institutional systems. Most importantly, by underpinning technical work with a detailed and thorough understanding of the needs of real researchers in a variety of discipline-based, institutional, and organisational contexts, the VRE will significantly reduce time to adoption and advance our collective understanding of the requirements of environments to support research and how to embed the Grid into complex and collaborative research processes.

3 Project Description and Benefits to NW-GRID Community

We will create a VRE by deploying a flexible integration framework for multi-institutional research and research-related administrative services, validating this framework and its components against the requirements and practice of several discrete end-user research communities. This will act to both guarantee existing investment made in components – particularly the significant range of available portlets – by NW-GRID and the broader UK and international community, and to reduce time the to implementation of potential solutions for the research communities. It will therefore be possible to complete three full iterations around the figure of eight lifecycle model indicated in the Circular 04/06 Appendix F during the life of the project, together with a fourth, partial iteration. In addition, further iterations of environment instantiations for particular research communities will be provided according to the needs of those communities.

Deployment of a VRE for NW-GRID will test and extend our understanding of practical deployments in the following areas:

- How can portal frameworks be configured to best suit the expectations and work practices of different research user communities and institutional or organisational contexts?
- Can tools from multiple institutions and organisations be brought together coherently to enable sharing of information, processes and collaboration?
- Can community-specific tools be integrated meaningfully alongside generic and remotely-hosted Web tools?

- Can a portal based approach provide the flexibility to enable effective use by both researchers and administrators?
- At what points are desktop tools or those provided by a mobile platform, more effective?
- How might these be best integrated within a meaningful user experience?

A deeper and more precise understanding of VRE requirements will be gained from addressing this range of questions, together with a validation of the applicability of both particular tool combinations and general approach to specific research groups. Whilst not the primary focus of the project, this may contribute evidence to assist the assessment of whether the same framework based approach can be used for teaching (or training) and research.

1. Analysis of the requirements and desires of the research community participants. It is anticipated that some groups will bring relatively well-formed ideas of their collective needs to this process, whilst others, perhaps less familiar with technology development, may require greater effort. In order to initiate development and integration activity more rapidly, outcomes of several previous NW-GRID VRE projects will be used in this strand, including (in details): EVIE (Leeds), Sakai Demonstrator and GROWL (Daresbury and Lancaster) and History VRE (UEA and Hull).
2. The initial technology effort will focus on providing a detailed architecture overview including; component and service cataloguing, registration, initial integration testing, and further work to decompose Sakai collaborative tools to enable their rendering and presentation within the uPortal and Luminis frameworks.

This first project phase (WPs 1, 1a and 2) will culminate in the deployment of "first cut" integrated VREs for the research communities. The subsequent project phases will focus on the iterated development of these environments, and their potential extension by the addition of further validating communities funded under the next tranche of VRE Programme funding. Further detail is provided in the work packages outlined below.

4 Track Record and Projects to be Leveraged

This section could mention a list of related projects and references and what we will draw from. Also note that personnel have been engaged in these projects. What benefit did they bring to their user communities?

The partners have complementary experiences and capabilities and different user communities. This diversity will enable rapid deployment of a VRE meeting a wide reange of requirements. THe available skilled staff will facilitate adapting existing tools to meet requirements of new user groups.

4.1 Previous Funded Projects

Below we list projects which have yielded software and tools which will be integrated into the NW-GRID VRE.

Sakai VRE: A JISC-funded project involving developers at Daresbury and the UNiversities of Lancaster, Oxford and Reading. Delivered ... ran from 2005-2007.

GROWL VRE: A JISC-funded project involving developers at Daresbury and Universities of Cambridge and Lancaster.

NGS Portal: Daresbury Laboratory staff are developing generic and application-specific portal interfaces for users of the National Grid Service. JISC and EPSRC funded.

OMII Grid Portlets: Daresbury Laboratory staff are exteinding their work on the NGS Portal to provide portlets which can be distributed and used in other projects. OMII funded.

ShibGrid: Developers at Daresbury and University of Oxford have integrated the Shibboleth security services with the NGS Portal. JISC funded.

SheBangS: Developers at Universities of Manchester and ??? have integrated the Shibboleth security services with other Portal services. JISC funded.

Integrative Biology: EPSRC-funded project lead by University of Oxford and involving Daresbury and Rutherford Appleton Laboratories and a number of other universitis. Staff at Daresbury set up the portal used for the JISC-funded IB VRE. IB has deployed and uses a data management and modelling infrastructure in common with several of the other projects (based on work from e-Minerals).

e-HTPX: BBSC/ DTI funded project which has run for 4 years and produced a data collection and analysis pipeline for high-throughput protein crystallography. This system was deployed on the Daresbury Synchrotron Radiation Facility and user by both academic researchers and drug companies such as Pfizer. Its portal interface and services are currently bing re-factored for use on the Diamond Light Source. The data analysis applications developed by CCP4 and used in this project run on NW-GRID.

e-Minerals: NERC-funded project

Reality Grid: steering portal

myGrid: myGrid workflow portal

: Manchester bioinformatics?

: Liverpool?

: Other Lancaster projects?

ESRC e-Infrastructure: University of Manchester, through its National Centre for e-Social Science (NCeSS) is leading a project to deploy a Grid infrastructure. This includes deployment of a Sakai-based portal and GROWL middleware from Daresbury and University of Lancaster.

:

4.2 VRE Components and Functionality

Work in the above projects has yielded a number of components which will be integrated into a fully functional VRE to be described below.

A major goal is to use the VRE to integrate various institutional systems and services to closely match the needs of real researchers. To do this, the following technologies will be investigated and used:

1. A highly-configurable and customisable portal with as many portlet-based tools as possible and available in different frameworks including Sakai and uPortal;
2. WSRP, Web services etc. for SOA and remote access;
3. Grid technologies including Globus, WS-RF, Condor, JSDL;
4. Registries using UDDI, or a semantic registry such as Grimoires or AgentX to work alongside the JISC IESR;
5. Disaggregated or decomposed Sakai worksite tools;
6. Other bridges to integrate tools written in languages other than Java such as PHP, Python etc.
7. Bridge between MS Exchange servers (Calendar, e-Mail etc.). We will investigate commercial solutions first, e.g. BEA Weblogic or SyncEx portlets (a preliminary evaluation document is available);
8. Authentication plugin based on JAAS for Shibboleth, Active Directory/ Kerberos, GSI/ X.509;
9. Information aggregators such as RSS or other fast event-based systems such as NaradaBroker or Tycho;
10. Audio-visual technologies, including bridging tools, for person-to-person and group communications.

5 Description and Timetable of Work Packages

The sections below provides a summary of VRE Work Packages, timetable and deliverables and indicates how the Work Packages inform each other in four phases throughout the project lifetime.

There will necessarily be a strong emphasis on integration through available standards and APIs.

The VRE will initialy address the following areas, but will be extended to cover others.

5.1 Sakai as a VO Enabling and Management Tool

Sakai has been adopted by several UK e-Science projects including the JISC-funded Sakai VRE Demonstrator Project. Besides functionalities like project management, Sakai is an ideal candidate for building up community portals. We aim to bring Sakai to Grid users in a multi-institution VRE.

The Grid and accociated middleware exist to enable multiple institutions to share resources. Sakai adds to Grid middleware by providing a user interface and a useful range of collaboration and administration tools. It also has the built in capability of managing groups of users and therefore supports the concept of Virtual Organisation (or VRC in the language of the OSI). Sakai is therefore an ideal candidate for building community-driven portals for Grid projects.

Sakai now also combines a number of ways of integrating distributed services into user-facing tools. Integration of Grid portlets with Sakai can be through: 1) Web Services for Remote Portlets (WSRP), and 2) Pluto integration inside Sakai. Both these approaches enable us to re-use Grid portlets developed for the UK National Grid Service (NGS) Portal and other project- or application-specific portlets. In addition, we have developed additional tools for Sakai including document a management tool for organising conferences and mashup tools for building Web 2.0 applications inside Sakai.

Such bespoke tools are made available to the VOs that require them. Each VO is provided with a "worksite" which can be customised to have a specific look-and-feel and configured to contain just the tools that are required. This can include interfaces to distributed services managed by a particular project or hosted as part of a Grid resource.

5.2 Project Management and Administration Tools

Existing Sakai tools will be immediately available. These are already being used fro project management and administration around NW-GRID. Each board (Project Board, Technical Board and Operations Board) has its own "worksite" with members able to see one or more as appropriate. These sites are managed by secretarial staff who can administred users, add papers, etc.

Link into institutional admin processes?

Adoption of the current production portals focussing on institutional/ research tools. Integration of new tools to meet the emerging requirements.

Deliverables: Central Sakai, uPortal and Luminis production servers established. All servers populated with portlets from VRE1 projects plus others identified in the public domain including: Agora communications portlet, CREE, SPP links to institutional repositories, Shibboleth/ Kerberos enabled login portlet, EVIE administration portlets.

5.3 Grid Resource Management Tools

Mention Ganglia, Gold, MOAB etc.

Adoption of the current production portals focussing on institutional/ research tools. Integration of new tools to meet the emerging requirements.

Deliverables: Deploy enhanced research Portals to: NCeSS, NGS, NW-Grid and Diamond Light Source. MyNGS and MyDLS on CD versions. Enhanced institutional Portals at CCLRC, Leeds, Lancaster and Hull. Software to include project- and institution-specific portlet tools to be identified, such as DataPortal meta-data search portlets, DLS User Office portlets.

5.4 Collaboration Tools

Adoption of the current production portals focussing on bridging tools. Integration of new tools to meet the emerging requirements.

Deliverables: asynchronous communications portlets. Other institutional tools (PHP, and Python) included via bridging technology to be identified, such as Exchange.

Adoption of the current production portals focusing on Web 2.0 tools. Integration of new tools to meet the emerging requirements.

Deliverables: Improved portals which have been evaluated by the user communities and which have some additional dynamic capabilities. Software to include links to remotely-hosted Web 2.0 tools to be identified, such as Gmail, Writely, etc.

5.5 Conferencing Tools

Lancaster staff have experience of Sakai and GROWL VRE projects, rollout to Lancaster University and National Centre for e-Social Science. The research group is in the e-Science Centre and is actively involved in the development of Sakai. For instance, Lancaster has written a collection of collaboration tools (Agora=video and voice over IP, shared desktop and white board) which will be distributed with the 2.4 release of Sakai in April 2007.

Lancaster will lead this work and will take responsibility for and contribute to the further development of audio-visual communication tools within the VRE.

Agora

This work will provide portal availability of open source, user friendly video and audio conferencing functionality, functionality which can currently only be found in complex and highly expensive systems such as Access Grid. OMII users will be able to seamlessly switch between video conferencing and any of the other portlet tools from the OMII stack. Agora, currently funded under the JISC-funded Sakai VRE Demonstrator project, is already very feature rich with video conferencing, shared desktop, shared whiteboard, movie casting, conference recording and chat, although some of the functionality is still being tested with user groups at Lancaster.

Agora is aimed at office, laboratory and home users and has been developed from the start to be very simple to setup; the media capturing software is simple to install, there is no need for expensive hardware (a £40 web cam and headset combination is all you need), you don't need multicast enabled routing in your institution and the user interface is very clean and intuitive. These features make it perfect for use by non-technical users in offices or laboratories or over broadband connections at home. Agora doesn't compete with Access Grid despite offering similar capabilities; it is complimentary. Access Grid is designed primarily for the scenario where small groups of people want to conference with other small groups and each group has access to a purpose built room from which to participate.

the motivating use cases or applications that will be used for testing ??????????

Milestones

1. The Agora software multicasting service needs to be factored out of Sakai's spring repository and into its own web application. This will mean that installation of the service will mean simple dropping the war file into Tomcat's webapps directory.
2. All the Agora service conference control APIs need to be exposed as web services
3. The Sakai JSF tool needs to be re-written as a JSR 168 portlet. Instead of loading the Spring service and directly calling the conference control APIs in memory, the portlet code will call the web services created in step 2 above.

5.6 Grid Computing

Daresbury staff have experience of Sakai and GROWL VRE projects, HPCPortal and NGS Portal, ShibGrid, ePubs and JISCmail, rollout to CCLRC departments, users of the National Grid Service, NW-Grid and Diamond Light Source. The Grid Technology Group which is to carry out this work sits in the e-Science Centre.

[Mention NGS Portal and OMII]

Daresbury will take responsibility for and contribute to the further development of Grid computing, services for experimental facilities, authentication and office integration tools. CCLRC also has its own open-archive ePubs service which will be made accessible to facility users via the portal.

5.7 Bioinformatics Portlets

[These are being developed at Manchester]

5.8 NW-GRID Flood-Modelling Portal

As was discussed at the NW-GRID meeting on 7th December 2006, in order to progress deliverable 3.2, Lancaster, Daresbury and Manchester will collaborate in producing a portal which provides access to the results of our flood modelling work. Visualising results from the model we use, Trim2D, is straightforward with little research potential. Therefore, we discussed implementing a mobile portal, which could also provide on-site functionality. This document provides a high-level sketch of this tool, along with a brief discussion of the challenges in implementing this functionality.

Proposed Application Overview

Our discussion of the portal application centred on implementing a location-aware mobile portal to our flood monitoring infrastructure, which seemed to be a good way of fulfilling deliverable 3.2, while producing novel research. There are three broad tasks that such a mobile portal could support: i) data recovery and repair, ii) real-time flood support and iii) data entry.

Data Recovery and Repair

While the wireless sensor network (WSN) is designed to be resilient and can be controlled from off-site, it is possible that the link between some areas of site and the lab can be broken, especially given the hostile conditions of a flood event. A mobile portal could be used to support data recovery and repair in the aftermath of such system failures:

Scenario: Users are equipped with location-aware mobile devices which interface with the WSN using WiFi or Bluetooth. The mobile portal shows a visualisation of the WSN and shows the nodes from which data can currently be recovered. As the mobile portal is location-aware, the system can instruct the user as to which location to go to next to continue data recovery, and thus implicitly illustrate broken links in the system.

Figure 1 – Data Recovery and Repair Scenario

This functionality occurred to me while thinking about scenario 3 – data entry. I think this is probably the weakest of the three suggested functions; however, automating data collection from environmental sensors could be useful in many situations, especially if data can be acquired over long ranges.

Challenges: decentralised network mapping, efficient data dissemination from the WSN to the mobile portal, acquiring data over long ranges, interfacing with sensor and networking middleware components and the inherent challenges of implementing a mobile/ embedded application.

Real-time Flood Support

It would be possible to provide support for evacuation and rescue in real-time using a location-aware mobile portal. Here, the portal device would show predicted flood-water levels. This data could be used by residents wishing to evacuate safely, or for other forms of real-time flood support.

Scenario: Users are equipped with the previously described mobile portal devices. The mobile portal would show a map of the site, overlaid with predicted flood conditions. The mobile portal could also provide a range of utility-features such as a) plotting an optimal path to safety and b) predicting the time remaining before a point-of-interest will be flooded.

Figure 2 – Flood Response Scenarios

This functionality was discussed early in the project as a way to highlight the benefits of having an intelligent and grid-enabled WSN. I think this scenario is interesting from a computing perspective as we can show grid technologies being used in a novel and real-time manner; however, it wouldn't be of much real-world use as the site has no residences or businesses on the flood plain.

Challenges: visualising Trim2D output, efficient dissemination of Trim2D results to the WSN, implementation of simple query processing on the WSN.

Data Entry

A mobile portal device could also be used to report additional data to the system. In this case, the device would be used to report flooding data using manual data entry from un-instrumented locations on-site and thus improve the accuracy of flood predictions.

Scenario: Users are equipped with the previously described mobile portal devices, however; rather than using these devices to interrogate the flood model, they are used to enter additional data. While

our scenario uses a dense sensor deployment, it is impossible to place sensors at every desirable location and for those areas that are un-instrumented; even simple manually-entered data would be useful. To support this, the mobile portal could show a map of the site divided up into a grid, for which the user can enter simple depth data, which will then be fed into the Trim2D model running on the cluster.

Figure 3 – Data Entry

Conversely to real-time flood support, this functionality is very useful from a real-world perspective, however, from a computing perspective, aside from the challenges of implementing this in a mobile embedded system, there probably isn't as much research depth.

Challenges: efficient "injection" of data from the user via the WSN to the cluster.

Going Forward

This document has presented a few ideas that could be incorporated into an access portal for the NWG flood modelling work. The proposed functionality could be implemented in a standard web-interface in the first instance and perhaps more interestingly on a location-aware mobile device to support a novel range of features: i) long-range data recovery, ii) real-time flood-responder support and iii) data entry for un-instrumented locations.

The Critical Path

It will be difficult for us to provide real-word site data in the short term as we are currently focused on deploying additional depth sensors and networking equipment. This is likely to remain the case for at least the next two months. In the interim, we can probably provide access to the output of predictions derived from existing data, which will be used on the cluster to test the use of population mote carlo sampling to sensor data for Trim2D. As the flood responder support functionality does not require extensive integration with the real-world WSN for meaningful evaluation, this might be a good place to start, moving on to the data-entry and data-recovery functionality as the WSN infrastructure is completed.

Hardware Platform

In terms of a hardware platform to support the mobile portal, I would recommend a number of GumStix equipped with VGA screens and GPS units. This setup has the advantage of being inherently compatible with all of the software developed at Lancaster along with being completely open and having a full Java Virtual Machine and standard C cross-compiler.

To maintain a low power profile, while achieving a reasonable range for the mobile portal, directional networking would be required. This would result in a range of a few KM rather than tens of meters and does not increase power consumption. To give an idea of what such a device might look like, see Figure 4 opposite, which shows a GumStix-based direction Bluetooth "sniper rifle". The mobile portal device would also of course require a display and we have a pair of suitable 7 VGA-resolution touch-screens.

In terms of user interaction, a directional device would require that users turn and point the rifle in the general direction of the WSN to input data or receive an update; however, it would mean that users could roam over 1KM from the system, while being able to connect.

Figure 4 – Directional BT Rifle

Two additional advantages of using our existing hardware platform are i) our pool of expertise with this platform and ii) many of the items used in the portal system have a dual-use from a WSN perspective.

Attached along with this document is a shopping list of the items that would be required to build a small number of mobile portal prototypes.

5.9 Lancaster

Other requirements from University of Lancaster Departments are as follows.

Portlet inteface to Java code for single sign on to the software environment we want to use on the NW-GRID. This will be based on work in the JISC ShibGrid and SheBangS projects and the Sakai VRE project.

Portlet interfaces are then required to the followin applications:

- Sabre (OSS Social Science, statistical computing, Economics, e-Sci)
- Environmental modellers (simulation for flood prediction, Computing Science/ Env Science, in house code)
- Siesta (licensed), Smeagol (OSS) (Molecular Engineers), Colin Lambert Physics (Condensed matter or Nano Physics)
- Non linear dynamics (home grown code) (Physics, Nano) (Theoretical Physics)
- Engineering and Computer Sci/Communication Systems network simulator (in house code called NS written in C++)
- Env Sci Flood model analysis (in house code)
- Env Sci weather modelling (in house code)
- Geography air pollution modelling (in house code)
- Statistics, Simulation and statistical computing (in house code)
- Medical Statistics Disease contagion modelling (in house code)

In addition to the main VRE portal Grid plugins are required for Matlab, (Commercial), used on the HPC by Env Sci, Management school (not sure of the applications), Psychology use it for the analysis of EEG (brain waves))

5.10 Liverpool

POL group - ensemble modelling of coupled ocean circulation / ecosystem (parameterisation study) - (in house, joint with MIT) - runs take weeks / months. NW-GRID use will be based on day-long, checkpoint restart jobs controlled using DAGMan

Earth Sciences - seismic modelling (several in-house codes) - again checkpoint / restart within Condor-G sounded interesting.

There is an EU project looking at capturing circa 1200 on-line streams of seismic sensor data that wants to be analysed in real-time - might benefit from a well-configured SRB.

Also after a meeting yesterday with the QCD people - hadronic physics - analysis of QCD data sets generated on supercomputers (in house) =¿ data staging from international sites is an interesting issue here so again a good SRB tie-in to the VRE would be important).

In addition, we have the usual suspects:

Chemistry - GAMESS-UK (widely used in Condor / Condor-G at present), pcgamess

Molecular Dynamics - VASP (also often used in parameterisation studies)

Nano technology - BSKAN (local code(?) for Scanning Tunnel stuff)

Surface Science - CRYSTAL - on the Surface Science to be used soon list

Chemistry/ Materials - NW-CHEM - mentioned as desirable by Werner.

Paul Watry's group in Library e-Services has an attractive annotation tool for collaborative working that would be worth while having included in the VRE tool set. Some of the semantic searching capabilities of Cheshire for things like searching abstracts/ papers in a domain of interest might also be a good inclusion. These are in use by many groups worldwide including the National Centre for Text Mining, NaCTeM.

5.11 National Centre for e-Social Science – Manchester Hub

The NCeSS Hub in Manchester is responsible for coordinating the ESRC e-Infrastructure project which is currently starting up. WP3.3 of this project is to develop and extend portal interfaces to relevant social science applications. The following text is taken from the work plan.

A comprehensive user interface is required for the various services to be deployed on the e-Infrastructure. This could be command line/ script based, a programming function library to link into desktop applications, or Web browser based. The GROWL middleware (WP 3.2) will be used for the former 2 cases. In WP3.3 we will provide Web interfaces capable of being integrated into a wide variety of portal frameworks alongside other familiar tools. We will provide a CD-based portal toolkit which projects (e.g. at the nodes) can install themselves. This will include a number of tested portlet interfaces to the e-Infrastructure, including job submission on NW-GRID and NGS and resource discovery. We will then work with node and other project staff to develop application-specific portlets for their

research. An alternative to local installation will be to host portlet-based tools in a centrally hosted portal at the NCeSS hub or at CQeSS on the Daresbury BladeCenter attached to NW-GRID.

- Deploy uPortal and Sakai as required – initially CQeSS (Lancaster), NCeSS (Manchester) and MOSES (Leeds));
- Provide CD with portal toolkit;
- Provide a 1-day course for portal developers and users in collaboration with ReDRESS and NGS Training Group;
- Collect additional user requirements;
- Work to help developers understand how to write portal interfaces;
- Help to develop application-specific portlets for nodes and other projects.

Deliverables will include: (i) Portal support/ deployment service for NCeSS Hub and Nodes; (ii) Usability requirements for portals; (iii) A hosted and maintained portal at CQeSS and NCeSS hub; (iv) Hosted and maintained portlet repository for tools and services.

The UK research community is particularly active in using portal technology to support collaborative activities and end-user access to distributed resources in an e-Infrastructure. Some portal development work currently funded by the JISC VRE and e-Science programmes will be used to provide a portal platform for the ESRC e-Infrastructure. This WP will build on the work being undertaken by CQeSS. CQeSS will work with application developers and data providers to create portal interfaces (using established standards, Java JSR-168 and WSRP) for the tools and services used and produced by NCeSS Hub and Nodes. These interfaces will be designed to meet appropriate usability requirements. Additional services will be provided to host portals as interfaces to the e-Infrastructure at CQeSS and on the NCeSS Hub. The preferred framework when the proposal was written was Sakai+uPortal, but in fact all tools can be provided through Sakai v2.4 which has full JSR-168 standard-compliant porlet support. Other frameworks such as GridSphere, LifeRay or eXo could be used if preferred by individual projects, or commercial platforms such as Oracle, Sun, BEA or WebSphere. CQeSS partners (Lancaster/ Daresbury) will share expertise and work together to set up the portal framework and establish a repository of re-usable portlet interfaces as part of these ongoing activities. This can include existing open source collaboration and Grid tools developed by the UK community, such as those already in use on the NGS.