White Rose Grid at Leeds:
e-Infrastructure for Research and Innovation

Figure 1: The WRG e-Science Centre's exhibition at the All Hands Meeting

Introduction

Since 2002 the White Rose Grid (WRG) consortium, which includes members of the Universities of Leeds, Sheffield and York, has been actively building and continuously enhancing the WRG e-Infrastructure with over £5m of investment. Increasingly, this newly established e-Infrastructure provides underpinnings for novel multidisciplinary and national and international collaborative research as well as cross site activities.

This leaflet summarises the White Rose Grid e-Infrastructure available for e-Research and innovative projects at the University of Leeds. Information on many other related e-Research activities, including current and past e-Science projects, is available from the WRG web site [1].

WRG e-Science Centre

The WRG e-Science Centre (Figure 1) is part of the UK e-Science Core Programme operated by EPSRC on behalf of all the Research Councils. The goal of this programme is to achieve a stable e-Science infrastructure by the end of 2007-08 [2]. The WRG e-Science Centre is one of seven such Centres across the UK (along with Belfast, Lancaster, Manchester, Newcastle, Oxford and the National e-Science Centre in Edinburgh and Glasgow) to receive further funding from EPSRC to continue their work for the next two years.

The WRG e-Science Centre's aims are to facilitate and be a focus for e-Research activities in the region.

Core Node Services

- Computation: (i) high performance computing (HPC) delivered via traditional means; and (ii) more innovative grid computing
- Data (and metadata) storage management using the San Diego Storage Resource Broker (SRB)
- Queuing and resource scheduling systems: Sun Grid Engine (SGE), Portable Batch System (PBS)
- Software environments
  - Compilers: Sun Forte Developer, PGI and Intel for C, C++, Fortran & Java
  - Libraries and tools: OpenMP, MPI, Performance analysis & debugging tools, NAG Numerical Libraries
  - Scientific visualisation: IRIS Explorer
- User management
- User support and training

Grid Technologies

- Authentication using PKI with X.509 digital certificates
- Certificate management through a MyProxy server
- Job submission/batch service - Globus Toolkit 2.4.3 and 4.0.3 /SGE and PBS
- Information service – MDS
- Grid monitoring- Ganglia, Inca
- Data access/integration services (GridFTP, SRB)
- Accounting (via SGE and PBS)
- WRG portal (the DAME portal utilising the Struts framework)

Registration Authority

The University of Leeds operates a Registration Authority (RA) for the UK Certification Authority. RA vets users' identities for the purpose of issuing UK e-Science X.509 digital certificates used by users to prove their identity to remote services available over the grid.

Compute Resources

The WRG core nodes hosted at the University of Leeds are key components of a coherent e-Infrastructure for research. These compute resources, supported by Information Systems Services, deliver huge amounts of computer power to a wide range of projects, including those in earth sciences, engineering, particle physics, life sciences, mathematics, food science, geography, computing, and medicine; as well as to many multi-disciplinary and e-Science projects.

The WRG's distinct approach to the exploitation of computer resources is to allocate these resources in such a way as to

Figure 2: Inside the NGS node hosted at the University of Leeds

Figure 3: Collaboration over Access Grid

bring together the provision of HPC services (with 75% resource allocation) and the emerging grid technology (25% resources).

Leeds Grid Node 1 - known as MAXIMA - is a general purpose facility - serving programs requiring large amounts of memory (up to 48GB), providing high throughput, as well as modest parallel capabilities. This is a constellation of shared-memory systems based on Sun Fire 6800 and V880 systems configured with UltraSPARC III Cu 900MHz processors and large physical memory.

Leeds Grid Node 2 offers capabilities for running highly parallel programs exploiting up to 256 processors as well as serving as a platform for developing knowledge of parallel programming techniques, tools and methods. It comprises two Beowulf type clusters (named SNOWDON and SNOWFLAKE) based on 2.2 & 2.4 GHz Intel Xeon processors interconnected with Myrinet 2000 networks, and in total delivering 292 CPUs.

e-Research Node - named EVEREST - constitutes a platform for highly-parallel programs using MPI, modest shared-memory parallel codes utilising OpenMP, and programs requiring large amounts of memory (up to 32 GB). This cluster is based on the latest AMD Opteron technology, and incorporates Sun Microsystems servers (V40z and V20z with 2.2 GHz dual-core processors) integrated by Streamline Computing, in total delivering 444 processor cores.

National Grid Service Node

The National Grid Service (NGS) [3] forms the foundation of the UK e-Infrastructure which underpins e-Research within the UK, providing standardised access to data management and compute resources and supporting collaborative computing across the UK, as well as offering a "national gateway" to related e-Infrastructures internationally.
The University of Leeds hosts one of the four core nodes of NGS (Figure 2). The service has recently progressed into the second phase which will enable researchers to take full advantage of the grid to enhance national and international e-Research and collaborations.

Visualisation Node

This facility is designed to support e-Science projects, and deliver visualisation across the network to users' current desktop systems. It comprises 16 Intel dual-processor nodes with nvidia graphics cards.

Storage Infrastructure

The Storage Area Network (SAN) provides 12 TB of resilient storage for WRG users.

Archiver

The archiving facility (12TB) based on EMC Centera servers offers a robust solution with two nodes; a primary unit at Leeds replicates data to a secondary device at the University of Sheffield. Access to the archiver is via an SRB interface.

UKlight

The WRG has a connection to the national UKLight [4] infrastructure which enables international collaborations by supporting experiments with high-bandwidth applications connecting to the US Starlight and the Netherlands NetherLight networks.

Access Grid

The Access Grid enables collaboration with the two other WRG partners and is used to support group-to-group interactions across the Internet. It is an essential means for many large-scale events such as e-Science Centres' Directors' meetings (Figure 3), collaborative work sessions with Beihang University in China, as well as the Worldwide Universities Network's seminars such as the one on SRB with the San Diego Supercomputer Centre.

Node of the National Centre for e-Social Science

This research node is part of the National Centre for e-Social science funded by the Economic and Social Research Council to investigate and promote the use of e-Science to benefit social science research. The project is developing a suite of modelling and simulation tools which are to exploit the capability of grid computing for e-Social science [5].

Training and Support

Training is delivered through courses in the area of HPC such as MPI or parallel programming as well as e-Science seminars, workshops and talks, for example on the Globus Toolkit 4.

Further Information

Contact:
Dr Joanna Schmidt, WRG e-Science Coordinator, email: j.g.schmidt@leeds.ac.uk

References:

[1] http://www.wrgrid.org.uk
[2] http://www.rcuk.ac.uk/escience
[3] http://www.grid-support.ac.uk/
[4] http://www.uklight.ac.uk/