The National Grid Service of the UK

Neil Geddes
CCLRC Rutherford Appleton laboratory
N.I.Geddes@rl.ac.uk
On behalf of the UK National Grid Service

Abstract

This paper describes the current status and programme of continued development and support for the UK's National Grid Service (NGS). Through the activities presented in this paper, the NGS will evolve to form the central component of the UK's e-Infrastructure for research, providing integrated access to research facilities and resources across the UK and internationally.

1. Introduction

The goal of the NGS is to provide coherent electronic access for UK researchers and their collaborators to all resources and facilities that they require to carry out their research. This access will be independent of location of resource or researcher. The focus of the NGS is on support for a production quality service to support research rather than grid research itself. Ultimately the National Grid Service will provide integrated, single sign on, access to the full range of the UK's computation and data based research facilities, together with a range of sophisticated services to support novel collaborative and cross resource activities. Through the NGS Partnership programme, the resources will span the complete space from advanced real time facilities such as the synchrotrons and neutron sources available to UK researchers [1] through to complex queries of historical data stored in national or institutional data centres.

The NGS will provide an infrastructure for combining computation and information from multiple data sources, reducing the need for specific arrangements for access and bespoke software or intensive researcher effort. This common infrastructure will be increasingly important as the number, scale and variety of resources and users develops. The NGS will operate the core services required to exploit National and International Partner facilities, will provide support to the users of the service, and will monitor and troubleshoot the services provided by all partners and the underlying infrastructure.

The NGS must pioneer systematic arrangements that relieve the load on both resource providers and resource users.

As well as supporting access to regional, national and international facilities, the partnership and affiliation programme of the NGS will provide a mechanism whereby local institutional or even departmental resources can be fully integrated into the NGS infrastructure. The goal is to support straightforward migration of work from local to national or international resources as required (and as authorised). This integration will provide a standard set of interfaces and services whereby local systems can be shared across institutions to optimise use and return on investment.

Overall the services supported by the NGS will support a National Grid infrastructure which will allow researchers to:

- systematically collect, process, preserve and make available digital information;
- easily navigate through the available resources;
- be confident in the quality of the services available;
- tie into international efforts

Examples of the services that will be available to researchers are

- Location independent data storage and replication
- Location independent access to institutional repositories and certified long term archival
- Location independent access to local, regional, national and international data resources
- Access to local, regional, national and international computational resources
- Internationally recognized electronic identity and identity management tools

- Tools for managing collaborative project or Virtual Organisation based authentication and authorisation
- Co-scheduling and operation of a wide range of national and international resources.
- More general tools to support distributed collaborative working e.g. portal framework and pluggable portal services to support a wide range of functions from mailing lists and meetings to advanced computation.

In addition the NGS will support

- 24 hour monitoring of the UK's grid infrastructure
- The policy framework for operation of the NGS (policies for partnership, user access, operations, development, security…)
- The ongoing review of partner services to both ensure that partnership remains an improving recognition of quality and to disseminate best practice.
- A central UK help desk and support centre, integrated into related international developments
- A repository of appropriate user and system documentation, FAQ's, best practice guides and training materials

These are key steps towards a consistent provision that will empower all researchers to access, integrate and analyse information. An important trigger towards democratisation of e-Research, where free-lance researchers, small groups in all laboratories and HEI institutions, as well as leading international groups, can mine the nation's combined information resources.

2. Current Status

The current National Grid Service, based around four cluster based computer systems and the UK's High Performance Computing services, entered full production in September 2004, following a pre-production test-phase that started in March 2004. During the period from March 2004 to June 2006, the NGS has continued to grow both in numbers of users and partner institutions [14].

The NGS is made up of different physical types of nodes but it can be useful to classify as one of two types, compute nodes and data nodes. Compute nodes are that are powerful in terms of computation, data nodes are less computationally powerful but have more storage available for data. Each of the compute nodes has a submission queue that is based on the number of CPU's required.

Each node also has a useful software environment that enables high performance parallel programs to be developed and executed.

The core node clusters are identical at present, but they do not need to be. The core compute nodes comprise 64 dual Intel Xeon 3.06 GHz CPUs, each with 2 GB of memory, two 123 GB internal disks and interconnected internally using high speed, low latency Myrinet message passing interconnects. They are very formidable computational resources for researchers without access to their own computational clusters with up to 128 processors available at each of the compute sites. The core data nodes each comprise 20 dual CPU Intel 3.06 GHz nodes with 4GB memory per node and connected to an 18TB Fibre SAN system for data storage.

2.1 Users

The NGS currently has almost 400 users, from a wide range of disciplines. The growth in new users has remained essentially linear since the start of the service, as shown in figure 1.

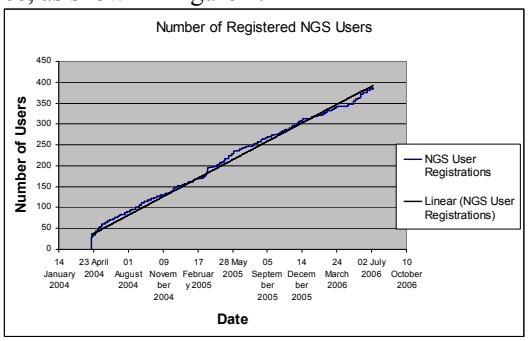

Figure 1. NGS user growth

Examples of current work include projects in chemistry (e-Minerals [2], e-Materials[3]), engineering (GEODISE [4]), Census data analysis, archaeology, medical imaging (MIAKT [5]), molecular dynamics (RealityGrid [6]), integrated biology and biological processing (Integrative Biology [7], e-HTPX [8] and others). The NGS also supports access through specific community or project "portals" (e.g. GEODISE and Bridges [9]). Users come from over 25 different institutions, mostly, but not exclusively from the UK. Most early NGS users focused on computational developments, leading to results published in Philosophical Transactions of the Royal Society, Physical Review Letters and numerous conferences.

2.2 Support centre

The Grid Operations Support Centre is a distributed virtual centre supporting the deployment and operation of the NGS. The GOSC comprises CCLRC, The White

Rose Grid at the University of Leeds, The University of Manchester and The University of Oxford. The National e-Science Centre at the University of Edinburgh also provides training support. The GOSC partners run the core nodes of the National Grid Service. The core NGS clusters support a diverse range of software applications. Some applications are available across all clusters, while some are only available on one or more clusters.

The GOSC operates a central helpdesk for the NGS, offering a one-stop-shop for help on all aspects of the grid service. Site specific questions are routed through to specific sites as required. The NGS helpdesk is integrated into the Europe wide support structure coordinated by EGEE.

2.3 Partners

In addition to the core NGS sites at CCLRC, Leeds, Oxford and Manchester, the UK's HPC facilities; CSAR and HPC-x, are also partners in the NGS. In addition, four new partner sites have joined the NGS since October 2004: Bristol, Cardiff, Lancaster and Westminster Universities. Following the specification of the NGS minimal software stack the limiting factor in this expansion is the level of core support that can currently be provided to integrate new partners. An additional seven sites are currently in negotiation to join the NGS. Careful specification of the NGS interfaces has allowed a diversity of sites to join the NGS.

Figure 2. Current NGS Partners

2.4 Procedures

A wide range of procedures have been developed to facilitate operation of the NGS. This covers user policy, site joining policy and security policy. Operations policies such as (security) incident response are integrated with similar activities in EGEE[1] and worldwide.

2.5 Documentation and training

NGS leverages training and related materials developed for the EGEE project. This has proved more difficult than hoped due to a range of incompatibilities between EGEE and NGS. The EGEE infrastructure has initially focused on a rather uniform software and hardware architecture. This, combined with the tight EGEE timescales which are driven in part by the real needs of the particle physics community, has led to technical choices and software products which are difficult to deploy and operate in a heterogeneous production service. EGEE is working to overcome these deficiencies as it's user base expands. Equally, EGEE and NGS experience is teaching us which of the EGEE decisions are absolutely vital for large scale production grids. From the user perspective differences between NGS and EGEE are gradually reducing. Despite the detailed differences a successful series of training events has been run at a number of locations in the UK and an increasing amount of NGS specific training material is available. Beyond induction training, user documentation is also available on the NGS website. Adequate and appropriate level user documentation remains a challenge, however, in the rapidly evolving grid world. This is an issue which confronts all grid infrastructure projects and will continue to be a problem until the technological and complexity gaps between the underlying infrastructure and the end users can be bridged. Grid infrastructures are still complex to use and increased basic training materials will be a required as the user base expands.

2.6 Stability

Basic stability and operational efficiency requirements have been specified for the NGS partner sites. To date these targets have easily been achieved but several monitoring requirements have been identified before more stringent requirements can usefully be defined. The core software stack deployed on the NGS is based on the Virtual Data Toolkit distribution. [12]. This has proved very reliable during the first 24 months production service with no significant service breaks being attributed to this area. This was confirmed during recent reviews of the services provided by the partner nodes. These reviews

[1] Enabling Grids for E-sciencE: a European grid infrastructure for research, connecting national e-infrastructure projects across Europe and beyond. . http://public.eu-egee.org/

also highlighted several areas of additional commonality between sites and requirements on central services.

As a result of the technology deployed and the procedures adopted, the NGS has achieved a production quality stable Grid resource for use by UK academics.

3.0 Future Development

Early in 2006, the NGS secured further funding to upgrade the core computational infrastructure, expand the partnership programme and operate the services through to 2009.

The goals for the NGS over the next 3 years can be summarised as:

1. Support the current service functionality and user base
2. Expansion to include new partners, increasing the capacity or range of services operated by the NGS
3. Improve interoperability with EGEE, TeraGrid[2] and DEISA[3]
4. Convergence with GridPP [10], other project/community infrastructure and Campus developments.
5. Work towards provision of value added services on top of the basic NGS infrastructure.

Progress on the NGS should wherever possible proceed by deployment and support of new services/functionality alongside existing ones with the latter phased out (replaced) only when replacements are robust. Existing services should be phased out on conservative timescales compatible with user account refresh but the NGS must be proactive in moving people on where improved services become available.

Our ultimate goal must be that as a default there is only one underlying UK e-infrastructure for research, with different logical views of this infrastructure from each of the user (VO view), resource provider and operational perspectives. In the same sense that there is only one internet this does not necessarily imply a single centrally administered and controlled infrastructure.

Architecturally we expect the services provided to gradually migrate towards a collection of interoperable open standard Web Service interfaces. Full support of more advanced services will be dependent upon provision of robust solutions that can effectively interoperate with the deployed infrastructure, for example, those emerging from the OMII [11] managed programme.

Detailed planning for the next 3 years is in progress, but several key developments are already identified which are outlined below.

3.1. Baseline services

The NGS will gradually support a complete range of "Baseline Services" which can be composed as required by users. These services will need to be compatible and interoperable with those offered on partner infrastructures. The range of services provided will include

1. Data storage
2. Data transfer
3. Data catalogue/management
4. Data integration
5. Computational resources
6. Virtual organisation management
7. Information registries

3.2 New partners

The NGS will expand through integration and support of new partners. Extension of the integration work will include access to data from National data centres [13], key data resources from bio-medical and other domains and research facilities (including ISIS, DIAMOND, LHC). In addition we will work with the VizNET Visualization Network to integrate visualization services with the NGS.

Ultimately the NGS partnership model aims to be "operator free". Automated compliance testing and monitoring will allow partners sites to fully integrate into the National infrastructure themselves, without requiring significant hand-holding. Only in this way can the infrastructure scale to a fully national service.

Over the next three years, the NGS will target specific examples of each of the following resources, to serve as exemplars and test cases for further partners.

1. Next generation HPC system
2. University campus grid
3. Dynamic service hosting environment

---

[2] TeraGrid is the US National Science Foundation's effort to build and deploy the world's largest distributed infrastructure for open scientific research. http://www.teragrid.org/

[3] DEISA – The Distributed European Infrastructure for Supercomputing Appications is a consortium of leading national supercomputing centres that currently deploys and operates a persistent, production quality, distributed supercomputing environment with continental scope http://www.deisa.org/

4. Large scientific instrument
5. National data centre
6. Institutional data repository

3.3 New user communities

The benefits of a shared e-Infrastructure are twofold: the amortisation of establishment and operational costs, and the development of a culture based on the common infrastructure. Success depends upon a wide variety of disciplines, groups and institutions joining in. A key driver will be the lowering of the threshold to transfer and develop applications and research procedures. To lower this threshold the NGS, working in concert with OMII-UK and others, will provide frameworks for applications, portals and virtual organisation administration, so that the researchers, application developers and service providers are led towards effective software and practices, and are provided with exemplars of best practice and ready made solutions to common requirements.

Engaging with large coordinated communities has proved most successful to date. For more disparate and loosely coupled activities the initial step to exploiting any advanced infrastructure can prove a significant barrier as no one group can make the significant initial investment required. To address this, opportunities to directly fund community projects or community "portals" will be made available. The goal will be to directly engage with new user communities by developing targeted applications which support access to the NGS. Common tools and services which emerge from these developments will then be supported by the NGS.

3.4 Training and outreach

The emergence of e-Infrastructure as national and campus facilities is a major evolution in the context for research and for the provision of ICT. For several years this will continue to be a rapid and disruptive change bringing with it both challenges and opportunities. For the UK to make the best use of this investment it is necessary to alert researchers and practitioners to the challenges and opportunities through outreach and to enable them to respond effectively through training. In addition to helpdesk support there is a continuum from the help desk through to training and a spectrum of activities from dissemination, through training to consultancy. Any provision must integrate across these dimensions and adapt to the varying technical contexts, the various communities and the geographic distribution involved in providing and using e-Infrastructure. The range of issues that must be covered across these dimensions are:

1. Support for decision makers – informing senior staff so that they can make decisions for their institutions or communities about when and how to engage in providing or using national e-Infrastructure.
2. Support for e-Infrastructure providers – assistance in developing the skills and knowledge necessary to set up and operate safely e-Infrastructure nodes or to connect campus e-Infrastructure with the national systems.
3. Support for users – attracting and informing new users as to the opportunities and how best to exploit them and refreshing and advancing their skills and knowledge as the e-Infrastructure capabilities progress – this often has to be targeted at specific discipline communities.
4. Support for application, portal and tool developers – this community requires expert and specific technical skill development – training has to present the rapidly emerging technologies and enable rapid growth in this community so that they can provide the tools for the user communities.
5. Support for educators – the educators in our universities need to equip their graduates and post graduates with the judgement and skills that will enable them to take full advantage of the e-Infrastructure context – they need help transforming their curricula and organising this new conceptual and practical environment so that it can be incorporated in their courses and hands-on classes.

3.5 Integration with computing service providers

Development of effective accounting and auditing will be central to the integration of computing service providers into the NGS. Initially, service providers are likely to be partner institutions or facilities funded explicitly to support researchers (e.g. NGS core nodes, LHC tier-1, National data centres). However, we expect 3rd party suppliers of both data and compute services to become more important both as the NGS interfaces mature and as explicit funding (whether direct or indirect) for services used by researchers becomes more common. Important components in the

closer integration with computing service providers are:

1. Well defined interfaces and procedures that allow a clear understanding of the roles and commitments of service providers re the NGS and vice versa.
2. A well understood core software stack with components of known provenance, maturity and support.
3. Effective metering and auditing of services and users
4. For Universities in particular, integration of the NGS authentication and authorisation framework with the standard national and campus systems (e.g. Shibboleth)
5. Training, documentation and awareness raising targeted at computing services leadership and staff.

As the number of partners grows and partners become more integrated into the NGS it will be essential to leverage exiting organisations such as University computing services to develop appropriate awareness raising, training and self-help/support infrastructures that will allow the NGS to scale effectively to cover the whole UK. Improving the identification of partners with the NGS can be achieved e.g. though focus community meetings and by supporting partners in representing the NGS at appropriate conferences and workshops.

4. Summary

The National Grid Service of the UK is already providing a production level service integrating access to computational and data resources across the UK The open standards based infrastructure provides a solid basis for the continued development of a collaborative shared infrastructure to support UK researchers. Over the coming 3 years the Service will expand to include a greater range and number of research related IT resources. The pace with which the service expands and, perhaps more critically, with which value added services can be built on top of the basic infrastructure, will depend in large part on the pace with which robust tools can be provided by both industry and academe. This is itself encouraged by the provision and use of a common e-Infrastructure which provides the reference points for discussion, facilitates the exchange of technical solutions and the shaping of policy and priorities. A coherent UK voice will strengthen the position of the UK in negotiating international tool provision and e-Infrastructure inter-working so that UK researchers using international facilities will benefit and imported software will be better suited to UK requirements. The common e-Infrastructure and a common culture will mean that the informal learning prevalent in the development of research and innovative skills will retain its value as researchers move between organisations as their career progresses.

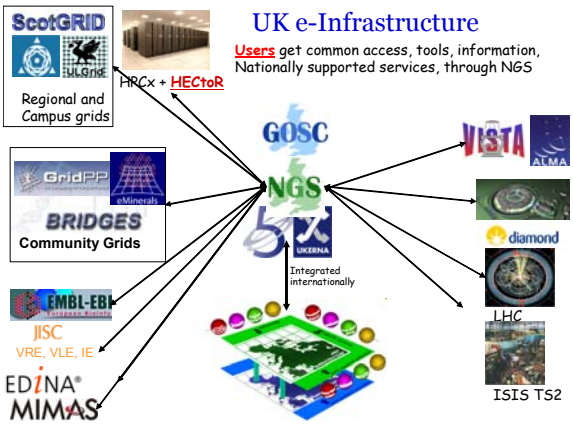

Figure 3. NGS: the heart of the UK's e-infrastructure

5. References

[1] For information on the range of facilities supported, see www.cclrc.ac.uk

[2] www.esc.cam.ac.uk/eminerals/

[3] For further details on this project, see www.e-science.clrc.ac.uk/web/projects/complexmaterials

[4] www.geodise.org/

[5] www.aktors.org/miakt/

[6] www.realitygrid.org/

[7] www.integrativebiology.ac.uk/

[8] clyde.dl.ac.uk/e-htpx/index.htm

[9] www.brc.dcs.gla.ac.uk/projects/bridges/

[10] GridPP in the UK Particle Physics Grid. See www.gridpp.ac.uk for further information.

[11] The Open Middlware Infrastructure Institute provides a web service infrastructure for building grid applications. See www.omii.ac.uk for further information.

[12] http://vdt.cs.wisc.edu/

[13] The UK supports several National Data Centres including EDINA (http://edina.ac.uk), MIMAS (http://www.mimas.ac.uk) and the Arts and Humanities Data Service (http://ahds.ac.uk)

[14] For further, current, information on the National Grid Service, see http://www.ngs.ac.uk