



Workshop on Sustainability and Future Business Models for the UK National Grid Service National e-Science Centre, Edinburgh

Neil Geddes, Andrew Richards, Science and Technology Facilities Council
Malcolm Atkinson, University of Edinburgh
Steve Chidlow, White Rose Grid, University of Leeds
W. T. Hewitt, Stephen Pickles, University of Manchester
Paul Jeffreys, University of Oxford

25-05-2007

Abstract:

This workshop brought together key representatives from a range of UK (e-infrastructure) service providers and funding organisations to investigate some of the key requirements that must be met to enable the UK's National Grid Service, or its successor, to evolve into a sustainable persistent infrastructure for the UK. Input was provided by a number of participants representing the National Grid Service, University computing services, National HPC services, National data centres, and International grid infrastructures. The workshop emphasized the importance of users and being user led, including:

- The need for well defined, robust services.
- The role of "brokering" between the different service providers and between service providers and user communities.
- Providing help and recommendations on standardisation and running production (grid) services
- Evangelism and outreach
- Engagement with Institutions, both for service provision and to support effective user authentication.

Specific opportunities identified included

- The drive towards shared services across the public sector, including Universities.
- Ensuring effective exploitation of the large SRIF3 investments in computing.
- Support tools for distributed collaborations.

A response by the National Grid Service to the points raised in the workshop is included in an appendix.

UK e-Science Technical Report Series

Report UKeS-2007-05

Available from http://www.nesc.ac.uk/technical_papers/UKeS-2007-05.pdf

Copyright © 2007 The University of Edinburgh. All rights reserved.

Workshop on Sustainability and Future Business Models for the UK National Grid Service

National e-Science Centre, Edinburgh

February 22-23, 2007

1. Executive Summary

This workshop brought together key representatives from a range of UK (e-infrastructure) service providers and funding organisations to investigate some of the key requirements that must be met to enable the UK's National Grid Service, or its successor, to evolve into a sustainable persistent infrastructure for the UK. Input was provided by a number participants representing; the National Grid Service, University computing services, National HPC services, National data centres, and International grid infrastructures. The workshop emphasized the importance of users and being user led, including:

- The need for well defined, robust services.
- The role of “brokering” between the different service providers and between service providers and user communities.
- Providing help and recommendations on standardisation and running production (grid) services
- Evangelism and outreach
- Engagement with Institutions, both for service provision and to support effective user authentication.

Specific opportunities identified included

- The drive towards shared services across the public sector, including Universities.
- Ensuring effective exploitation of the large SRIF3 investments in computing.
- Support tools for distributed collaborations.

Ultimately users and user need ensure sustainability.

2. Introduction

This workshop brought together key representatives from a range of UK (e-infrastructure) service providers and funding organisations. The aim of the workshop was to investigate some of the key requirements that must be met to enable the UK's National Grid Service, or its successor, to evolve into a sustainable persistent

infrastructure for the UK. The workshop focussed on the issues and problems around providing, funding and sustaining services in an integrated infrastructure.

The specific objectives of the workshop were to understand what the represented bodies want a national grid service to be and not be and what is needed to achieve this.

The workshop programme was as follows:

Day 1, 14:00-17:30

Introduction	Neil Geddes
NGS Partnership	Stephen Pickles
University resources and Full Economic Costing	Pete Clarke
National Computational Services	David Henty
National Data Centres	Peter Burnhill
D-Grid progress towards sustainability	Wolfgang Gentzsch
Commercial Services	Terry Hewitt
outsourcing	Rhys Newman

Dinner discussion of issues and requirements for a national grid service.

Day 2, 09:00

Breakout groups to discuss

- Requirements and issues from the perspective of service providers
- Requirements and issues from the perspective of users and funders

Summary and next steps

The goal of the first day was to provide input to the discussions of the second day. Speakers were given a very broad brief to explain their perspective, expose their requirements and/or constraints on any national infrastructure, and raise any general issues and challenges. We have not attempted to reproduce the full presentations here. These should be available from the NeSC website at the following URL:

Original Event Page: <http://www.nesc.ac.uk/esi/events/731/>

Event Material : <http://www.nesc.ac.uk/action/esi/contribution.cfm?Title=731>

In this report we only give a very brief summary of day 1, focussing primarily on the discussions from day 2.

3. Day 1

Introduction - Neil Geddes

Neil gave an introduction to the Workshop and a brief overview of the mission and goals of the current National Grid Service (NGS).

NGS Partnership - Stephen Pickles

Stephen gave an overview of the current NGS partnership programme. He summarised the current and possible future partners, providing justifications as to why NGS partnership was valuable for the partners.

University resources and Full Economic Costing - Pete Clarke

Peter explained the requirements that “Full Economic Cost” (FEC) were now placing upon University computing services. He suggested ways in which a National Grid Service could fit into this model, both as an enabling centre of expertise and as a forum for sharing University resources. Additionally such a service could provide access to paid for services. The FEC regime presents some significant challenges to Campus computing. Peter suggested that computing should be an overhead activity, but to achieve this required that services became much more general and agile (like the telephone systems). There is a need to break the link between “computing service” and “owning a computer”. If this can not happen there is a risk that FEC could kill off much University funded computing. Universities need to provide “free at the point of use” computing for many of their members in addition to the ability to support guaranteed services directly allocated to research projects. Many universities were now coming to mixed models supporting both.

National Computational Services - David Henty

David summarised the history of UK HPC services over the last 12 years. Went on to describe why grid computing is not more prevalent in HPC. This includes:

- Systems are specialised
- Projects allocated substantial CPU/disk budgets at grant stage and support large multi user collaborations running for many years on the same system.
- Applications are not standard packages. Projects typically spend a lot of time developing and debugging software.
- most users are experts (and code developers)
- Data volume is a huge issue

David went on to describe the DEISA project which is addressing grid style metacomputing across HPC sites. Interestingly there is work underway within DEISA to abstract the HPC layers under OGF/SAGA compliant interface layers which allow for use with more sophisticated grid computing tools such as workflow management.

National Data Centres - Peter Burnhill

Peter gave a brief overview of the National Data Centres, Edina and MIMAS. He explained the user (service) focus of the centres and the emphasized the importance of a robust and integrated user authentication and authorisation system supporting open and licensed access Any new or emerging National infrastructure would have to provide full integration with the systems supported by and for the National Data Centres. This is Athens at present and Shibboleth in the future. An important feature here is that the user authentication lies with the Universities. It is important to engage users through these institutions. Infrastructure (user) communities take a long time to grow. Therefore, it is important that any new infrastructure has some strategic security of funding. This must be accompanied, however, by a robust review of the services provided. Peter Indicated that the Data Centres have a long history of providing user focused services, which could be valuable to the NGS.

D-Grid progress towards sustainability - Wolfgang Gentzsch

Wolfgang gave an overview of the D-Grid project and its progress towards becoming a sustainable infrastructure in Germany. A workshop on D-Grid sustainability was held in October 2006. Wolfgang emphasized that “need” ensures sustainability. In general users like stability and do not like change. The current, second, phase of D-Grid includes commercial service providers in order to test new business models; looking towards sustainability. Inevitably, the Government will continue to fund the Grid resources themselves. D-Grid is based on creating virtual competence centres based around existing expertise. Historically too many attempts to set up new “green field” organisations have failed, primarily because they did not value and manage the existing people expertise effectively. Sustainability needs a professional service, a professional service needs experts, and expertise needs long term funding. D-Grid plans virtual centres for generic middleware, grid resources, user support and knowledge (research). Applications are a strong driver for D-Grid with little support for the belief that if “you build it they will come”. A Service Level Agreement (SLA) layer is being planned between the user communities and the generic middleware and grid services. D-Grid have invested in a downloadable “D-Grid software stack”, where quality, integration and interoperability is assured by the project.

Commercial Services - Terry Hewitt

Terry gave an overview of the CSAR service which ran from 1998 to 2006 and in particular of the novel (at the time) funding model used for this service. He emphasized that there were already organisations who are prepared to offer guaranteed computational services via the grid on a commercial basis. A major hurdle is a suitable funding model for users. In particular one which would be fully integrated into that for other national research resources such as the National HPC services.

Outsourcing - Rhys Newman

Rhys explained how, having been responsible for coordinating the provision of a new university computer building, he was now convinced that Universities should outsource their large scale research computing. The costs associated with commissioning, constructing and operating modern large scale computing installations are not justifiable in the era of FEC. So far, however, no major Universities had successfully outsourced significant computing. Rhys also reported on a project he is involved with which aims to provide a virtualisation layer to support creation and management of “cycle scavenging” grids. An important component of the project is the architecture independence of the infrastructure as viewed by the users. This project initiated much discussion, accompanied by some scepticism of cycle scavenging grids in general.

4. Day 2

The discussions are reported in summary. There was no attempt to reach a formal agreement on the points raised, however, many of them reflect a consensus which emerged from subgroup discussions during the meeting. Here we simply report the comments made and where there was strong agreement or disagreement.

Breakout Session: Requirements and issues from the perspective of service providers

A Key aspect to understand in defining the framework for any national grid service is the focus on services and service provision. It is important to understand better the nature of the service or services offered by (or through) a national grid.

It is clear that there are and will always be a number of different classes of user. Different users will have very different requirements for authentication and authorisation. This is particularly relevant where we expect services to be used e.g. in turn by university ICT service providers who will wish systems to be integrated with their existing authentication systems which give them local control.

There are a number of separate classes of service provider; National providers such as UK HPC, local HPC systems and high throughput computing systems based on campus grids and Condor. Access to these systems may require very different levels of authentication and they may provide very different levels of user service.

There are significant challenges around licensing of commercial software (and data) in grid environments. Issues arise around who the (potential) users are, which machines the licenses will be used on and who owns these machines. Many of these questions can not be answered a priori and current commercial licensing models rarely work for grid computing. Much relevant licensing already takes place at an institutional level which shows that academic community licensing models are possible. It was agreed that addressing issues around licensing can benefit from larger scale aggregated negotiations. Developments on a European scale could be helpful here. It is not just software which is licensed, data is licensed too. Should the NGS be a gateway to licensed software? Ultimately this will depend upon the available licensing models and whether these would allow negotiation or management on behalf of user groups. In any case Service Providers are conscious of terms of licenses they hold and are generally held responsible for enforcing the license terms.

To be successful a grid infrastructure needs to be rich in services that people want. This want or need is ultimately what will ensure stability. Is there a range of existing services that would benefit from grid integration or delivery? Or where their users would benefit from such integration? What new services can be conceived?

“Data” is clearly an important area. Accessible storage is needed for an increasing range of primary research data. There is also a growing need to store secondary user generated or derived data. Funding models are deficient in this area at present. In practice, there is value added data, which is expensive to re-create (e.g. bio-informatics, large facility) and is valuable for secondary use. All of these are general data issues. What NGS should do for data? Should NGS “cache” data? Provide places for replicas? Should NGS backup data? Should NGS curate data? Should NGS preserve data? Probably not for the latter two, but this depends upon the detailed definition of the NGS. There are clear roles for “partners” in all of these areas with access possible through an integrating “NGS”.

Users and organisations increasingly are looking to outsource off-site backups/replicas. How does NGS fit into this? Is there a brokerage role or service role

here? Can NGS recommend what the appropriate interfaces should be, develop appropriate SLAs, or possibly act as the centre of a support network?

In each of these data related areas there is a clear brokering role between institutions. NGS should work with data service providers to develop suitable models. Institutions are the guardians of user identity and are also candidate service providers.

In the longer term, should NGS be a service provider to end users, or provider to provider broker? NGS should focus on things that must be done on the network level and can't be done at the institutional level.

Looking towards the future the NGS could provide a platform for developing and expressing demands on service providers at institutions and elsewhere. There may be an opportunity to develop a marketplace for specialist services that not all institutions can afford. NGS needs to identify and advertise places where these exist. Cost re-allocation between institutions would be a real coup! There is a fear that local university services are dumbing down, becoming more generic, and creating gaps.

There are opportunities for the NGS as advocate, engager of service providers and recommender of interfaces.

Breakout Session: Requirements and issues from the perspective of Users

Full Economic Costing is necessarily bringing the full cost of computing/IT services under scrutiny. Universities are seeking to recoup the costs of investments wherever possible, for example, through charging for use of their services, particularly to external users. This will reduce the availability of "free" service offered by the NGS to those from the core nodes, or services explicitly funded to provide free at the point of use facilities. The NGS should move towards an open market where partners are free to declare their own requirements and restrictions on (charging for) services.

Users of the grid currently suffer from poor levels of control and visibility through the various layers of software/middleware. The grid is currently complex, often counter intuitive and the complexity is not hidden from the users. Users often receive poor and confusing feedback.

What is the "NGS Service" from a user perspective? There needs to be more emphasis on getting the key messages across and on identifying exactly what these are, e.g. stability and data storage. Outreach and evangelising are vital for the NGS. There should be a dedicated "Outreach" post. The NGS should place more emphasis on data storage rather than compute.

There will be increasing demand for Windows access. More and more (commercial) programmes are moving the Windows O/S as their primary platform. This will continue with the emergence of the Windows HPC platform.

The NGS aims to support distributed collaboration. However, there is little or nothing in the way of collaboration tools explicitly supported by the NGS. This is an area which could easily be improved, not necessarily by the NGS developing or supporting its own tools, but by sharing expertise and experience.

5. Specific Additional Discussions

What is the role for a national grid service?

The consensus view was a role providing coordination of UK efforts in network ICT, including, for example:

- Liaison with other “National Grid” efforts
- Blessing of key central services, not necessarily provided by the NGS
- Operating as a benchmark of software interfaces
- Operating central services necessary to create the distributed communities
- Access to specialised or novel facilities not available at all universities
- A forum for sharing expertise and best practice

In addition there are clear needs and opportunities in providing data storage, facilitating inter-group collaborations and in providing advice and support for effective exploitation of the large SRIF3 investment in computing infrastructure.

What should a national grid service not be/do?

The NGS should not compete with things better done elsewhere or attempt to define the underlying software. The NGS should focus on only the interfaces.

How should a national grid service be governed?

Any National Grid Service should be user led with an appropriate steering committee reflecting this. Ultimately it should be an independent body which provides a (paid for) service that the “users” can influence and/or walk away from.

How should a national grid service be funded?

A wide range of ideas and requirements were identified here, including:

- Compute resources should be through university or national resources (i.e. not funded by the NGS itself).
- The NGS should be funded in 5 year tranches with rigorous review of outputs before renewal.
- The “shared services” agenda presents an opportunity for the NGS (and others) to remove the need for things to be done locally.
- Several funding streams are necessary to avoid over dependence upon any single body or award.
- Lottery grant (after-dinner suggestion!)

6. Summary and Next Steps

Key themes emerging during the discussion and summary sessions were:

- The importance of users and being user led.
- The need for well defined, robust services.
- The importance of data and the need to define the roles of the NGS and its partners in this area.
- The role of “brokering” between the different service providers and between service providers and user communities.
- Providing help and recommendations on standardisation and running production (grid) services
- Evangelism and outreach and the importance of dedicated effort here.

In addition, several specific opportunities were identified, including:

- The drive towards shared services across the public sector, including Universities.
- Ensuring effective exploitation of the large SRIF3 investments in computing.
- Support tools for distributed collaborations.

In wrapping up, the question was asked “If we didn’t have a national grid service today would we want to create one?” A range of responses was forthcoming; from “definitely yes” through to “probably not”.

Follow up actions from the workshop are:

1. Reinstate plans to form a user led steering group for the current NGS.
2. Improve the information available on the services offered by the NGS.
3. Appoint an outreach officer for the NGS.
4. Support for collaborations
5. Write up a (brief) summary of the workshop.
6. Focus on services

Appendix: A Response from the National Grid Service to the NGS Sustainability Workshop

Neil Geddes, Science and Technology Facilities Council

Background

Intentionally, this workshop did not begin with a comprehensive vision or overview of the current UK National Grid Service (NGS). This was to encourage participants to bring to the workshop their own interpretations of what is meant by such a service, including their unconstrained aspirations and their prejudices. In this way we, the NGS, might benefit from their vision, while being able to address misinterpretations and prejudices following the workshop. This response in part addresses this, and in part offers an initial formal response to the workshop.

The vision of the National Grid Service is to provide coherent electronic access for UK researchers to all the computational and data based resources and facilities required to carry out their research, independent of resource or researcher location. To achieve this, the NGS is leading the deployment of a common grid infrastructure for combining services and information from multiple sources. The goal is integration of services and resources to support straightforward exploitation of local, national and international resources as required (and authorised). The NGS aims to provide a standard set of interfaces and tools whereby local systems can be shared across institutions to optimise both use and return on investment. The NGS must pioneer systematic arrangements that relieve the development load on both resource providers and resource users in order to reduce the need for specific arrangements for access, bespoke software, or intensive researcher effort. Through the NGS Partnership programme, NGS seeks to integrate services to access a growing number, scale and variety of resources spanning the complete space from advanced real time facilities such as synchrotrons and telescopes through to complex queries of historical data stored in national or institutional data centres.

Partnership is a key theme. The NGS itself operates the core services required to exploit local, National and International Partner facilities, provides support to the users of the service, and will monitor and help troubleshoot the services provided by all partners and the underlying infrastructure. However, the NGS itself can not directly control or operate all of the international resources that researchers require. It is extremely unlikely that the NGS, or any successor infrastructure, would ever “own” or control all of the resources that researchers would need to access even considering the UK alone. Therefore, a federated approach such as the NGS partnership model is central to the successful future evolution. In some sense Partnership continually changes, and broadens, the scope of the NGS. This is the intention. The scope of the NGS is and should be defined by its partners within the overall vision of an integrated infrastructure.

To exploit fully the potential of a shared infrastructure requires that key stake holders, including resource owners and funders as well as researchers, develop and agree a shared vision of the future and of their own responsibilities in developing and supporting it. Any vision which does not meet the needs of these stakeholders will be

a largely pointless exercise, but equally any infrastructure missing key existing and planned resources through lack of engagement will be a sorely missed opportunity. This is particularly true in emerging inter-disciplinary fields which, by definition, will bring the need to work across traditionally separated resources. Many existing and planned research facilities, of all types, have previously not identified the implied integration as a benefit and may see it as a threat. Arguments against such integration, e.g. from specialised services or facilities, miss the point. Many services are specialised, be they the UKs most powerful HPC system, the data from unique instruments such as Diamond, data from national data services, or systems where specialist software is installed. The debate should rather be whether or not the research community can benefit from more integrated access to and exploitation of the full spectrum of such resources nationally and internationally. Staying out of this (“grid”) game would appear to restrict rather than enhance the potential exploitation of the resources and thereby encourage rather introverted or closed communities.

Key Points Raised During the Workshop

The “brokering” and championing role of the NGS was identified as important in two ways; as an arbiter of best practice and approved mechanisms/interfaces, and as a neutral match-maker between service providers and users. The emphasis at the workshop implied that this represents an opportunity at least as much as a current status. This in turn reflects the lack of a truly shared vision across the workshop participants. Further work is clearly needed to develop and support such a vision. The workshop clearly identified the National Grid Service as one potential mechanism for achieving this but is equally seen as a threat or competitor in some areas represented: e.g. *“The NGS should not compete with things better done elsewhere or attempt to define the underlying software”*. Clearly there is a need for a body to take on this broking and coordinating activity. **NGS will work to further develop interface/standards recommendations and a brokering role and is keen to work with others to define and achieve the goals of a broader infrastructure.**

The NGS sites now have extensive shared experience running production grids which integrate small, medium and large HPC systems together with data storage systems. This experience will be valuable to the increasing number of institutions investing significantly in HPC for the first time. **NGS will investigate how best to exploit this expertise and will seek to work with groups such as HPC-SIG and HPCSC to achieve this.**

The issue of “Shared Services” across the public sector was mentioned on several occasions. Common or centralised provision of existing or standard services (related to computing and data management), while possible, is not a current driver for the NGS itself. Candidates for such services could emerge naturally from the model of robust services that the NGS aims to deploy/support, but equally, relevant services could be defined elsewhere. Wherever such a service was developed, there should clearly be a strong requirement for integration with the NGS through formal partnership.

A recurring theme during the workshop discussion was a focus on the self-sustainability of a set/system of well defined services. In practice, many of these in fact may already exist. The key is to access and/or integrate them. This requires commitment from the services which in turn requires at least acknowledgement and

probably commitment from the funding bodies. Services are at the heart of the NGS. Not simply reproducing existing features that can be bought “off the shelf” but addressing current and future needs of researchers collaborating across boundaries. The boundaries may be subject, geographical, institutional, or associated with particular resources or facilities. At all times, however, the NGS remains careful to deploy functionality supported by robust managed services. The balance between providing well defined stable services, versus driving the development of innovative new integrating activities is a key component of the shared vision that is needed. **The NGS will continue to deploy robust managed services based on open interoperable standards.**

The importance of outreach and user focus was emphasised at the workshop. Both of these are recognised by the NGS as vital for the success of the UK’s e-Infrastructure. The NGS web site and support material is currently undergoing a detailed reorganisation in order to better meet the needs of users. At the same time **NGS are working with JISC to develop a programme of development activities targeted at developing and supporting new user communities.** The initial components of this work have been funded under the JISC e-Infrastructure programme. Additional opportunities for new communities should be announced shortly under this same programme. For outreach activities, **NGS is currently seeking to appoint a dedicated outreach/liason officer** with full authority/responsibility for developing new user communities. This post will also work closely with related activities in GridPP (<http://www.gridpp.ac.uk>) and EGEE (<http://public.eu-egee.org>). **In addition, NGS will set up a formal User Board to advise on support of the existing user community.**