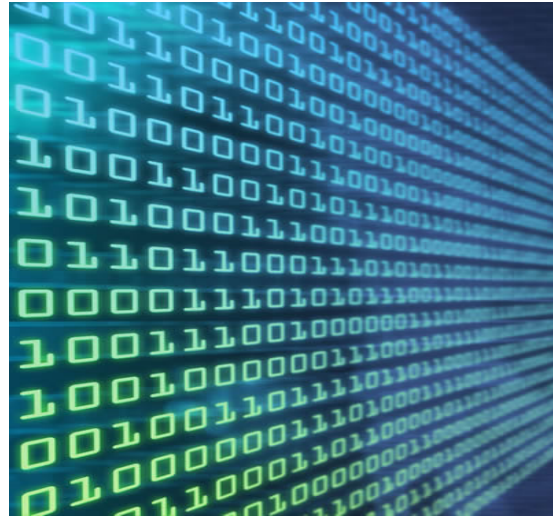


## Joined up HPC Accounting: Grid-SAFE and the NGS

*"Grid-SAFE software has the potential to allow the NGS to account across a broader range of institutes than at present, whilst maintaining a standards based approach that does not commit institutes to use a specific software solution on their resources."*

**Andrew Richards, Executive  
Director National Grid Service**



### ***Introduction***

National Grid Service (NGS) is the national e-Infrastructure which is distributed across the UK Higher Education Institutes. It supports academic research by providing researchers collaborative access to high performance computing (HPC) and data resources. It was founded in 2004 and is funded by the JISC and EPSRC. It comprises of 25 sites: 10 partner sites with four of them being 'core sites' which host dedicated hardware; and a further 15 affiliate sites. As of May 2010, there are over 350 users who run eight thousand jobs per day.

Grid-SAFE developed at EPCC with JISC funding, handles accounting, reporting and usage monitoring for advanced computing facilities. It is a software framework, which comprises of a number of modularised components which can be assembled to provide end-to-end HPC or Grid service management.

The NGS and EPCC are involved in an on-going collaboration to provide a solution to enable an integrated accounting solution across all NGS sites.

### ***HPC Accounting Challenges***

Any organisation running an HPC system needs to know about its usage and users – what they are doing, how much they are doing it and when they are doing it. This allows administrators to charge for usage accordingly; and gives project leaders and investigators an insight into what they are using so they can plan and budget. This is easily done using EPCC's SAFE software suite on sites such as HECToR.

In a large distributed HPC collaboration such as the NGS, this is more challenging – how does a central administrator know if a user at site A is using all the resource at

site B without centrally accessible data? At the NGS, this is presently carried out by "pushing" and "pulling" accounting data from individual HPC sites.

The 'push' sites have Perl scripts which parse the local batch system logs. PBS and LSF are supported and as of May 2010, SGE and Condor scripts are being brought online. These sites generate usage records in NGS's Open Grid Forum Usage Record (OGF-UR) format which are posted using https to a Central Accounting Service. The Perl scripts also parse the Globus logs to extract the X509 user mapping and the Local Centre Authorisation Service (LCAS) and Local Credential MAPPING Service (LCMAPS) logs to extract the VO mappings. The Central Accounting Service is a Resource Usage Service (RUS) web service which stores the records in an Oracle database.

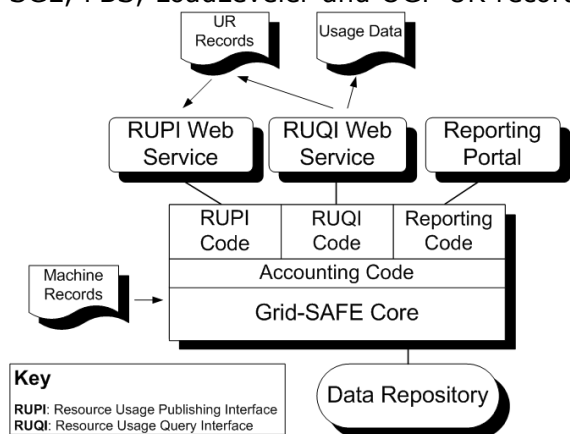
The 'pull' sites use the gLite APEL which pushes usage records to a central APEL database. A daily aggregate of the NGS usage can be pulled from this APEL database by the NGS. The aggregate data is stored in a separate table from the core sites usage records.

For usage querying – of both 'push' and 'pull' sites - an Oracle Application Express (APEX) application is used. It is also used for managing user accounts. It does not have build-in support for either X509 authentication or VOMS authorisation.

### Grid-SAFE and the NGS

Grid-SAFE supports both X509 certificate authentication and VOMS authorisation. It can support publishing and querying of records in different formats -NGS OGF-UR and APEL aggregate OGF-UR formats - via a maintained RUS web service. It can support other underlying databases, including Oracle. At present, Grid-SAFE supports SGE, PBS, LoadLeveler and OGF-UR records, but support for new batch systems can

be provided by the Grid-SAFE development team at EPCC. It offers a configurable reporting service, WebAcct Portal, which could replace the APEX reporting application, and which can optionally be installed at the NGS sites for the generation of local accounting reports. As Grid-SAFE supports aggregate records, real-time accounting reports can be generated of for the vast number of usage records stored by the NGS. Grid-SAFE meets the accounting requirements of the NGS.



EPCC and the NGS are extending Grid-SAFE so the NGS can complement parts of their accounting system and bring benefits to the Service:

- Central Accounting Service can be complemented with an Oracle-based instance of Grid-SAFE.
- A Grid-SAFE command-line usage record publishing tool can be used by sites to push their usage records to the new central Grid-SAFE instance.
- 'Push' and 'pull' scripts are replaced by a standards based software solution maintained and developed by a team of software engineers.

Andrew Richards, Executive Director, NGS comments that, "The NGS has had a long standing requirement for accounting software, that can be deployed at a local and national level to enable the NGS to provide inter-institutional accounting of resources used by projects. The Grid-SAFE software has the potential to allow the NGS to account across a broader range of institutes than at present, whilst maintaining a standards based approach that does not commit institutes to use a specific software solution on their resources. Ongoing collaboration with the Grid-SAFE team at EPCC is allowing the NGS to evaluate the software and its ability to integrate it into the NGS accounting service."

Furthermore, Grid-SAFE also offers individual NGS sites benefits as the technology deployed to provide the accounting data for the NGS from each local site can be easily augmented to provide those sites with local reporting and accounting functionality. These services, outlined in some of our other case studies (such as the STFC Daresbury Laboratory, ECDF, and Bristol University case studies), are valuable to local site managers and administrators as they simplify the process of managing and administering the computing resources as well as offering the potential for allowing users to query their usage and gain a deeper understanding of the resources they are using.

### **Technical Achievements**

- **External authentication services** are supported; for NGS this is X509 certificate authentication.
- **External authorisation** is supported; for the NGS this is VOMS authorisation.
- **Oracle database support**; which for the NGS contains all their existing usage record data.
- **Dynamic reports** defined by the service providers; for NGS, the service team defined a variety of reports for different users from the NGS users to the NGS Service Director.
- **Restricted access to reports** prohibits unauthorised users for viewing a report; For NGS, only the 'NGS Site Administrator' could access all the that NGS Site report.
- **Restricted data within a report** prohibits unauthorised users from generating reports on data which they are not authorised to view; for NGS, a NGS user report cover only their usage.
- **Scalable reports** are possible by the use of daily aggregate records; for NGS as of May 2010 there are approximately three million usage records in the oracle database.
- **Command line Parser** which parses the standard batch system logs file and generated output in XML Usage Records.
- **Command line Report Generator** which can generate HTML, PDF or csv reports.

### **Further information**

NGS: <http://www.ngs.ac.uk/>

Grid-SAFE : <http://www.epcc.ed.ac.uk/projects/grid-safe>