



Grid-SAFE Final Report

Project Information			
Project Acronym	Grid-SAFE		
Project Title	Grid-SAFE		
Start Date	01/09/2008	End Date	31/6/2010 – was 31/12/2009
Lead Institution	The University of Edinburgh		
Project Director			
Project Manager & contact details	Davy Virdee		d.virdee@epcc.ed.ac.uk 0131 650 6485
Partner Institutions	None		
Project Web URL	www.epcc.ed.ac.uk/projects/gridsafe gridsafe.forge.nesc.ac.uk		
Programme Name (and number)	e-Research		
Programme Manager	James Farnhill		

Document Name			
Document Title	Grid-SAFE Final Report		
Reporting Period			
Author(s) & project role	Davy Virdee (Project Manager), Stephen Booth (Principal Investigator), (Project Manager until May 2009),		
Date	25/6/2010	Filename	Grid-SAFE-FinalReport.doc
URL			
Access	<input checked="" type="checkbox"/> Project and JISC internal		<input type="checkbox"/> General dissemination

Document History		
Version	Date	Comments
1.0	25/6/2010	First Version

Table of Contents

Acknowledgements	2
Executive Summary	3
1. Background	4
2. Aims and Objectives	4
3. Methodology	5
4. Implementation	6
5. Outputs and Results	9
6. Outcomes.....	13
7. Conclusions	15
8. Implications	15
Appendix A : List of Deliverables	16
Appendix B : Description of Grid-SAFE architecture	17

Acknowledgements

This project was funded by a JISC Grant under part of JISC's e-Research Programme.

The project team would like to thank Andrew Richards and Jason Lander at the NGS for their collaboration; as well as Tavia Stone at STFC for her feedback when deploying Grid-SAFE. In addition, the project team would like to the programme manager James Farnhill, for his help and advice.

Executive Summary

The Grid-SAFE project has provided a modular and customisable software solution for High Performance and Grid Computing resource usage, monitoring and accounting. It has provided consultancy and software to UK institutions and European projects – both as end users and collaborating partners.

Resource usage monitoring and accounting has long been standard practice on high-end compute resources, such as the national HPC facilities. Such compute resources are expensive to procure and run, and as such necessitate the need to be able to re-coup or assign costs to individual users, groups, projects, institutions, or other such groupings. Whilst it has long been a requirement to undertake this type of activity for high-end computing resources, there was identified a lack of tools or products available that provide all three functions for HPC and Grid facilities. The Grid-SAFE project built on the success of EPCC's SAFE (**S**ervice **A**dministration **F**rom **E**PCC) to make SAFE available as a customisable, modular accounting solution for the UK's HPC and Grid communities.

The Grid-SAFE project followed three main strands of work:

- Accounting Framework
- Client/Usage Monitoring
- Outreach and Consultancy.

Developing the **accounting framework** has involved actively working with the Open Grid Forum's Resource Usage Service working group to progress the standards to provide usable and useful service definitions and standards for accounting and resource usage services. This has culminated in the production Resource Usage Publishing Interface (RUPI) and the Resource Usage Query Interface (RUQI). These interfaces allow a standard way for users/managers to query and upload their usage data. To aid deployment of the Grid-SAFE at any HPC or Grid site, the project developed a series of modular components to allow Grid-SAFE to access most machine usage logs, and then publish the data in a database for access by both RUPI and RUQI interfaces.

The project developed a reporting portal and web-service access that enables **usage monitoring** by users/administrators. Reports on usage/accounting data can be generated via a web interface. This may be done by using some pre-defined report templates or via the production of bespoke reports to query the specific data required.

Putting together the Accounting Framework and Client/Usage monitoring software, Grid-SAFE provides a suite of modular and customisable software components to allow HPC sites and Grid services the ability to monitor and account their systems.

The third strand of the project was making the Grid-SAFE solution available to stakeholders and users via **outreach and consultancy**. We worked with our major stakeholders, the National Grid Service and ECDF at the University of Edinburgh, to deploy Grid-SAFE at HPC sites and enable integration of accounting and usage data and ensure Grid-SAFE meets their needs and requirements. In addition, we have worked with the Universities of Bristol and Bath; the Science and Technology Facilities Council; and two European projects – DEISA2 and EUFORIA. All of these stakeholders and users have evaluated Grid-SAFE with the view of adoption of all or certain modules of Grid-SAFE to provide usage monitoring and accounting solutions for their Grids.

1. Background

At the start of 2007 a report was commissioned by the JISC to examine the current status of developments in accounting and usage in the UK academic community. The objectives were to address the following topics:

- An analysis of stakeholders and their requirements.
- The identification of key metrics and a framework within which they could fit.
- An assessment of current tools and how they could work or be re-tasked to work within the framework.
- Recommendations for further work to create an eventual solution.

The stakeholder review was conducted through visits, teleconferences, email, and via a questionnaire. Feedback was received from over forty people from various stakeholder groups including national, regional and campus grid services, the high performance computing community (including HPCx and HECTOR), EDINA, MIMAS, NeSC, OMII-UK, OMII-Europe, and grid accounting software developers. From these discussions, the major requirements for accounting and usage were deemed to be:

- Standards should be used where possible to maximise interoperability. However deficiencies identified by the stakeholders must be addressed both through extensions to the standards and by other means if necessary.
- The solution should be lightweight and not prescriptive, and should not compromise performance.
- Further development is required to address the longer term requirements for data & service accounting.

From this analysis of stakeholder requirements, it was decided that the project would proceed by developing a software solution based on EPCC's SAFE software. The development would use OGF standards and drive these standards by being actively involved in the working group. The project would engage with the stakeholder to ensure that their requirements were being met as the software evolved.

2. Aims and Objectives

The main recommendations in the accounting and usage report identified two key components that need to be developed in parallel:

- Accounting Framework – an accounting framework that will assist deployments based on the standards specifications.
- Client/Usage Monitoring Framework – a client based usage monitoring/management framework to provide a common basis for monitoring related activities.

The main aim of this project was to meet these needs by developing a software framework to support accounting, reporting and usage monitoring, and resource management on advanced computing facilities. To this end two tasks were outlined:

- Provide a modular accounting solution for HPC systems.
- Provide help and support for users wishing to use some or all of the Grid-SAFE solutions on their systems

The software frameworks comprise of a number of modularised components which can be assembled to provide end-to-end applications for monitoring resource use on all HPC systems.

This work was based on the EPCC SAFE system that provides these functions for the HPCx and HECToR national HPC services. The results of this project will be applicable to both Grid and non-Grid operating environments.

3. Methodology

Overview

The work of Grid-SAFE was structured as a typical EPCC project and ran according to EPCC's common project management methodology and agile software development with a strong emphasis on developer test.

We used a tailored version of PRINCE2 for project management. Its key features were:

- A staged approach to delivery, with stages planned in detail as required.
- Active risk management based on a combined risks and issues log which was reviewed as required. Risks perceived to have a high likelihood were managed and discussed with the project team at weekly meetings.
- A lightweight yet powerful change process based on project issues.
- Documented quality control procedures and a focus on continuous integration and developer-led testing.
- Product-based planning to maintain focus on project deliverables.
- A team-based approach to decision making.

Project execution

Grid-SAFE followed the standard EPCC project execution pattern, which is:

- The team holds weekly checkpoint meetings to define work packages and tasks, accept completed tasks, check progress, issues and risks (in the Risks and Issues Log), and take corrective actions if needed.
- In addition, informal ad-hoc meetings were encouraged and frequent.
- The team accepts, executes and delivers completed tasks.
- The project manager (PM) recorded the amount of time spend each month spent by each team-member.
- The PM sends Highlight Reports to the Board (in this case JISC Programme Management) at regular intervals. Programme Management provides advice as needed.
- The PM escalated issues that might push the Phase outside its tolerance limits, either to EPCC Management or Programme Management.
- Near the end of each Phase, the PM updates the overall Project Plan and Risk and Issues Log.

Project team

The following table describes the project team structure at EPCC.

	Role	Name
Project Manager [0.2 FTE]	Day-to-day management; project tracking; risk management; deliverable sign-off; out-reach and consultancy	Davy Virdee d.virdee@epcc.ed.ac.uk 0131 650 6085
Technical Lead [0.5 FTE]	Technical strategy; software architecture. Principal Investigator; out-reach and consultancy	Stephen Booth s.booth@epcc.ed.ac.uk
Developer [0.8 FTE]	Software design; software development; software test; out-reach and consultancy	Nix McDonnell
Developer [1.0 FTE]	Software design; software development; software test	Malcolm Illingworth (replaced Josh Green)

Technical Reviewer [0.1 FTE unfunded]	Document quality control; software test review; out-reach and consultancy	Adrian Jackson
--	--	----------------

Technical Development

We used the principles of agile software development in developing software within the project. We used three key features of agile development to ensure higher-quality software:

- Continuous requirements analysis: engagement with key end-user stakeholders throughout the development phases of the project, adopting and adapting development to match requirements one by one.
- Developer testing: Encouraging development staff to test as they go and thus build quality control into the development process.
- Continuous integration: The incremental addressing of requirements and writing of developer tests leads naturally into a process of continuous integration. Software is stored under version control and regular “checkout/build/test” cycles were used to cement the quality control process.

The pre-existing software base of SAFE was written in **Java** and designed for deployment within a standard **Java Application Server** environment. Therefore all the Grid-SAFE software was developed on this architecture.

The SAFE software also makes extensive use of the **JUnit** unit test framework; this framework was extended for Grid-SAFE.

4. Implementation

Work began on the project in September 2008, and was originally planned to span over three phases:

Phase I

This phase ran from the project beginning in September 2008 until April 2009. In this first phase of the project, work began on the accounting frameworks and the client usage monitoring. Specifications and standards were gathered from our stakeholders and the OGF RUS Working group to allow us to put together a series of requirements to build on SAFE.

Phase II

The second phase ran from April 2009 until the original project end date on the 31/12/2009. This phase used the feedback from our key external stakeholders – ECDF and the NGS, along with findings from OGF working groups to dictate how the software would evolve. Towards the end of this phase in November 2009, we had a meeting with all interested parties to provide a final steer for the final phase of the project which was enabled by a no-cost extension until the 30th June 2010.

Phase III

The final phase of the project focused on outreach and meeting key stakeholder requirements. With several releases of the software made during phases I and II, we found the software now in a state where interested parties could download, install, and evaluate the software. Hence, this phase was more about enhancements to the software to enable users to use Grid-SAFE, and supporting stakeholders, than about further developing the underlying software.

The phases were broken down into a number of activities to aid tracking and help plan the use of resources.

Phase I

A1 Accounting-Framework I 1/9/2008 - 6/4/2009

This activity developed an implementation of specifications from the OGF RUS-WG and OGF UR-WG working groups. The main output of this activity was:

- UR provider implementations that allow resource providers to generate OGF-UR records for their resources and upload these to any RUS repository;
- A RUS repository implementation.

As part of this work, we produced deliverable D1 – a report into database performance issues. The first release of the accounting framework was made available on the project website on the 31/3/2009 as deliverable D2

A3 Client/Usage Monitoring I 1/9/2008 - 6/4/2009

This activity built on the existing SAFE technology to produce technology that generates resource usage reports for both charging and for usage analysis. The reporting portal can be directly attached to the RUS implementation as described above.

This activity ran in parallel to activity A1 and produced a Framework architecture document D5 and a release of software on the 6/4/2009 as deliverable D6.

Phase II

A plan for this phase was released as D3 on the 30/4/2009.

A2 Accounting Framework II 6/4/2009 - 31/12/2010

This activity extended the basic RUS service primarily to support query operations on the usage data via a web service. In addition, this activity will defined a common programming interface for querying the data held in a remote repository of usage records. This was identified as a priority at 5/5/2009 meeting attended by representatives of Grid-Safe, NGS and EGEE. These two interfaces were released as the Resource Usage Publishing Interface – RUPI; and the Resource Usage Query Interface – RUQI.

A Specification document for a web query interfaces was released as D8 on the 31/10/2009 after we consulted with both the RUS WG and stakeholders as to what a query interface should include. The software was implemented and released as D9 on the 30/11/2009.

Finally, the activity produced incremental releases of the accounting framework and the stand-alone accounting portal. This was driven by feedback and additional requirements from the early adopters of the software. Software released during this activity was delivered as D4, D7 and D10a

A3 Client/Usage Monitoring II 6/4/2009 - 31/12/2009

We worked closely with our early adopters of the Grid-SAFE software to-date; the ECDF at the University of Edinburgh and the EUFORIA project - both added value and real use-cases to the software development. Deploying the software at these sites provided live-testing for the releases, and shaped what enhancements we added and any defects we fixed.

This activity extended the reporting framework.

- Integration with VOMS. VOMS seems to be the most important mechanism for managing a virtual organisation. By integrating with VOMS we can identify usage by VO and define access lists based on VO membership and roles. Development work was undertaken in the EUFORIA

project and merged into the Grid-SAFE work, which it is required by communities such as the NGS and other Grid users.

- Extensions to the Framework were undertaken including modules for handling resource allocations. This handles allocations to user's projects or groups optionally over pre-defined periods of time. Use against a given allocation is automatically tracked as records are inserted into the system.

Code developed during this activity was released incrementally – a major release was packaged and uploaded as D10b - an integrated RUS service on the 30/12/2009.

Phase III

A4 Grid-SAFE Enhancements 1/1/2010 -30/6/2010

This activity was introduced to the project plan to incorporate the stakeholders' requirements we gathered during the initial phases of the project from September 2008 and in particular in December 2009. The release of both RUPI and RUQI has given the community new ways to utilise the Grid-SAFE solutions.

This activity provided incremental releases of the Grid-SAFE software at regular intervals, including the following features:

- Additional functionality for storage accounting.
- Allocation tracking.
- Non-accounting upload features.
- Additional Databases connectivity – for example, Oracle.
- Adding real-time data.
- Simplified configuration of Grid-SAFE functionality

Incremental releases of the software, which included all new functionality and enhancements were uploaded as D12 on the 30/3/2010 and a final release as D13 on 30/6/2010

A5 Outreach, Consultation and Support 1/1/2010 -30/6/2010

It was crucial to the success of Grid-SAFE that stakeholders and other HPC users could benefit from the solutions Grid-SAFE provides for HPC and Grid accounting. It was seen as vital that the project communicate the message that Grid-SAFE solutions can increase efficiency, allow institutions to interoperate and hence reduce their cost base, and increase the range of services they can offer to users. We developed an Outreach Evaluation report which aided discussions with the programme manager to aid outreach and take-up in the final part of the project.

Specific areas of outreach, although by no means exhaustive, that were undertaken are:

- Attendance at HPC-SIG and Campus Grid events in early 2010, and throughout the remainder of the project.
- Deploying Grid-SAFE at the Hartree centre.
- Ongoing work with ECDF and EUFORIA projects
- NGS work
 - Importing NGS data into Grid-SAFE
 - Deploying Grid-SAFE modules at NGS sites to aid their accounting solution.
- Providing Grid-SAFE instances to Bristol and Bath Universities.

In addition, this activity produced a set of instructional videos and case studies to act as both information points and demonstrations for potential users. These videos were available both from the project website and from YouTube (D11).

A series of case-studies was produced and made available on the project website. The purpose of these studies was to convey to readers how Grid-SAFE has (or will) benefited a user group – and from this communicate how Grid-SAFE can aid any HPC/Grid service.

5. Outputs and Results

Software and documentation from the project can be found on the project website at

<http://gridsafe.forge.nesc.ac.uk/>

This includes links to download all software produced as well as presentations and papers produced during the project. This website remains live at the end of the project as EPCC continue to develop and maintain Grid-SAFE.

The main output of the project was the software we produced, but it is more useful to consider the impact of the project on the users who decided to install, evaluate and use the software. We produced software releases at regular intervals which were made available on our website to download and evaluate. A list of our deliverables and software releases can be seen in Appendix A.

The following sections describe the work we have carried out with our major stakeholders – the NGS and ECDF; as well as other organisations.

ECDF at the University of Edinburgh

ECDF adopted Grid-SAFE about a year after the launch of the *Eddie* – their HPC cluster. The Sun Grid Engine parser used for Eddie is one of batch parser already supported by Grid-SAFE. The dynamic reporting component, developed as part of the JISC project, not only allowed ECDF administrators to generate service reports, but also allowed the Heads of School, project PIs and the users up-to-date visibility of their usage.

Grid-SAFE already supported the concept of *groups* and *roles* within a group; however this was extensively re-factored to make it configurable on a per site bases. The relationships between users and groups could be defined using ‘classifiers’. For ECDF this allowed the School, Project and Project Member relationships to be classified, along with the School and User relationships. Users could hold multiple roles such as ‘Support Staff’, ‘School Administrator’ and ‘Project PI’.

Eddie has a huge number of small jobs; approximately 24 million jobs per annum had been run since October 2007. This differs to the HPC service for which Grid-SAFE was originally developed. A service report for ECDF might report over all records for a year. To be able to generate these reports on-the-fly, daily aggregate records are used. This feature was added for ECDF and it reduced the number of records parsed for a service report from 55 million to 35 thousand.

ECDF and EPCC are currently negotiating a support package in which ECDF will provide some funding to continue to support and maintain Grid-SAFE on Eddie.

The National Grid Service

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.

The primary requirement from the NGS was to provide a tool to allow NGS member sites to produce OGF-UR formatted usage data for upload into the NGS RUS service. The NGS has adopted the use of virtual organisations via COMS certificates and the LCAS/LCMAPS plug-ins and any accounting solution needs to be able to extract virtual organisation information from the log files generated by these plug-ins.

We implemented the following features in support of this work:

- XML formatting code to format Grid-SAFE data as OGF-UR
- Parser for LCAS/LCMAPS voms_accounting file format.
- Parser for the Globus job-manager logfile format in use by the NGS (This differs from the EGEE format previously used by the EUFORIA project)
- Parser for LSF batch data (One of the batch formats in use within the NGS not previously supported by GridSAFE).
- Parser for GridMap files.
- Command-line interface to Grid-SAFE code. This allows access to the Grid-SAFE core code from scripts without the need for a running web container – Grid-SAFE uses Tomcat. This was required to allow OGF-UR records to be automatically generated/uploaded within a script and also simplifies deployment for sites that wish to generate OGF-UR records without running a local accounting system.

A small amount of additional work is still needed to support sites that are not yet using the recommended LCAS/LCMAPS system to map pool accounts to virtual organisations. We aim to have this complete by the end of July 2010.

EPCC and the NGS will continue to collaborate to evaluate and deploy Grid-SAFE on NFS sites. Both ECDF and STFC are NGS sites that have Grid-SAFE deployed as a local accounting solution are being used as pilot sites test integration of a service-wide accounting solution based on Grid-SAFE.

University of Bristol

The Advanced Computing Research Centre (ACRC) at Bristol University has two main HPC systems, BlueCrystal Phase 1 and BlueCrystal Phase 2, where Phase 1 (consisting of 384 cores) is intended for serial and small scale parallel jobs, and Phase 2 (consisting of 3328 cores) is used for larger parallel jobs.

The two HPC systems use the PBS batch system for the submission of user jobs. PBS records usage "events" in a log file. Currently, the system administrators manually produce a monthly report of the usage of the systems by processing the relevant PBS logs with scripts and obtaining the number of jobs submitted, the average and maximum wait time for jobs, the total number of processor hours used and a break down of usage over a number of basic, defined, groups.

In order to more efficiently manage and report upon the usage of the HPC systems the University of Bristol needs to be able to report on the usage of individual users and of specific projects. They also want to provide users themselves with a mechanism for querying their usage. Finally, they are looking for some level of automation for the reporting task to reduce the amount of effort which is required to undertake reporting, especially if new reports are required, that is reports on usage metrics that have not previously been analysed.

The Grid-SAFE project created an example Grid-SAFE service for Bristol using real usage data (PBS logs) provided by the system administrators. This service was setup at EPCC for Bristol to access and evaluate. This included configuring a custom parser to process project data (which was provided separately to the usage data), as well allowing data from multiple machines to be uploaded and categorised by machine, ensuring usage can be reported on a per machine basis if required, or over all machines if appropriate.

Bristol are very happy with the test service provided and we are now working with them to deploy it locally so it can be used as a production service within the next 6 to 12 months.

Science and Technology Facilities Council Daresbury

STFC Daresbury Laboratory has a long history of operating and providing HPC services for UK academics, including the previous national HPC service (HPCx). They currently have a number of different HPC systems, from a cluster provided for the NW-GRID (consisting of 128 cores) to a 1024 node BlueGene/L to a GPGPU cluster and others. The machines are provided for a range of users from many different communities and consortium.

With a range of HPC systems there are a number of different usage logs, in a variety of formats. These are generally produced by the specific batch systems of the machines and stored locally. Currently, reports are hand generated using these logs either quarterly or six monthly for each of the systems operated at STFC Daresbury Laboratory. These reports contain overall system usage and the number of jobs and the time used by each user. This is a time consuming process which is not well suited to changes in the reporting metrics or format (i.e. it is difficult to change what is reported on). This process also requires manual intervention to report usage across all HPC systems, and has no mechanism for investigating specific project or group usage.

System administrators at STFC Daresbury Laboratory obtained the Grid-SAFE software from the project website and, following the documentation, installed Grid-SAFE on a local system. The Grid-SAFE instance is being evaluated along side their existing systems, providing added value to enable better tracking of resource usage, monitoring of systems, and helping them determine the best method for managing systems, schedulers, and resources.

One of the key benefits they see from using Grid-SAFE will be removing the task of generating report data from all the individual system administrators, enabling them to focus on other work. Grid-SAFE

will both centralise and standardise the collection of reporting data and enable reporting to be undertaken with minimal effort once the system has been installed and configured.

At the time of writing, STFC are beginning to deploy Grid-SAFE as production service.

EUFORIA

The EUFORIA project is a European Union funded project to provide Grid and HPC resources and expertise to the European fusion research community. Part of this task involves providing a dedicated Grid for fusion, the EUFORIA Grid. This is built from the standard EGEE/gLite Grid software components and hosted by 4 partners around Europe, each providing computational resources and storage for Grid users. In total the EUFORIA grid contains over 2,000 processors and 50 GB of data storage.

The EGEE/gLite software provides a minimum reporting component to allow managers to evaluate how much total user of the Grid there has been but does not provide the level of detail that the project management required.

Grid-SAFE was used to provide an accounting solution for the EGEE/gLite environment. A tool was provided which, at each Grid site, takes those two sets of logs and periodically uploads the data to a central Grid-SAFE service. This Grid-SAFE service processes the logs and matches job manager entries (which specify who submitted a particular job) with the batch system entries to create a full record of resources used - resource managers can analyse the usage of particular users, particular sites and list all user usage – as well as other metrics and data.

Access control and authorisation is performed through the Grid-SAFE service using Grid certificates and the VOMS service that the EUFORIA Grid utilised. This is a separate service which manages the permissions of users to access Grid resources and Grid-SAFE requires users to provide a Grid certificate which is authorised in the VOMS service before they can access the reporting/accounting functionality that Grid-SAFE provides. Both the development and integration of this VOMS functionality and the functionality to integrate the job manager logs with the batch system logs were undertaken jointly between the Grid-SAFE and EUFORIA projects.

The ability for users to be able to interrogate and produce reports on their usage was considered important in ensuring that the EUFORIA Grid was a production computing service that users could understand and trust. The ability for managers to accurately interrogate the usage of the Grid system being provided is vital to ensure that the project can correctly manage the resources under its control and evaluate both the user experience and the cost-benefit issues surrounding those resources. Furthermore, it is vital that the funder of these resources (in this case the European Commission) is able to accurately understand and evaluate the computational resource it is funding and the cost-performance profile they have. This is particularly important in the current era where there is a range of computational resources available; Grid, Cloud, and HPC. Being able to understand the usage of a resource, reconcile it with the cost of that resource, and then compare it with other available technologies is very valuable for funders and managers alike.

We will be working with the European Middleware Initiative (which maintains and develops gLite/EGEE amongst other things) to evaluate where our Grid-SAFE solution to Grid accounting and reporting can be integrated with the standard release of the middleware, and whether it is something other users of this middleware are interested in.

DEISA2

DEISA operates some of the largest HPC resources in Europe across international boundaries, with many partners and international collaborators. Many DEISA users access HPC resources at multiple sites across several countries. Currently the usage information is collected at each site. An existing reporting tool, DART (DEISA Reporting Tool), can query each individual DEISA site for usage

information. However, the usage information is not currently held centrally and there is no dedicated repository for this information. This means that summary reporting of DEISA usage is currently not optimal and there is no central archive of DEISA usage records.

The project began collaboration with the DEISA project in November 2009 as part of Work Package 4 of the DESIA2 project. The motivation was to evaluate Grid-SAFE as a solution to the accounting challenges of the DEISA project. From January 2010, funding was available for 0.8-1.0 FTE working Grid-SAFE maintenance and customisation for DEISA.

To date, we have modified Grid-SAFE to work with the DEISA project by providing:

- Summary usage information is centralised instead of being held at each site.
- Reporting across all DEISA sites and users can be simplified.
- Sites can automatically push usage data to the central service instead of the current requirement to query each site.
- Role-based security can be provided to ensure that appropriate levels of confidentiality are maintained.

Funding will continue until October 2010, by which time it is envisaged that Grid-SAFE will be deployed and used across the DEISA consortium as part of their accounting solution.

6. Outcomes

Grid-SAFE's original aim was to provide an accounting solution for the Grid community. We achieved this by producing several releases of Grid-SAFE over the life-time of the project. In addition, by engaging with other projects – in particular EUFORIA and DEISA, we gained feedback from outwith of our original stake-holders in the UK, and this led to general improvements and ideas that then benefited the community as whole. Most of the success in bringing new users on board and meeting the requirements of our original stakeholders came about because the project was able to extend from December 2009 until 30th June 2010. Late in 2009 when it became clear an extension was possible, the meeting with stakeholders allowed us to gather key requirements from them to drive development in early phase III. It allowed the project to focus on key enhancements to the software to aid end-users, and also focus on bringing the software to potential users. Attending Campus-SIG and HPC-SIG meetings were crucial in informing interested parties where the project was going, how it was going to get there and when it was able to meet users and potential users' needs. The third-phase of the project also allowed us to build on our stakeholder relationship with the NGS and bring them on board more as a collaborating partner rather than an end-user.

The table below is from the original project plan - it identifies our key stakeholders and summarises their up take of Grid-SAFE at the end of the project. It also lists new stakeholders and collaborators we identified throughout the project.

Stakeholder	Interest / stake	Importance	Status
Originally listed Stakeholders			
UOEHPCx Ltd Managing agent of UK National HPCx and HECToR services	End-user	High	Continue to use Grid-SAFE as management and accounting solution
NGS UK National Grid Service	End-user / Collaborator	High	EPCC have been collaborating with NGS during project duration and will continue to do so. NGS

			piloting Grid-SAFE solution at two NGS sites
ECDF Edinburgh Compute and Data Facility – University of Edinburgh central research computing facility	End-user	High	Continue to use Grid-SAFE as management and accounting solution
DEISA-2 EU FP7 project and consortium of European high-performance computing centres	End-user	Medium	Evaluation of software with a view to deploy. Funding available to modify and expand.
Additional stakeholders			
University of Bath	Evaluator	Low	Evaluation of software
University of Bristol	End-user	Medium	Evaluation of software using their usage data
STFC Daresbury Laboratories	End-User	High	Evaluation of software on site on their HPC/Grid service. STFC is an NGS site and part of the part of the NGS pilot
EUFORIA	End-User	High	European Middleware Initiative will see Grid-SAFE being used to aid accounting solutions in EGEE.

Take up by our original stakeholders has been good, and the addition of new stakeholders and end-users has demonstrated that Grid-SAFE can and does provide an accounting solution for HPC and Grid services.

The key benefit that Grid-SAFE offers to users is the ability to have an “off-the-shelf” accounting solution which is highly customisable. Potential users can download source code, pre-compiled binaries or a virtual machine image in which everything is set up and “ready to go”. Once deployed and running, administrators have access to accounting, usage and budgeting information at their finger-tips – highly beneficial in the current financial climate.

Grid-SAFE should not be viewed as just a software solution. EPCC staff have provided consultancy on accounting policies for HPC and Grid services, as well as support for setting up, deploying and customising Grid-SAFE instances. The project has also helped to progress international standards in this area.

7. Conclusions

- Producing an accounting solution based on standards allows take-up by HPC and Grid communities. Being active in standards development by developing software that uses these standards drives forward the standard. Indeed – the software *becomes* the standard.
- Modular software allows for customisation and tailoring to many systems without the need for heavy-weight software development.
- Emerging software products need frequent feature steer from their stakeholders and potential end-users.
- Key stakeholders should be viewed as, or indeed encouraged to become collaborative partners.
- The timescales for adoption of new software or technologies for production HPC and Grid services are relatively long. Most end users require a significant number of months (often 6 or more) to evaluate the software and another period of time to make the policy and management decision to adoption the software.

8. Implications

The Grid-SAFE software is currently used as a key management tools for a number of HPC and Grid systems. It is actively under integration as a component to the NGS software stack, and is likely to be adopted by many more sites in the future. The open source software, which is freely available for download, installation, and further development, will be continually updated at EPCC based on requirements and identified issues from the use at the national HPC services, NGS sites, and the other places it is used. The software is also under consideration by the DEISA project to provide account and reporting, and as such may be a solution for such functionality for the European PRACE infrastructure.

Appendix A : List of Deliverables

Table of deliverables.

D	Deliverable name	Form	Due	Status
D3	Accounting Framework phase II development plan	Plan	30/04/09	Delivered
D7	Integrated RUS service Brought Forward from 24/8/2009	Software	04/06/09	Delivered
D4	Second Accounting Framework release and Architecture Document– Incremental release Remains the same	Software and Document	24/08/09	Delivered
M2	<i>Management Report II</i> Remains the same	<i>Report</i>	16/10/09	Delivered
D8	Specification for Query Interface New Deliverable	Document	31/10/2009	Delivered
D9	Web Service Query Interface New Deliverable – Replaces D9 Resource Management Framework	Software	30/11/09	Delivered
D10	Incremental releases of frameworks	Software	31/12/09	Delivered
D11	Grid-SAFE information/tutorial videos	Videos	31/1/2010	Delivered
M3	Review and evaluation of outreach	Document	31/3/2010	Delivered
D12	Incremental releases of Software	Software	30/3/2010	Delivered
D13	Final incremental releases of frameworks	Software	30/6/2010	Delivered
M4	Final Report	Report	30/6/2010	Delivered

Appendix B : Description of Grid-SAFE architecture

Grid-SAFE Architecture

Grid-SAFE, a software suite developed at EPCC, The University of Edinburgh, is based on the well established SAFE (Service Administration From EPCC) software which is used to help administer and manage a range of HPC resources. SAFE provides project and resource accounting, reporting, and management. Grid-SAFE builds on this functionality to undertake these service administration tasks for a wider range of resources and deployment scenarios, from campus HPC systems to national and international Grids.

The fundamental aim behind Grid-SAFE is to provide a modular and customisable set of software that resource or project administrators/managers can easily deploy to provide the reporting and accounting functionality they require. The components may function as separate applications; or be assembled to provide end-to-end resource monitoring for HPC and Grid systems. The full components and architecture of Grid-SAFE are shown in Figure 1.

Services provided by Grid-SAFE are accessed through the web, utilising well established java technologies to provide the efficient, secure, and user friendly service required. This web-based approach simplifies the access requirements and usability issues associated with running a multi-user resource, especially as often users are from different locations or organizations.

There are currently three separate components provided:

- RUPI¹: A service to allow usage records to be added
- RUQI²: A service which allows usage records to be queried
- Reporting: A service which allows customisable reports to run

All these components build on a common code base which supports reporting and accounting and the required access to a relational database to maintain usage data.

The main data element that is used for reporting and accounting in Grid-SAFE is a usage record. This represents the consumption of some resource for some period of time. Usage records used with Grid-SAFE are generated by processing usage records/data provided from a resource or group of resources. This processing extracts relevant information and applies any local

¹ RUPI: Resource Usage Publishing Interface

² RUQI: Resource Usage Query Interface

policies that have been defined by the administrators of the Grid-SAFE service to generate an appropriate usage record. Both the parsing of the raw usage data and the policies that are applied to the parsed data can be customised to match the HPC resources they are using and their particular policies on resource usage, etc...

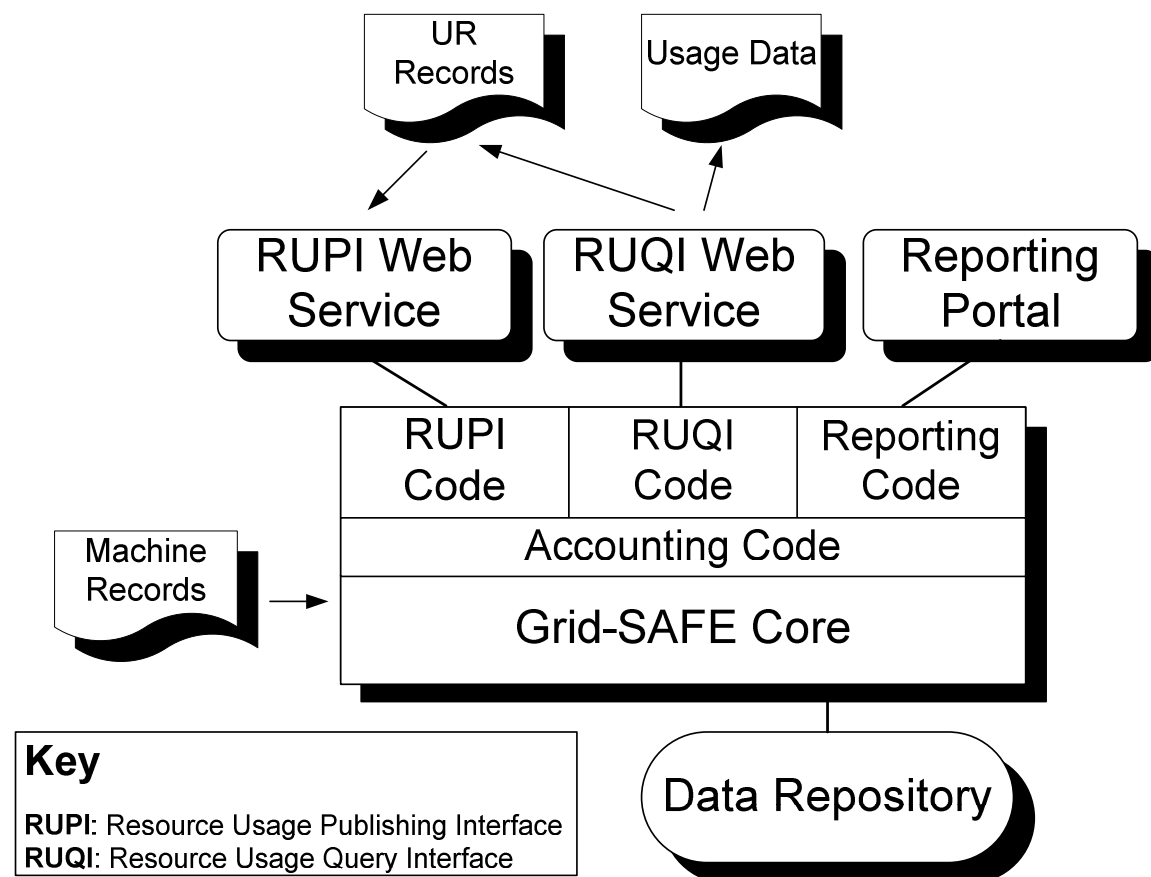


Figure 1: Grid-SAFE Architecture

Grid-SAFE offers two ways which data can be added to the data repository (the relational database). One is through the RUPI service, where data conforming to the OGF-UR standard can be uploaded. The other is to add machine records using a servlet that accesses the Grid-SAFE core code base. This second mechanism is the one generally used for Grid-SAFE instances that are monitoring local HPC machines such as a campus HPC resource or a national HPC machine.

As previously mentioned, the Grid-SAFE components can be used separately or together for any given deployment of the Grid-SAFE software. There are many different scenarios for the use of the software, from a single HPC system that simply uses the Reporting portal, to an international Grid which uses the RUQI service as an interface for its own stored data and conversely someone who uses the RUPI service simply as an upload interface to add usage data to service.

Project Acronym: Grid-SAFE
Version: 1.0
Contact: d.virdee@epcc.ed.ac.uk
Date: 25/6/2010

Further documentation of each component within Grid-SAFE and deploy details can be found at the Grid-SAFE website.