

T2K LCG Portal

Gidon Moont, Yoshi Uchida, David Colling. Imperial College London, UK

Abstract

A working prototype portal for the LHC Computing Grid (LCG) is being customised for use by the T2K 280m Near Detector software group. This portal is capable of submitting jobs to the LCG and retrieving the output on behalf of the user.

The T2K specific development of the portal will create customised submission systems for the suites of production and analysis software being written by the T2K software team. These software suites are computationally intensive, and therefore warrant utilisation of the LCG.

The portal runs on an Apache server with the GridSite module. It is accessed over https, identifying users by their Certificate Authority signed Grid certificate. A user can upload files to the portal, as well as edit them, using the GridSite CGI. Proxy certificates are created on a users desktop/laptop machine using a JavaWebStart program that does the equivalent of a voms-proxy-init using the user's Grid certificate, and this limited time proxy is then securely put on the portal. Once there, the proxy is available exclusively to that user to utilise in submitting jobs to the LCG.

The portal may also be used as a joint collaborative site for the experiment. GridSite makes it easy to have joint responsibility for maintaining public web pages spread amongst collaboration members. Other collaborative tools such as diaries and lists of publications and submitted abstracts are also easily implementable.

Please address all correspondence to g.moont@imperial.ac.uk

WHAT IS A PORTAL

In a non-Grid computing environment, we would understand a portal to mean an access point to some kind of online information. Examples would be sites such as Google, which not only provide a start point to the web through its search engine, but also provides access to usenet postings, both present and past, and more recently to an online map service. More specialised sites include SPIRES, BIDS and MEDLINE, all of which provide information on academic publications, and sometime links through to online journals. It should be noted that these services are not always only accessible through a web browser. Google Earth is an application that uses the same web based information as the web based Google Local, and Endnotes bibliography software can link into Medline (given the proper add-on and account authorisation).

In a Grid environment, a portal should mean something

that provides at least one of

- Access to computational resources
- Access to distributed storage - both read and write

The something could mean either a web page or an application. The emergence of standardisation means that web services can be accessed by a variety of methods, and so both a web page based portal or an application base portal could use the same underlying web services.

A portal can also integrate several components in order to provide a one stop solution to a user. Hence a Grid portal may also provide a distributed work environment for an experiment, such as an online calendar and document repository.

DESIGN

The design for the portal can be seen in Figure 1.

The only way to currently use the LCG is with a User Interface (UI); a computer with all the necessary software installed. This software provides all the standard LCG commands to do tasks such as send and retrieve jobs, as well as interact with the LCG storage systems.

The UI is split from the portal by implementing a Simple Object Access Protocol (SOAP) services layer to the LCG commands, described by a web services description language (wsdl) file.

The front end of the portal is currently on a GridSite [1] enhanced Apache web server. It uses CGI/Perl to provide the user with forms that then access the LCG SOAP services described above.

In the future, a Java application could do the same, as the SOAP services are described by a wsdl file, so making it easy for multiple platforms to use the services. Future middleware is expected to provide alternatives to using a UI, as LCG services would be directly accessible via their own web services interfaces

SECURITY

We want to provide a secure "one click" login method to the Portal. The Portal uses GridSite, and this can easily be configured to limit access to users with a Certificate Authority signed Certificate in their browser.

To use any Grid resource, a user needs to have a limited lifetime proxy certificate that identifies them. The Portal provides a Java Web Start application to generate a Virtual Organisation Membership Services (VOMS) proxy - a proxy that defines which Virtual Organisation the user is a member of - without the need for a UI. This proxy is then

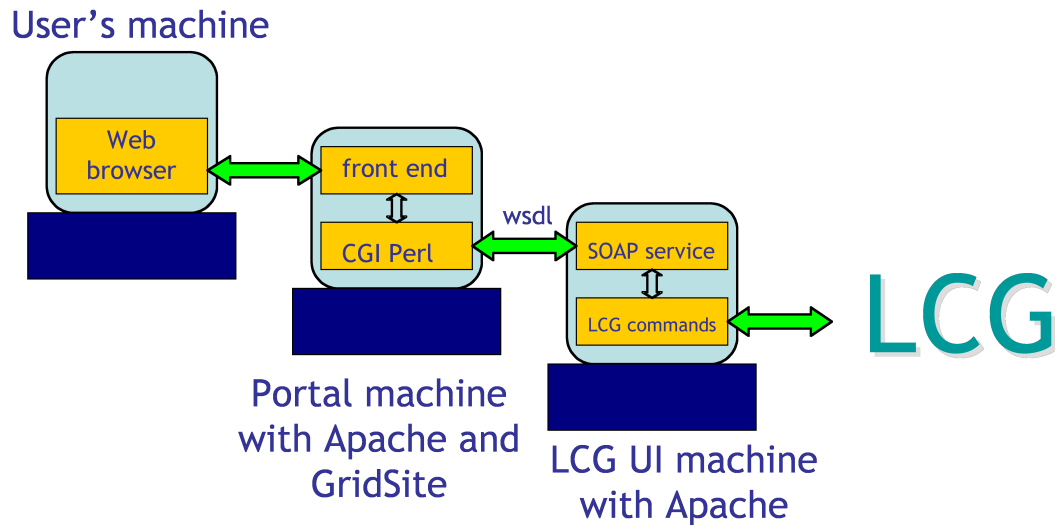


Figure 1: The Current Design of the Portal.

moved securely to the Portal. The proxy is then used by the Portal on behalf of the user. The portal can handle a user having multiple proxies for different Virtual Organisations.

The screenshot of the front page (Figure 2) shows how the portal can be a one stop solution in that other non-Grid specific information can be accessed.

PROTOTYPE

The first Prototype was finished in April 2005. It was not VOMS enabled, and the interface was quite low level (though not as low level as using LCG command line programs). It was tested by the MICE experiment, who found it useful. However, the experiment was not at a stage where they needed to use the Grid, and work was put on hold.

The VOMS enabled prototype was finished in January 2006 [2]. It currently shows an example of sending a Computing Element (CE) test job to a specified CE. It uses the Real Time Monitor (RTM) to determine the status of a job. This has the advantage of reducing the load on the Resource Brokers (RBs) of the Grid with continual status requests. A job's outputs are now automatically retrieved once the status shows the job to have completed. The portal can also handle multiple proxies for different Virtual Organisations owned by one user.

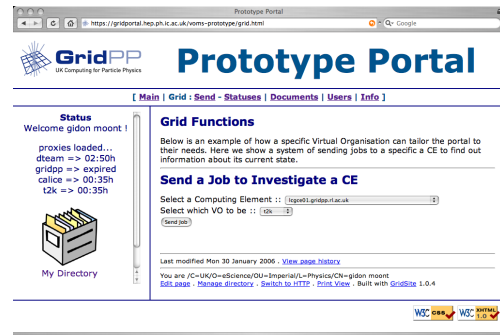


Figure 3: Example of Sending a Job Using the Portal.

The prototype portal demonstrates sending a grid job and the automatic retrieval of jobs. Screenshots in Figures 3 and 4 show what this can look like.

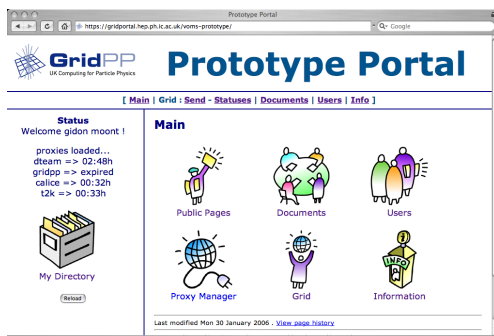


Figure 2: The Main Portal Window.

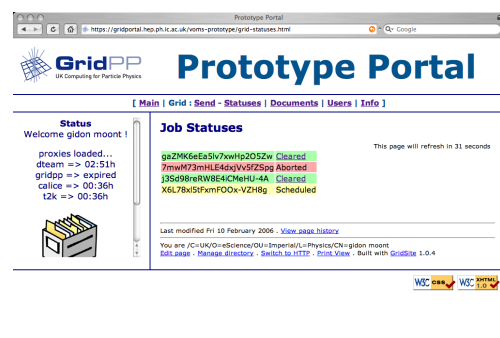


Figure 4: Job Statuses shown on the Portal.

T2K

For T2K to use the LCG, a VO had to first be created for them. This was done, and is managed by GridPP (Manchester). A handful of experimenters have Grid Certificates and are members of the VO.

A basic web portal exists with public pages and internal documentation and help [3]. These pages are maintained by T2K experimenters.

The experiment software is currently undergoing changes from using

- ROOT 4.00.08 → 5.08.00
- CLHEP 1.8.2 → 2.0.2.2
- GEANT 4.6.2 → 4.8.x (newest at time of upgrade)

This has caused delays in getting software ready to run on the Grid. However, the infrastructure is ready, so jobs will be running as soon as software is compiled.

OTHER SMALL EXPERIMENTS

One of the aims of the GridPP Portal Project is to encourage small HEP and non-HEP experiments to use the LCG.

As described above, we have worked with MICE and T2K. We are currently also working with CALICE. They have a VO with several experimenters in the UK and elsewhere as members. They are currently developing software to run on the Grid. Some people at DESY who work in the ILC collaboration have also expressed an interest. The ILC and CALICE software is very similar, so they are looking to see how CALICE benefits from the portal.

REAL TIME MONITOR

The Real Time Monitor [4] has developed from a demonstration to show real time usage of the LCG. It gets its information by direct querying of around 30 Resource Brokers in the LCG. It has a graphical display via an applet that can be seen in Figure 5.

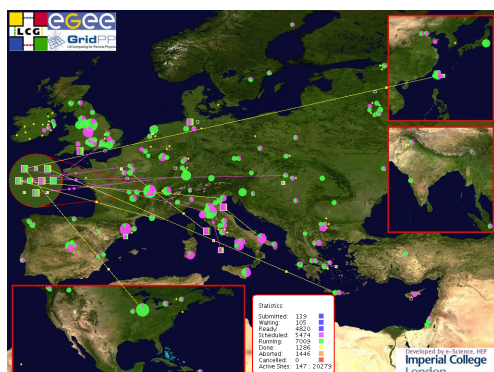


Figure 5: The Real Time Monitor.

It is used by the portal to determine job statuses. It also provides daily summary reports (including per VO)

and further development will provide real time triggers for problematic behaviour. Real time XML files are publicly available.

CONCLUSIONS

We have a working VOMS aware portal with one-click log in. This portal has been used to run jobs by non Grid savvy experimenters from MICE, T2K and CALICE.

The moving target of underlying LCG middleware means that no portal is going to work for long without maintenance. However, such a Portal is essential as the LCG becomes the dominant means by which small experiments can have access to large computing resources. System managers prefer to have only one entry point to their resource, and do not have to manage individual accounts.

The T2K software should be running on the LCG through the Portal very soon. Also, the current document archive of T2K may move to the Portal, once enough T2K experimenters have Grid Certificates.

Other small experiments, such as CALICE, should follow soon after.

REFERENCES

- [1] <http://www.gridsite.org/>
- [2] <http://gridportal.hep.ph.ic.ac.uk/>
- [3] <http://gridportal.hep.ph.ic.ac.uk/t2k/>
- [4] <http://gridportal.hep.ph.ic.ac.uk/rtm/>