Mathematical Problems of Computer Science 29, 2007, 21 – 25.

Yerevan Physics Institute Steps Towards an LHC Computing Grid[1]

Karen S. Mkoyan

Yerevan Physics Institute after A.I. Alikhanyan,
karen@yerphi.am

Abstract

The main goal of this paper is to develop a User Interface (UI) for scientists from Yerevan Physics Institute to have access to LCG infrastructure. The structure and mechanisms of the suggested interface, which uses DESY Grid site for accessing to LHC Grid are given.

1. Introduction

The term "Grid" was coined in the mid1990s to denote a proposed distributed computing infrastructure for advanced science and engineering. Considerable progress has since been made on the construction of such an infrastructure, but the term "Grid" has also been conflated, at least in popular perception, to embrace everything from advanced networking to artificial intelligence [1,2]. It is an architecture designed to unify a large amount of different computing resources within one virtual computing resource. Grid enables work to be distributed on each of these resources (which can be servers, storage devices, databases, etc.). Such architecture enables high availability and a big computing capacity at low cost. There are two major types of Grids. Computational grids define an infrastructure which will solve complex computational problems and provide coverage for "number crunching" and other intensive workloads. They are often used for very large problems needing a lot of CPU and memory. And Data Grids that give a common interface for all data repositories through which large amount of distributed data can be queried, managed and secured. They are often combined with computational grids.

The case of the LHC [3] experiments at CERN [4] and HERA [5] experiments at DESY (Deutsches Elektronen-Synchrotron) [6] illustrates well the motivation behind the Grid technology. The LHC accelerator already started operation in 2007, and the experiments that use it (ALICE[7], ATLAS[8], CMS[9]) will generate enormous amounts of data. The processing of this data requires large computational and storage resources and the associated human resources for operation and support. It was not considered feasible to fund all of the resources at one site, and so it was agreed that the LCG [10] computing service would be implemented as a geographically distributed Computational Data Grid. This means that the service will use computing and storage resources, installed at a large number of computing sites in many different countries, interconnected by fast networks. LCG-2 Grid middleware will hide much of the complexity of this environment from the user, giving the impression that all of these resources are available in a coherent virtual computer centre. The LCG-2 Grid middleware comes from a number of Grid development projects, like DataGrid, DataTag, Globus[11], GriPhyN, and the EU project EGEE (Enabling Grids for E-sciencE) [12].

[1] This work is a part of INTAS Young Scientists Fellowship grant Ref. 05-110-4812

The users of a Grid infrastructure are divided into Virtual Organizations (VO), abstract entities grouping users, institutions, and resources in the same administrative domain. The LCG-2 VOs correspond to real organizations or projects, such as the four LHC experiments; other VOs exist in the context of EGEE as well. Besides the VO membership user must have a digital certificate, signed by LCG trusted Certification Authority (CA). We well discuss techniques of getting a certificate in details below.

At the same time Yerevan Physics Institute is participating in many international HEP (High Energy Physics) collaborations, like ATLAS, CMS and ALICE at CERN, H1 and HERMES [13] at DESY. YerPhI does not dispose of Grid infrastructure but in the framework of H1 and HERMES collaboration DESY shares its Grid resources with YerPhI. The Grid facilities deployed at YerPhI will allow the physicists to start with the massive Monte Carlo production, necessary to perform the physics analysis, particularly after the end of HERA running, when the analysis will become the main goal for collaborating institutions. To use those resources, a User Interface has been deployed at YerPhI in a manner that it uses the infrastructure at DESY to submit the job to the Grid.

2. User Interface Description

The point of access to the LCG-2 Grid is the User Interface (UI). A UI has client functionality and can be setup at any host. This is a machine where users have a personal account and where the user certificate is installed. This is the gateway to Grid services. From the UI, a user can be authenticated and authorized to use the Grid resources. It is the component that allows users to access the functionalities offered by the Information, Workload and Data management systems. It provides a command line interface to perform Grid operations like:

- submitting jobs for execution;
- listing all the resources suitable to execute a given job;
- canceling jobs;
- retrieving the output of finished jobs;
- showing the status of submitted jobs;
- retrieving the logging and bookkeeping information of jobs;
- copying, replicating and deleting files from the Grid.

The individual steps are described as follows.

1. After obtaining a digital certificate from one of the LCG trusted CAs, registering with an LCG VO and obtaining an account on an LCG-2 UI, the user is ready to use the LCG Grid. (S)he logs to the UI machine and creates a proxy certificate to authenticate himself in every secure interaction.
2. The user submits the job from the UI to the **Resource Broker (RB)** node. In the job description file one or more files to be copied from the UI to the RB node can be specified. This set of files is called Input Sandbox.
3. The **Workload Management System (WMS)** looks for the best available **Computing Element (CE)** to execute the job. To do so, it interrogates the **Berkeley Database Information Index (BDII)** to query the status of computational and storage resources and the File Catalogue to find the location of required data. The BDII knows about all hosts and their status.
4. The RB prepares the job for submission creating a wrapper script that will be passed together with other parameters to the selected CE.

5. The CE receives the request and sends the job for execution to the local **Local Resource Management System (LRMS).**
6. The LRMS handles the job execution on the available local farm **Worker Nodes (WN).** User files are copied from the RB to the WN where the job is executed.
7. While the job runs, Grid files can be accessed from a close **Storage Element (SE).**
8. The job can produce new output data that can be uploaded to the Grid and made available for other Grid users to use. This can be achieved using the Data Management tools. Uploading a file to the Grid means copying it on a Storage Element and registering it in the file catalogues. At the same time, during job execution or from the UI, data files can be replicated between two SEs using again the Data Management tools.
9. If the job reaches the end without errors, the output (not large data files, but just small output files specified by the user in the so called Output Sandbox) is transferred back to the RB node.
10. At this point, the user can retrieve the output of his job from the UI.
11. Queries of the job status are addressed to the **Logging and Bookkeeping Database (LB)** database from the UI machine. Also, from the UI it is possible to query the BDII for a status of the resources.
12. If the site where the job is unable to accept it, the job will be automatically resubmitted to another CE that still satisfies the user requirements. After the maximum allowed number of resubmissions is reached, the job will be marked as aborted. Users can get information about what happened by querying the LB service. [14]

As we see it is an infrastructure which consists of User Interface **(UI)**, Resource Broker **(RB)**, Workload Management System **(WMS)**, Computing Element **(CE)**, Berkeley Database Information Index **(BDII)**, Local Resource Management System **(LRMS)**, Worker Nodes **(WN)**, Storage Element **(SE)**, Logging and Bookkeeping Database **(LB)**.

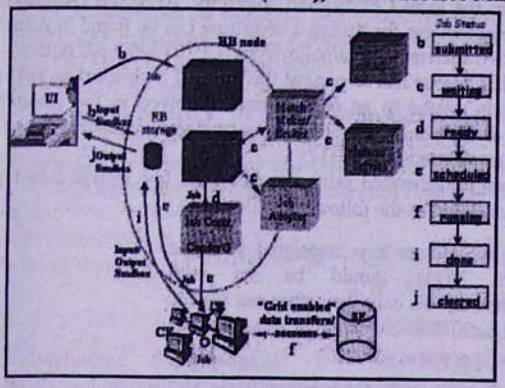
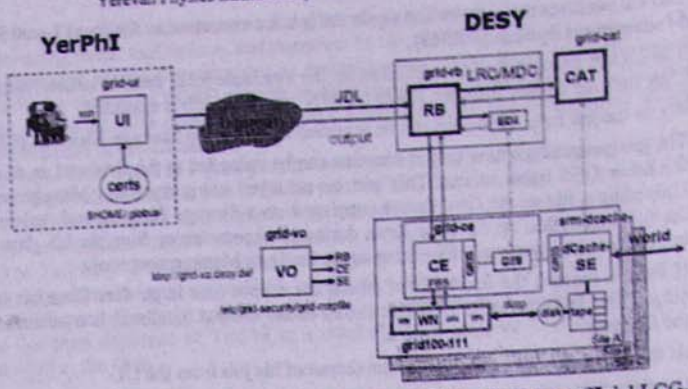


Figure 1 illustrates the process that takes place when a job is submitted to the Grid.

User Interface has been deployed at YerPhI in a manner that uses the infrastructure at DESY to submit the job to the Grid. YerPhI users now can submit the job from UI at YerPhI and after a while get the output. What actually happens is that the jobs were submitted to the RB to DESY and the output retrieved by the UI at YerPhI. The transport between UI and RB is internet. Having stable connection to DESY, now it is not necessary to build a full featured infrastructure including resource broker, computing element, etc.

Figure 2 illustrates Grid infrastructure at DESY and the User Interface at YerPhI



Since August 2004 DESY operates the DESY Production Grid which is an official LCG-2-site in the LCG TestZone and is monitored by the Grid Operation Centre (GOC). The Grid activities are carried out in the context of the EU-project Enabling Grids for E-sienciE (EGEE) which started on April 1st, 2004. DESY is one of 70 partners mainly in Europe, aiming at installing and operating a Grid infrastructure for e-science. DESY is also founding partner of the German initiative to exploit Grids for e-science D-GRID. DESY hosts a number of VOs like 'ilc', 'zeus', 'hone', 'hermes', etc..[15, 16]

In order to start using the Grid at YerPhI scientist needs:

- To request an account from YerPhI IT on grid-ui.yerphi.am
- Login using ssh client
- Request a Grid certificate, from the local Certification Authority, ArmeSFo (Armenian e-Science Foundation). All the requirements for getting a certificate can be found in ArmeSFo CA CP/CPS document http://www.escience.am/ca/policy/Arm_CP_CPS-0.2.pdf points 3.1.9, 4.1 etc). In order to get a certificate one has to present the copy of a passport as well as a document proving that he/she is related to an organization "involved in the research or deployment of multi-domain distributed computing infrastructures, intended for cross-organizational sharing of resources" (point 1.3.3). [17]
- The certificate requests can then be generated using the command line interface by typing `"grid-cert-request"`, which will create the following 3 files:

`userkey.pem`	contains the private key associated with the certificate. (This should be set with permissions so that only the owner can read it, and kept safe) (i.e. `chmod 400 userkey.pem` and copy to safe dir.).
`userreq.pem`	contains the request for the user certificate.
`usercert.pem`	should be replaced by the actual certificate when sent by the CA. (This should be readable by everyone) (i.e. `chmod 444 usercert.pem`).

- Then the `userreq.pem` file is sent (usually by e-mail using a particular format) to the ArmeSFo CA.
- After receiving a certificate from ArmeSFo copy it to `$HOME/.globus/usercert.pem` on the UI.

- Finally register to a VO. http://grid-voms.desy.de:8443/voms/hermes

3. Conclusion

The suggested infrastructure and environment are acceptable to any organization that does not dispose of high performance computing infrastructure to access resources using Grid technologies available at well equipped organizations. This is a low-cost solution to use shared Grid resources.

4. References

[1] The Anatomy of the Grid- Enabling Scalable Virtual Organizations I. Foster, C. Kesselman, S. Tuecke. *International J. Supercomputer Applications*, 15(3), 2001.

[2] Computational Grids., I. Foster, C. Kesselman. *Chapter 2 of "The Grid: Blueprint for a New Computing Infrastructure"*, Morgan-Kaufman, 1999

[3] LHC – The Large Hadron Collider *http://lhc.web.cern.ch/lhc/*

[4] CERN - European Organization for Nuclear Research *http://www.cern.ch*

[5] HERA Results on Physic A.F. Zarnecki Institute of Experimental Physics, Warsaw University

[6] Deutsches Elektronen-Synchrotron *http://www.desy.de*

[7] A Large Ion Collider Experiment (ALICE) *http://aliceinfo.cern.ch/*

[8] A Toroidal LHC ApparatuS (ATLAS) *http://atlas.web.cern.ch/*

[9] The Compact Muon Solenoid (CMS) *http://www-cms.desy.de/*

[10] LCG - LHC Computing Grid Project *http://lcg.web.cern.ch/LCG/*

[11] Globus Toolkit Version 4: Software for Service-Oriented Systems. I. Foster. IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779, pp 2-13, 2006.

[12] Enabling Grids for E-sciencE *http://public.eu-egee.org/*

[13] HERMES experiment at DESY *http://www-hermes.desy.de/*

[14] LCG-2 Users Guide, Manual Series

[15] GRID@DESY *http://grid.desy.de/*

[16] A. Gellrich, DESY. 1st EGEE Course at DESY, Introduction to Grid Computing with the LCG-2 middleware, DESY, Hamburg, Germany

[17] ArmeSFo Certificate Policy and Certification Practice Statement Document *http://www.escience.am/ca/policy/Arm_CP_CPS-0.2.pdf*

Քայլեր դեպի Հաշվողական Grid ցանցեր

Կ. Ս. Մկոյան

Ամփոփում

Հոդվածում ներկայացված է հաշվողական Grid-ի տրամաբանությունը և ճարտարապետությունը, հաշվողական Grid-ը բարձր էներգիայի ֆիզիկայի ոլորտում: Մանրամասն նկարագրված է User Interface (UI) սերվերի աշխատանքի սկզբունքը , այն հիմնական գործընթացները, որ տեղի են ունենում ցանցի տարբեր հանգույցներում, երբ օգտագործողը Grid ցանցում մի որևէ առաջադրանք (խնդիր) ներկայացնում է կատարման:

Որպես օրինակ դիտարկվում են Երֆի-ում տեղադրված User Interface սերվերը և այն քայլերը, որ պետք է կատարի Հայաստանի գիտակրթական կոմպյուտերային ցանցի օգտագործողը DESY գիտական կենտրոնների Grid ցանց մուտք ունենալու և իր խնդիրները լուծելու համար: