GRID COMPUTING: INFRASTRUCTURE, DEVELOPMENT AND USAGE IN BULGARIA

MOMCHIL DECHEV[1], GEORGI PETROV[1], EMANUIL ATANASSOV[2]

[1]Instutute of Astronomy, BAS, Sofia, 72, Tsarigradsko Chaussee Blvd.
[2]Institute of Paralel Proceesing, BAS, Sofia, 25A Acad. G. Boncheff Str.
E-mail: mdechev@astro.bas.bg

Abstract. What's GRID? What is the structure and how to access to GRID resources? Here we try to answer to these questions. We also discuss the development of GRID infrastructure in Bulgaria and the main steps to become a GRID user. A simple example of submitting a job to the GRID is shown.

1.WHAT IS GRID?

1. Maybe the simplest explanation what is grid is done in http://ezinearticles.com/?Introduction-to-Cloud-Computing&id=4845331

- "Grid computing is a form of distributed computing and parallel computing, whereby a 'super and virtual computer' is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks"

That means we use large set of computer resources to process a task. The computer resources can be geographically interspersed, different types hardware and so on, but they work as one virtual computer. More, we can extend resources when we need more computing resources simply by adding new work nodes. That concept permits to solve very resource consuming tasks or large number of simultaneous running applications.

2.GRID STRUCTURE

The organization of the GRID can be divided as a simplification in logical and software structure. The logical structure is shown on Figure 1. The smallest logical unit is called site. One site contains the computers, disk missives, etc., i. e. all hardware.
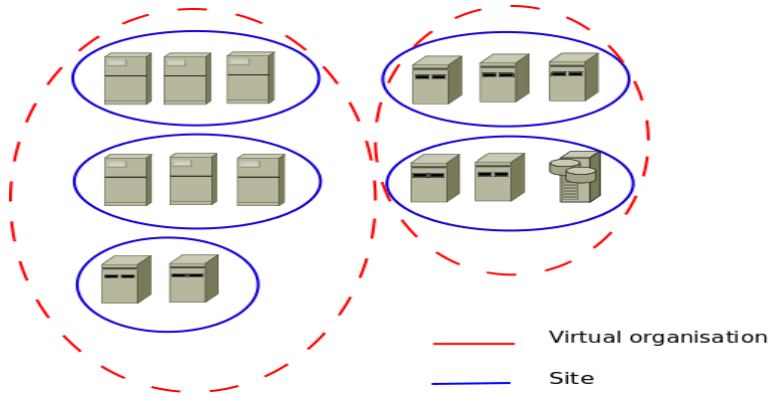
Figure 1: GRID Structure.

Some site form a bigger logical structure called Virtual Organization (VO). It is important to note that a site can be member of many VOs. As it concerns software structure, it can be divided in 4 layers:

- network layer: This is lowest layer which connects grid resources.

- resource layer: Actual grid resources, such as computers, storages and so on.

- middleware layer: This layer provides tools that enable various elements to form the grid.

- application layer: This layer supports various scientific applications. Users see and interact with.

Users and applications do not interact with the last layer and the middleware translates task to grid and give back results.

3. EGEE

The **Enabling Grids for E-sciencE** (EGEE) is one of the biggest grid projects. EGEE project is funded by the European Commission and aims to build on recent advances in grid technology and develop a service grid infrastructure, which is available to scientists 24 hours a day (http://public.eu-egee.org/). The project will concentrates on three core areas:

- The first area is to build a consistent, robust and secure Grid network that will attract additional computing resources.

- The second area is to continuously improve and maintain the middleware in order to deliver a reliable service to users.

- The third area is to attract new users from industry as well as science and ensure they receive the high standard of training and support they need.

EGEE provides a computing support infrastructure for over 14 000 researchers worldwide, from fields as diverse as high-energy physics, earth and life sciences. It brings together 300 sites, 90 organizations, 50 countries, more than 80 000 CPU, 20 PB disk storage and 14 000 users. As a sum we have 300 000 jobs per day with data transfer > 1.5 TB/s.

The EGEE project officially ended on the 31 March 2006. EGEE II started on 1 April 2006 and officially ended on April 30 2010.

4. BULGARIAN GRID

For development of grid infrastructure in Bulgaria a GRID consortium was founded. Members of the consortium are:

Institute for Parallel Processing,

Institute for Nuclear Research and Nuclear Energy,

Institute of Mechanics,

Institute of Electrochemistry and Energy Systems,

Institute of Astronomy,

Institute of Mathematics and Informatics,

Geophysical Institute,

Faculty of Mathematics and Informatics - Plovdiv University,

Faculty of Mathematics and Informatics - Sofia University,

Faculty of Physics - Sofia University,

University of Mining and Geology "St. Ivan Rilski",

Institute of Organic Chemistry with centre of Phytochemistry,

University of National and World Economics.

Soon after the GRID consortium establishment some grid sites were built.

Fig. 2 represents some statistics of these sites. Nowadays we have 5 fully functional grid sites and some more in test phase.

BA-04-PMFSA	Bosnia and Herzegovina	1	Certified	Production	SE: se - error	3.9 d
					SRM: se - error	3.9 d
					SRMv2: se - warn	3.9 d
BG01-IPP	Bulgaria	1	Certified	Production	CE: ce002 - ok	4.6 d
					SE: se001 - ok	21.6 h
					SRM: se001 - ok	5.2 d
					SRMv2: se001 - ok	1.6 d
BG03-NGCC	Bulgaria	1	Certified	Production	CE: ce - ok	13.0 d
					SE: se - ok	13.0 d
					SRM: se - ok	13.0 d
					SRMv2: se - ok	1.6 d
					MPI: ce - ok	31+
BG04-ACAD	Bulgaria	1	Certified	Production	CE: ce02 - ok	10.1 d
					SE: se03 - ok	17.1 d
					SRM: se03 - ok	17.1 d
					MPI: ce02 - ok	31+
BG05-SUGrid	Bulgaria	1	Certified	Production	CE: ce001 - ok	4.1 d
					SE: se001 - ok	21.6 h
					SRM: se001 - ok	18.1 h
BG06-GPHI	Bulgaria	2	Certified	Production	CE: ce1 - ok	20.6 h
					SE: se1 - ok	20.6 h
					SRM: se1 - ok	20.6 h
					SRMv2: se1 - ok	20.5 h
GE-01-GRENA	Georgia	2	Certified	Production	CE: ce - ok	11.6 d
					SE: se - ok	21.6 h
					MPI: ce - error	31+
GE-02-HEPI	Georgia	2	Uncertified	Candidate		
HG-01-GRNET	Greece	1	Certified	Production	CE: ce01 - ok	6.8 d
					SE: se01 - ok	15.6 h
					SRM: se01 - ok	2.8 d
					MPI: ce01 - ok	31+

Figure 2: Certified Bulgarian GRID sites.

5. HOW TO BECOME A USER?

The procedure to become a user of the grid structure is simple, but very strict. First step is to contact one of our Registration Authorities (RA) (http://ca.acad.bg) to make an appointment. Second - Meet the RA in person. You must bring with you:

- valid Identification Document - Identity Card, Driver's License, or Passport
- the declaration (official note) from your Institute/Employing Organization certifying that you are an employee of that organization.
- the removable media on which you store your certificate request (CR). It will be copied by the RA and the removable media will be returned to you. When you meet the RA, you will have to sign a statement that you have read the CP/CPS, so please make sure you have done so.

Figure 3 shows the Bulgarian Academic Certification Authority.

A good explanation of the role of Certification Authority for grid security is done in Pocotilenko et al. (2008).

BG.ACAD | CA
Bulgarian Academic Certification Authority

Registration Authorities

Location	Operated by	Jurisdiction	Manager	Telephone
Sofia	IPP-BAS	All of Bulgaria	Vladimir G. DIMITROV*	+359 2 9796615
Sofia	IPP-BAS	All of Bulgaria	Mariya K. DURCHOVA	+359 2 9796793
Sofia	IPP-BAS	All of Bulgaria	Rayna S. GEORGIEVA	+359 2 9796608
Sofia	IPP-BAS	All of Bulgaria	Sofiya L. IVANOVSKA	+359 2 9796608
Sofia	IM-BAS	BAS, Campus "4th Kilometre"	Kiril S. SHTEREV	+359 2 9796488
Sofia	GPhI-BAS	BAS, Campus "4th Kilometre"	Georgi K. GADZHEV	+359 2 9793708
Sofia	IA-BAS	BAS, Campus "7th Kilometre"	Momchil Ts. DECHEV	+359 2 9795929
Sofia	INRNE-BAS	BAS, Campus "7th Kilometre"	Elena P. PUNCHEVA	+359 2 9795532
Sofia	UNI-SOFIA|FMI	Sofia University	Radoslava D. GORANOVA	+359 2 8161508
Plovdiv	UNI-PLOVDIV	All of Bulgaria	Atanas T. TERZIYSKI*	+359 32 261449
Rousse	UNI-ROUSSE	Town of Rousse	Plamena S. NENKOVA	+359 2 9796626
Sofia	IOCCP-BAS	IOCCP-BAS	Nikolay G. Vassilev	+359 2 9606172

* Denotes a CA staff member.

Figure 3: Bulgarian Academic Certification Authority.

6. RUNNING A TASK ON GRID

Every task on GRID is called job. Executing a task on GRID has some stages:

1. Grid job submission, certification and resources allocation
2. Upload of the computing application and the data
3. Queuing in the local queue
4. Initiation of the computing application
5. Processing
6. Monitoring
7. Download of the results

Stage 6 is not obligatory.

To run a job you have to be a certified user and you have to have so called user interface (UI). UI is part of the grid middleware and is your "door" to the grid resources. Here we summarize basic commands on UI to run a job:

First you need to prepare a jdl file, which will run your job. The structure of the jdl file is as follows:

hello.jdl

```
--------------------------------------------------------------------------------l
Executable = "hello.sh";
StdOutput = "hello.out";
StdError = "hello.err";
InputSandbox = {"hello.sh"};
OutputSandbox = {"hello.out","hello.err"};
#Requirements = other.GlueCEUniqueID ==
"ce001.grid.bas.bg:2119/jobmanager-lcgpbs-
seegrid";
#Requirements = Member("GLITE-
3_0_1",other.GlueHostApplicationSoftwareRunTimeEnv
ironment);
RetryCount=0;
Arguments="test";
#Environment={"…","…"};
-------------------------------------------------------------
```

An example of a JDL file:

Red variables are mandatory. **Executable** is your program. As example hello.sh can be:

hello.sh

```
----------------------------------------------------------------------
#!/bin/bash
sleep 100
echo HELLO GRID
-----------------------------------------------------------------
```

1. StdOutput is where to be written the result of the hello.sh (or your job). Other variables are well explained in https://grid.ct.infn.it/twiki/bin/view/GILDA/SimpleJobSubmission.

To submit this jdl file you use:

glite-wms-job-submit -a hello.jdl,

to check the current status of a job:

glite-wms-job-status --noint -i /home/username/hello.jdl

and to retrieve the output files from job

glite-wms-job-output --dir <directorypath> -i hello.jdl.

To see information about current state of grid resources available for the given VO you should use:

lcg-infosites --vo $LCG_GFAL_VO   ce
lcg-infosites --vo $LCG_GFAL_VO   se
lcg-infosites --vo $LCG_GFAL_VO   lfc,

here $LCG_GFAL_VO should be set to your Virtual Organization (VO)

And instead conclusion some sources of useful information are given:
- https://grid.ct.infn.it/twiki/bin/view/GILDA/UserTutorials
- https://grid.ct.infn.it/twiki/bin/view/GILDA/SimpleJobSubmission
- http://griddeployment.web.cern.ch/griddeployment/documentation/LFC_DPM/lcg_util/html/
- http://ppewww.physics.gla.ac.uk/~fergusjk/howtolowlev.html
- http://training.omii-europe.org/Tutorials/glite3.0/Tutorials.html

Acknowledgements

References

Pocotilenco, V., Altuhov, V. A., Bogatencov, P. and Sidorenco, V.: 2008, Proceedings of the 2nd International Conference on Telecommunications, Electronics and Informatics, May 15-18, 2008 Chisinau, Moldova,- Ch.: UTM, 2008.-ISBN 978-9975-45-082-9. Vol.1, 2008, ISBN 978-9975-45-083-6, p. 389-394.