Acta Physica Universitatis Comenianae
Vol. XLVIII XLIX, Number 1&2 (2007 2008) 65 69

GRID and FMPhI-UNIBA

M. Bábik[1], T. Darányi[1], V. Fekete[2], P. Šťavina[2], M. Zagiba[2], T. Ženiš[2]

[1]Institute of Experimental Physics, Slovak Academy of Sciences, Košice
[2]Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava

About GRID (GRID Networks)

The word GRID has several meanings, so it is not an abbreviation. All of them have in common description of GRID as a form of hardware and software solution for distributive computing. Additionaly, word GRID is also used for distributive computing of many computers and not one super computer with several processors. It, of course, does not mean that such a supercomputer cannot be a part of the GRID. Typical task for GRID is computer programs execution and to data storage.

GRID vs. Supercomputers

GRID computing [2] is a special type of parallel computing which uses classical PC-like computers - so called *nodes* (mainboard, CPU, hard drive, power supply, network card, memory, etc.), connected together with standard network link (Ethernet). It is a world- wide heterogenous network of computing and data resources which works for end user like one entity - one "supercomputer". This is different in classic supercomputers. They have many processors connected together with a high speed local bus. One of the advantages of GRID against supercomputers is its low cost. The disadvantage is that link between computers is relatively slow. That also points to the main use of GRID networks - parallel computing of algorithm which does not depend on each other. Most common application of this technology is in scientific and mathematical computing.

GRID in details

GRID should provide an easy access to computational power and data storage resources. It should free the user from all technical details such as where the program was executed, where temporary data are stored etc. It should behave as a homogeneous environment which knows about the running process (programs), knows its life cycle - from the submition of "execute program request" up to final data retrieval. It knows where the results of the program are stored so user can download them and store them locally on his/her computer. All the steps necessary to take for program execution such as search for platform, which fits programs needs or finding out the propper computers which fulfill

Devoted to Prof. Pavel Povinec 65-th anniversary program's minimal requirements or sending this program to specific location where the code will be finaly executed remaining hidden for user.

User of the grid is identied by his/her GRID certificate and by belongings to *virtual organization*. Virtual organization is solution of "many users" problem. According to a plan there will be too many users for transparent user management, so they are grouped into so called Virtual Organizations witch grants specific rights to their members in program execution and data store. The trust to user is change for trust to virtual organization. Users are usually grouped to collaborate on mutual project.

When a user logs on to a *User Interface* (dedicated computer for GRID job submitting and checking the user's current job status), it sets the name of the program to run, gives path for input data, and sends the request for execution to GRID network. After that it is only up to GRID network to find a propper *site* where the job will be done.

Under the term *site* one should imagine a group of computers where so called services are running. They are responsible for:

- acceptance of job by machine which governs computing (Computing Element) and its transfer to computer where the job will be executed
- Execution of the job - (Worker Node)
- Data store - temporary/permanent (Storage Element)
- Monitoring of job's life cycle - (Monitoring Element)

There are plenty of other servies (for example Logging and Book-keeping it is responsible for checking of jobs status in any part of the submition/execution/termination process and provides these pieces of information to the user; Resource Broker searches for proper site for program execution according it's requirements). Details can be found in the documentation [4, 3].

From the software point of view the idea of GRID is made out of group of server/client programs (in case of gLite) and web services. In present time the only supported platform for gLite is Linux. Of course, it is Linux distribution independent but the preferred one is the Scientific Linux (for now it is recommended to use version SLC/SL 3.x).

Projects and collaborations which created GRID

One of the biggest is EGEE Enabling Grids for E-sciencE [1]. The project is mostly funded by the European Union. Its goal is to create a functional grid infrastructure, which could be available to scientists for 24 hours per day in such a way that it joins together already existing national and institutional grid infrastructures into one homogeneous entity. The project ended on March, 31st 2006. EGEE II started in April, 1st 2006. The biggest user and founder of GRID are people from particle physics - section of science represented by European laboratory for particle physics CERN and its grid infrastracture related to LHC project LCG (LHC Grid) [3]. The main goals of LHC grid are simulations of experiments on the Large Hadron Collider LHC as well as simulations of LHC itself. There are virtual organizations defined for LHC, one for each experiment.

Another big collaboration is The Globus Consortium (*http://www.globusconsortium.org*). It is an open-source grid infrastructure solution for enterprises and corporations. The name of software developed by them is *The Globus toolkit*.

GRID site at FMPhI CU - FMPhI-UNIBA

People from Nuclear physics and biophysics department are members of LCG virtual organizations. There were and still are efforts to compute problems related to experimental physics on equipment located in Slovak Republic and therefore to be independent on availability or inavailability of computation resources at European laboratory for particle physics. Even more! The site built in Slovak Republic under control of department of nuclear physics and biophysics could be understood as slovak contribution to CERN's experiments. That was the starting signal for members of department of nuclear physics and biophysics working for ATLAS and ALICE experiments to build such a site at Faculty of Mathematics, Physics and Informatics. Project has been written and work on it started soon after its acceptance. The main phase of the project to build such a site and make it operational is finished now. Current efforts are spent to enhance computing power as well as to increase reliability of the services provided by the site. The name of the site is FMPhI-UNIBA. Logo of the site is:

In present time our site is used for ATLAS and ALICE experiment problems. Because the site is a part of LCG grid infrastructure it has been included in the project of grid sites monitor in museum of sciences in London. List of sites an be found at this URL (You have to push the expand button near Slovakia):

http://gridportal-ws01.hep.ph.ic.ac.uk/dynamic_information/egee-locations.xml

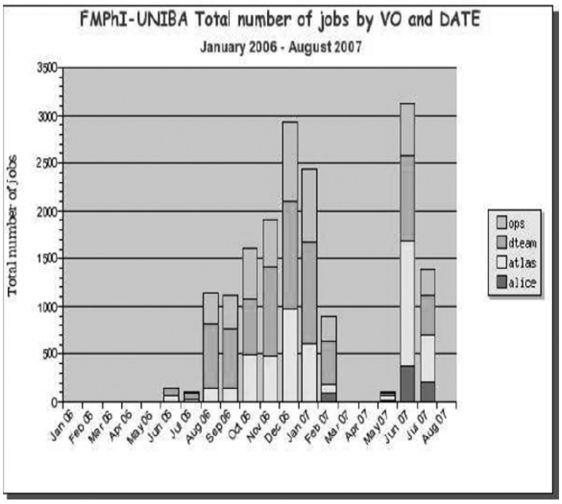
In present time, basic services of the site run at the following hardware (you can see in the table which configuration we chose for a given service, as well as the technical details of the computer).

Type of service	**CPU cores**	**Hardware**
Computing Element	2	two CPUs AMD Opteron 2 GHz, 4 GB RAM, 2.8 TB disk space
Worker Node	56	14 are AMD athlon64X2 DUALCORE 4200+ 2 GHz, 2 GB RAM 14 are Intel Core2 2.4 GHz, 4 GB RAM
Storage Element	2	This service run on the same computer as Computing Element (2 services at the same PC), we also have 2 disk arrays. Total capacity is 10 TB
User Interface	1	AMD Athlon64 3000+ 2 GHz, 1 GB RAM
Other Supporting Servers	9	AMD Athlon64 3000+ 2 GHz, 1 GB RAM

Operating system is Scientific Linux and grid services are provided via gLite software.

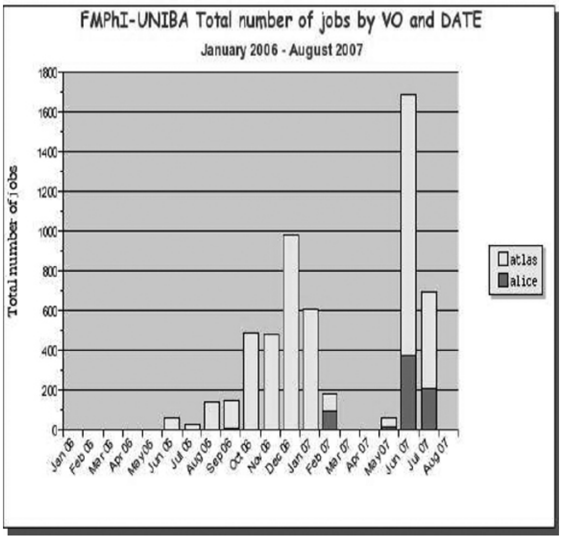
Overall results (finished jobs) on site since september 2006 to july 2007[1]

	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07	Total
alice	0	0	0	0	0	0	0	0	6	1	0	1	1	92	0	0	11	373	204	0	689
atlas	0	0	0	0	0	59	24	139	138	489	478	977	607	86	0	0	50	1,311	490	0	4,848
dteam	0	0	0	0	0	78	62	682	620	586	939	1,119	1,067	457	0	0	28	899	427	0	6,964
ops	0	0	0	0	0	2	17	323	355	531	490	827	767	261	0	0	11	541	262	0	4,387
Total	0	0	0	0	0	139	103	1,144	1,119	1,607	1,907	2,924	2,442	896	0	0	100	3,124	1,383	0	16,888



Results (finished jobs) on site since september 2006 to july 2007 for ATLAS and ALICE VOs

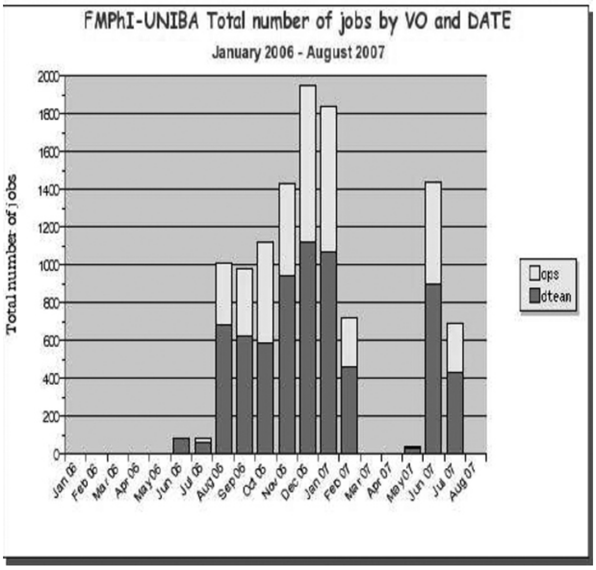
	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07	Total
alice	0	0	0	0	0	0	0	0	6	1	0	1	1	92	0	0	11	373	204	0	689
atlas	0	0	0	0	0	59	24	139	138	489	478	977	607	86	0	0	50	1,311	490	0	4,848
Total	0	0	0	0	0	59	24	139	144	490	478	978	608	178	0	0	61	1,684	694	0	5,537



[1] Pictures were made and taken from EGEE monitoring system
https://monitoring.egee.man.poznan.pl/index.php?go=26

Results (finished jobs) on site since september 2006 to july 2007 for DTEAM and OPS VOs

	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07	Total
dteam	0	0	0	0	0	78	62	682	620	586	939	1,119	1,067	457	0	0	28	899	427	0	6,964
ops	0	0	0	0	0	2	17	323	355	531	490	827	767	261	0	0	11	541	262	0	4,387
Total	0	0	0	0	0	80	79	1,005	975	1,117	1,429	1,946	1,834	718	0	0	39	1,440	689	0	11,351



References

[1] EGEE Frequently Asked Questions, *http://www.eu-egee.org/faq-folder/contents/faq-folder/faq1.*
[2] GRID Computing, *http://en.wikipedia.org/wiki/Grid_computing.*
[3] J. Knobloch: LCG Technical Design Report, CERN-TDR-01 CERN-LHCC-2005-024, 2005, *http://lcg.web.cern.ch/LCG/tdr/LCG_TDR_v1_04.pdf.*
[4] EGEE Collaboration: gLite Installation and Configuration Guide v3.0 rev. 2, 2006 *http://glite.web.cern.ch/glite/packages/R3.0/R20060502/doc/installation_guide_3.0-2.html.*