IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
6-8 September 2007, Dortmund, Germany

Ukrainian Grid Infrastructure: Practical Experience

Mykhaylo Zynovyev [1], Sergiy Svistunov [2], Oleksandr Sudakov [3], Yuriy Boyko [4]

1) Gesellschaft für Schwerionenforschung mbH, 1 Planckstr., 64291 Darmstadt, Germany; Bogolyubov Institute for Theoretical Physics of National Academy of Sciences of Ukraine, 14b Metrologichna str., 03680, Kyiv, Ukraine; Mykhaylo.Zynovyev@cern.ch

2) Senior Searcher, Bogolyubov Institute for Theoretical Physics of National Academy of Sciences of Ukraine, 14b Metrologichna str., 03680, Kyiv, Ukraine, svistunov@bitp.kiev.ua, http://www.bitp.kiev.ua

3) Head of Sector, Information&Computer Center, National Taras Shevchenko University of Kyiv, 64, Volodymyrska str., 01033, Kyiv, Ukraine, saa@grid.org.ua, http://cluster.kiev.ua

4) Head of Information&Computer Center, National Taras Shevchenko University of Kyiv, 64, Volodymyrska str., 01033, Kyiv, Ukraine, boyko@univ.kiev.ua, http://icc.univ.kiev.ua

***Abstract** – In 2006 National Academy of Sciences of Ukraine approved the project for development of Grid technologies in Ukraine. Since then an extensive work is carried out in order to build a full-scale Grid infrastructure for scientific and educational institutions. This article deals with the aspects of building such infrastructure, its organization and peculiarities of implementation in Ukraine, shows the current status of project, describes scientific tasks which benefit from using Grid.*

***Keywords** - Grid, ARC, AliEn, Cluster, Docking*

I. Introduction

The project to create Ukrainian Grid infrastructure is aimed to integrate distributed computing resources of the member institutions of National Academy of Sciences of Ukraine (NASU), as well as of other Ukrainian educational organizations which require high-performance computing to solve their tasks. Building high-speed fiber-optic network, designing and setting up computing clusters, and installing Grid middleware allowed to create full-scale experimental Grid-environment which integrates and provides a uniform user access to computing and data storage resources.

This article examines conception behind the project, describes its implementation specificity, lists the reasons for choosing particular middleware and shows the first results of using Ukrainian Grid to solve scientific tasks.

II. Building Ukrainian Grid Infrastructure

Almost all of the world's developed countries are carrying out their national Grid-projects to build an applicable infrastructure and develop technologies which provide remote access to various computing resources regardless of user's location. EGEE project (Enabling Grids for E-Science in Europe) is coordinating these activities in Europe.

Grid technology has only recently been introduced in Ukraine. Although the first Grid sites were set up in 2005 due to initiative of Bogolyubov Institute for Theoretical Physics (BITP) of NASU and Information & Computer Center (ICC) of National Taras Shevchenko University of Kyiv (KNU) in the frame of supporting ALICE (A Large Ion Collider Experiment) [1] experiment in CERN (European Center for Nuclear Research), Geneva, Switzerland. These sites are active members of ALICE virtual organization in AliEn Grid environment and are used for running jobs of simulation, reconstruction, and analysis of physics data of the ALICE experiment. KNU cluster [2] started to run in 1999 and was the first high-performance computing system to have Internet connection and to serve Ukrainian scientific and educational institutions for solving tasks in physics, chemistry, biology, and informatics.

Gained experience in maintenance of Grid-sites and clusters, and understanding the necessity of Grid technologies development in Ukraine have led to the start of designing NASU project for building its own Grid infrastructure to serve the leading research institutes of Ukraine. The design for the project and the Grid development initiative as a whole were approved by Presidium of NASU in 2006. As a result the clusters in Institute of Molecular Biology and Genetics (IMBG) of NASU, Institute of Cell Biology and Genetic Engineering (ICBGE) of NASU, and Main Astronomical Observatory (MAO) of NASU were to be constructed; the clusters in BITP, Institute for Condensed Matter Physics (ICMP) of NASU and Kharkiv Physical-Technical Institute (KPTI) were to be modernized. All the computing resources of these institutes were to be integrated in Ukrainian Grid infrastructure.

It is widely known that Grid environment is built on top of computing resources, high-speed and reliable Internet connection, and middleware. At the start of Ukrainian Grid project only middleware was available. Clusters' setup and network construction were performed on the run.

In December of 2006 Grid environment which consisted of NASU and KNU resources was set up for experimental exploitation. In March 2007 the resources of Institute of Cybernetics (IC) and Space Research Institute of NASU (IKD) were added to the NASU Grid infrastructure.

As you see from this historical reference the Grid initiative in Ukraine is in its development phase. Therefore to acquire exhaustive information on Ukrainian Grid application is not yet possible and a certain period of exploitation has to pass before the results could be analyzed and presented.

To design computing clusters of NASU the Beowulf concept with the following key features was chosen:

- Computing servers are of standard PC architecture;
- Gigabit Ethernet is used as a network equipment to connect servers;
- Free and open-source software, and Unix-like operating systems, such as Linux, are used.

That's why all computing clusters of NASU have x86, x86_64 architecture for dual-, quad-processor servers with 1-4 GB RAM and 36-80 GB hard drives. For inter-server data transfer Gigabit Ethernet is used with 1 Gb/s switches, except for two clusters where InfiniBand is used.

Computing clusters' characteristics are shown in Table I.

It is important to notice that hard drive space capacity is used for storing OS files, software packages, and temporary files and is not available for user data. To store user data every cluster has a disk array. Various Linux versions of operating system are installed. TORQUE version of PBS and also open-source SGE system are used for job management.

Performance benchmarks of some of the clusters produced the following results:

- IC (62 nodes) – 2235 GFLOPS;
- ICMP – 680 GFLOPS;
- KNU – 105 GFLOPS.

TABLE I
COMPUTING CLUSTERS' CHARACTERISTICS

#	CLUSTERS	NODES	PROCESSORS	RAM	HARD DRIVES	NETWORK	STORAGE	OS
1	KNU	24	8 nodes-2xIntel Xeon CPU 2.40 GHz 7 nodes 2xIntel Xeon CPU 3.2 GHz 5 nodes 2xIntel Xeon 5130 2Ghz 4 nodes 2xIntel Pentium III CPU 1 GHz	512 MB - 4 GB	36-72 GB	Gigabit Ethernet	6 TB	Fedora Core 1.x-3.x
2	BITP	20	10 nodes - 2xIntel Xeon, CPU 2.80GHz 10 nodes - 2xIntel Xeon, CPU 3.2GHz	4 GB	36 GB	Gigabit Ethernet	12 TB	Scientific Linux 4 2.6.9-34.ELsmp
3	MAO	8	2 x Dual Core Intel Xeon CPU 5130, CPU 2.00GHz	2 GB	80 GB	InfiniBand	3 TB	Debian GNU/Linux 4.0
4	ICBGE	8	2 x Intel Xeon, CPU 3.20GHz	4 GB	36 GB	Gigabit Ethernet	3 TB	Scientific Linux 4 2.6.9-34.ELsmp
5	IMBG	8	2 x Dual Core AMD Opteron, CPU 2.0 GHz	4 GB	80 GB	Gigabit Ethernet	1 TB	Scientific Linux 4 2.6.9-34.ELsmp
6	ICMP	36	17 nodes - 2 x Dual Core Intel Xeon 5130 2Ghz, 13 nodes – 2x AMD Athlon MP 2200+, CPU 1.8 GHz 6 nodes - 2 x AMD Opteron 246, CPU 1.8 GHz	4 GB	40-80 GB	InfiniBand Gigabit Ethernet	6 TB	Linux CentOS 4.0-4.4
7	IC	74	2x Dual Core Intel Xeon 5160 CPU 3. 0GHz	8 GB	80 GB	InfiniBand	15 TB	Linux CentOS -4.2

Reliable Internet connection is one of the cornerstones of Grid computing. All mentioned institutions had access to Internet before, but the quality of connection was not sufficient for reliable data exchange. This obstacle has been only recently overcome. By request from NASU Ukrainian Academic and Research Network (UARnet) company has built during the last two years an infrastructure which connects academic institutions with 100 Mb/s fiber-optic channels to the network backbone Lviv-Kyiv-Kharkiv (2.5 Gb/s) with subsequent connection to Exchange Points in Slovakia and Poland (155 Mb/s) "Fig. 1."

Various ARC middleware components [3,4,5] are built on top of Globus Toolkit (GT), but the Grid Manager, gridftpd (the ARC GridFTP server), the information model and providers (NorduGrid schema), User Interface and broker (integrated into the user interface), extended Resource Specification Language (xRSL), and the monitoring system are truly innovative solutions.

ARC information system makes use of MDS system of GT2, which is based on OpenLDAP backend. Using dynamic indexing, it allows constructing different topologies of Grid-connected sites.

For indexing and displaying monitoring information of Ukrainian Grid a separate server (lcg.bitp.kiev.ua) was set up in BITP. Indexing is done by GT's Grid Information Index Service (GIIS), which makes use of NorduGrid schema, and monitoring is done by ARC Grid Monitor. In the beginning GIIS in BITP was chosen as the only index service to register all resources for Ukrainian Grid environment. But this topology will evolve in the near future to increase fault-tolerance.

The following ARC services are installed on front-end machines of each cluster:

- Grid Manager takes care of user jobs, session directories and the input data cache area.
- GridFTP Server (gridftpd) provides mechanism for job submission, FTP server functionality for job data files, and access control based on user certificates. It also supports virtual directory trees.
- Downloader and Uploader make possible the transfer of job files to and from user's computer;
- Virtual Organization Membership Service (VOMS) is an extension for managing user membership in virtual organizations.

It is worth to mention that installation of ARC middleware has not caused any significant problems and required little intervention from system administrators. Installation was performed using pre-built binaries as well as by building from sources. ARC configuration is simplified and boils down to modification of only one file arc.conf. Our experience has confirmed that ARC is a lightweight, non-intrusive, portable, out-of-the-box middleware solution.

Fig. 1. UARnet network.

At present Grid environment of NASU is running in experimental mode and users are educated in using it. The monitoring information on resources is available at http://lcg.bitp.kiev.ua/loadmon.php "Fig. 2."

To submit jobs to Grid environment users must install a User Interface (UI) package on their PCs. UI command line utilities can be used to do the following: to submit jobs to the Grid, to query status of jobs, to cancel jobs, to retrieve job output files, to check cluster status and availability.

Security for the Grid environment is ensured by certificates based on X.509 standard for public key infrastructure, which are issued by Certification Authority (CA). Experimental CA for Ukrainian Grid of NASU is situated in BITP. Its certificates are not recognized by any other Grid environments and are used only for NASU. The setup of a true CA of the European standard is a current task which is planned for completion on the next stage of the project.

III. Application

Ukrainian Grid infrastructure is aimed to create a powerful distributed computing environment, which can offer resources for solving various computational tasks in different fields of applied sciences. The current research activities of institutions which can benefit from using Ukrainian Grid infrastructure are specified below.

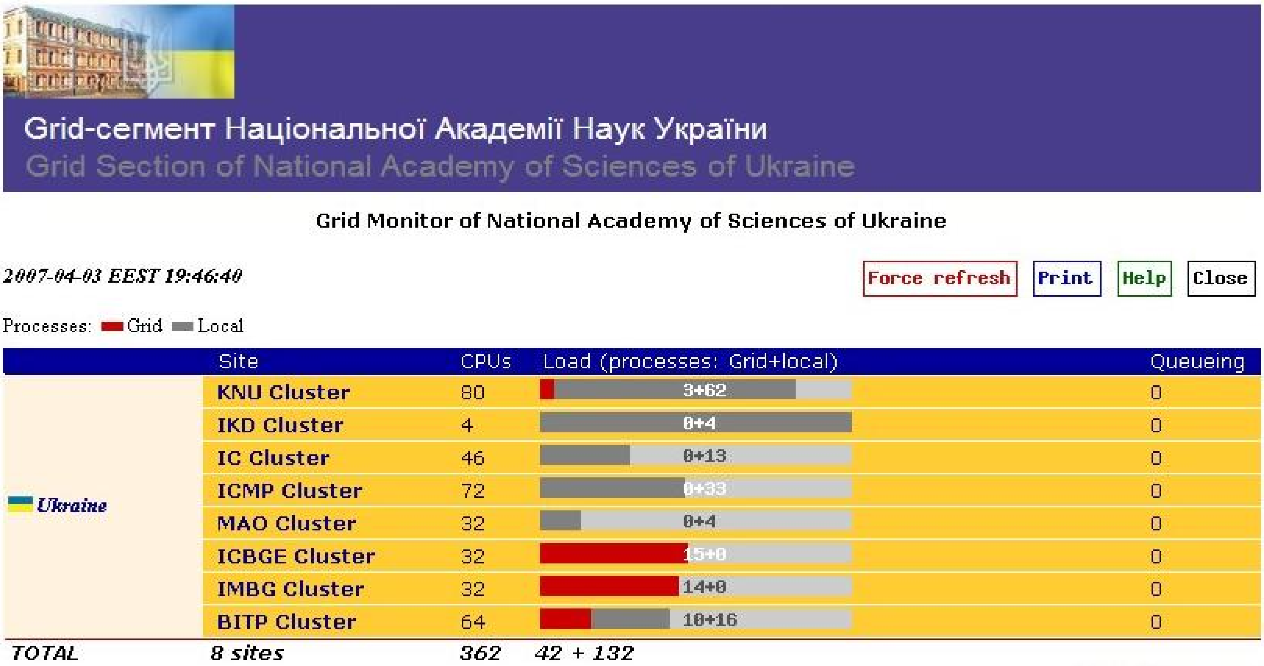
BITP is planning to participate in the analysis of proton-proton, proton-nuclei and nuclei-nuclei collisions data from the LHC detectors ALICE, CMS and TOTEM. Today the Institute theorists are dealing with the data coming from all the RHIC Collaborations pursuing the idea of creating quark-gluon plasma in the ultrarelativistic heavy ion collisions.

The scientists of IMBG are performing simulation of proteins and calculations of the molecular dynamics of certain proteins, targeting future development of new medicines, including HIV- protease and its mutant forms, resistant to inhibitors.

ICBGE is carrying out analysis of biological sequences, modelling of space structure of biological macromolecules and their interactions with other macromolecules and low molecular substances.

MAO scientists are taking part in the following studies: a study of the spectra of stars and sub-starry objects on different evolutionary stages; study of physics of active processes on the surface of comet nuclei; the simulation of the processes of formation and evolution of the large-scale structures of the universe, such as galaxies and the nucleus of galaxies (black holes), clusters of galaxies and processes of formation and evolution of stars.

This incomplete list is a good illustration, that it’s impossible to use a single software solution for all these tasks. Every team uses its own either universally recognized (like EnFuzion by Axceleon) or unique self-elaborated software. Therefore ARC middleware is used for application-independent services like job submitting, job status query, job output retrieval, data transfer, monitoring. Software package manager which can be responsible for automatic remote installation of special software package has not been implemented yet.



	Site	CPUs	Load (processes: Grid+local)	Queueing
Ukraine	KNU Cluster	80	3+62	0
	IKD Cluster	4	0+4	0
	IC Cluster	46	0+13	0
	ICMP Cluster	72	0+33	0
	MAO Cluster	32	0+4	0
	ICBGE Cluster	32	15+0	0
	IMBG Cluster	32	14+0	0
	BITP Cluster	64	10+16	0
TOTAL	8 sites	362	42 + 132	

Fig.2. Example of monitoring information.

On this stage of the project jobs are ready-to-run modules, which do not require any external libraries for remote installation. Certainly such approach is limiting Grid functionality. To solve this issue some special software packages were preinstalled.

xRSL of ARC turned out to be a very convenient tool. Specifying such attributes as “inputFiles“ and “outputFiles” one can transfer a significant number of input files to the cluster where the job will run, and retrieve job results after Grid Manager transfers output files to the specified location (if location is not specified, the user can retrieve them later via UI). With a “cluster” attribute one can explicitly specify the list of clusters where the job can run and where the specific software is preinstalled.

These features helped to process the X-radiation data from INTEGRAL satellite (project of University of Lausanne, Switzerland) with the means of XMM SAS and

INTEGRAL OSA analysis packages.

Application of Grid is suitable for solving tasks of molecular simulation, which include computer development of medicines, study of physics processes in bio systems on molecular level. Simulation is done with the methods of quantum chemistry and molecular mechanics with the help of such software as Flo/qxp [6] GAMESS [7], GROMACS. These software packages are not Grid-enabled, and in order to use them on Grid special scripts were developed, that allow middleware access to transfer required files to the clusters. The packages were preinstalled at the temporary storage.

One of the examples of efficient Grid usage is high-performance molecular docking with Flo/qxp software. This task can be effectively parallelized with SIMD (Single Instruction, Multiple Data) technique. The same copy of program is submitted to every Grid node, but every copy has its own unique data to process. The data for the task is a part of compounds database, and is used for docking at the user specified binding site. Communication delays for such tasks are the consequence of data transfer to the site at the start and retrieval of results at the end. The size of the database of 1000000 compounds, which was used for the experiment, is 670 MB. 480 independent jobs were prepared for computing on the Grid, each having its own data to process. All the jobs were submitted to the Grid environment simultaneously. 120 processors in the Grid were allocated for these tasks. Thus on average 4 jobs per processor. Each job performed geometry optimization on 2000 compounds on average. Total execution time took 73 hours, that means 2 minutes per compound. For comparison, such computation on KNU cluster alone would take up to 2 weeks of full load.

Thus, maximum advantage of using Grid environment is taken by those tasks which require large data arrays processing and can be parallelized in SIMD manner. These include tasks in molecular dynamics (calculating trajectories for statistical analysis), algorithms for analysis of data taken from satellites, medical diagnosis applications.

IV. Conclusions

Experimental Grid Infrastructure of NASU started to operate only 6 months ago. Beginning with 2 clusters in BITP and KNU, nowadays it unites 8 clusters located in Kyiv and Lviv. In 2007 resources of Institute of Radio Astronomy of NASU and Institute of Scintillating Materials of NASU in Kharkiv, Institute of Geotechnical Mechanics of NASU in Dnipropetrovs'k, Institute of Physics of NASU in Kyiv are to be added. Increase in number of resources and increase in number of tasks for computation will pose new problems in ensuring stability and reliability of operation, monitoring, and Grid-enabling applications. However, first experiments of using Ukrainian Grid infrastructure assure that Ukrainian scientific institutions are capable of solving these problems. NASU is in the beginning phase of creating national Grid community. The closest perspective is to join EGEE as an official member. This will offer new opportunities in international cooperation on Grid development and in usage of Grid technologies for solving various tasks in different fields of science.

References

[1] P. Saiz, P. and others, *AliEn - ALICE environment on the GRID*, Nucl. Instrum. Meth., A502, 2003, pp. 437-440.

[2] Yu.V. Boyko, O.O. Vystoropsky, T.V. Nychyporuk, O.O. Sudakov "Kyiv National Taras Shevchenko University High Performance Computing Cluster," *In Proc. Third International Young Scientists Conference on Applied Physics*. June, 18-20, 2003, Kyiv, Ukraine, pp. 180-182

[3] M. Ellert, M. Grønager, A. Konstantinov, B. Kónya, et al. "Advanced resource connector middleware for lightweight computational grids". *M.Ellert et al., Future Generation Computer Systems*, 23, 2007, pp. 219-240.

[4] A.Konstantinov et al, "Data management services of NorduGrid"., *in Proceedings of CHEP 2004*, eds. A.Aimar, J.Harvey and N.Knoors, CERN-2005-002, vol. 2, 2005, p. 765.

[5] P.Eerola et al "The NorduGrid production Grid infrastructure, status and plans," *Proc. 4th Int'l Workshop on Grid Computing (GRID'03)*, IEEE CS Press, pp. 158-165

[6] C. McMartin, R. S. J. Bohacek Powerful, *Rapid Computer Algorithms for Structure-Based Drug Design*. J. Comput.-Aided Mol. Des. 1997, 11, pp. 333-344 .

[7] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Kosecki, N. Matsunaga, K.A. Nguyen, S.J.Su, T.L. Windus, M. Dupuis, J.A. Montgomery, J. Comput. *Chem*. 14, pp. 1347-1363 (1993)