TASK QUARTERLY 17 No 1–2, 119–125

PLGRID PLUS SUPPORTS POLISH SCIENCE

ZOFIA MOSURSKA AND ROBERT PAJĄK

AGH University of Science and Technology,
Academic Computer Centre Cyfronet AGH,
Nawojki 11, 30-950 Cracow, Poland
{Z.Mosurska, R.Pajak}@cyfronet.krakow.pl

1. Introduction

The IT infrastructure for science, Polish Grid (PL-Grid), was created within the PL-Grid project (`http://projekt.plgrid.pl/en`) during 2009–2012 period. The PL-Grid Infrastructure has got currently more than 1600 Polish scientists registered who actively use its vast computing resources – over 570 TFlops of computing power and more than 5.5 PBytes of storage capacity – and submit ca. 1 million jobs per month. Within the next project – PLGrid Plus (2011–2014, `http://www.plgrid.pl/en/plus`) it is planned to extend these resources up to 730 TFlops of computing power and more than 8 PBytes of storage capacity.

The PL-Grid Infrastructure resources are not just computers and storage resources – installed in the five partner centers that are part of the PL-Grid Consortium. The infrastructure also includes specialized software as well as the services and tools – developed by the PL-Grid Consortium – which support both users and administrators of computers.

Access to the PL-Grid Infrastructure enables scientists to increase the scale of the calculations carried out within the scientific research, which would be impossible to achieve using separate computers, and to place demands for resources and services available within the infrastructure.

2. The PLGrid Plus project

The e-infrastructure-related requirements of the scientific community are highly diversified and depend on the scientific field. The differences may relate to the computing power and type of a computing infrastructure, software resources and databases, as well as a unique measuring/research equipment, which is often also a part of the infrastructure. In order to effectively support, in terms of IT, the

development of scientific research in the various problem areas and researchers – it is necessary to fit the PL-Grid Computational Infrastructure to the problems being the subject under study.

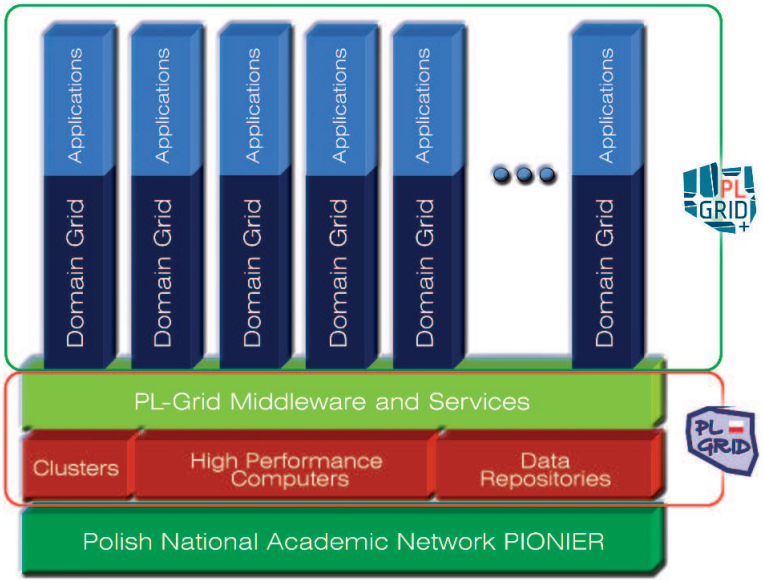
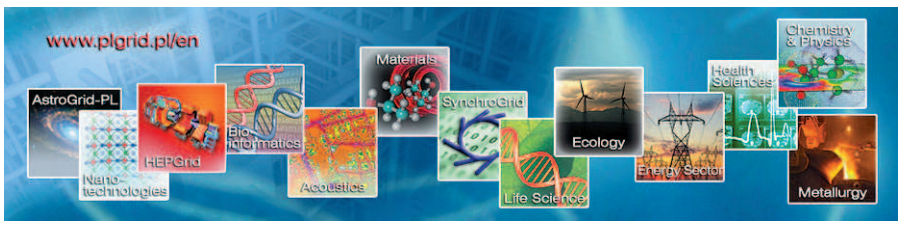


Figure 1. PL-Grid structure

Therefore, what has become the most important goal of the PLGrid Plus project, is expansion of the existing infrastructure towards domain-specific solutions for research teams, which will allow for conducting more effective research and facilitate wider international cooperation in the area of e-Science.

These domain-specific solutions are currently being created for 13 groups of users, representing strategic areas and important topics for the Polish and international science: AstroGrid, HEPGrid, Nanotechnologies, Acoustics, Life Science, Chemistry and Physics, Ecology, SynchroGrid, Energy Sector, Bioinformatics, Health Sciences, Materials, and Metallurgy. Examples of research carried out under the PLGrid Plus project, conducted by the scientists in these areas, are presented at the end of the article.



Subsequently, within the Project it is foreseen to launch more IT services, also for teams of scientists representing other scientific disciplines, who plan

experiments supported by large-scale simulations or work with large databases (collections) of data.

3. Computing services

The concept of domain-specific computing environments involves creation and provisioning of a set of specialized computing services that can cover many aspects of the infrastructure. Examples of categories of such services are as follows:

- access to specialized software,
- integration of domain data, on which calculations are based,
- user support in the use of computing resources,
- assistance in planning, executing and analyzing of complex, multi-step computing experiments.

4. User support

The offered services include a wide range of consulting and training. Users of the PL-Grid computing resources are provided with support and professional help in solving any problems related to access and effective execution of calculations on these resources. We offer assistance in launching the specialized scientific packages on distributed computational resources as well as technological and IT support when designing their own scientific applications or multi-step computing experiments to be launched on the PL-Grid Infrastructure.

Traditional and remote training of all levels are organized. Information on training courses is regularly published at the PL-Grid Portal. It is also possible to organize training courses on demand or at a place of the contracting authority.

For some time, on the PL-Grid computers we provide services and tools, designed also for novice users, being an excellent assistance for easy job submitting. One example of such a tool is QCG-Icon (`http://www.qoscosgrid.org/trac/qcg-icon`), installed on a user's computer. It allows the user to send tasks to be performed by QosCosGrid services in one of the computing centers included in the PL-Grid Consortium, without having to search for available resources. Then, this tool allows the user to follow the calculations, and finally, automatically downloads results to a local computer. Thus, it spares the user (even an unskillful one) performing many actions, which are not very interesting for him. Another example is the GridSpace2 Experiment Workbench (`https://gs2.cyfronet.pl/`), which is a novel virtual laboratory framework that enables researchers to conduct virtual experiments on grid-based resources and other HPC infrastructures.

5. Computational grants in the PL-Grid Infrastructure

Computational grants were implemented in the summer of 2011, as a response to the need for more efficient management of PL-Grid Infrastructure resources. To be able to perform scientific calculations, the user must have got an active computational grant.

With the registration in the PL-Grid Portal, the user is granted (without any formalities) a *personal grant*. Such a grant is intended for the start of using of resources – it is designed to provide resources to test or perform minor calculations. Within the personal grant, one can use the PL-Grid Infrastructure resources for 6 months, until reaching the limit: 1000 hours of computations, 40 GB of storage. The personal grant is automatically renewed after 6 months and is not subject to accounting.

In many cases, the resources assigned within the personal grant may be insufficient to perform scientific calculations. If the users expect that their personal grant offers too few resources to carry out the planned calculations, they should request through the PL-Grid Portal for additional resources within the *actual grant*.

The actual grant is an agreement between the user (or a team of users) and the provider of resources (*i.e.*, a computing center or centers in the PL-Grid Infrastructure) to provide computing power, storage space, or additional resources offered in a specified capacity and at a specific time. The actual grant is subject to accounting approximately every six months during its lifetime and after. The user then is obliged to provide information on the scientific results, including publications, realized within the grant.

6. How to start the cooperation?

Every scientist in Poland may use free of charge the resources and services of the PL-Grid Infrastructure. To obtain an account in the PL-Grid Infrastructure that enables access to its computing resources, one should register in the PL-Grid Portal, available from the `http://www.plgrid.pl/`. In addition, we also accept applications from teams of researchers (which represent areas other than the 13 pilot ones), regarding the need for new IT services that would substantially support their calculations. Such applications, as well as all inquiries regarding registration, access to the infrastructure, services and tasks execution, should be directed to: *helpdesk@plgrid.pl*.

7. Use case scenarios implemented in the PL-Grid Infrastructure

At the Gdansk University of Technology, a study on **new antifungal antibiotics** has been performed, using the PL-Grid Infrastructure computing resources. Powerful infrastructure supported the process of designing the antibiotics, whose use gives fewer side effects, but they are equally effective against fungi attacking the human body. More at: `http://www.isgtw.org/feature/designing-better-antibiotics`.

Another example of the use of the PL-Grid Infrastructure are actions in **the field of acoustics** undertaken by scientists from the Department of the Multimedia Systems of the Gdansk University of Technology, who – in the framework of the PLGrid Plus project – will develop and make available

a set of tools to carry out numerical calculations in the field of generation and propagation of noise in the urban environment, which comes from different sources (*e.g.*, industrial, road or rail noise). The developed computing services will enable acquisition of data (of the currently unprecedented accuracy) on noise pollution. The work of these scientists has already been awarded in the International Eureka Contest, where 236 inventions from 22 countries have been reported. Department of the Multimedia Systems of the Gdansk University of Technology received a silver medal for the System of Dynamic Creation of Noise Maps produced using the Supercomputing Platform, developed within the PLGrid Plus project. More at: `http://www.plgrid.pl/projekty/plus/materialy_promocyjne/newsletter/Newsletter_PLGrid_Plus-wrzesien_2012.pdf`.

A next example is the use of PL-Grid Infrastructure for the work associated with **the construction of the National Centre of the Synchrotron Radiation Solaris** – the first Synchrotron in Poland, which is being built in the 3rd Campus of the Jagiellonian University in Cracow and which will use computing resources in many ways. At the first stage, the primary result of the work carried out under SynchroGrid will be calculation and verification of the parameters of the built accelerator. Synchrotron model running on the infrastructure will also verify the synchrotron control applications, before they are executed in the real machine. The use of the infrastructure already allows for faster execution of design tasks, and it is planned that – thanks to the PL-Grid Infrastructure – the synchrotron will be run much faster than if one had to test everything yet on the "living organism". In the second step, the outcome will be the results of various tests carried out using synchrotron radiation, and analyzed using large-scale calculations.

Researchers from the Faculty of Metals Engineering and Industrial Computer Science of AGH – University of Science and Technology in Cracow (involved in the **metallurgy domain grid**) work on the reconstruction of material microstructure including single and multiply grains. Such a reconstruction can be used directly in numerical simulations of microstructure behavior under loading conditions, or it can be applied in the procedure of Statistically Similar Representative Volume Element (SSRVE) creation. The procedure of SSRVE creation is based on the Non-Uniform Relational B-Splines (NURBS) and optimization algorithm, which runs Abaqus software in each iteration, making this solution very time consuming and inefficient. Therefore, a parallel and distributed implementation of this algorithm was proposed for 2D SSRVE by using the grid infrastructure. The main aim of the work within PLGrid Plus project is to extend this approach by implementation of 3D interpolating NURBS and its application in 3D SSRVE procedure, which could be executed in parallel. The software implementing 3D NURBS for material grain reconstruction was realized for PL-Grid Infrastructure and initially tested by using the QCG Icon application. Tests were performed for 64 tasks of a separate startup optimization procedure, performed on a various

number of cluster nodes. Future work assumes development of this solution and its use to the modelling of the full microstructure of metallic materials.

The most famous example of the use of computing infrastructure are **High Energy Physics (HEP**) experiments, where computers have played an important role for many years, however, the awareness of the importance of IT was not common in the experimental teams and the IT infrastructure was treated as an important, but not a critical addition to the experimental equipment.

Changes came with the start of the LHC project, focusing on the construction and operation of the Large Hadron Collider. Starting with that moment a period of intense physical analyses – conducted using the WLCG (Worldwide LHC Computing Grid) – began. At the final stage, the analyses are carried out in the computing centers of the Tier-2 class (Polish Tier-2 is a federation of three centers: Cracow-Poznan-Warsaw) or on the local computer clusters of the research institutes of the Tier-3 class.

In various countries, scientific groups create their own processing models for the needs of the final analyses. In Poland, prior to the start of PLGrid Plus project, such a dedicated computing environment for physical analyses did not exist. The services developed within the *HEP domain grid* are about to fill this gap and multiply the productivity of the research conducted by the Polish HEP groups. At the beginning, the services will be prepared, useful for all the groups, built around the ROOT package (commonly used in HEP) that supports writing programs for the analysis of multi-dimensional variables for millions of cases. Currently, the implementation of its version offering parallel processing of data

Figure 2. "Zeus" in the ACC Cyfronet AGH – currently the most powerful computer in Poland (the TOP500 list, November 2013), being a part of the PL-Grid Infrastructure (*photo by Mark Magryś, ACC Cyfronet AGH*)

in a reduced format is being realized. It will speed up the data analysis cycle by a factor equal to the number of processor cores used in parallel.

In addition to the well-defined services, dedicated to specific applications, we also plan to implement the whole sets of tools, facilitating access from local computing clusters of Tier-3 class to the data stored on the servers of Polish Tier-2 class and to the data in the foreign WLCG centers. It is also planned to support the currently developing computational models of the new experiments (Belle II and SuperB) and the experiments aiming at the study of neutrinos, dark matter, and large atmospheric bunches.

8. The PL-Grid Consortium

The PLGrid Plus project is realized by the PL-Grid Consortium, formed in January 2007. The consortium includes five institutions: Academic Computer Centre Cyfronet AGH in Cracow (Project Coordinator), Interdisciplinary Centre for Mathematical and Computational Modelling in Warsaw, Poznan Supercomputing and Networking Center, Wroclaw Centre for Networking and Supercomputing, Academic Computer Centre in Gdansk.

The Project is co-funded by the European Regional Development Fund as part of the Innovative Economy program.

INNOVATIVE ECONOMY
NATIONAL COHESION STRATEGY

EUROPEAN UNION
EUROPEAN REGIONAL
DEVELOPMENT FUND