Virtual Laboratory as a Remote and Interactive Access to the Scientific Instrumentation Embedded in Grid Environment

M. Okoń, D. Kaliszan, M. Lawenda, D. Stokłosa, T. Rajtar, N. Meyer, M. Stroiński

All authors: Poznań Supercomputing and Networking Center
E-mail: Marcin.Okon@man.poznan.pl

Abstract

The following paper describes the key architecture elements of the Virtual Laboratory Project, which is developed in the Poznań Supercomputing and Networking Center. Project started in 2002, and since then the work was carried under different state-sponsored grants and now is carried under EU FP6 IST projects. The authors will explain briefly the main ideas and key design elements behind this original concept. The current state of the project will be presented, as the project reached its production phase, and plans for the near and far future will be revealed. A special attention will be given to the concept and implementation of Digital Science Library, as it plays an important role in the scope of the whole system. Although the various aspects of this system were presented on workshops and conferences, this paper brings together all the pieces, and presents the overall look on the entire scope of this Virtual Laboratory System.

1. Introduction

The Virtual Laboratory research project [1] has been developed in Poznań Supercomputing and Networking Center since the beginning of the year 2002. VLab was a part of the "High Performance Computations and Visualization for Virtual Laboratory Purposes with the Usage of SGI Cluster" project which was funded by the State Committee for Scientific Research. Furthermore, a parallel research grant has been launched by the Ministry of Scientific Research and Information Technology, more oriented towards the research approach.

The main goal of the VLab project was to design and develop an universal system which should provide a remote access to the wide variety of different laboratory equipment. Additionally, the concept of the VLab tasks, scientific instruments and computational tasks and resources in the Grid environment was generalized. The idea of Dynamic Measurement Scenarios [2][3] was introduced, and the users are capable of creating a dynamic workflows describing their research process. On those dynamic scenarios (or workflow diagrams) experiments with remote access to laboratory equipment are treated as any other "regular" tasks. User creates the scenario using the graphical tool, by placing appropriate boxes representing tasks and connecting them with arrows, defining the data flow between different experiments on scientific instruments and "regular" computational tasks, distributed across the Grid.

A lot of effort was also put on the aspect of sharing and reusing the experiments results. The Digital Science Library [4] – a distributed database system, was introduced, allowing users to store their large data files with experiments results, pictures, papers and any other kinds of material. The data can be described using the flexible meta-data system. Each element from the database can have a custom defined meta-data information, used for categorizing and effective search. The owners of the information stored in the Digital Science Library have full control over access rights to their material.

The pilot installation of this system was done for the domain of the NMR spectroscopy, with cooperation with Institute of Bioorganic Chemistry of Polish Academy of Science from Poznań, Poland [5]. The system was installed and configured for two NMR spectrometers, allowing its users to access Varian Unity 300MHz spectrometer and more advanced Bruker Avance 600MHz. After its successful deployment, the work has started on the remote access to the 32m radio telescope located in Piwnice, Poland. This work is currently under progress, and it is done with close cooperation with Radio Astronomy Department of the Nicolaus Copernicus University in Toruń, Poland.[6] Furthermore, the work has just

started to include the Freeze Atmospheric Dryer – a custom built device from the Faculty of Process and Environmental Engineering of the Technical University of Łódź, Poland. [7]

Furthermore, large parts of the Virtual Laboratory system will be used and modified for the purpose of EU FP6 project EXPReS, which was launched in March 2006. The VLab parts an important role in the task of creating the distributed correlator for the purpose of radio telescope VLBI observations. The work is being done with close cooperation with JIVE. [8]

There is also another European project, RINGrid, in which the Virtual Laboratory expertise is used for trend analysis and recommendations for designing next-generation Remote Instrumentation Services. This project covers also the topic of the synergy between Remote Instrumentation and next-generation high-speed communications networks and grid infrastructures.

2. General Virtual Laboratory Architecture

The Virtual Laboratory system has a modular architecture. The modules can be grouped into three main layers. The general diagram is presented on the figure 1 below[9]:
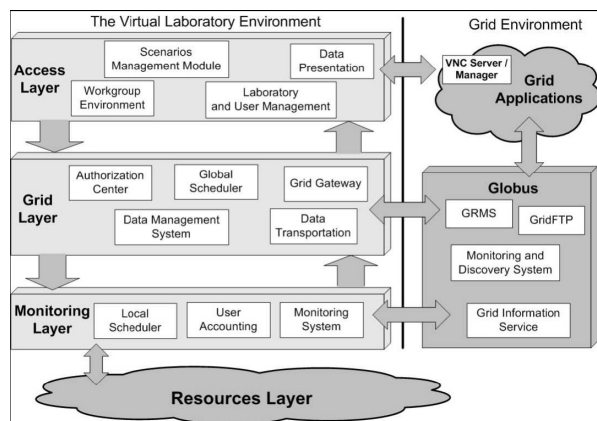

Fig. 1 The Virtual Laboratory System Architecture

An Access Layer is the top layer of this structure. Basically, it contains all modules and components responsible for a VLab user access and graphical interface to the system (including a web portal), and data input interface. Below there is a Grid Layer, which communicates with external grid environment. It is responsible for user authorization and authentication, data management, general task scheduling. Modules from this layer also handle the transfer of the computational tasks to the Globus [10]system, and gather feedback data and the computed results. The Monitoring Layer consists of lower-level modules, such as hardware dedicated schedulers, system monitoring, gathering accounting data etc. All the psychical scientific devices are located in the Resources Layer, as well as modules responsible for their direct control. On the Grid Environment side the most important element is the Globus system, with all its components (GRMS [11], GridFTP etc). Globus also allows to execute computational tasks (batch and interactive) on a wide variety of grid applications.

3. Dynamic Measurement Scenarios

The term "Dynamic Measurement Scenario" [2][3] describes a Virtual Laboratory workflow, containing the set of computational tasks and experiments performed with the usage of remotely accessed scientific instruments. Special XML language was introduced in order to be able to describe this process, all necessary dependencies and other parameters, related to the remote access issues. This allows the Scenario Management Module (workflow manager) to effectively control the execution of the entire scientific process. It is "dynamic" - this means that on the various stages the execution path is determined considering the set of condition parameters, which are defined together with the definition of tasks. One of the most important role of measurements scenarios is automating the data flow between different steps (the data from the instrument are transferred transparently to the computational server for the post processing). Special application was created to facilitate the creation of Dynamic Measurement Scenarios. It is called the Scenario Submission Application. It is the integral part of the VLab system, communicating directly with VLab Scenario Management, Data Management System and other VLab subsystems. The sample screenshot from this application is presented on the figure 2 below:
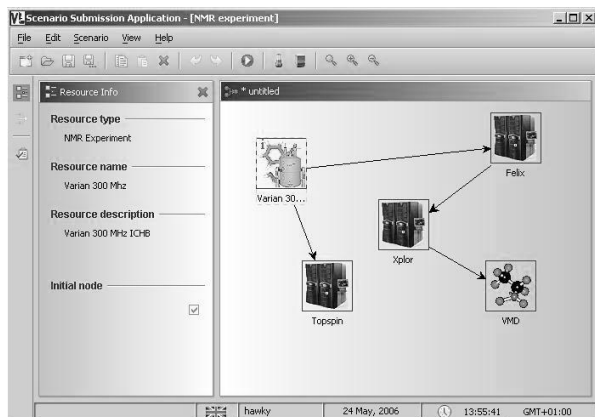
Fig.2 The Scenario Submission Application

The measurement scenario shown on the above picture is from the domain of the NMR Spectroscopy.

4. Experiments and Interactive Tasks

As it was mentioned before, the Virtual Laboratory System distinguishes between two major types of tasks. The first and most important are the experiments using remote access to scientific instruments. The system architecture is universal for different types of instruments. The remote access is realized by presenting the user with the GUI of a instrument control application. The VLab is a multi-user system, and special customized schedulers were implemented in order to effectively manage the submitted tasks.

The schema of the implemented solution is presented on the figure below:
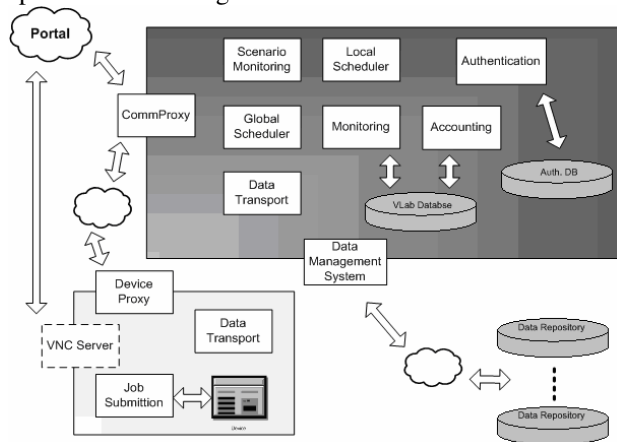

Fig. 3. Remote access to the scientific instruments

The second type of tasks are the interactive computational tasks. The main difference is that those tasks do not require the access to the remote instruments. But what they share in common with

experiments, is the fact that they also require the GUI interface to be presented to the user, in order to perform the computations. This problem did not allow to apply a standard grid-based solution for task submission and execution, but forced the authors to develop a custom solution, which would deal with the problem of remote access to the computational programs GUI. As it was mentioned in paragraph 2, the GRMS system was used, and modified for this purpose. The mechanism for scheduling interactive computational tasks is presented on the figure 4 [12]:
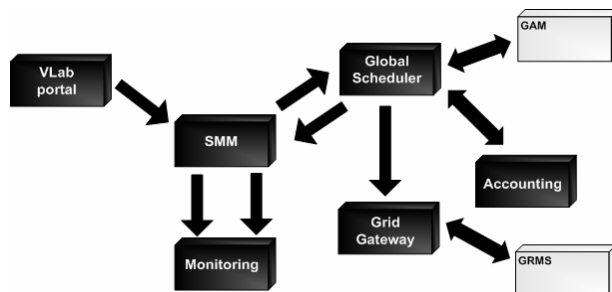

Fig. 4 Task submission for interactive computational tasks

5. Digital Science Library

The Digital Science Library (DSL) [4] is a system that was built on top of the Data Management System (DMS) [14] which was implemented for the project PROGRESS [13]. DMS main functionality - to store data in the distributed fashion over the grid environment – was supplemented with the specific functionality designed for the purpose of Virtual Laboratory.

The main role of the DSL in the VLab system is to collect the data from every step of the measurement scenario, and to transfer the data to the next-stage node on which the computational task was scheduled. This process is controlled by the Scenario Management Module, and it is transparent to the user. Additionally, all the data is stored in the repository for an easy access. The flexible meta-data system is available, and the users have the possibility to define their meta-schemas for their files. This allows for a very accurate search and data categorization. The DSL can be customized and easily adopted to hosting the data from almost any scientific domain.

The sample screen shot of the DSL interface for the NMR library is presented on the figure 5.
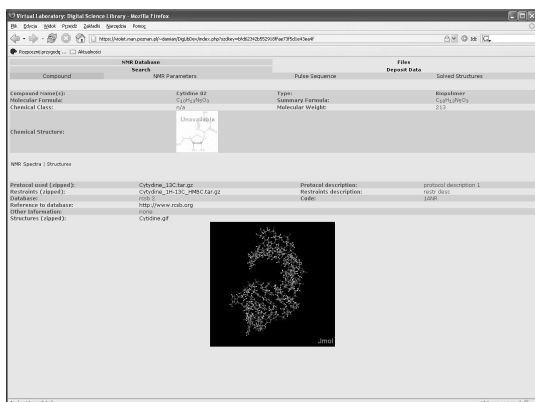
Fig.5 The interface of Digital Science Library for Nuclear Magnetic Resonance

6. Current developments and plans for the future

As it was mentioned in the paragraph 1, the work is in progress in order to launch the Virtual Laboratory of radio astronomy. The plans include the remote access to the 32-meter radio telescope located in Piwnice, Poland, owned by Radio Astronomy Department of the Nicolaus Copernicus University in Toruń.

Also from the radio astronomy domain, some elements of the Virtual Laboratory will be used in European project EXPReS, with the purpose of building the distributed , grid-based data correlator for the VLBI observation.

The short-term plans also include the custom built device, Freeze Atmospheric Dryer, to be included into the Virtual Laboratory environment. This work is being done with the cooperation with the Faculty of Process and Environmental Engineering of the Technical University of Łódź, Poland.

There is also another recently accepted European FP6 project in which the expertise from the four years of work on Virtual Laboratory will be utilized. Project name is "Remote Instrumentation in the Next-Generation Grids" (RINGrid). The scope of RINGrid project is to analyze the current trends and present recommendations for design specification of the next-generation Remote Instrumentation Services, or future Virtual Laboratories. This project will identify new possibilities for remote laboratories and remote instrumentation from the synergy between Remote Instrumentation and next-generation high-speed communication networks and grid infrastructures.

7. Summary

Authors are aware that most of the topics discussed in this paper are presented briefly, and some are deliberately omitted (such as security aspects). This is partly because of the limited size of this paper, and partly because this work was ment to give just an overall impression on the size and scope of the problems and issues which needed to be address during the Virtual Laboratory design and development.

8. References

[1] Virtual Laboratory Website http://vlab.psnc.pl
[2] Lawenda, M., Meyer, N., Rajtar, T., Okon, M.et al.: Workflow with Dynamic Measurement Scenarios in the Virtual Laboratory. 6th CARNet Users Conference, ISBN 953-6802-04-X, Zagreb, Croatia, 2004
[3] Lawenda, M., Meyer, N., Rajtar, T., Okon,et al.: Dynamic Measurement Scenarios in the Virtual Laboratory system. 5th IEEE/ACM International Workshop on Grid Computing, IEEE Computer Society Order Number P2256, ISBN 0-7695-2256-4, ISSN 1550-5510, pp. 355-359, Pittsburgh, USA, 2004
[4] Rek, P., Kopec, M., Gdaniec, Z., Popenda, L., et al.: Digital Science Library for Nuclear Magnetic Resonance Spectroscopy. The 4th Cracow Grid Workshop, ISBN 83-915141-4-5, pp. 404-411, Cracow, Poland, December 12-15, 2004
[5] Institute of Bioorganic Chemistry PAS, Poznań: http://ibch.poznan.pl
[6] Radio Astronomy Departament NCU Toruń: http://astro.uni.torun.pl
[7] Faculty of Process and Environmental Engineering of the Technical University of Łódź http://wipos.p.lodz.pl/index.php?lang=en
[8] JIVE http://www.jive.nl
[9] Lawenda, M., Meyer, N., Rajtar, T., Okon, M., et.al: General Conception of the Virtual Laboratory. International Conference on Computational Science 2004, LNCS 3038, pp. 1013-1016, Cracow, Poland, June 6-9, 2004
[10] Globus website http://www.globus.org
[11] GRMS - WP9 Resource Management - GridLab Project http://www.gridlab.org/WorkPackages/wp-9/
[12] M. Okoń, M. Lawenda, N. Meyer, et al.: Scheduling Interactive Tasks in the Grid-based Systems TERENA Networking Conference 2005, Poznan, 2005.
[13] PROGRESS - Polish Research on Grid Environment for SUN Servers: http://progress.psnc.pl/English/index.htm

[14] Grzybowski P., Mazurek C., Spychała P., Wolski M.: Data Management System for grid and portal services. Submitted to Grid Computing: Infrastructure and Applications special issue of The International Journal of High Performance Computing Applications (IJHPCA), Cardiff University, UK, http://progress.psnc.pl/English/DMS.pdf