Development of an Information Platform for New Grid Users in the Biomedical Field

Daniela SKROWNY [a,1], Frank DICKMANN [a], Benjamin LÖHNHARDT [a], Tobias A. KNOCH [b,c] and Ulrich SAX [d]
[a] *Department of Medical Informatics, University of Göttingen, Robert-Koch-Straße 40, 37075 Göttingen, Germany.*
[b] *Biophysical Genomics, Dept. Cell Biology & Genetics, Erasmus MC, Dr. Molewaterplein 50, 3015 GE Rotterdam, The Netherlands.*
[c] *Biophysical Genomics, Genome Organization & Function, BioQuant Center / German Cancer Research Center, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany.*
[a,d] *Department of Information Technology, University Medicine Göttingen, Robert-Koch-Straße 40, 37075 Göttingen, Germany*

Abstract. Bringing new users into grids is a top priority for all grid initiatives and one of the most challenging tasks. Especially in life sciences it is essential to have a certain amount of users to establish a critical mass for a sustainable grid and give feedback back to the technological middleware layer. Based on the presumable lack of grid IT knowledge it is notably more arduous to satisfy user demands although here the requirements are especially demanding. Therefore, the development of an information- and learning platform could support the efforts of grid experts to guide new users. By providing a platform about grid technology and their feasibilities for users of the community of biomedicine potential, users could be supported using the high potential of their discipline.
Keywords. grid computing, healthgrid, learning platform, information platform, sustainability, accessibility

Introduction

The usage of grid technology is getting more and more important. With the impact of the Large Hadron Collider of CERN in Geneva using grid, it has even rayed out to communities like the humanities (SUB, TextGrid [1]). The biomedical sciences community, more precisely MediGRID [2] was the first biomedical approach to grid technology within the German Grid Initiative (D-Grid) [3]. However, most biomedical scientists are not familiar with the basics of distributed computing. But since there is a promising future using grid for medical and biomedical research, it is

[1] Corresponding Author: Email: dskrowny@med.uni-goettingen.de.

important to introduce this topic to scientists. Existing institutions, like the recent German projects Services@MediGRID and WissGrid[5, 6], are going to support new users with their decision to use grid technology. In fact, both appendages concentrate on recruiting new grid user, but are not concentrating on a special discipline and their requirements. Most importantly they try to solve these problems centrally without user participation and without local secretariats with local renowned expertise.

But exactly that may be very important. Often requirements vary on a broad range, subject to a respective discipline. Thus, it is essential to give new users exactly the information they need. A trusted information- and learning platform will support the users getting all the important information at one place instead out of a pool of too many contact points, before dedicated local and specific support might be necessary.

1. Materials and methods

Building up an information- and learning platform requires taking notice of the needs and requirements of the scientists. Based on the analysis of existing projects and their challenges, as well as a survey of experts of grid, a concept and the realization of an information- and learning platform was developed.

1.1. Analysis of existing grid projects

After the German D-Grid initiative was established in 2004, one year later the first projects started [7]. The biomedical community, represented by the MediGRID project [8], was one of those first.

Table 1: Overview of the results of the respontents' answers. The numbers given are equivalent to the number of answers to that multiple choice question.

Information	Essential	Neutral	Less important
Describing the architecture of grids	1	4	1
Describing existing grid projects	3	3	1
Giving direct examples	7	-	-
Demonstration of applications	7	-	-
Colorful figures	5	2	-
Giving contact information	4	3	-
Providing a glossary	3	2	2
Illustrate characteristics of the medical field	3	3	-
Complemented information about privacy	5	1	-
Specify existing solutions	2	3	-
Specify possible scenarios	3	2	-
Information about gridifying applications	6	-	-
Providing a short self-test	1	3	-

There is a high potential of grid usage in the biomedical discipline but one of the main challenges is to bring new users into the grid. At the HealthGrid 2009 Conference in Berlin the sustainability of grid projects was one of the key topics. Experts allude that applications have to be implemented in an easy to use way. Another important result was that developers have to concentrate on usability and open access to grid technology. [9] The described challenges are known by the communities as well as the developers. There are some projects that try to overcome the difficulties – all using certain information- and learning approaches. Examples are Open Science Grid [4], GridCafé [10] and WissGrid [6].

Analysis have shown the special requirements of the health grid communities as e.g. the high privacy regulations, potential new users with less knowledge about distributed computing as well as the heterogeneity of the data is essential to support especially these communities. Whilst the existing projects follow all different ways of supporting new grid users, a need for a centralized and tailored to particular needs of the medical sector is still seen. Existing ideas like building up a team of experts to support new users or giving information about the usage of grid technology are of promising potential when combined with the development of an information- and leaning platform.

1.2. Survey with experts

Developing a platform for new grid users it is important to know exactly where problems in health grid projects exist. Thus a survey was undertaken with a selective group of grid experts. All of the experts have several years of experience with grid computing in the area of biomedicine and related topics. In addition, many of the respondents have first-hand experience in supporting new grid users and thus, they can provide their knowledge about the most challenging tasks. The survey was concentrated on the future content of the information- and learning platform. (see Table 1). The different information where chosen on basis of the challenges of existing projects. The results showed that the experts' positions corresponded with these challenges. Giving direct examples on how to use grid computing and showing demonstrations of applications were considered as the most important information for the information- and learning platform. In addition, it is seen as essential to provide details about the process of gridification (adapting applications for the grid usage) as well as illustrate the content in a descriptive way. Most of the experts gave supplemental advises about further content e.g. getting a grid certificate, becoming a member in a virtual organization or getting in contact with the D-Grid User Support.

2. Results

Based on the investigation and survey a concept for the platform was designed. Adapted on the target area, the platform has to match the following characteristics: First it has to be quickly and easily accessible to a large number of users. It should also be possible to use the platform as a total layperson as well as a sophisticated beginner. One of the most important goals is an easy operation and a straightforward access. It should be interoperable with the users' soft- and hardware and free of charge. A fast moving technology like grid computing requires an easy way to administrate the platform to keep it up to date. All in all, the final result has to be user oriented and an easy to operate solution.

Based on these demands, the development of a XHTML Website was chosen out of a pool of other possible solutions.

2.1. Development of a platform

After the decision for a XHTML website a concept for the design and content delivery was made. The design should be very clear and easy to work with. It is geared to common designs, that nowadays, can be found widely spread throughout internet.

The content presentation should give users an easy entry into the topic, it should not bore a sophisticated beginner and there should always be a possibility to get further information, in case a user gets stuck.

Therefore, the content is divided into five major parts: introduction, information, learning, contact and links (see Figure 1). The content is based on investigations on existing projects, existing solutions, and, of course, the survey. The introduction is designed to give first time users an overview about the intention and basic principles. The two major parts, information and learning, are the most important ones. They offer users the possibility to get necessary information and guidance without waiting for reply by an advisor. Especially for new grid users the presentation of existing grid solution in the field of medicine is very helpful. The learning part is destined for all potential new users that already have an idea about what to do with grid technology and now want to get information about the usage of grid in detail. The last two parts are administrative information: The contact part offers users the ability to get in touch with the information provider. Links is a collection of references of useful other website to keep the platform clean of information that are provided somewhere else in an adequate manner.

3. Discussion

Whereas the first years of grid computing were characterized by building up the infrastructure and middleware of grids, investigations have shown that research in the

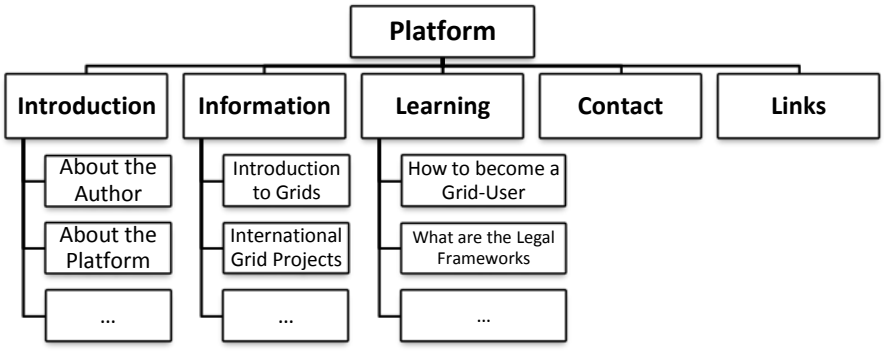

Figure 1: Examples of the Content of the platform

field of grids nowadays concentrates on the development of easy to use applications as well as a sustainable infrastructure including business models. The sponsorship of the Germany Federal Ministry of Education and Research (BMBF)

will focus on new grid applications and improvements that are based on the existing infrastructure. But to prove the benefits of grid technology, it is essential to produce sustainable and financially independent grid projects. The key factor to this challenge is to bring new users to grids. Thus, it is necessary not only to gain in the number of available experts but also to reach information technology laymen.

An information- and learning platform for new grid users of the biomedical field supports and follows the movement of improving the grid access and lowering the threshold of using this technology. Based on the investigations on existing grid projects as well as a survey on experts of the biomedical field it has been shown that there is a need of such a learning platform and that it could help to enhance the acceptability of grids. Althought the number of questionnaires was rather low, the return rate of 100% within the survey on experts illustrates the demand for a solution for the challenge of obtaining new users. The further work must now concentrate on the formalization of the administration procedures which will be addressed in the outlook. Not only the experts are interested in the concept but also existing projects like Services@MediGrid and WissGrid. Therefore, projects follow slightly different objects. One is to lower the entry level of getting a grid user like easing the security procedures the other is to support the users to get around existing obstacles. The information- and learning platform concentrates mainly on the second idea but is also interested to work together with the experts that are going to reduce the entry level in order to exchange experiences and requirements of the users. The platform offers the ability to recommend additional topics and manuals necessary. These requirements will be forwarded to developers to get professional feedback on their applications and to support the development of user-friendly applications. The most important result is that the users want to have a centralized platform rather than contacting many different people. Their desire is to have a kind of local secretary or contact point, like the information- and learning platform, to find out exactly who the right contact person or expert is.

4. Outlook

As grid computing is a fast moving research field, it is important to offer up to date information. Therefore, ongoing maintenance and support of the platform is necessary. Furthermore, it is essential to establish a communication concept as well as a concept for updates. The communication concept can be implemented in cooperation with the WissGrid concept and their team of advisors, especially with the advisors that are responsible for the medical grid community. The information- and learning platform could support the team of advisors with a centralized drop-in center. Because the platform offers already information about grids and becoming a grid user it could give advisors the ability to point to it and so save time to help new users with individual problems. In a possible cooperation with WissGrid the platform could offer information about the team of advisors as well as contact information.

A content lifecycle concept is necessary to keep information- and learning subjects always on a current status. Together with WissGrid synergies can achieved in this context. In order to be reachable by a large number of interested parties a good dissemination is necessary. This could be done by integrating the platform into an existing support platform and by involving grid users into further development. The support team could inform new users about the platform as an access point of getting

introductive information. Thus the advisors can concentrate on the non-standard and struggling requests by the users.

Next steps will concentrate on an intensive scientific and information exchange with WissGrid to optimize the concept and adapt it for future work.

Acknowledgements

This publication was supported by the alliance project Services@MediGRID funded by the German Federal Ministry of Education and Research (BMBF), FKZ 01IG07015A-G and the alliance project WissGrid funded by the German Federal Ministry of Education and Research (BMBF), FKZ 01IG09005A-L.

References

[1] P. Gietz, A. Aschenbrenner, S. Büdenbender, F. Jannidis, M. Küster, C. Ludwig, W. Pempe, T. Vitt, W. Wegstein, and A. Zielinski, TextGrid and eHumanities. *E-SCIENCE* **6**, 133–141.

[2] D. Krefting, J. Bart, K. Beronov, O. Dzhimova, J. Falkner, M. Hartung, A. Hoheisel, T. Knoch, T. Lingner, and Y. Mohammed, MediGRID: Towards a user friendly secured grid infrastructure. *Future Generation Computer Systems* **25** (2009), 326-336.

[3] H. Neuroth, M. Kerzel, and W. Gentzsch, *German Grid Initiative D-Grid*. Universitätsverlag Göttingen, 2007.

[4] A. Bejan, M. Wilde, and B. Clifford, Open Science Grid. Education, Outreach & Training. in: *Working Towards Global Shared Cyber-Infrastructure for Science*, Open Science Grid - University of Chicago, Computation Institute and Argonne National Lab, Chicago, 2007.

[5] D-Grid Projektantrag, Grid für die Wissenschaft -WissGrid-. 2009.

[6] WissGrid, WissGrid - Grid for Science. Access date: 01.03.2010, URL: http://www.wissgrid.de/index_en.html.

[7] H. Hegering, D-Grid: Schritte zu einer nationalen e-Science-Initiative. *E-Science and Grid–Ad-hoc-Netze–Medienintegration* **18** (2004), 285-292.

[8] S. Kottha, K. Peter, T. Steinke, J. Bart, J. Falkner, A. Weisbecker, F. Viezens, Y. Mohammed, U. Sax, and A. Hoheisel, Medical image processing in MediGRID. 2007.

[9] TMF, Grid-Experten mahnen: Anwendungen sorgfältig wählen und benutzerfreundlich implementieren. 2009, URL: http://www.tmf-ev.de/News/tabid/108/ArticleType/ArticleView/ArticleID/518/PageID/448/Default.aspx.

[10] F. Grey, M. Heikkurinen, R. Mondardini, and R. Prabhu, Grid Cafe: A Palace for Everybody to Learn About Grid. Access date: 01.03.2010, URL: http://www.gridcafe.org/.