



COLLABORATION STRUCTURE AND FUNCTIONALITY FOR SCIENTIFIC/ACADEMIC COMMUNITIES IN BALTIC STATES AND BELARUS

Document Filename:	BGII-DNA3.4-VU-NA3Collaboration.doc
Activity:	NA3
Partner(s):	EENET, KBFI, IMCSUL, VU, RTU, ITPA, BNTU, UIIP, VGTU
Lead Partner:	VU
Document classification:	PUBLIC

Abstract: The deliverable DNA3.4 “Collaboration structure and functionality for scientific/academic communities in Baltic States and Belarus” presents the concepts, measures, variables, attributes for the identification of research applications’ collaboration and it’s functionality, preferably for academic communities of the Baltic States and Belarus. The content of this deliverable fulfils the milestone MNA3.2 “Collaboration structure of applications”.



**Document review and moderation**

	Name	Partner	Date	Signature
Released for moderation to	Ake Edlund	KTH	30/04/2009	
Approved for delivery by	Project Management Board	All partners	29/04/2009	

Document Log

Version	Date	Summary of changes	Author
0.1	20/04/2009	Draft version	Algimantas Juozapavicius
1	22/04/2009	Draft version	Algimantas Juozapavicius, Linas Butenas
1.1	28/04/2009	Draft version	Algimantas Juozapavicius, Hardi Teder, Kalle Keskrand, Ihar Miklashevich, Sergey Feranchuk



CONTENTS

1	LIST OF ACRONYMS	4
2	INTRODUCTION	5
2.1	PURPOSE	5
2.2	OBJECTIVES, MEASURES AND INSTRUMENTS FOR NA3 ACTIVITY.....	6
3	COLLABORATION FRAMEWORK	7
3.1	FOUR PERSPECTIVES OF COLLABORATION.....	8
3.2	REQUIREMENTS FOR RESEARCH SECTOR IDENTIFICATION	9
3.3	TEMPLATE FOR THE COLLABORATION WITHIN SECTOR	10
4	POTENTIAL FOR COLLABORATION OF RESEARCHERS	13
5	SUMMARY	14



1 LIST OF ACRONYMS

ANSYS	Ansys software
BG-I	BalticGrid, EU FP6 project
BG-II	BalticGrid Second Phase, EU FP7 project, continuation of the BalticGrid-I project
BNTU	Belarus National Technical University
CAE	Computer Aided Engineering
CFD	Computational Fluid Dynamics
CMS	Compact Muon Solenoid
DEM	Discrete Element Method
EGEE	European Grid Enabled E-Science
EM	Engineering & Manufacturing
EENET	Estonian NREN
GRIA	Specific Grid Middleware
HaaS	Hardware-as-a-Service
IMCSUL	Institute of Mathematics and Computer Science, Latvian University
ITER	International Thermonuclear Fusion Reactor Project
ITPA	Institute of Theoretical Physics and Astronomy, Lithuania
KBFI	Estonian Institute of Chemical Physics and Biophysics
LHC	Large Hadron Collider
NEURO	Modelling of neurology processes
NS2	Discrete event simulator
NWCHEM	NWChem software
PAR	Particle Technology
PAYG	Pay-As-You-Go
RTU	Riga Technical University
SaaS	Software-as-a-Service
UIIP	United Institute of Informatics Problems, Belarus
VG TU	Vilnius Gediminas Technical University
VU	Vilnius University



2 INTRODUCTION

2.1 PURPOSE

The purpose of this document (deliverable DNA3.4 “Collaboration structure and functionality for scientific/academic communities in Baltic States and Belarus”) is to describe how the goal of BG-II to become the reference point for high quality grid-related research-oriented applications, the knowledge and software will be achieved. Also BG-II has to be envisaged as a e-Science channel, as well as to join the European Marketplace for Grid technologies and e-Science. This deliverable is exploring principles and strategies that will allow offering of added-value services to both Grid providers and potential users or buyers.

BalticGrid-II, however, is an open and interactive site. It will host Grid-related research-oriented results of other scientists, their groups, substantial projects and from organizations. We will have internal and external sources and introductory information about Grid computing. There will be information and code for software solutions to real-life problems. In fact, anybody can contribute to BalticGrid-II, provided that the contribution is of technical merit and high research relevance. This will be the basic policy governing collaboration, its functionality and BalticGrid-II content.

The content of this deliverable fulfils the milestone MNA3.2 “Collaboration structure of applications”.

The applications in the project BG-II are foreseen in seven research areas, as starting sectors for collaboration:

- High-energy physics
- Experimental and Astrophysical Thermonuclear Fusion
- Material sciences and quantum chemistry
- Framework for Engineering Modelling Tasks
- Operational Modelling of Baltic Sea Ecosystem
- Linguistics
- Bioinformatics and Biomedical Imaging

In chapter 3 a collaboration framework based on scientific, conceptual and technical backgrounds in the Grid sector has been developed. This approach will confirm that each contribution, which is valid for upload, can be included into e-Science infrastructure of the Baltic States and Belarus. In the paragraphs of the chapter 3 the set of perspectives, parameters and attributes, needed for applications’ collaboration identification are presented and discussed. These parameters and attributes form suitable basis for the development of collaboration framework, which is presented also there.

In chapter 4 the potential functionality for application collaboration is described and the guidelines for it are given. This functionality includes research activities of grid users in programs of National academies of science, science funding agencies, in international research programs or other internationally spanned research activities. In addition it include organizational and technical means for the promotion of scientific collaborations, technical/material facilities, various software packages, used in applications mentioned above, and also perspective ones. The functionality will be granted by suitable VOs and their sophisticated management in different Grids, by grid communities in the external high-level scientific events (participation in international scientific conferences, in EGEE User Forums, etc.).



The analytical summaries, arranged in suitable workshops, will give analysis and evaluation of collaboration process efficiency. The results of analysis and measures will be summarized also in joint document, clarifying the role of NGIs in European Grid Initiative, in close cooperation with the project “EGI Design Study”

The summary giving an overview of collaboration activities is placed in chapter 5.

2.2 OBJECTIVES, MEASURES AND INSTRUMENTS FOR NA3 ACTIVITY

As BG-II is to become the reference point for high quality grid-related research-oriented applications, as well as to be envisaged as a e-Science channel, joining the European Marketplace for Grid technologies and e-Science, the research sectors will be major playing units. These units will be defined by help of frameworks, from technical point of view, and they will be focused on the key areas of Grid, including:

- Data management
- License management
- Security: service level agreements
- Portals
- VO management and programming environment

Also BalticGrid-II targets the audience encompassing:

- Key decision makers
- Potential end users
- Researchers investigating new and emerging models

The formation of collaboration as an e-Science paradigm for the Baltic States and Belarus will use such tools as analytical summaries and frameworks. Objectives and measures included into their content is presented in the next chapters.



3 COLLABORATION FRAMEWORK

The BalticGrid-II project has to achieve a goal of becoming the reference point for high quality grid-related research-oriented applications, the knowledge and software. But BalticGrid-II also has to be envisaged as a e-Science channel, as well as to join the European Marketplace for Grid technologies and e-Science, where vendors of Grid solutions can meet buyers from the research and commercial sector; where consultants, suppliers and programmers can exhibit and trade their offerings and where potential buyers of all levels can find the information and products they require. This deliverable is exploring strategies that will allow offering of added-value services to both Grid providers and potential buyers.

BalticGrid-II, however, is an open and interactive site. It will host Grid-related research-oriented results of other projects and from organizations. We will have internal and external sources and introductory information about Grid computing. There will be information and code for software solutions to real-life problems. In fact, anybody can contribute to BalticGrid-II, provided that the contribution is of technical merit and high research relevance. This will be the basic policy governing collaboration and BalticGrid-II content.

As planned, the applications in the project BG-II are foreseen in seven research areas:

- High-energy physics
- Experimental and Astrophysical Thermonuclear Fusion
- Material sciences and quantum chemistry
- Framework for Engineering Modelling Tasks
- Operational Modelling of Baltic Sea Ecosystem
- Linguistics
- Bioinformatics and Biomedical Imaging

In order to ensure that all the content meets these criteria, we have to develop a collaboration framework, especially based on technical backgrounds in the Grid sector. This approach will confirm that each contribution, which is valid for upload, can be included into e-Science infrastructure of the Baltic States and Belarus.

By developing this framework, from technical point of view, we have to focus on the key areas of Grid, including:

- Data management
- License management
- Security: service level agreements
- Portals
- VO management and programming environment

With it's focus on research value and practical technical solutions, BalticGrid-II target audience encompasses:

- Key decision makers in both research and technical roles who will oversee the implementation of Grid in their businesses or who are looking to further develop their existing Grid capabilities
- Potential end users attracted by the benefits that Grid technology can bring to their work such as reduced time to market, new service offerings, reduced costs, improved quality and greater flexibility



- Researchers investigating new and emerging models that Grid technology brings, such as the software-as-a-service (SaaS) and hardware-as-a-service (Haas).

3.1 FOUR PERSPECTIVES OF COLLABORATION

Four perspectives of the successful BalticGrid-II collaboration framework:

Research sectors. They cover a set of representative scientific and research sectors that address concrete issues and in which the main actors of the research sector are represented. Instead of dealing with many specific sectors, we'll group them in aggregated verticals based on similar behaviour and/or common characteristics

Grid middleware. The scientific applications are using different Grid foundation middlewares, evaluating their suitability for solving specific real-world problems. Middlewares can be classified based on the ones that have been adopted for specific solutions (gLite, Globus Toolkit, Unicore, GRIA, etc.).

Business models. The applications above represent a number of different industrial sectors and their players and therefore their business models differ depending on the characteristics of the sector concerned. Applications have been categorized using a framework suitable for current and future models based on three distinct criteria that take into account their value propositions, their technological and economically associated incentives, and the future of Grid services:

- *Grid applications/experiments with a clear performance associated benefit.* This group comprises those cases that primarily aim at addressing problems or limitations (e.g., additional CPU power, huge amount of data storage /memory, and access to heterogeneous, geographically distributed data resources)
- *Grid applications/experiments with a highly collaborative benefit.* This comprises the applications in which the resulting benefit from Grid adoption comes from sharing data, power and resources utilised for a common scope.
- *Grid applications/experiments exploiting new software paradigms.* This group includes those scenarios exploiting new software paradigms such as the service oriented architectures or component-based development along with new models for providing services such as PAYG (Pay-As-You-Go) or SaaS (Software as a Service).

Technical solutions. The grid projects have defined a set of technical areas, named clusters, that are key for the implementation of the solutions based on Grid technology:

- *General Security:* Trust and security are fundamental for Grid technology because a perceived lack of security is a substantial obstacle for Grid computing to be used outside of the boundaries of a single enterprise in a networked economy or research
- *Licence Management:* Licence management is essential for adoption of business models based on PAYG but currently its support in Grid environments is sadly lacking.
- *Data Management* is concerned with the storage, access, translation and integration of data.
- *Virtual Organisation Management* deals with the secure federation of autonomous administrative domains and the adoption of semantic models and languages to improve the processing of resources and services in a virtual organisation.
- *Portals* enhance the business aspect of the Grid and provide a less complicated working environment for the end user. They have become a key enabler for helping users move more deeply into Grid-based applications where the technology can demonstrate its superior utility.



- *Service Level Agreements* specify the services, responsibilities and guarantee in the provision of the Grid service.

3.2 REQUIREMENTS FOR RESEARCH SECTOR IDENTIFICATION

The requirements of the research sector are exposed by specific topics, listed below.

About the sector

The advanced sector (which corresponds to specific research area) has to comprise a range of research applications, representing group of scientists, distributed geographically. The needs across this broad range have to be similar:

- Groups of researchers/companies must find ways to deliver superior research schemes as quickly as possible in order to get products and results to publicity and market faster than the competition whilst keeping production costs low
- Computational power is essential for the timely completion of these products.

Preference of Grid usage

In dynamic and competitive sector, getting to research quickly with the right product is key. Companies and groups of scientists are under extreme pressure to design, simulate and test in tight time lines that signify massive requirements for computing capacity. It's also critical to reduce costly phases of design and coding. Traditionally, some institutions have met these increasing demands for power through large, continual investments in hardware. This has resulted in data centers that are underutilized the majority of the time but struggle with peak loads.

Research relationships are becoming increasingly complex, with a trend towards modular sourcing. This model requires extremely close relationships between the manufacturer, the supplier providing the subsection, and the lower-tier suppliers. The relationships grow ever more complex, with increased interaction between participants of different tiers that must have access to complex analyses involving massive data sets and e-business tools such as procurement and inventory to automate flows and speed up collaborative processes.

Grid can accelerate product development. It improves computing resource availability and maximizes the utilisation of IT assets and expensive software application licenses. It also improves workload throughput for more simulations, better testing, and faster model development cycles leading to superior products offering new and improved solutions. Institutions can take advantage of grid service providers to get flexible computing capacity on demand that adapts to their requirements.

Moreover, Grids can be used to provide the secure data storage required by industrial partners in such contexts.

Grid impact and perspectives

The advantages of grid computing give a good solution for applications and requirements in the sector. However, the involved parties may use or may not use a collaborative manner. Additionally, they do not have the required Grid infrastructure, and offerings in the form of Software as a Service (SaaS) are necessary. Existing application providers with centralized solutions have to have the



opportunity to provide, in addition to these applications, support for the computing and data-intensive part of the work with an SaaS. Thus application providers get the opportunity to differentiate from competitors.

Specialized Grid service providers that are able to provide specific solutions for fast and reliable set-up of software, may differ from those, which only provide basic Grid resources, and therefore the approach must be to differentiate by adding higher value services to basic Grid infrastructure, such as security, in order to gain an advantage in the SaaS market as a foundation for a sustainable business model in the future.

After initial offerings for Grid computing resources are established in an SaaS manner, a consolidation and commoditization of the services might take place. In this case, coverage of niche markets or providing differentiating services will provide a better position in the market compared to competitors offering generic Grid services.

3.3 TEMPLATE FOR THE COLLABORATION WITHIN RESEARCH SECTOR

This section presents justification of template, suggested to form the sector for collaborating applications. It gives an example of research application from the sector of engineering and advanced manufacturing. Such sector has important potential to the scientists and specialists of the Baltic states (Lithuania, Latvia) and Belarus. To describe a sector and applications within it, the template for CFD is suggested. Main parts of this template are described below.

Engineering modelling (EM) applications include modelling and simulation of problems from fundamental sciences and engineering technology mainly, but not limited to. The grid will give a possibility for scientists to join efforts and make the software more available by creating software infrastructure for engineering and using it on demand.

Parts of the template are presented below.

Situation

Computational Fluid Dynamics (CFD) is currently the fastest growing segment in Computer Aided Engineering (CAE) simulation. The adoption of the technology is continually increasing and users have a latent need for access to scalable, cost-effective CFD.

Grid-enabled technologies can fulfill the latent need by not only increasing the cost effectiveness but accelerating the uptake and application of more advanced CFD.

Grid-enabled solution for performing CFD design gives a new approach for solving CFD design optimization problems and new ones have been achieved for real industrial applications.

The partners involved in the sector include: (list and description of partners research tasks)

Challenges

CFD is an example of an application requiring great computational power. Design problems often have very complex geometry and physics, which require the use of high resolution computational models and powerful parallel computing. As a result, hardware costs become high to achieve these results. In addition, designers increasingly rely on automated methods to reduce the product development time and stay competitive in the market. These automatic optimization processes also



require many calculations to reach an optimum design.

CFD engineers in small to large labs in various sectors often express the wish to carry out more sophisticated, larger computational models, in a quicker timeframe to impact their businesses. Most are delayed or prevented from doing so by the challenging expense of the hardware but most of all the software. Computer clusters provide an attractive price/performance compromise for high-performance computing. The effort of machine setup and maintenance can however be a deterrent for most SME and occasional users.

The benefits for companies wishing to perform such simulations include:

- Availability of more computer processors for time-critical calculations
- On-demand payment for hardware and software for easier budgeting
- Ability to perform calculations with larger computational meshes

Solutions

Grid offers itself as a unique opportunity where computing power, state-of-the-art software and human expertise can be found on demand. Grid computing is an alternative to using in-house hardware resources for performing CFD calculations. The case for Grid computing in CFD simulation is clear:

- it not only increases the cost effectiveness but accelerates the application of more complex CFD, which is required in many cases to capture the science influencing product development.

In the case study carried out, the results obtained running the OpenFOAM CFD toolbox in a remote Grid environment.

A first investigation has shown that the overall times per calculation in the Grid are comparable to execution in a local server with similar CPU power. Some differences were attributed to data transfer, ADSL internet connection speeds and limited CPU power on the remote cluster. Recommendations were given for performing many concurrent simulations on a large state-of-the-art remote cluster.

Research impact

The research impact aims to evaluate the efficiency of the use of Grid in the sector, id est what kind of qualitative improvement is given to application. To show such impact, the following options have to be evaluated:

- Improved productivity
- Increased flexibility
- Reduced cost
- Improved quality

Perspectives

This task will enable the researchers to enter a new research segment where computing on-demand (demand for computing resources needed by application efficiently to run) and is nearly unknown. The showcases demonstrate the potential of Grid technologies for relevant applications in the sector. It is expected that the results will be of interest not only to scientists, but also to specialists and even industries.



TEMPLATE FOR THE COLLABORATION WITHIN RESEARCH SECTOR

Name of application	
Contact person, institution, phone, e-mail	
The group of researchers using application (description)	
Situation: research task, it's status, computing procedures and services used, achievements, grid-enabled technologies needed	
<i>(1000 symbols)</i>	
Challenges: requirement of great computational power, complex problems to be solved, benefits for performance of simulations, including: i) availability of more computer processors for time-critical calculations, ii) on-demand payment for hardware and software for easier budgeting, iii) ability to perform calculations with larger computational meshes	
<i>(1000 symbols)</i>	
Solutions: computing power needed, state-of-the-art software needed, human expertise needed	
<i>(1000 symbols)</i>	
Research impact: improved productivity, increased flexibility, reduced cost, improved quality	
<i>(1000 symbols)</i>	
Perspectives: description of new research segment, grid technologies for relevant applications in the sector, results interested not only to scientists, but also to specialists and even industries	
<i>(1000 symbols)</i>	
Involvement in research programs (description)	
<i>(1000 symbols)</i>	



4 FUNCTIONALITY FOR COLLABORATION OF RESEARCHERS

The potential functionality for application collaboration may be considered as based on following activities:

- The involvement of research group or grid user in programs of the National academies of science and governmental science funding agencies in the region
- The involvement of research group or grid user in international research programs or other internationally spanned research activities
- Organizational and technical means for the promotion of scientific collaborations provided by SIG portals, Migrating Desktop and Gridcom

The following technical/material facilities will be used to foster application collaboration:

- NWChem – powerful and precisely described software for molecular dynamics and quantum chemistry which can be extended by additional molecular dynamics software (AMBER, GROMACS), quantum chemistry software (GAMESS), other computational biology programs and user-developed blocks of common interest for similar scientific groups in material science, physics and bioinformatics.
- the FEMTOOL, OpenFOAM, LSDYNA etc. - computational software for CFD, interests are to use for explosion modelling, explosive welding, dynamical and impuls fracture of building and construction, also in design and optimisation the construction of microjets on the base of activated silicium.
- ANSYS – the powerful software for engineering tasks, used as basis for many applications
- the composite material and the computational mechanics software – has a potential to create the sustainable multidisciplinary research cooperation
- Particle Technology and DEM software – has a potential to generate great scientific and industrial interests (now the scientific group covers partners from Germany, Poland, United Kingdom, Lithuania, this international network is a part of the corresponding excellence center)
- Text corpora of Belarus, of Lithuania, of Latvia, of Estonia – has great potential to foster collaboration of scientists and to combine their efforts

Additional measures will include:

- VOs in different European Grids will be selected and contacted to extend collaboration ties as wide as possible
- The participation of grid communities in the external high-level scientific events (participation in international scientific conferences, in EGEE User Forums, etc.) will be arranged - constantly.

The results and descriptions of these activities will be also included into analytical summaries. The detailed discussion will be arranged in suitable workshops.

The results of analysis and measures will be summarized also in joint document, clarifying the role of NGIs in European Grid Initiative, in close cooperation with the project “EGI Design Study”



5 SUMMARY

The NA3 activity “Application Identification and Collaboration” in BG-II project has an important task to validate and prospect the efficient usage of BG-II infrastructure, to adjust and validate this infrastructure to Baltic States and Belarus research ecosystem, to develop ways for collaboration of applications within BalticGrid-II infrastructure, as well as in the infrastructure of European Grid ecosystem.

The deliverable DNA3.4 “Collaboration structure and functionality for scientific/academic communities in Baltic States and Belarus” describes how to transfer BG-II project to the reference point for high quality grid-related research-oriented applications, to the knowledge and software. BG-II also has to be envisaged as e-Science channel, as well as to join the European Marketplace for Grid technologies and e-Science.

The content of this deliverable partly fulfils the content of milestone MNA3.2 “Collaboration structure of applications”.

In this deliverable the set of perspectives, parameters and attributes, needed for applications’ collaboration identification are presented and discussed. These parameters and attributes form suitable basis for the development of collaboration framework, which is presented also there.

Also the potential functionality for application collaboration is described and structured, this functionality includes research activities of grid users in programs of National academies of science, science funding agencies, in international research programs or other internationally spanned research activities. In addition it include organizational and technical means for the promotion of scientific collaborations, technical/material facilities, various software packages, used in applications mentioned above, and also perspective ones. The functionality will be granted by suitable VOs and their sophisticated management in different Grids, by grid communities in the external high-level scientific events (participation in international scientific conferences, in EGEE User Forums, etc.), by analytical summaries, arranged in suitable workshops.