Participation of MD-Grid JRU Consortium in SEE-GRID-SCI Project

Peter Bogatencov, Veaceslav Sidorenco
RENAM
bogatenc@renam.md

Abstract

New EU-supported SEE-GRID-SCI project stimulates widespread eInfrastructure uptake by new user groups extending over the region, fostering collaboration and providing advanced capabilities to more researchers, with an emphasis on strategic groups in seismology, meteorology and environmental protection. The initiative enlarges the regional eInfrastructure to cater for demands of the communities by increasing the computing and storage resources and involving Moldova together with new partner countries in the region. The project will aim to attract local political support for materializing the eInfrastructure vision and aspires to contribute to the stabilization and development of South-East Europe, thus enabling collaborative high-quality research across target scientific fields.

1. Introduction

Enabling large-scale innovative research to be conducted through collaboration of distributed teams of scientist across the European Research Area (ERA) paves the way towards a long-term vision of a sustainable, transparent, ubiquitous electronic infrastructure (eInfrastructure) open to a wide range of scientific user communities providing the development of Information Society in Europe. In this context, the inclusion of less-developed regions of Europe, which suffer from the digital divide and brain-drain in all fields and especially high-technology, into the wider European Research Area is an aim closely aligned to the general policy of expansion of the European Union.

Advancing the Information Society in areas such as South-East Europe (SEE), strengthening of the local eInfrastructures, activating new user communities and enabling collaborative research across a number of fields, would strongly contribute to closing the existing technological and scientific gap, and thus bridging the digital divide, stimulating research and consequently alleviating the brain drain in the region.

In the past 5 years, a number of targeted initiatives funded by the European Commission via its RTD programmes have contributed to ameliorating the state of eInfrastructures in the region. The SEEREN initiative project, through its two phases, established the SEE segment of the pan-European GÉANT network and successfully connected the research and scientific communities in the region [1].

Currently, the SEE-LIGHT project is working towards establishing a dark-fibre backbone that will interconnect most national Research and Education networks in the region.

The SEE-GRID project, similarly through its two phases, has established a strong human network in the area of scientific computing and has set up a powerful regional Grid infrastructure, and attracted a number of applications from diverse fields from countries throughout the South-East Europe. Furthermore, the SEE-GRID initiative in its current second phase is strongly involved in incubating the nascent National Grid Initiatives of the countries in the region, bringing together under one umbrella, on the national level, all interested parties involved in provisioning of scientific computing infrastructure as well as its use. This is very much in line with the European vision of paving the way towards a long-term sustainable European Grid Initiative through strong support of National Grid Initiatives, and, in this aspect, SEE-GRID is to some extent leading the way by its successful establishment of NGIs and by producing NGI-related recommendations that are propagated to the European level.

Strong liaisons have been established with core European eInfrastructure initiatives like GN2 project / GÉANT2 network, and EGEE [2]. A number of countries in the region have already joined the GÉANT and EGEE communities as peers. The SEE regional development model has been successfully disseminated and used as a role model for developing other regions such as Mediterranean, Latin America, etc. - where the

European Commission has also supported a number of equivalent eInfrastructure projects.

2. Project Objectives

The regional vision of establishing a transparent eInfrastructure open to a wide range of scientific communities is materialising, and a decisive step towards its sustainability will be taken with the SEE-GRID-SCI (SEE-GRID eInfrastructure for regional eScience) project [3]. The project will work across several strategic lines of action.

First and foremost, the project will stimulate the use of the regional eInfrastructure by opening up its use to target scientific communities and specifically encouraging the cross-border user communities i.e. scientific collaboration and use of the eInfrastructure on the level of South East Europe as a whole. In this context, the initiative will aim to have a catalytic and structuring effect on a number of SEE user groups, with a strong focus on the key seismological, meteorological, and environmental communities.

The second line of expansion towards new user communities is the inclusion of the Caucasus region in the project – namely Georgia and Armenia.

Finally, the existing user communities in the region will be strengthened at the national level by specifically supporting the non-EGEE countries through maturing their NGIs for inclusion in the envisaged European Grid Initiative foreseen for 2010 onwards.

SEE-GRID-SCI project will involve three strategic international scientific communities/Virtual Organisations (seismology, meteorology, environmental protection) and thus further stimulate the use and expansion of the existing regional eInfrastructure and its services, and capitalize on the existing human network to further strengthen scientific collaboration and cooperation among participating SEE communities in the area of eInfrastructures. The inclusion of the new scientific communities and the expansion of the infrastructure in terms of both size and geographical spread, together with a set of coordinated actions aimed at strengthening the National Grid Infrastructures in the region, will ensure that at the end of the project each country in the region will be ready to join the long-term, sustainable European Initiatives as a full-fledged peer.

Figure 1 depicts the layered regional eInfrastructure where SEE-GRID-SCI effectively adds the eInfrastructure knowledge layer on top of the existing network and Grid planes (i.e. including integration of sensors and databases/repositories for some user communities), thus enabling a range of new cross-border eScience applications to be deployed over the regional eInfrastructure. This approach effectively creates an integrated eInfrastructure for new target scientific communities.

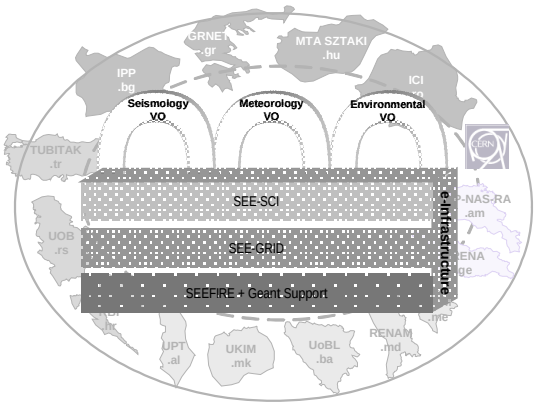

Fig. 1. The SEE-GRID-SCI eInfrastructure and new user communities

3. Project Participants

The SEE-GRID-SCI consortium consists of fifteen beneficiaries, representing twelve National Grid Initiatives from Albania, Armenia, Bosnia and Herzegovina, Bulgaria, Croatia, Former Yugoslav Republic of Macedonia, Georgia, Greece, Hungary, Moldova, Montenegro, Romania, Serbia, and Turkey, plus CERN in an advisory and liaison role with the pan-European grid initiatives. Of the aforementioned SEE NGIs, four are coming from EU member-states (Bulgaria, Greece, Hungary, Romania), two from candidate countries (Croatia and Turkey), five are coming from International Co-operation Participant Countries (ICPC) - Western Balkans (Albania, Bosnia-Herzegovina, Former Yugoslav Republic of Macedonia, Montenegro and Serbia), and 3 from ICPC – Eastern European and Central Asia countries (Armenia, Georgia and Moldova). Armenia, Georgia and Moldova are also part of the European Neighborhood Policy (ENP). The consortium thus represents full coverage of the region with a good balance between European member states and their neighbors.

By using the participation in the EGEE production-level infrastructure as a point of reference to judge a participant's maturity in deploying eInfrastructures, three layers can be identified (as portrayed in the figure below): Bulgaria, Greece, Hungary, and Romania were members of EGEE and will carry on future activities; Croatia, Serbia, and Turkey have joined EGEE-II and will carry on, and Albania, Bosnia-Herzegovina, Former Yugoslav Republic of Macedonia, Moldova, and Montenegro are not yet part of the European and

World-wide eInfrastructures. Armenia and Georgia are new to EC eInfrastructure initiatives. The consortium is complemented with CERN that is also the coordinating organization of EGEE project and will be an indispensable liaison with the Pan-European and Global grid initiatives.

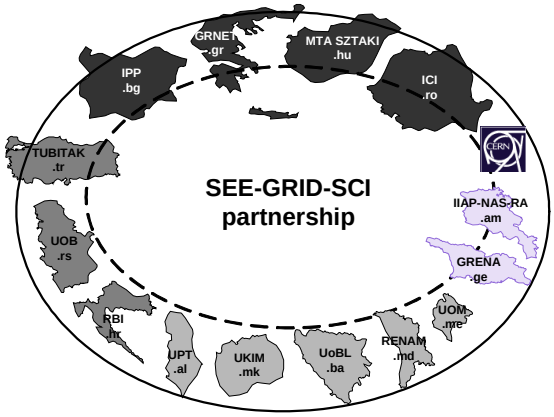

Figure 2. The SEE-GRID-SCI partnership

Moreover to further assist the engagement of wide range of new user communities and to boost development of a nation-wide grid infrastructure, the scheme of joint research unit (JRU) and third-parties will be used in order to expand Grid usage in SEE-wide communities, further strengthen organization participation at national level in each SEE country and contribute to the sustainability plan. Of special importance is the inclusion of crucial international scientific communities from different disciplines as part of a group of institutions directly collaborating with the project as 3rd parties. SEE-GRID-SCI is endorsed and supported by Ministries of participating countries; as well as flagship European Grid infrastructure project EGEE, related regional initiatives such as EUMEDGRID, EUCHINAGRID, EUINDIAGRID and EELA, the BELIEF project; UNESCO, and industrial players IBM, HP, Microsoft and Sun.

Moldova participates in the project as MD-Grid JRU Consortium with RENAM as Contractor and following third parties - JRU members [4]:

- Faculty of Radioelectronics and Telecommunications of Technical University of Moldova (FRT-TUM)
- Institute of Geology and Seismology of Academy of Sciences of Moldova (IGS ASM)
- State Hydrometeorological Service of Moldova (SHMS).

RENAM will be responsible for NGI development and its sustainable operation, extension of the National Grid infrastructure by installation of new clusters and their inclusion into common Grid segment of Moldova and into the SEE-GRID infrastructure. RENAM will assure installation and operation of four Grid sites, support of development and functioning of applications in the fields of seismology, climate modeling, ecology and mathematics. RENAM will be responsible for users and Grid sites engineers training and will take part in dissemination and public relations activities.

FRT-TUM cooperates with a significant number of universities and centres from different countries, participate in SEE-GRID-2 project as third party behind RENAM Association. The grid node installed at the faculty together with its grid infrastructure and specialists will participate in SEE-GRID-SCI project offering resources for weather environment and earthquake monitoring applications development and use support in collaboration with regional neighbours from Romania, Bulgaria and Turkey.

IGS ASM will provide contributions on real-time monitoring of earthquakes and development of regional monitoring capacities, implementation of real-time seismic data processing and real-time data exchange at regional and international level. Despite the territory having high earthquake hazard and risk, its seismic activity remains poorly monitored. Moreover, cross-border data exchange and regional applications for data accumulating and processing, which are essential for good quality monitoring, are very limited. As a consequence, it is constraining ability to cooperate with the international research and engineering community.

SHMS will contribute to the environmental VO by testing and deployment of a pilot application. The service infrastructure and specialists will be involved into MD-Grid JRU scientific research and production grid activities in requirements definition, input data provision and interpretation of results on national, regional and international levels.

4. Applications area

4.1. Seismology VO

Seismology Virtual Organization consists of neither a single application nor a single regional data repository. Seismology VO is an environment that enables Grid infrastructure for all regional scientists. Except development of proposed applications and adapting routine analysis software for the Grid seismology VO, a high level programming interface for seismic files will be offered to researchers for handling huge amounts of data. This high level programming interface can be used as a wrapper for seismic data analysis applications or can be integrated in those

applications. The initial version of the wrapper which will provide transparent access to data has already been developed by Bogazici University Computer Engineering Department. In seismology VO, the iterators which have been developed can be used in order to speed up file access.

Following applications were identified at the proposal stage for the seismology VO, more applications could be added in the course of the project if identified.

Seismic Risk Assessment: This application is a regional collaboration between Middle East Technical University Department of Statistics (Turkey) and Geophysical Institute, Bulgarian Academy of Sciences. The Institute of Geophysics and Seismology (IGS) of the Academy of Sciences of Moldova is also interested in developing this application. Seismic risk assessment is very important for public safety and hazards mitigation. It is also important for the correct determination of earthquake insurance premiums and also for understanding the social and psychological effects of earthquakes. The reliability of the seismic risk assessment depends on complete data sets. The collected seismic data from many different sources should be made available for processing, and important events are then located automatically. This application will provide seismic risk map for all the participating countries to be used in social, psychological studies and determination of insurance premiums.

Massive Digital Seismological Signal Processing with the Wavelet Analysis: Seismological Observatory in FYR of Macedonia proposes another application which will extensively use regional seismology data repository. Several topics of interest can be identified in this context:

- Common characteristics identification for the registration of the same epicentral zone.
- Procedure for automatic location of the earthquake based on just one registration.
- Identification of the key differences for the registrations from the different epicentral zones.
- Characteristics of one geological profile in the vector space of its wavelet representation.
- A representative wavelet model for registration of a given epicentral zone. This could be used for producing artificial seismograms.
- Empirical Green function for the given epicentral zone.

Center of Experimental Seismology of IGS (the national seismology data Center of Moldova) is also ready to provide the seismic data for this regional application and contribute in development of specific parts of the application referring to the some of the above topics of interest.

Numerical Modelling of Mantle Convection: This application, proposed by Geodetic and Geophysical Research Institute of the Hungarian Academy of Sciences Seismological Observatory is a modelling application, different from data analysis-based applications which have been proposed so far. It is aiming to make a quantitative study of the structure and surface manifestation of mantle plumes and to make systematic investigation of the parameters influencing the character of mantle convection in 3D. In previous calculations the institute studied the areal density, the diameter, the temperature anomaly, the topographic and the geoid anomalies of mantle plumes in steady state convection. It is possible to monitor the characteristics of individual plumes even in chaotic convection system, in large provinces, at high resolution, but this needs large computational capacity. SEE-GRID-SCI infrastructure will provide the required computing resources.

Seismology virtual organization involves researchers from FYR of Macedonia, Moldova, Bulgaria, Albania, Hungary, Greece and Turkey. The partnership will be coordinated by TUBITAK. Members of the seismology virtual organization are also members of the ORFEUS ("Observatories and Research Facilities for EUropean Seismology") non-profit organisation that aims at coordinating and promoting digital, broad-band seismology in Europe. ORFEUS foundation also initiated the NERIES ("Network of Research Infrastructures for European Seismology") FP6 project. NERIES project aims to integrate EIDA within a world-wide datagrid and SEE-GRID-SCI project will collaborate with NERIES to share know-how for the benefit of seismology researchers. SEE-GRID-SCI has established contact with NERIES and identified areas of collaboration.

4.2. Meteorology VO

Meteorology virtual organization involves researchers from Greece, Serbia, Montenegro, Croatia, and Bosnia and Herzegovina. The partnership will be coordinated by the National Observatory of Athens (NOA). Members of the meteorology virtual organization will be the South Environment and Weather Agency of Serbia, the Hydrometeorological Institute of Montenegro, the 'Ruđer Bošković' Institute of Zagreb, the department of Geophysics - 'Andrija Mohorovičić' Geophysical Institute of the University of Zagreb, the faculty of Graphical Art of the

University of Zagreb, and the Federal Hydro-Meteorological Institute of Bosnia and Herzegovina.

There are two sets of applications that will be initially deployed on the SEE-GRID-SCI infrastructure:

(a) Regional scale Multi-model, Multi-analysis ensemble forecasting system. For the deployment of this application the following steps will be undertaken:

- The development of the multi-model, multi-analysis ensemble weather forecasting system. This system will comprise the use of four different weather prediction models (multi-model system). Namely the state-of-the-art numerical weather prediction models BOLAM, MM5, NCEP/Eta, and NCEP/WRF-NMM will be ported on the Grid infrastructure. The above models will be run for the same region many times, each initialized with various initial conditions (multi-analysis).
- The development of a procedure based on advanced capabilities of the Grid infrastructure that will coordinate, collect and analyze the outputs from all models for the generation of probabilistic forecasts over the area of central and eastern Mediterranean.

(b) Study of the interaction of airflow with complex terrain. This second application will be focused over Croatia and Bosnia and Herzegovina, using the WRF (Weather Research & Forecasting) model. Both countries have large areas covered with terrain obstacles and it is essential to obtain high-resolution information. The application of the results will be in estimating the effect of improved resolution on the numerical weather prediction quality, and consequently in improving the forecasting skill, and also in the air-pollution dispersion modelling over complex terrain. For the latter purpose, it is intended to further develop the Lagrangian Particle Stochastic model ALPS and to couple it with the WRF model.

The results produced within the meteorology VO related to the Regional Multi-model, Multi-analysis Ensemble Prediction System will allow the meteorological entities participating in the project to assess the probability of a particular weather event to occur and to provide this information to the authorities, the general public, etc, in order to help them to make the necessary decisions based on this probabilistic information. Special focus will be given at the SEE-wide scale where detailed forecasts are necessary for the protection of life and property.

4.3. Environmental VO

Environmental VO bring together scientists and research institutions, working in the domains of Environmental modelling (including Geomagnetism), Environmental security, Environment evolution supervision and impact of climate-related events in the SEE region, with the aim to leverage the resources of the SEE-GRID-SCI infrastructure for developing powerful applications in this field and achieving results with sufficient temporal and spatial resolution. An important development in this direction will be the creation of a Grid-based platform for environment oriented satellite image processing, which makes possible for a large number of users to access the environment information by a set of available Internet services. Environmental VO covers main applications set listed below.

Modelling System for Emergency Response to the Release of Harmful Substances in the Atmosphere (MSERRHSA): The aim of this application is to develop and deploy, on the SEE-GRID-SCI infrastructure, a modelling system for emergency response to the release of harmful substances in the atmosphere, targeted at the SEE and more specifically Balkan region. This system will provide operational response to accidental releases of harmful gases in the atmosphere, whether as a result of terrorist attack or industrial accident. The system must be able to perform highly accurate and reliable risk analysis and assessment for selected “hot spots” and provide the national authorities and the international community with short-term regional scale forecast of the propagation of harmful gases.

Monte Carlo Sensitivity Analysis for Environmental Systems (MCSAES): The aim of the application is to develop an efficient Grid implementation of a Monte Carlo technique for sensitivity studies in the domains of Environmental modelling and Environmental security. The developed application will be applied for studying the damaging effects that can be caused by high pollution levels (especially effects on human health), when the main tool will be the Danish Eulerian Model (DEM). Sensitivity simulations lead to huge computational tasks (systems with up to 4x10^9 equations at every time-step, and the number of time-steps can be more than a million). An important part of the application is the visualization.

Multi-scale atmospheric composition modelling (MSACM): The aim of this application is to use the Grid environment to produce an integrated, multi-scale Balkan region oriented modelling system, able to interface the scales of the problem from emissions on the urban scale to their transport and transformation on the local and regional scales. This system should be able to study the atmospheric pollution transport and transformation processes (accounting also for heterogeneous chemistry and the importance of

aerosols for air quality and climate) from urban to local to regional (Balkan) scales, to track and characterize the main pathways and processes that lead to atmospheric composition formation in different scales, account for the biosphere-atmosphere exchange as a source and receptor of atmospheric chemical species, and to provide high quality scientifically robust assessments of the air quality and its origin. It is envisaged the application to be based on US EPA Models-3 system, which is known to be one of the best modelling tools that continues to be developed intensively by the efforts of a big community of scientists both in the US and Europe.

Regional Modelling of the Geomagnetism: The main aim is to apply Spherical Cap Harmonic Analyses, SCHA, with physical regularization to synthetic series obtained from the Comprehensive model CM4, at 46 European observatory locations and additionally 11 'virtual observatories' chosen to improve the initial data distribution. Different model parameters will be tested (spherical cap angle, maximal spherical cap harmonics order, number of splines, norms). The misfit of the final model will be tested using different criteria: the rms values, the time evolution of the coefficients and the behaviour of the original versus modelled time series at each location. Comprehensive model CM4 will be widely used to calculate both the internal, as well as the external geomagnetic field contributions at different locations. The X (northward), Y (eastward) and Z (vertically downward) components of the main field represent the input dataset for the model. Simultaneous modelling in spatio-temporal domain will be done by using cubic splines. The study of the modelling results will open a way to describe in detail regional geomagnetic main field and its secular variation.

Environment oriented Satellite Data Processing (ESIP): main objectives are (i) to provide a Grid based software platform for satellite image processing; (II) to develop a methodological approach for application development based on this platform; (iii) to experiment this approach on a pilot application. First two activities will be carried out within JRA1. The basic approaches and algorithms have been experimented on MedioGrid project academic and research infrastructure (www.mediogrid.utcluj.ro) and will be extended to a larger production Grid infrastructure by processing real data for SEE regions. The proposed software platform has a service-oriented architecture and provides the basic functionality required to Grid and Web application developers.

The pilot application for the ESIP platform: The aim is a refinement of surface- and vegetation parameters in SEE region based on satellite images. In the frame of this project, construction, usage and comparison of diverse satellite datasets will be performed. High resolution satellite imagery time series can be used for numerous environmental studies (climate-related or air pollution modelling). Using the sophisticated environmental data the change of the vegetation distribution in the Carpathian Basin and its climate-related causes will be investigated. It may also be used to extend the urban climate related research. The dataset may also contribute to the development of air pollution models, which describe the dispersion of tracers or their exchange between the surface and the atmosphere. These transport and deposition/exchange models require detailed input fields about surface and vegetation (for example vegetation types, leaf-area index, albedo). Satellite data could be an appropriate source to create these input datasets for the regional models.

RENAM and State Hydrometeorological Service of Moldova will support pilot application testing and deployment.

5. Conclusion

SEE-GRID-SCI project helps to deploy a sustainable Grid infrastructure, which, coupled with human networking activities and outreach to and engagement of SEE-wide user communities, will be of particular societal, educational, and political importance to the region and Moldova. The process of preparing for EU accession is closely linked to modernization of the candidate country's economy. Such modernization is a key aim of eEurope - which aims at accelerating the development of the information society in Europe and ensuring its availability to everyone - and SEE-GRID-SCI can have a significant impact in the SEE region by helping the participating countries achieve this goal and align their national priorities with the EU recommendations and guidelines.

6. References

[1] South-East European Research and Education Networking initiative. http://www.seeren.org/

[2] Enabling Grids for E-sciencE. http://public.eu-egee.org/

[3] SEE-GRID eInfrastructure for regional eScience. http://www.see-grid-sci.eu/

[4] MD-Grid. National Grid Initiative. http://www.grid.md/