Basic Set of Applications of SEE-GRID-SCI Project

Veaceslav Sidorenco
Technical University of Moldova
Chisinau, Moldova
e-mail: svv@renam.md

Peter Bogatencov
RENAM
Chisinau, Moldova
e-mail: bogatenc@renam.md

ABSTRACT
EU FP7 SEE-GRID-SCI project applications are presented and briefly described. The application belongs to three types of Virtual organizations: Seismology, Meteorology and Environmental. Every application has its specific goals and particularities of grid implementation. The analysis of applications development status in first year of project demonstrates that a number of applications already have some versions deployed or being tested successfully on the SEE-GRID infrastructure.

Keywords
grid computing, grid applications, virtual organizations, meteorology, seismology, environment protection.

1. INTRODUCTION
Main objectives of the SEE-GRID-SCI project consist in engaging international user communities and providing application-specific service extensions, so as to explore application similarities and abstract out common elements with respect to the approaches to actual Grid Computing usage by analyses in detail the applications under deployment [1]. It is necessary to capture their grid-specific requirements in a formal way so as to enable identification of common areas of development for specific application-level services and their implementation. It is important to provide and operate the next-generation eInfrastructure for the SEE region [2]. In SEE-GRID-SCI context this refers to operating the Grid infrastructure and specific end-user services for the benefit of new user communities.

To achieve this, three regional communities are supported, structured in the form of Virtual Organizations (VO):

- Seismology VO has six applications ranging from Seismic Data Service to Earthquake Location Finding, from Numerical Modelling of Mantle Convection to Seismic Risk Assessment.
- Meteorology VO, with two comprehensive applications, follows an innovative approach to weather forecasting that uses a multitude of weather models and bases the final forecast on an ensemble of weather model outputs. The other problem attacked is the reproduction/forecasting of the airflow over complex terrain.
- Environmental (Protection) VO supports six applications focusing on environmental protection/response and environment-oriented satellite image processing.

Each application has made its choice about application specific services offered by the SEE-GRID-SCI that it will use. In the development stage, these services are not critical for the applications, but they will be important for efficient move towards production utilization.

2. SEISMOLOGY VO APPLICATIONS
Seismology VO has two kinds of assets: Seismology data collected from partner countries and applications. Whereas applications are quite important as the assets of the Seismology VO, seismology data is even more important since data assets can help us in attracting more applications and users to the Seismology VO. For this reason, special attention is given for the collection, storage and retrieval mechanisms of data. To provide data upload and access services Seismic Data Server Application Service (SDSAS) is being developed. Seismology VO involves researchers from Albania, Armenia, Bulgaria, Hungary, FYR of Macedonia, Moldova, Serbia and Turkey. The partnership is coordinated by the Boğaziçi University (TR).

As applications, the following are developing:

- Seismic Risk Assessment (SRA)
- Seismic Data Server (SDS): web interface to SDS application service
- Earthquake Location Finding (ELF)
- Fault Plane Solution (FPS)
- Numerical Modeling of Mantle Convection (NMMC3D)
- Massive Digital Seismological Signal Processing with the Wavelet Analysis (MDSSP-WA).

SRA application 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 aim is to develop an application framework to allow us to embed alternative (deterministic, probabilistic etc.) models. SRA application can be grouped into four main categories: Accessing Earthquake Catalogue, Earthquake Source Model, Seismic Hazard Models, Producing Seismic Hazard Maps.

SDS will have a web interface to seismic data archive organized by the SDS Application Service. It will offer only access capabilities and let the users visually display various kinds of information about the seismology waveform data files, earthquakes and stations. The application will be written as a set of executables that generate html files and/or kml files for Google Earth. The inputs to these executables will come from web forms filled by the user.

ELF application is based on HYPO71 program and finds the location of earthquakes by scanning seismic waveform data presented in SAC format. It outputs coordinates of the location of the earthquake. This application is not compute intensive, but it is data intensive. The application can be parallelized by scanning data files in parallel by multiple using worker nodes. A workflow can be generated automatically by a program corresponding to the time intervals in which to look for earthquake.

FPS application computes source parameters (strike, slip, dip) and moment magnitude of an earthquake. It inputs a crust model: layer thicknesses, seismic velocities, densities, q-factor and actual seismic waveform data (in SAC format). It outputs fault parameters of the earthquake. The application can be useful for identifying tectonic structures that are not visible on Earth's surface.

NMMC3D provides model of the thermal convection in the Earth's mantle. The application solves numerically the equations of thermal convection (the conservation laws of mass, momentum and energy) in a Cartesian 3D model-box. The results will help to understand better the dynamics of the Earth. The outer part of the Earth consist of moving, rotating and interacting plates. The motion of these plates suggest a large convective system in the Earth's 2900 thick layer, the mantle. The numerical calculations suggested that the convective cells are formed by sheet-like elongated downwellings (subuction zones) and narrow, cylindrical upwellings (mantle plumes, at the hotspots). The main goal of research is the 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.

MDSSP-WA is based on wavelet theory that has become mature in past years as a new mathematical tool for time series analysis. However, although it seems quite natural to be applied in seismology, it seems that it is in its initial stage. Making the continuous or discrete wavelet transforms and plotting the results in coordinate system, scales versus time, one could notice striking similarity of the wavelet images, between different seismic records, coming from the same source region or noticeable difference for records of earthquakes occurred in different source region. It is assumed that those similar image patterns are due to same underlying geological setting while the differences (usually for smaller scale) are due to different source mechanism and finer geological structures. In the first approximation of geological structure, similarities of the image patterns in domain of large scale are noticeable even for the records from different source regions. With massive processing of earthquake records one could define: Common features of the propagation path for the given seismic source region or to define empirical transfer function of the media (Green function); Calculation of the artificial seismograms; Determine the source region based on a single earthquakes record; Determine the more realistic attenuation curve of the selected feature (parameter), very much needed in seismic hazard and risk analysis; Mapping (coding) of the given earthquake prone region in terms of selected parameters; Seismic source parameters. Apart from above stated goals other possible benefits could be better noise removal from the seismic records, possible better and more accurate different phase identification.

The territory of the Republic of Moldova is influenced by earthquakes of intermediate depth from seismic zone Vrancea, located d on the territory of Romania. The strongest of them are distributed in the depth interval of 80-150 km, and their maximal magnitude (Mmax) according to different estimations achieves 7.5-7.8 on Richter scale. The resulted seismic effects from this zone are influencing a big territory, including Romania, Moldova, Ukraine, Bulgaria and Serbia. The most significant seismic effect is observed in Romania, Moldova and Bulgaria. The maximal seismic intensity in the territory of Moldova is evaluated at the level of 8-9 degrees according to 12-degrees scale. Despite this territory having high earthquake hazard and risk, its seismicity remains poorly monitored. Cross-border data exchanges, which are essential for good quality monitoring, are very limited. As a consequence, it contributes in an inability to connect with the international research and engineering community. Institute of Geology and Seismology of the Academy of Sciences of Moldova and RENAM Association are active participants of Seismology VO SRA and SDS applications development and deployment.

3. Meteorology VO Applications

The results produced within the meteorology VO 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 is given at the SEE-wide scale where detailed forecasts are necessary for the protection of life and property. Meteorology VO involves researchers from Greece, Serbia, Moldova, Montenegro, Croatia, and Bosnia and Herzegovina. The partnership is coordinated by the National Observatory of Athens (NOA).

Advances in numerical weather prediction (NWP) and related applications have been always very closely related with advances in computing sciences as NWP requires numerical calculations that are also parallelizable. The computer resources needed for NWP applications are important both in terms of CPU usage and disk storage. Although many institutions are working/have experience on NWP, they may not have access to the necessary computer resources for operational implementation of such applications, so the porting of any NWP application to the grid is a natural choice. There are two sets of applications that will be deployed on the SEE-GRID-SCI infrastructure:

- REFS: Regional scale multi-model, multi-analysis Ensemble Forecasting System
- WRF-ARW: Weather Research Forecast – Advanced Research WRF

The REFS application consists in 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). With this procedure not only one deterministic forecast but a multitude of forecasts will be produced.

The WRF-ARW application will be focused over Croatia and Bosnia and Herzegovina, using the WRF 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 modeling over complex terrain.

4.Environmental Protection VO Applications

Environmental VO brings together scientists and research institutions, working in the domains of environmental modeling, environment-oriented satellite image processing, environmental security, environment evolution supervision and impact of climate-related events in the SEE region.

Environmental VO involves researchers from Albania, Bulgaria, Hungary, Romania, Serbia and Moldova.

Following applications are under development:

- Monte Carlo Sensitivity Analysis for Environmental Systems (MCSAES)
- Study of CHanges of Environment with Remote Sensing (CHERS)
- Refinement of surface and vegetation parameters in SEE region based on satellite images (GreenView)
- Groundwater Flow Simulation System (Lizza-PAKP)
- Multi-scale atmospheric composition modelling (MSACM)
- Modelling System for Emergency Response to the Release of Harmful Substances in the Atmosphere (MSERRHSA)

The Environmental VO applications can be divided in two subgroups. At one side Lizza-PAKP, MCSAES, MSACM and MSERRHSA deal with environmental modeling. The air pollution modeling applications use two types of models – larger scale Danish Eulerian Model, which covers whole Europe, parts of the Atlantic Ocean and Africa, and medium scale models, which are more appropriate for regional modeling. In some scenarios, the output of the DEM can serve as boundary condition for the medium size models. These models are mostly based on Finite Element Methods, and require MPI or large number of concurrent serial jobs.
The other sub-group is related to satellite images processing, where the GreenView application has the role of a pilot application and other applications like CHERS will be included in the future.

The following modeling tools will be used:

- DEM - Danish Eulerian Model;
- MM5 - 5-th generation PSU/NCAR Meso-meteorological Model used as meteorological pre-processor;
- CMAQ - Community Multiscale Air Quality System (transport and transformation of the pollutants);
- SMOKE - the Sparse Matrix Operator Kernel Emissions Modeling System (emission processing);
- WRF-Chem - a new generation joint Meteorological & Air Quality – future development of the system; collaboration with Meteorological VO regarding the use of WRF is to be investigated;
- ESIP - platform for satellite images data processing.

The main application-specific services offered by the SEE-GRID-SCI project that will be used are: TCP Binder - interface for managing collections of jobs; JTS - Job tracking services; UPM - Performance monitoring from user perspective.

RENAM, FRT TUM and State Hydrometeorological Service of Moldova support pilot application for the ESIP/GreenView development, platform testing and deployment.
Main goal of the MCSAES 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 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). Using MPI enabled GRID clusters will provide precise computation of sensitivity coefficient which is important for further improvement of the model.

CHERS objective is running in grid environment remote sensing packages for evaluation of environmental changes. Use of high resolution LANDSAT, MODIS satellite images to analyze environmental changes. In particular the use of Synthetic Aperture Radar (SAR) data for evaluation of vertical movements of the ground in environmentally hot areas in Adriatic Sea coast. More complex processing of images on grid permits the quantitative analysis of dynamics of different environmental parameters, as well as their correlations. Evaluation of geological parameters is possible, in particular movements of water shores (seas, lakes, rivers) and vertical movements of the ground (the latter using SAR interferometry).

The aim of the GreenView application is a refinement of surface- and vegetation parameters in SEE region based on satellite images. Construction, usage and comparison of diverse satellite datasets will be performed. High resolution satellite measurements can be used for numerous environmental studies (climate-related or air pollution modeling). 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 [3].

GreenView is a web application, and the user accesses by Internet the computation resources and environmental data that are provided by the Grid infrastructure. The GreenView application is based on the ESIP platform and related methodology, which are going to be provided by JRA1 of the SEE-GRID-SCI project. ESIP supports the development and the execution of the Grid based applications concerning particularly with the processing of satellite images and generally with environmental related processing and studies.
ESIP is based on the gProcess platform that provides the user with the possibility to explore the optimal solutions for Grid processing and information searching in the multispectral bands of the satellite images. The gProcess platform is an interactive toolset supporting the flexible description, instantiation, scheduling and execution of the Grid processing. The gProcess platform provides a flexible diagrammatical description solution for image processing workflows in the Earth Observation. At the conceptual level the algorithms are described by processing acyclic graphs, in which the nodes represent operators, services, subgraphs and input data (e.g. satellite image bandwidths), and the arcs represent the execution dependencies between nodes.

Lizza-PAKP is a groundwater flow simulation system containing solver based on finite element method (PAKP) and user interface (Lizza). It was developed in cooperation between Institute for Water Resources "Jaroslav Cerni", Belgrade and CSASA Kragujevac, Serbia. It provides full 3D modeling capabilities, stationary and non-stationary simulations, saturated and non-saturated environment calculation, as well as mass and heat transport handling. PAKP module incorporates solver based on MPI paradigm,

making its presence a grid site requirement. The application also provides TCP binder interface (service developed by UOB as multi-application interactive job environment), which helps hiding grid complexity from the end user.

The aim of MSACM application is to use the Grid environment to produce an integrated, multi-scale Balkan region oriented modeling 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:

1) 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 scales;

2) to track and characterize the main pathways and processes that lead to atmospheric composition formation in different scales, to account for the biosphere-atmosphere exchange as a source and receptor of atmospheric chemical species; and

3) 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 modeling tools that continues to be developed intensively by the efforts of a big community of scientists both in the US and Europe.

The aim of MSERRHSA application is to develop and deploy, on the SEEGRID-SCI infrastructure, a modeling system for emergency response to the release of harmful substances in the atmosphere, targeted at the SEE and more specifically Balkan region, which would be able to:

1) Perform highly accurate and reliable risk analysis and assessment for selected “hot spots”;

2) At a warning signal from the measuring network, by using the adjoin functions technique, to detect (if not known) the harmful release location and evaluate the nature and the amount of the released harmful gases;

3) Provide the national authorities and the international community with short-term regional scale forecast of the propagation of harmful gases;

4) Perform, in an off-line mode, a more detailed and comprehensive analysis of the possible longer-term impacts on the environment and human health in the Balkan region and make the results available to the authorities and the public.

Conclusions

The analysis of applications development status in first year of project demonstrates that a number of applications already have some versions deployed or being tested successfully on the SEE-GRID infrastructure. In addition to the number of applications deployed, the other mid-project targets – numbers of users, research groups and countries involved – are also met. These results are convincing in terms of not only stability of the regional infrastructure but also maturity of the infrastructure and SEE grid community. From the analysis it could also be concluded that project VOs are also getting mature considerably with their applications, repositories and researchers. On April 30 2009 first national training workshop on applications development and deployment was held in SHMS premises (Moldova) [4].

REFERENCES

[1] SEE-GRID-SCI Project. http://www.see-grid-sci.eu/

[2] Bogatencov P., Sidorenco V. Participation of MD-Grid JRU Consortium in SEE-GRID-SCI Project. *7th RoEduNet International Conference 2008. Networking for research and education: 28-30 August, Cluj-Napca, Romania: Proceedings.*- Cluj-Napca: Editura U.T.Press, 2008. ISBN 978-973-662-395-5. P.65-70.

[3] Bacu V., Gorgan D. Graph Based Evaluation of Satellite Imagery Processing over Grid. *International Symposium on Parallel and Distributed Computing,* 2008. ISPDC, ISBN 978-0-7695-3472-5. P.147-154.

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