УДК 681.63

ENVIRONMENT MODELLING IN GRIDS

Nataliia Kussul[1], Guoqing Li[2], Andrii Shelestov[1], Sergii Skakun[1], Ievgen Zagorodnii[1], Yuliia Grypcyh[1]

[1] Space Research Institute NASU-NSAU, Ukraine
[2] Center for Earth Observations and Digital Earth CAS, China

У статті описане рішення щодо створення розподіленого середовища, призначеного для екологічного моделювання. Воно являє собою інфраструктуру Grid, що розробляється Інститутом космічних досліджень НАНУ-НКАУ та є поєднаною з іншими географічно розподіленими ресурсами в межах проекту по розгортанню Wide Area Grid (WAG).

Introduction

At present, global climate changes in the world made a rational land use, environmental monitoring and prediction of natural and technological disasters the tasks of a great importance. The basis for the solution of these problems is the integrated use of data of different nature: modelling data, in situ measurements, and indirect observations such as airborne and spaceborne remote sensing data.

Satellite observations have an advantage of acquiring data for the large and hard-to-reach territories, as well as providing continuous and human-independent measurements. Such important applications as monitoring and predictions of floods, droughts, vegetation state assessment etc. heavily rely on the use of Earth observation (EO) data from space. It should be noted that the same EO data sets and derived products can be used for a number of applications. That is, once we develop interface to access required data and products, it can be used in a uniform way for different purposes and applications. The EO domain is also characterized by the large volumes of data that should be processed, catalogued, and archived [1, 2]. For example, GOME instrument onboard Envisat satellite generates nearly 400 Tb data per year [3]. The processing of satellite data is carried out not by the single application with monolithic code, but by the distributed applications.

Environment modelling tasks have essential demands for underlying computational and storage resources, which hardly could be satisfied by the individual hardware solutions. Furthermore, effective data usage requires high accessibility of data that are used for different tasks.

Mentioned factors leaded to a decision of using a distributed modelling environment (including computation and data storage resources). Considering the fact that several concerned organizations already use Grid systems to provide resources for their own computational tasks, we made efforts to merge partner's resources into a federation utilizing the concept of Wide Area Grig (WAG). WAG is a international effort of several space agency that integrate their resources for the solution of complex problems of environmental modelling.

Each resource provider is supposed to retain the ownership of the provided resource within the federation. Such organizational decentralization improves robustness of the whole infrastructure, making it less dependent on local administrative issues.

We focus on the description of the Grid environment that is under development within the WAG project.

Description of the Grid infrastructure

The Grid infrastructure integrates the resources of geographically distributed organisations, in particular:

− Space Research Institute NASU-NSAU (Ukraine) with deployed computational and storage nodes based on Globus Toolkit 4 (http://www.globus.org) and gLite 3 (http://glite.web.cern.ch) middleware (including Virtual Organization Management Service node, Computing Element node powered by CREAM middleware, and Storage Element node which provides access to data via GSIFTP), access to geospatial data and Grid portal;

− Institute of Cybernetics of NASU (Ukraine) with deployed computational and storage nodes based on Globus Toolkit 4 middleware and access to computational resources (approximately 500 processors);

− Center for Earth Observations and Digital Earth of Chinese Academy of Sciences (CAS) with deployed computational nodes based on Globus Toolkit 4 middleware and access to geospatial data (approximately 16 processors).

It is also worth mentioning that satellite data are distributed through the Grid environment. For example, ENVISAT WSM data (that are used within the flood application) are stored on the ESA's rolling archive and routinely downloaded for the Ukrainian territory. Then, they are stored in the Space Research Institute archive that is accessible via the InterGrid. MODIS data from Terra and Aqua satellites that are used in flood, crop yield and biodiversity assessment applications are routinely downloaded from the USGS' archives and stored in the Space Research Institute NASU-NSAU and Institute of Cybernetics of NASU.

Access to the resources of the Grid environment is organised via a high-level Grid portal that have been deployed using GridSphere framework [9]. Through the portal, users can access the required satellite data and submit jobs to the computing resources of the Grid in order to process satellite imagery.

The existing architecture of the Grid is shown in Fig. 1.

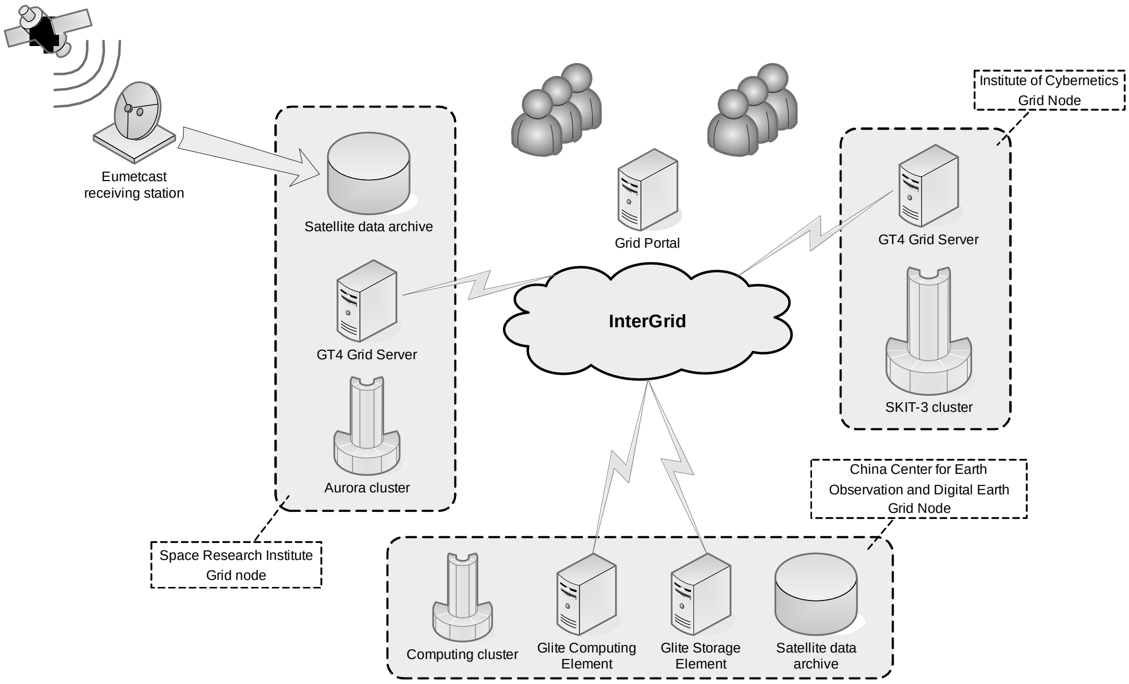

Fig. 1. Architecture of the Grid infrastructure

Applied services within Grid

In this section we review applications that are solved using the resources of the Grid system.

Weather prediction. Weather forecast data is used in the core models of flood monitoring and crop state prediction applications in the Grid environment. We use numeric weather prediction model WRF (Weather Research & Forecasting Model, http://www.wrf-model.org) that was configured and adapted to the territory of Ukraine [6].

Flood monitoring. We developed a neural network approach to flood extent extraction from satellite synthetic-aperture radar (SAR) imagery [2]. We developed a parallel version of our method that can be run on several computational nodes. The use of the Grids allowed us to considerably reduce the time required for image processing. In particular, it took approximately 10 min to process a single SAR image on a single workstation. The use of Grid computing resources allowed us to reduce the time to less than 1 min. The developed Web service is accessible via Internet through the address http://floods.ikd.kiev.ua (Fig. 2).

Biodiversity assessment. We, in cooperation with scientist from the Centre for Aerospace Research of the Earth (Ukraine), developed an approach for land biodiversity assessment and mapping using Earth observations data [8]. The proposed approach was developed for the Pre-Black Sea region, but, in general, can be extended to any other region. The developed Web service is accessible via Internet through the address http://biodiv.ikd.kiev.

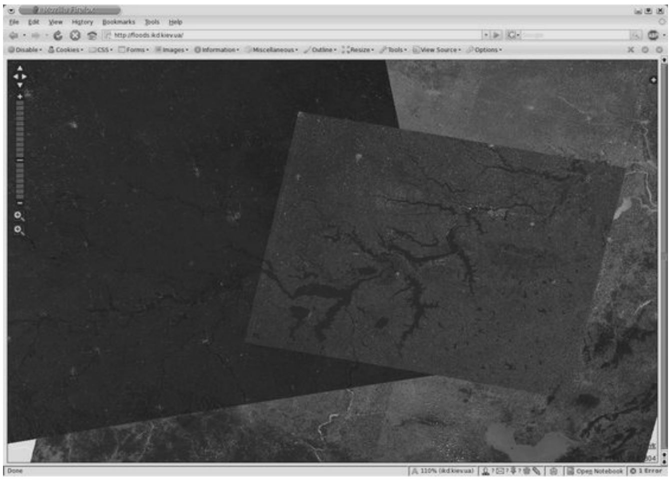

Fig. 2. Flood application within the Grid infrastructure. Flood event: River Huaihe, China, July, 2007. Data sources: Envisat/ASAR (© ESA, 2007) and RADARSAT-1 (© CSA, 2007)

Crop yield prediction. We implemented a time series analysis of vegetation index approach for yield prediction and vegetation state assessment [9]. As a basis, we use the enhanced vegetation index (EVI). The WOFOST model is used with the assimilation of Leaf Area Index (LAI) derived from satellite observations. The visualization interface for the developed services is depicted in Fig. 3.

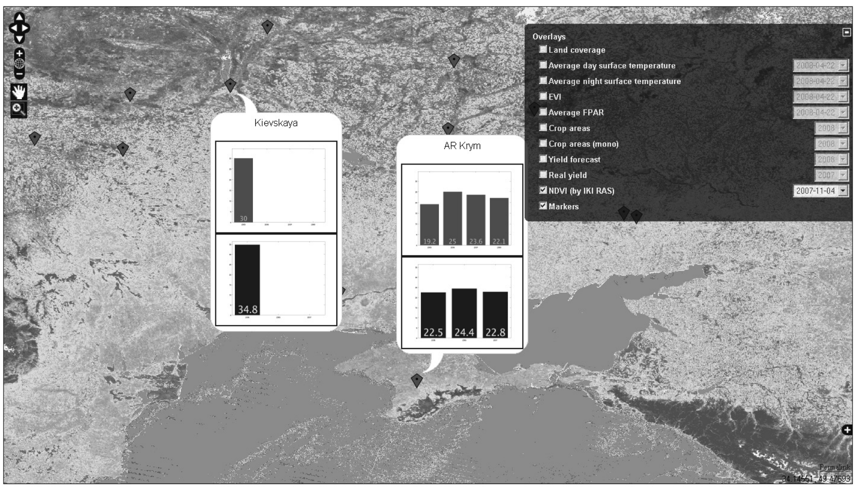

Fig. 3. Wheat yield forecast (Ukraine, 2008)

Conclusions

We presented a Grid infrastructure for environmental modelling that is being developed in the Space Research Institute NASU-NSAU and integrates the resources of several geographically distributed organizations: the Space Research Institute NASU-NSAU, the Institute of Cybernetics NASU and the China's Center for Earth Observations and Digital Earth of CAS. Currently, we are using a Grid portal solution based on GridSphere framework to integrate Grid systems with different middleware, such as GT4 and gLite 3.

Acknowledgments. This work is supported by ESA CAT-1 project "Wide Area Grid Testbed for Flood Monitoring using Spaceborne SAR and Optical Data" (No. 4181), joint project of the Science and Technology Center in Ukraine (STCU) and the National Academy of Sciences of Ukraine (NASU), "Grid Technologies for Multi-Source Data Integration" (No. 4928), project of the Ministry of Education and Science of Ukraine, "Development of Integrated Remote Sensing Data Processing System using Grid Technologies" (No. M/72-2008), and Ukrainian-Chinese bilateral project.

References

1. Fusco, L., Cossu, R., Retscher, C.: Open Grid Services for Envisat and Earth Observation Applications. In: Plaza, A.J., Chang, C.-I. (eds.) High performance computing in remote sensing, 1st edn. pp. 237-280. Taylor & Francis Group, NY (2007)
2. Shelestov, A., Kussul, N., Skakun, S.: Grid Technologies in Monitoring Systems Based on Satellite Data. J of Automation and Inf. Sci. 38(3), 69-80 (2006)
3. Fusco, L., et al.: Putting Earth-Observation on the Grid. ESA Bulletin, 114, 86-91 (2003).
4. Foster, I.: The Grid: A New Infrastructure for 21st Century Science. Physics Today 55(2), 42-47 (2002)
5. Novotny, J., Russell, M., Wehrens, O. GridSphere: An Advanced Portal Framework. In: 30th EUROMICRO Conference, pp. 412-419 (2004)
6. Kravchenko, O., Kussul, N.: Simulation cascade of NWP and land surface model for drought monitoring. European Geosciences Union General Assembly, Vienna, Austria (2008)
7. Kussul, N., Shelestov, A., Skakun, S., Kravchenko, O.: Data Assimilation Technique For Flood Monitoring and Prediction. Int. J. on Information Theory and Applications, 15(1), 76-84 (2008)
8. Kussul, N., Popov, M., Shelestov, A., Stankevich, S., Korbakov, M., Kravchenko O., Kozlova, A. Information service for biodeversity assesment in the context of Ukrainian segment of GEOSS. J. of Science and Innovations, Vol. 6, 13-25 (2007) (in Ukrainian)
9. Kussul, N., Ilin, M., Skakun, S., Lavrenuk, A.: Crop state estimation and winter wheat yield prediction using remote sensing data. In: Markov. K., Ivanova. K., Mitov. I. (eds.) Int. Book Series "Decision Making and Business Intelligence, Strategies and Techniques", Number 3, pp. 103-109. (2008) (in Russian)

Received 28.05.09