GRID APPROACH TO SATELLITE MONITORING SYSTEMS INTEGRATION

Nataliia Kussul, Andrii Shelestov, Serhiy Skakun

Abstract: This paper highlights the challenges of satellite monitoring systems integration, in particular based on Grid platform, and reviews possible solutions for these problems. We describe integration issues on different levels: data integration level and task management level (job submission in terms of Grid). We show example of described technologies for integration of monitoring systems of Ukraine (National Space Agency of Ukraine, NASU) and Russia (Space Research Institute RAS, IKI RAN). Another example refers to the development of InterGrid infrastructure that integrates several regional and national Grid systems: Ukrainian Academician Grid (with Satellite data processing Grid segment) and RSGS Grid (Chinese Academy of Sciences).

Keywords: Grid systems, satellite monitoring systems.

ACM Classification Keywords: H. Information Systems - H.1 Models and Principles

Conference: The paper is selected from Sixth International Conference on Information Research and Applications – i.Tech 2008, Varna, Bulgaria, June-July 2008

Introduction: Specifics of Earth Observation Problems

At present global climate changes in the world made rational land use, environmental monitoring and prediction of natural and technological disasters the tasks of a great importance. The basis for solving these problems is the use of data of different nature: modeling data, in situ measurements and indirect observations such as airborne and spaceborne remote sensing data. However, mutual disarrangement of heterogeneous data and measurement technologies, spatial and temporal inconsistency of measurements are limiting potentials of modern technologies for solving actual problems of environmental monitoring and forecasting of disasters. Thus, development of effective technologies for heterogeneous data integration is a very important issue.

Nowadays Earth Observation (EO) data play a major role in solving problems in different domains. Satellite observations enable acquisition of data for large and hard-to-reach territories, can provide continuous measurements and human-independent information, etc. EO domain, in turn, is characterized by large volumes of data that should be processed, catalogued, and archived. For example, GOME instrument onboard Envisat satellite generates nearly 400 Tb data per year [1]. EUMETCast system that is part of global GEONETCast system [2] of GEOSS enables acquisition of more than 50 Tb of processed and unprocessed information per year. Moreover, the processing of satellite data is carried out not by the single application with monolithic code, but by distributed applications. This process can be viewed as a complex workflow that is composed of many tasks: geometric and radiometric calibration, filtration, reprojection, composites construction, classification, products development, post-processing, visualization, etc. [3].

To enable processing and management of such volumes of data sets and information flows an appropriate infrastructure is needed that will support [1, 4]: access to distributed resources; high flexibility; portal enabling easy and homogeneous accessibility; collaborative work; seamless integration of resources and processes; allow processing of large historical archives; avoid unauthorised access to/use of resources.

Grid can provide appropriate facilities for high-performance computations and efficient data management in EO domain. Grid computing is an emerging paradigm for global computing and a very active research domain for complex, dynamic, distributed and flexible computing and resource sharing [5]. Grid computing belongs to main trends of on-line environment development among with web services, semantic web and peer-to-peer networking. The integration on these technologies is essential for the next generation networks.

Grid systems are recognized to be very efficient for EO and geospatial community for a number of reasons: geospatial data and associated computational resources are naturally distributed; the multi-discipline nature of geospatial research and applications requires the integrated analysis of huge volume of multi-source data from multiple data centres; most geospatial modelling and applications are both data and computational intensive. The aggregated computational power of Grid system can provide for the application.

In this paper we highlight the challenges of satellite monitoring systems integration, in particular based on Grid platform, and review possible solutions for these problems.

State of the art: Grid-based systems for EO data processing

At present, Grid technologies are widely applied in different domains, in particular EO domain. EU-funded European DataGrid Project (EDG) was one of the first Grid-enabled projects allowing European Space Agency (ESA) to gain firsthand experience in the use of emerging Grid technologies [1]. Based on the gained experience European Space Agency (ESA) and European Space Research Institute (ESRIN) are developing Grid Processing on Demand (G-POD) for Earth Observation Applications (http://gpod.eo.esa.int). Online access to different data is enabled within this project, in particular to data provided by various instruments on Envisat satellite (http://envisat.esa.int), SEVIRI instrument onboard MSG (Meteosat Second Generation) satellite [6], ozone profiles derived from GOME instrument, etc. One of the most important applications is the analysis long-term data. Grid Web Portal provides access to the “Grid-on-demand” resources enabling: personal certification, time/space selection of data directly from the ESA catalogue, data transfer, job selection, launching and live status, data visualization.

DEGREE (Dissemination and Exploitation of GRids in Earth science) project (http://www.eu-degree.eu) is initiated within EGEE/EGEE-II. A major challenge for DEGREE is to build a bridge linking the Earth Science and GRID communities throughout Europe, and focusing in particular on the EGEE-II Project. Grid provides appropriate infrastructure enabling international cooperation within GMES and GEOSS. The following problems are within the scope of DEGREE: earthquake analysis, floods modeling and forecasting, influence of climate changes on agriculture

Japan Aerospace eXploration Agency (JAXA) and KEIO University started establishing “Digital Asia” system aimed at semi-real time data processing and analyzing. They use GRID environment to accumulate knowledge and know-how to process remote sensing data. The Digital Asia project is the part of bigger Sentinel Asia project that is targeting on building natural disasters monitoring system [7].

CEOS Wide Area Grid (WAG) project is initiated by CEOS Working Group on Information Systems and Services (WGISS), and aims at providing horizontal infrastructure enabling efficient integration of resources of different space agencies. WAG testbed infrastructure is currently under development within ESA Cat-1 project “Wide Area Grid Testbed for Flood Monitoring Using Spaceborne SAR and Optical Data” (no. 4181) [8]. Within WAG project Space Research Institute NASU-NSAU have developed testbed that integrates resources of Ukrainian Grid segment (Ukrainian Academician Grid) with resources of international organisations (ESA, RSGS-CAS).

Tendencies of globalization and integration of satellite monitoring systems

Nowadays there is a trend for globalization of monitoring systems with purpose of solving more complex problems and reducing collaboration expenses. EO data are naturally distributed over many organizations involved in data receiving and processing. This leads to the need of integration of existing systems for solution of complex problems. The development of GEOSS (Global Earth Observation System of Systems) [9] is coordinated by Group on Earth Observations (GEO) [10] that was launched in response to calls for action by the 2002 World Summit on Sustainable Development and the G8 (Group of Eight) leading industrialized countries. GEO is a voluntary partnership of governments and international organizations that provides a framework within which these partners can develop new projects and coordinate their strategies and investments. It is recognised that GEOSS work with and build upon existing national, regional, and international systems to provide comprehensive, coordinated Earth observations from thousands of instruments worldwide, transforming the data collected into vital information for society.

Modern tendencies of globalization and development of "system of systems" GEOSS lead to the need of integration of heterogeneous satellite monitoring systems. Integration can be done on different levels: (i) data exchange level, (ii) task management level. Data exchange level is supposed to provide tools for sharing data and products. This infrastructure enables data integration where different entities provide various kinds of data to support joint solution of complex problems (Fig. 1). Task management level envisages running applications on distributed computational resources provided by different entities (Fig. 2). Since many of the existing satellite monitoring system rely on Grid technologies appropriate approaches and technologies should be evaluated and developed to enable Grid system integration (so called InterGrid).

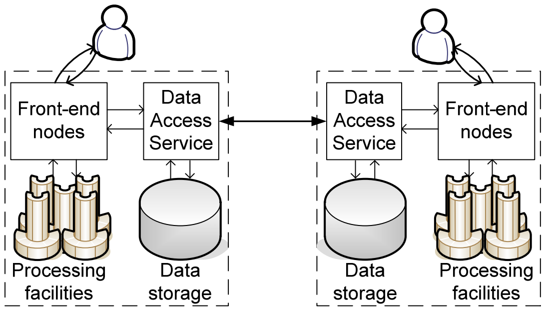

Fig 1. Data integration level

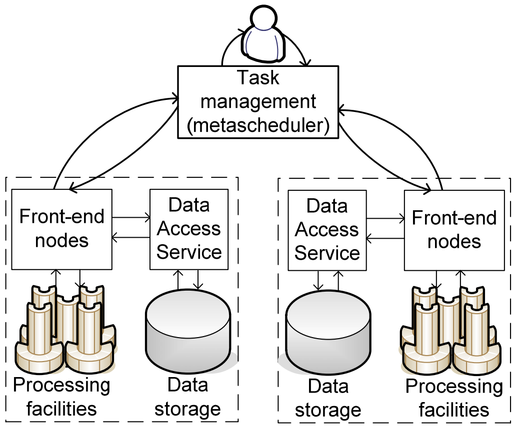

Fig 2. Task management level

The next sections will highlight main challenges and possible solutions for satellite monitoring systems integration on both levels, and provide description of case-studies for both cases.

Levels of integration: main problems and possible solutions

Integration on data exchange level could be done by using common standards of EO data exchange, common user interfaces, and common data and metadata catalog. As to task management level the following problems additionally should be solved: joint computational infrastructure setup; development of jobs submission and scheduling algorithms; load monitoring enabling; security policy enforcement.

Data exchange level. At present the most appropriate standards for data integration is OGC standards. Data visualization issues can be solved by using the following set of standards: WMS (Web Map Service), SLD (Style Layer Descriptors) and WMC (Web Map Context). OGC's WFS (Web Feature Service) and WCS (Web Coverage Service) standards provide uniform ways for data delivery. In order to provide interoperability on the level of catalogues CSW (Catalogue for Web) standard can be applied.

Since data are stored on geographically distributed sites there can be issues regarding optimization of visualization schemes. In general, there are two possible ways for distributed data visualization: centralized visualization scheme and distributed visualization scheme. Advantages and faults of each scheme were described in [11].

Task management level. In this subsection we present main issues and possible solutions for Grid-system integration. Main prerequisite of such kind of integration is certificates trust. It could be done, for example, through EGEE infrastructure that nowadays brings together the resources of more than 70 countries. Another problems concerned with different Grid systems integration are as follows: enabling data transfers and high-level access to geospatial data; development of common catalogues; enabling jobs submission and monitoring; enabling information exchange.

Data transfer. GridFTP is an appropriate and reliable solution for data transfer. The only limitation is the requirement of transparent LAN (local area network) infrastructure.

Access to geospatial data. High-level access to geospatial data can be organised in two possible ways: using pure WSRF services or using OGSA-DAI container. Each of this approach has its own advantages and weaknesses. Basic functionality for WSRF-based services can be easily implemented (with proper tools), packed and deployed. But advanced functionality such as security delegation, third-party transfers, indexing should be implemented by hands. WSRF-based services can also pose some difficulties if we need to integrate them with other data-oriented software.

OGSA-DAI framework provides uniform interfaces to heterogeneous data. This framework makes possible to create high-level interfaces to data abstracting hiding details of data formats and representation schemas. Most of problems in OGSA-DAI are handled automatically, e.g. delegation, reliable transfer, data flow between different sources and sinks. OGSA-DAI containers are easily extendable and embeddable. But comparing to WSRF basic functionality implementation of OGSA-DAI extensions is more difficult. Moreover, OGSA-DAI requires preliminary deployment of additional software components.

Task management. There are two possible approaches for task management. One of them is to use Grid portal (Fig. 3) supporting different middleware platforms, such as GT4, gLite, etc. Grid portal is an integrated platform to end-users that enables access to Grid services and resources via standard Web browser. Grid portal solution is easy to deploy and maintain, but it doesn't provide application interface and scheduling capabilities.

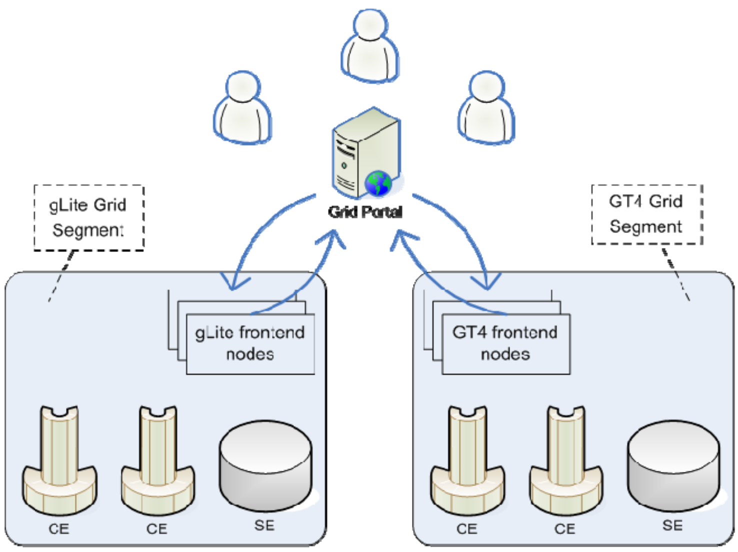

Fig. 3. Portal approach to Grid system integration

Another approach is to develop high-level Grid scheduler (Fig. 4) that will support different middleware by providing some standard interfaces. Such metascheduler interacts with low-level schedulers (used in different Grid systems) enabling in such way system interoperability. Metascheduler approach is much more difficult to maintain comparing to portals; however, it provides API with advanced scheduling and load-balancing capabilities. At present, the most comprehensive implementation for the metascheduler is a GridWay system. The GridWay metascheduler is compatibility with both Globus and gLite middlewares. Starting from Globus Toolkit v4.0.5 GridWay become standard part of its distribution. GridWay system provides comprehensive documentation for both users and developers that is a important point for implementing new features.

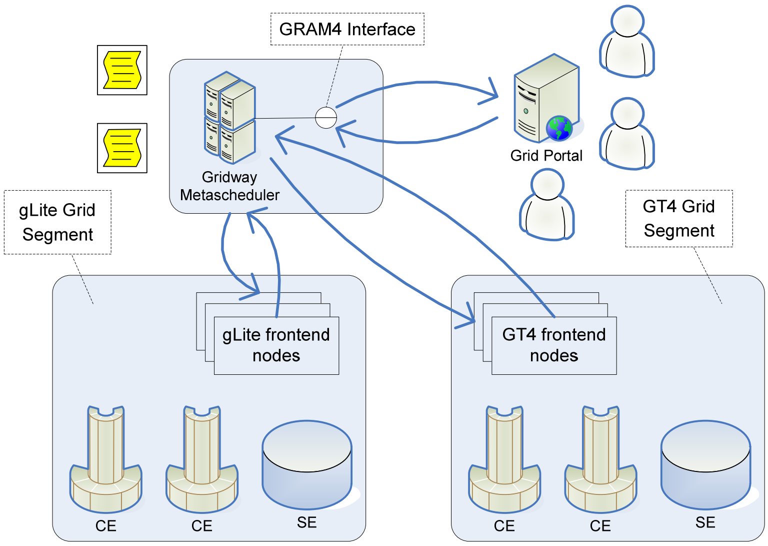

Fig. 4. Metascheduler approach

In the next section we show the examples of application of described approaches to integration of satellite monitoring systems and development of InterGrid environment.

Implementation: lessons learned

Integration of satellite monitoring systems. The first case-study refers to the integration of satellite monitoring systems of NSAU (Ukraine) and IKI RAN (Russia). The overall architecture for integration of data provided by two organizations is depicted in Fig. 5. The proposed approach is applied for the solution of problems for agriculture resources monitoring and crop yield prediction. Within integration NSAU provides WMS interfaces to NWP modelling data (using WRF model) [12], in-situ observations from meteorological ground stations in Ukraine, and land parameters (such as temperature, vegetation indices, soil moisture) derived from satellite observations from MODIS instrument onboard Terra satellite. IKI RAN provides WMS interfaces to operational land and disaster monitoring system. Both NSAU and IKI RAN provides user Web-interfaces to monitoring systems that support OGC WMS standards.

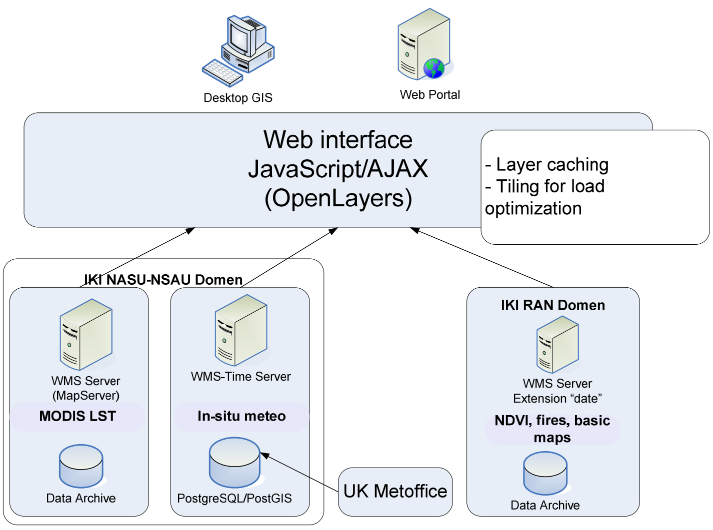

Fig. 5. Architecture of satellite monitoring system integration

In order to provide user interface that will enable visualization of data from multiple sources we use open-source OpenLayers framework (http://www.openlayers.org). OpenLayers is “thick client” software based on JavaScript/AJAX and fully operational on client side. Main OpenLayers features also include: support for several WMS servers, support for different OGC standards (WMS, WFS), cache and tiling support to optimize visualization, support for of both raster and vector data. The provided data and products are accessible via Internet http://land.ikd.kiev.ua. The example of OpenLayers visualization of data from multiple sources is depicted in Fig. 6.

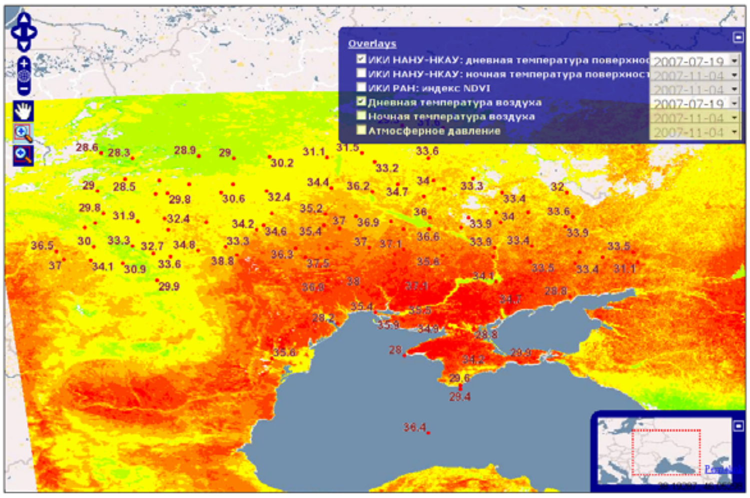

Fig. 6. OpenLayers interface to multiple data

InterGrid testbed development. The second case-study refers to the development of InterGrid for environmental and natural disaster monitoring. InterGrid integrates Ukrainian Academician Grid (with Satellite data processing Grid segment) and RSGS Grid (Chinese Academy of Sciences) and is considered as a testbed for Wide Area Grid (WAG) implementation—a project initiated within CEOS Working Group on Information Systems and Services (WGISS).

The important application that is being solved within InterGrid environment is flood monitoring and prediction. This task requires adaptation and tuning of existing hydrological and hydraulic models for corresponding territories and the use of heterogeneous data stored on multiple sites. Flood monitoring and prediction requires the use of the following data sets: NWP modelling data (provided by Satellite data processing Grid segment), SAR imagery from Envisat/ASAR and ERS-2/SAR satellites (provided by ESA), products derived from optical and microwave satellite data such as soil moisture, precipitation, flood extent etc., in-situ observations from meteorological ground stations and digital elevation model (DEM). The process of model adaptation can be viewed as a complex workflow and requires the solution of optimization problems (so called parametric study). Satellite data processing and products generation tasks also represent complex workflow and require intensive computations. All these factors lead to the need of using computational and informational resources of different organizations and their resources into joint InterGrid infrastructure. The architecture of proposed InterGrid is depicted in Fig. 7.

GridFTP was chosen to provide data transfer between Grid systems. In order to enable interoperability between different middleware (for example, Satellite data processing Grid segment is using GT4; RSGS Grid is using gLite 3.x; Ukrainian Academician Grid is based on NorduGrid) we developed Grid portal that is based on GridSphere portal framework (http:// www.gridsphere.org). The developed Grid portal allows users to transfer data between different nodes and submit jobs on computational resources of the InterGrid environment. The portal also provides facilities to monitor statistics of the resources such as CPU load, memory usage, etc. The further works on providing interoperability between different middlewares are directed to the development of metascheduler using GridWay system. In the nearest future we are intended to provide integration with ESA's EO Grid-on-Demand infrastructure.

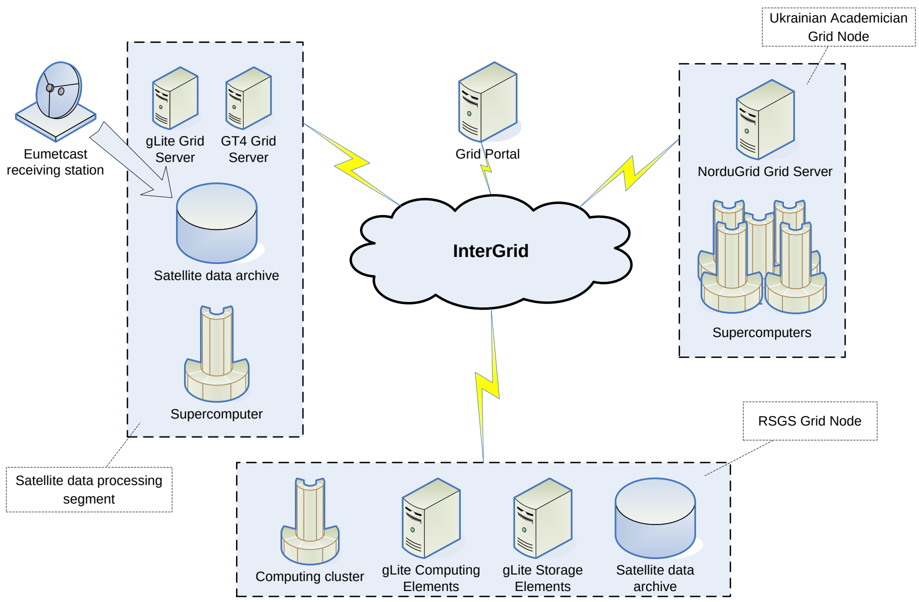

Fig. 7. InterGrid architecture

Conclusions

This paper focuses on the problems of integration of satellite monitoring systems, in particular those using Grid platform. We described two possible levels of integration (data level and task management level) reviewing possible solution for implementing each of them. Considering data integration level we found that integration could be provided by using existing standards for geospatial data, in particular OGC standards. We demonstrated applicability and usability of this approach for integrating existing satellite monitoring systems of Ukraine and Russia for agriculture applications. The use of standard OGC interfaces makes it possible to standardise and facilitate development of integrated satellite monitoring systems (based on existing ones) to exploit synergy and acquire information of new quality.

As to integration on task management level we reviewed two solutions: portal-based and metascheduling approach. We implemented portal solution based on GridSphere framework for the InterGrid environment that integrates several regional and national Grid systems. In order to provide advanced scheduling and load-balancing capabilities the further works will be directed to the implementation of metascheduler based on GridWay system. Also we are intended to provide integration with ESA's G-POD. Further investigations will be directed to integration of distributed monitoring systems with SensorWeb networks to provide automatic delivery of data from heterogeneous sources.

Acknowledgment

The work is supported by ESA CAT-1 project "Wide Area Grid Testbed for Flood Monitoring using Spaceborne SAR and Optical Data" (#4181) and by INTAS-CNES-NSAU project "Data Fusion Grid Infrastructure" (Ref. Nr 06-1000024-9154).

Bibliography

1. Fusco L., Goncalves P., Linford J., Fulcoli M., Terracina A., D'Acunzo G. Putting Earth-Observation on the Grid. In: ESA Bulletin, 2003, 114, pp. 86-91.
2. GEONETCast, http://www.earthobservations.org/progress/GEONETCast.html.
3. Rees W.G. Physical Principles of Remote Sensing, Cambridge University Press, 2001.
4. Shelestov Andrey Yu., Kussul Nataliya N., Skakun Sergey V. Grid Technologies in Monitoring Systems Based on Satellite Data. J. of Automation and Information Science, 2006, vol. 38, issue 3, pp. 69-80.
5. Foster I., Kesselman C. The Grid: Blueprint for a New Computing Infrastructure. 2nd Edition, Morgan Kaufmann, 2004.
6. EUMETSAT, http://www.eumetsat.int.
7. Fukui H. Sentinel Asia/Digital Asia: Building Information Sharing Platform by Geo web services and contributing to Disaster Management Support in the Asia-Pacific Region.
8. Kopp P., Petiteville I., Shelestov A., Li G. Wide Area Grid (WAG). In: Proc. The 7th Ukrainian Conference on Space Research, National Flight and Control Center, Evpatoria, Ukraine, 2007, p. 209.
9. GEOSS Portal, http://www.geossportal.com.
10. Group on Earth Observations (GEO), http://www.earthobservations.org.
11. Shelestov A., Kravchenko O., Ilin M. Distributed visualization systems in remote sensing data processing GRID, International Journal "Information Technologies and Knowledge" - volume 2 / 2008. (in print)
12. Kussul N., Shelestov A., Skakun S., Kravchenko O. Data Assimilation Technique For Flood Monitoring and Prediction. International Journal on Information Theory and Applications, 2008, volume 15. (in print).

Authors' Information

Nataliia Kussul – Prof., Dr., Head of Department of Space Information Technologies and Systems, Space Research Institute of NASU-NSAU, Glushkov Ave 40, build. 4/1, Kyiv-187, 03680 Ukraine,
e-mail: inform@ikd.kiev.ua

Andrii Yu. Shelestov – PhD, Senior Researcher, Department of Space Information Technologies and Systems, Space Research Institute of NASU-NSAU, Glushkov Ave 40, build. 4/1, Kyiv-187, 03680 Ukraine,
e-mail: inform@ikd.kiev.ua

Serhiy V. Skakun – PhD, Research Assistant, Department of Space Information Technologies and Systems, Space Research Institute of NASU-NSAU, Glushkov Ave 40, build. 4/1, Kyiv-187, 03680 Ukraine,
e-mail: serhiy.skakun@ikd.kiev.ua