e-Campus System Based on the Pervasive Grid

Ruay-Shiung Chang and Jih-Sheng Chang
Network Innovation Technology Laboratory
Department of Computer Science and Information Engineering
National Dong Hwa University, Hualien, TAIWAN
rschang@mail.ndhu.edu.tw

Abstract

With the advancements of computer system and communication technologies, Grid computing can be seen as the popular technology for the next generation distributed computing application. Grid computing using the broadband networks, efficiently integrates various computing devices, databases, software, instruments, and even professional expertise. In this paper, we propose a pervasive grid environment based on the Grid Computing technology for integration of all wireless and wired computing devices into a Grid Computing environment. The e-Campus system is implemented on the basis of pervasive grid architecture in order to improve the efficiency of data access and computation.

Two applications, including e-Ecology and e-Classroom, based on the pervasive grid platform are proposed.

Keywords: distributed computing, e-Campus

1. Introduction

Grid computing [3, 4] is a burgeoning technology with the capability of integrating a variety of computing resources as well as scheduling jobs from various sites, in order to supply a number of users with breakthrough computing power at low cost. With these critical features of grid computing involving resources sharing, parallel computing, distributed storage, and secure infrastructure, it can contribute to the development of digitalized platform to facilitate the efficiency in campus data access and computation. For the reasons mentioned above, we propose an effective platform called pervasive grid to make resources available as conveniently as possible. The pervasive grid platform integrates all of the wired and mobile devices into a uniform resource on the grounds of the existing grid infrastructure. Resources can be accessed easily anytime and anywhere through the pervasive grid platform.

Figure 1 reveals the overall architecture of e-Campus with the hierarchical components. The underlying grid middleware is deployed by the Globus Toolkit [5]. Globus Toolkit is a noted middleware on grid computing developed by the Globus Alliance [5]. It provides several fundamental grid technologies along with an open source implementation of a series of grid services and libraries. Pervasive grid platform is implemented on the basis of the Globus Toolkit. A service-oriented provider, consisting of data, computation and information service, offers users a comprehensive computing environment. Data transmission and replication are the main operations in data service. We utilize GridFTP [7] as the underlying transmission protocol. Computation service provides the computing resources for job executions. Information service gives up-to-date resource information such as CPU frequency, available memory space and average system loading. Such information could be utilized by region job dispatcher during job submission in order to decide a proper grid site for execution. E-campus applications are built above the platform and services.

E-Campus system makes use of the advantages of pervasive grid to provide students and teachers with a digitalized education system. From the perspective of the most users, a friendly interface without complicated manipulations is necessary. In order to simplify the interconnection and operations, we have developed a user portal by means of Java CoG Toolkit [6]. Due to the nature of cross-platform execution of Java, our portal solution can run on various operating systems. A user could connect and access the e-campus services in an unsophisticated way via our client portal.

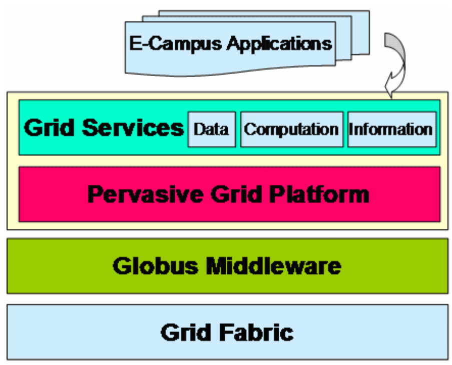

Figure1. Components of e-Campus

2. Related Work

In our implementation, we have made use of the Globus Toolkit [5] as our system infrastructure. It provides several fundamental grid technologies along with an open source implementation of a series of grid services and libraries.

A few critical components within Globus Toolkit are listed below:

- Security: GSI (Grid Security Infrastructure) [8] provides the authentication and authorization mechanisms for system protection according to X.509 proxy certificates.
- Data Management: It is utilized to manipulate data including GridFTP [7] and RLS (Replica Location Service) [10]. RLS maintains location information of replicas from logical file names (LFN) to physical file names (PFN).
- Execution Management: GRAM (The Grid Resource Allocation and Management) [9] provides a series of uniform interfaces to simplify the access of remote grid resources for job execution.
- Information Services: MDS (Monitoring and Discovery System) [11] enables the monitoring and discovery services for grid resources.

We have developed a portal program on client side by means of the CoG Toolkit [6]. The Java CoG Toolkit provides a series of programming interfaces as well as reusable objects in grid services, such as GSI, GRAM, GridFTP, MDS and so on. It presents programmers with a mapping between the Globus Toolkit and Java APIs, so as to ease the programming complexity. For further details about the CoG Toolkit, please refer to [6].

In[12], a proxy-based wireless grid architecture was proposed. A proxy component is deployed as a interface between computing resources and mobile devices for service managements and QoS requirements. Having built the proxy, a mobile user can connect a grid environment with ease, without taking care of the differences between various mobile devices, for attaining to heterogenous interworking and pervasive computing.

A conception of pervasive wireless grid was presented in [13]. The whole computing environment consists of a backbone grid and access grids. Mobile devices are considered a terminal to connect the backbone grid. Most of computing jobs are dispatched to the backbone grid. In addition, the impact of service handoff for mobile users is disscussed in this paer.

In Taiwan, eTaiwan program is the main project to improve the national information infrastructure, focusing on digital entertainment, digital archives[15], e-learning[16], and so forth. As regards the Government and enterprises, several digital applications are promoted, including e-Government[17], e-Transportation[18], and e-Business[19].

Digital applications have been applied to medicine [20] in California. As for remote districts, the medical resources are always limited. With the aid of digtal medical system and the Internet, a doctor can heal the sick online. It is very helpful to advance the medical quality.

In Europe, the European Union actively develops digital system. In eEurope plan [21], there are several digital system projects, comprising e-Learning, e-Government, e-Health, e-Business and e-Inclusion. Take e-Learning for example, it involves three components, includingvirtual campus, boradband connections and Grids for e-Learning, for the improvement in learning efficiency by means of digital contents.

3. e-Campus System

3.1. System Overview

This research is based on the grid technology in support of pervasive computing for digitalized platform in a campus. We attempt to develop a pervasive grid environment based on the Grid Computing technology to coordinate all of wireless and wired computing devices within a Grid Computing environment. From the standpoint of users, all the resources are considered a uniform type regardless of the type of resources. A user can access a variety of resources conveniently through the Web Services deployed in our system.

We have adapted the layered-design approach to implement the e-Campus system. The design framework of the e-Campus system appears in Figure 2. The layered-design makes e-Campus more flexible if new services are added to the system as needed. Based on the Web Services architecture, we develop a service-oriented provider, which offers users a comprehensive grid computing service, including data, computation and information service. It provides flexibility for future services that support the e-Campus system.

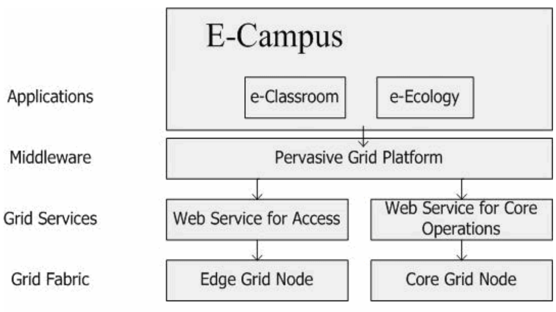

Figure 2. Framework of E-Campus System

There are five components within e-Campus as show below:

- Core computing infrastructure
- Edge grid node
- Web Services
- Pervasive grid platform
- Applications

The core computing infrastructure is the main computing and storage resource. It provides a computing platform with a capability of storage elements, scheduling system, workflow management. In opposition to the core computing infrastructure, the edge grid node is a terminal, such as notebooks, PDAs, and personal computers, for connection with the core grid infrastructure. An edge grid node could access the grid services as well as publish services to the public. Web Services [22, 23, 24] is a popular technology based on XML and HTTP for the construction of distributed computing applications. It works at an open-architecture with the capability of bridging any programming language, computing hardware or operating system. Accordingly, we adopt Web Services as our software interface, in order to build a uniform entry between an edge grid node and our grid services. As for the pervasive grid platform, it is a middleware to provide the basic grid services for users, such as location management, service handoff, and personal information management. In the applications layer, we develop some useful applications based on the pervasive grid platform, including e-Classroom and e-Ecology.

3.2. Pervasive Grid Platform

In the mater of the core computing infrastructure, it contains the computing power and storage capability, to provide mobile or wired users with grid services. An edge grid node is just a terminal between a user and the core computing infrastructure. It is necessary to provide users with an efficient interface in a seamless and transparent way by the core computing infrastructure. Consequently, it is essential to develop a high-performance platform to process the requirements of users.

There are several works to address as given below in terms of the pervasive grid platform:

- To process the join and leave of edge computing nodes: Our system follows GSI (Grid Security Infrastructure) [25] to design a user authentication/authorization mechanism for adapting to our environment.
- Managing the interconnection between the core computing infrastructure and edge grid nodes: There are several differences and limitations among various edge grid devices. The pervasive grid platform is capable of manage these differences as well as fit in with user's QoS (Quality of Service).
- The interconnection between the pervasive grid platform and core computing infrastructure: As presented in Figure 3, we implement the interface to handle the interconnection between the interconnection between the pervasive grid platform and core computing infrastructure through the Globus APIs. The corresponding algorithm is developed to cope with user's jobs via Globus as well.
- Job dispatch, management, and QoS: We are concerned with the development of the flexible, high-performance, and reliable dispatcher and scheduler within the pervasive grid platform, in order to suffice for the requirements of users. A user with different priority could obtain the corresponding service level.

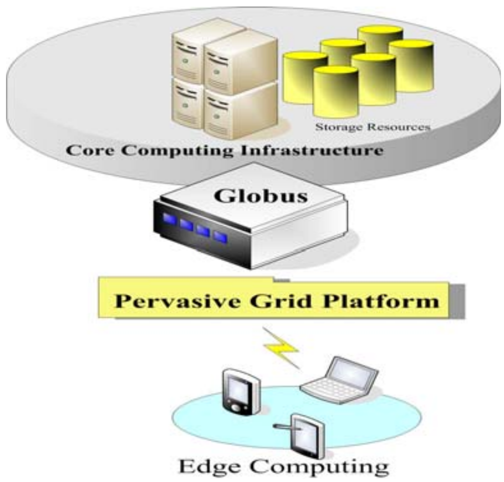

Figure 3. Interconnection between the core computing infrastructure and edge nodes

3.3. Grid Service Provider

On the basis of the pervasive grid platform, we would like to implement a service-oriented provider in a module-design way based on the Web Services technology. It is easy to add or remove a service without taking great pains to maintain system services. For example, Data Grid [26, 27, 1] service is intended to provide a large amount of storage resources and distributed access technology for data-intensive applications. There are three grid service modules within our system, including Data-Grid service, Computational service, and Information service. The computational service is to supply users with computing services for job execution. The information service has the capability of gathering the information about hardware resources.

With the exception of the core grid infrastructure, an edge node could also publish and provide some specified services. For instance, a PDA (Personal Digital Assistant) may publish the GPS service to the public. Other edge grid nodes could access the GPS service provided by the PDA. Through the share of services, our system has not merely a better service-oriented architecture, but a complete and diverse service provider. Therefore, as shown in Figure 4, it is reasonable to deploy and build a service repository system for maintaining all registered services dynamically such as query, joint, remove for services.

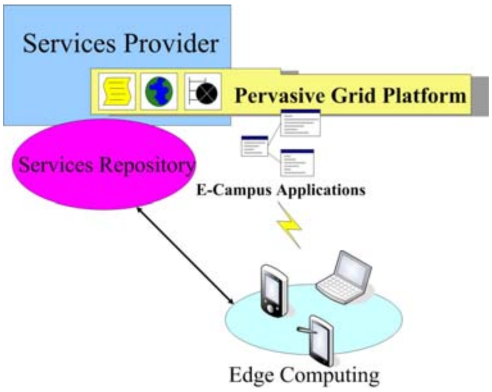

Figure 4. Service Provider and Repository

3.4. e-Campus Applications

In the light of the development of the pervasive grid platform and service-oriented provider system, we make a study of academic applications within a campus. There are two applications for e-Campus system, including e-Ecology and e-Classroom.

The National Dong Hwa University (NDHU) has a well-preserved natural ecosystem. It is a precious treasure for teaching and education. In addition, with regard to visitors, they may like to understand and observe the natural environment within NDHU. For this reason above, we are attempted to develop the e-Ecology application, as shown in Figure 5, by keeping records of the daily activities of the natural ecology system as video files over a long period of time in NDHU. It is obvious that the data size of video files must be very tremendous.

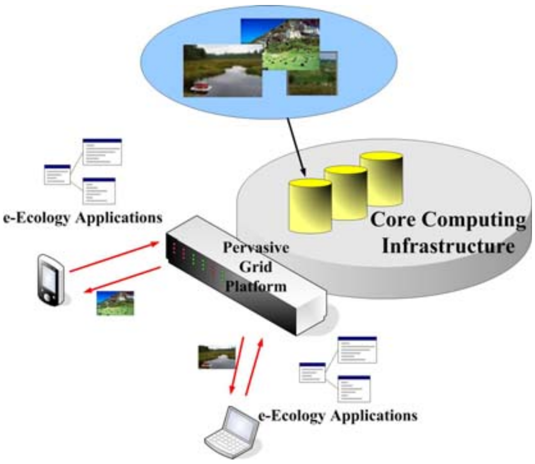

Figure 5. e-Ecology Architecture

As presented in Figure 6, in order to cope with such large amount data, we have implemented a storage broker based on the Data-Grid technology in support of the e-Ecology system. The overall components of the storage broker are presented in Figure 7. The file mover uses GridFTP as its transmission protocol to copy files between two grid nodes. The upload processing engine gets the space information of each storage node from MDS. We adapt the roulette wheel algorithm [2] to the storage broker for choosing a node to upload a file. According to the roulette wheel algorithm, the larger capability of storage resources has the larger probability to be chose, with a view to achieve the system balance for storage resources. Download processing engine is an agent distributed in each node. When the broker gets a download request, it retrieves information from RLS database and redirects this request to the node that contains the file. The download agent receives the request and start transferring file. The Storage Broker would distribute the download jobs to each storage node, in order to shorten the download time. Therefore, users could browse the digital files smoothly without significant delay. The search engine helps us to look for some files by some keywords or properties.

With the efficient storage broker system, the visitors can joint the pervasive grid to access our ecological data via the e-Ecology system, as long as they are authorized. The students or teachers also can investigate into the ecology within NDHU for their researches.

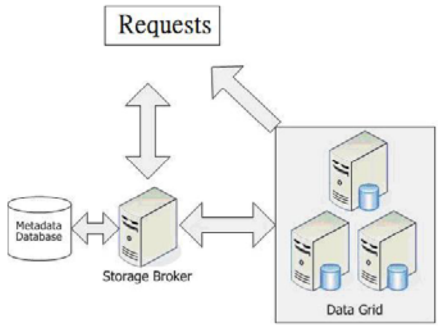

Figure 6. Storage Broker Architecture

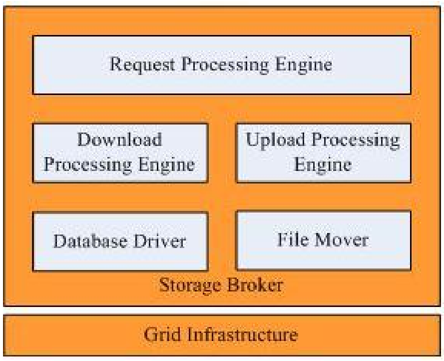

Figure 7. Storage Broker Components

In respect of e-Classroom, the video data for a course can be digitalized as well as stored in our system via

Data-Grid technology. The students can review the course by browsing the video data via e-Classroom. With the review of courses in multimedia, the teaching efficiency could be improved. In addition, the teaching data could be shared among various universities for distance learning, so as to achieve the objective of the share of education resources.

4. System Implementation and Evaluation

4.1. System Implementation

We have implemented an integrated portal of e-Campus with a kind user interface called NDHU Grid Client, as shown in Figure 8. NDHU Grid Client is very friendly towards students and teachers even if they have little knowledge of computer. Several functions are integrated in this portal, including user's certification tool, GridFTP transmission tool, grid job tool and e-Campus applications. One application is created by an internal frame as an independent thread. Each job will not influence other jobs by means of multi-threding programming model.

In the matter of e-Campus applications, take e-Ecology for example, if we feel like browsing an ecological video file via e-Ecology, we should connect the storage broker at first, as shown in Figure 9. Then we input the LFN (Logical File Name) of the video file. The storage broker will search an optimal site containing this file to download via GridFTP. GridFTP supports parallel data transfer using multiple TCP streams to improve the bandwidth over a single one. We make use of parallel data transfer to shorten the waiting time for users. After the transmission, the ecological file is presented through e-Ecology interface, as shown in Figure 10.

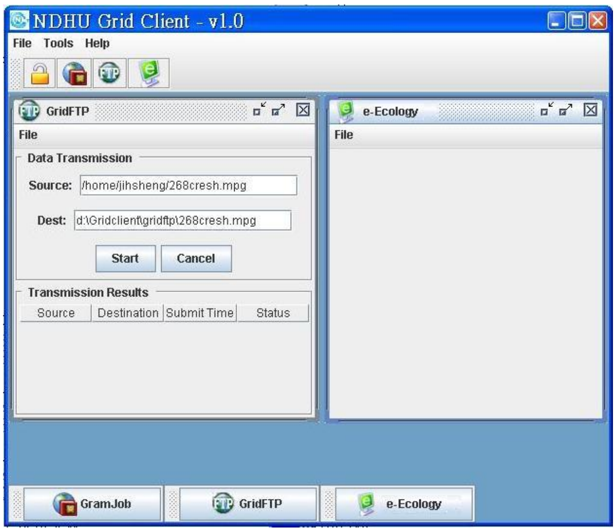

Figure 8. NDHU Grid Client

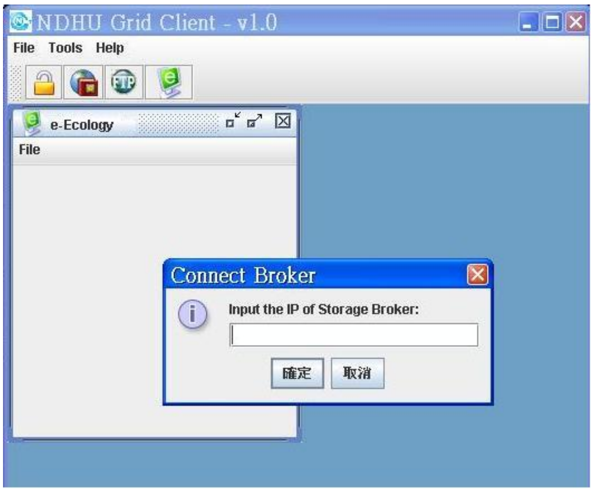

Figure 9. Connection with the Storage Broker

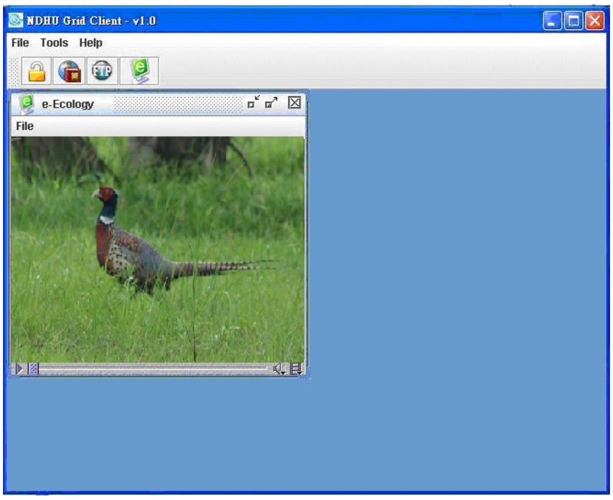

Figure 10. e-Ecology

4.2. Performance Evaluation and Analyses

GridFTP supports parallel data transfer using multiple TCP streams for better performance. We adapt the parallel data transfer to our system, in order to shorten the download time. Increasing the parallelism of transmissions seems to achieve better performance; it may lead to more computing overhead on account of too many working threads in your system. We have experimented on a variety of the number of TCP data streams from one stream to six, for downloading a video file with 500 Mega Bytes, with a view to determine the appropriate parallelism value. The result is shown in Figure 11. It is found from the result that data stream of three is superior to the others. Therefore, we adopt the parallelism as three data streams in our implementation.

We have also made experiments on the comparison between the conventional transmission with single stream and the parallel transmission. As shown in Figure 12, the result indicates that our transmission model can outperform the conventional one. Users can obtain excellent browsing quality for large size video data via e-Campus system.

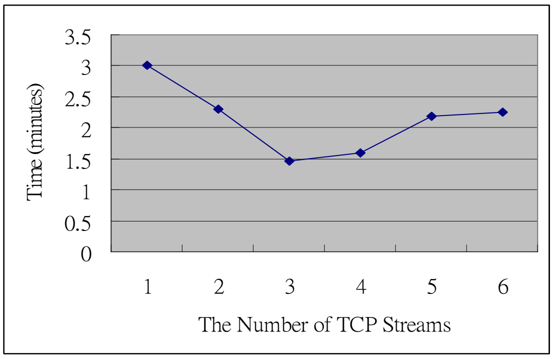

Figure 11. Evaluation of A Variety of The Number of TCP Data Streams

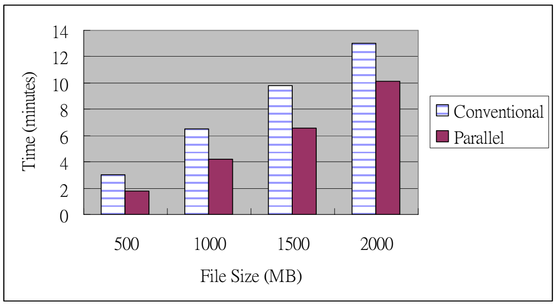

Figure 12. Comparison Between Conventional and Parallel Transmissions

5. Conclusions and Future Work

We have developed the e-Campus system based on the pervasive grid architecture, in order to build the digital grid applications, e-Ecology and e-Classroom. We also develop a user portal by means of Java CoG Toolkit to simplify the interconnection and operations for users.

As for our future work, we would like to develop more applications based on the e-Campus system in an attempt to make our system more useful and powerful.

Acknowledgements: This research is supported by NSC under contract numbers NSC94-2213-E-259-005 and NSC95-2422-H-259-001.

References

[1] W. Hoschek, J. Jaen-Martinez, A. Samar, H. Stockinger, K.Stockinger, “Data Management in an International Data Grid Project”, *IEEE/ACM International Workshop on Grid Computing (Grid’2000)*, London, UK, December 2000.

[2] Goldberg, D. E., "Genetic Algorithm: In Search, Optimization and Machine Learning," Addison-Wesley, 1989.

[3] Ian Foster, Carl Kesselman, S. Tuecke, “The Anatomy of the the Grid: Enabling Scalable Virtual Organization,” *International Journal of Supercomputer Applications*, Vol. 15, NO. 3, pp.200-222, 2001.

[4] IBM Grid Computing, http://www-1.ibm.com/grid/

[5] The Globus Alliance, http://www.globus.org/

[6] CoG Toolkit, http://www.cogkit.org/

[7] GridFTP, http://www.globus.org/toolkit/docs/4.0/data/gridftp/

[8] Globus: Grid Security Infrastructure (GSI), http://www.globus.org/security/

[9] Globus: The Grid Resource Allocation and Management (GRAM), http://www.globus.org/toolkit/docs/3.2/gram/

[10] Replica Location Service (RLS), http://www.globus.org/toolkit/docs/4.0/data/rls/

[11] Information Services, http://www.globus.org/toolkit/mds/

[12] Junseok Hwang, Praveen Arvamudham, “Middleware Services for P2P Computing in Wireless Grid Networks”, *IEEE Internet Computing*, pp.40-46, August 2004.

[13] S. H. Srinivasan, “Pervasive Wireless Grid Architecture”, *The Second Annual Conference on Wireless On-demand Network Systems and Services (WONS’05)*, 2005.

[14] eTaiwan Program, http://www.etaiwan.nat.gov.tw/

[15] National Digital Archives Program, http://www.ndap.org.tw/

[16] National Science and Technology Program for e-learning, http://elnp.ncu.edu.tw/

[17] e-Government, http://www.gov.tw/

[18] e-Transportation, http://e-trans.iot.gov.tw/query/

[19] e-Business, http://sme.nat.gov.tw/

[20] California, Blue Cross to implement telemedicine, http://www.govhealthit.com/article90967-09-30-05-Web

[21] eEurope 2005 action plan, http://europa.eu.int/information_society/eeurope/2005/index_en.htm

[22] Java Technology and Web Services, http://java.sun.com/webservices/

[23] Eric Newcomer, “Understanding Web Services”, 3th edition, Addison-Wesley, 2002.

[24] Steve Graham, Simeon Simeonov, Toufic Boubez, Doug Davis, Glen Daniels, Yuichi Nakamura, Ryo Neyama, “Building Web Services with Java”, SAMS, 2001.

[25] GSI (Globus Security Infrastructure), http://www.globus.org/Security/

[26] Ann Chervenak, Lan Foster, Carl Kesselman, Charles Salisbury and Steven Tueckem, “The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Data Sets”, *Journal on Network and Computer Applications*, Vol.23 (3), July 2000, pp.187-200.

[27] The EU Data Grid Project, http://www.eu-datagrid.org/.