Towards Building a Semantic Grid for E-Learning

Wenya Tian[1,2] and Huajun Chen[2]

[1] Information Technology Department, Zhejiang Economic & Trade Polytechnic, Hangzhou 310018, China
[2] Grid Computing Lab, College of Computer Science, Zhejiang University, Hangzhou 310027, China
Tianweny@163.com, huajunsir@zju.edu.cn

Abstract. In an E-learning scenario, educational resources, such as course documents, videos, test-bases, courseware, and teacher information etc., are shared across different schools. DartGrid is built upon several techniques from both Semantic Web and Grid research areas, and is intended to offer a semantic grid toolkit for data integration. In this paper, a Semantic Grid for E-leaning based on DartGrid is introduced, and it provides a Semantic-based distributed infrastructure for E-learning resource sharing. We explore the essential and fundamental roles played by RDF semantics for e-learning, and implement a set of semantically enabled tools and grid services for E-learning such as semantic browser, ontology service, semantic query service, and semantic registration service.

1 Introduction

Facilities to put machine-understandable data on the Web are becoming a high priority for many communities. The Semantic Web is an effort to improve the current Web by making Web resources machine-understandable because current Web resources do not reflect machine-understandable semantics [5,6]. The Semantic Web [3] provides a common framework that allows data to be shared and reused across applications, enterprises, and community boundaries. It is based on the Resource Description Framework (RDF), which integrates a variety of applications using XML as syntax and URIs for naming.

The Grid [1] is aimed to connect a wide variety of geographically distributed resources such as Personal Computers, workstations and clusters, storage systems, data sources, databases and special purpose scientific instruments and presents them as an integrated resource, and it is a technology that enables distributed computing resources to be shared, managed, coordinated, and controlled.

The Semantic Grid [4] is an Internet-centered interconnection environment that can effectively organizes, shares, clusters, fuses, and manages globally distributed versatile resources based on the interconnection semantics. In short, the Semantic Grid [7] vision is to achieve a high degree of easy-to-use and seamless automation to facilitate flexible collaborations and computations on a global scale, by means of machine-understandable knowledge both on and in the Grid.

F. Eliassen and A. Montresor (Eds.): DAIS 2006, LNCS 4025, pp. 128 – 137, 2006.

In an E-learning scenario nowadays, educational resources, such as course documents, videos, test-bases, courseware, and teacher information etc, are shared across different colleges. Typically, teachers from many colleges in different district collaborate with each other for teaching. E-learning is the result of the development of modern information technology, and it is the primary method of building life long people education system during the knowledge economy age. E-learning gives students the freedom to study anytime and anywhere and is widely developed and deployed in our country recently. In E-learning, we often need to integrate E-learning services across distributed, heterogeneous, dynamic "virtual organizations" formed by the disparate education resources within a single enterprise and/or from external education resource sharing via service provider relationships. This integration can be technically challenging because of the need to achieve various qualities of E-learning service when running on top of different scholastic platforms.

DartGrid[1] is a data integration toolkit using technologies from semantic web and grid, and is intended to offer a generic semantic infrastructure for building database grid applications. Roughly speaking, DartGrid is a set of semantically enabled tools and grid services such as *semantic browser, semantic mapping tools, ontology service, semantic query service, semantic registration service*, that support the development of database grid applications.

In this paper, a Semantic Grid for E-leaning based on DartGrid is introduced, and it provides a Semantically distributed infrastructure for E-learning scenarios aforementioned. We explore the essential and fundamental roles played by RDF semantics for E-learning grids, and implement a set of semantically enabled tools and grid services for E-learning resource sharing such as semantic browser, ontology service, semantic query service, and semantic registration service.

This paper is outlined as below: Section 2 introduces the architecture and the core components of a Semantic Grid for E-leaning from technical perspective. Section 3 introduces a working scenario for the E-learning grid application. Section 4 mentions some related works. Section 5 gives the summary.

2 Layered Architecture and Core Grid Services

2.1 Technical Approach

2.1.1 RDF

At the present time, the most popular languages for representing data semantics are RDF framework and OWL language, which is proposed in Semantic Web research area and standardize-ed by W3C organization. The Resource Description Framework (RDF) is a language for representing web information in a minimally constraining, extensible, but meaningful way.

The RDF structure is generic in the sense that it is based on the directed acyclic graph (DAG) model. RDF is based on the idea of identifying things using Web identifiers (called Uniform Resource Identifiers, or URIs), and describing resources in terms of simple statements about the properties of resources. Each statement is a

[1] DartGrid Official Website: http://ccnt.zju.edu.cn/projects/dartgrid

triplet consisting of a subject, a property and a property value (or object). For example, the triple ("http://example.org", ex:createdBy, "Wenya") has the meaning of "http://www.example has a creator whose value is Wenya".

RDF also provides a means of defining classes of resources and properties. These classes are used to build statements that assert facts about resources. While the grammar for XML documents is defined using DTD or XSchema, RDF uses its own syntax (RDF Schema or RDFS) for writing a schema for resources. RDFS is expressive and it includes subclass/superclass relationships as well as constraints on the statements that can be made in a document according to the schema. The generic structure of RDF makes data interoperability and evolution easier to handle different types of data can be represented using the common graph model, and it offers greater value for data integration over disparate web sources of information. OWL is an extension of RDF/RDFS and supports more sophisticated knowledge representation and inference.

In our work, RDF is used to describe E-Learning data semantics.

2.1.2 OGSA/WSRF and the Globus Toolkit

OGSA/Web Service Resource Framework focuses on service- oriented architecture for grid application. In a grid, computational resources, storage resources, networks, programs, databases, and the like are all represented as services. A service-oriented view allows us to address the need for standard interface definition mechanisms, local/remote transparency, and adaptation to local OS services, and uniform service semantics.

The open source Globus Toolkit [20] is a fundamental enabling technology for the "Grid," letting people share computing power, databases, and other tools securely online across corporate, institutional, and geographic boundaries without sacrificing local autonomy. The toolkit includes software services and libraries for resource monitoring, discovery, and management, plus security and file management.

The Globus Toolkit is built to remove obstacles that prevent seamless collaboration. Its core services, interfaces and protocols allow users to access remote resources as if they were within their own machine room while simultaneously preserving local control over who can use resources and when. The toolkit components that are most relevant to OGSA are the Grid Resource Allocation and Management (GRAM) protocol and its “gatekeeper” service, which provides for secure, reliable, service creation and management [22]; the Meta Directory Service (MDS-2) [21], which provides for information discovery through soft state registration [23, 11], data modeling, and a local registry (“GRAM reporter” [22]); and the Grid Security Infrastructure (GSI), which supports single sign on, delegation, and credential mapping.

In our work, the E-Learning services conform to the OGSA/WSRF specification, and are implemented upon Globus 4 toolkit. Globus 4 is also used as the service container for the E-learning grid application.

2.2 Layered Architecture

Fig.1. illustrates the layered architecture of E-learning Semantic Grid.

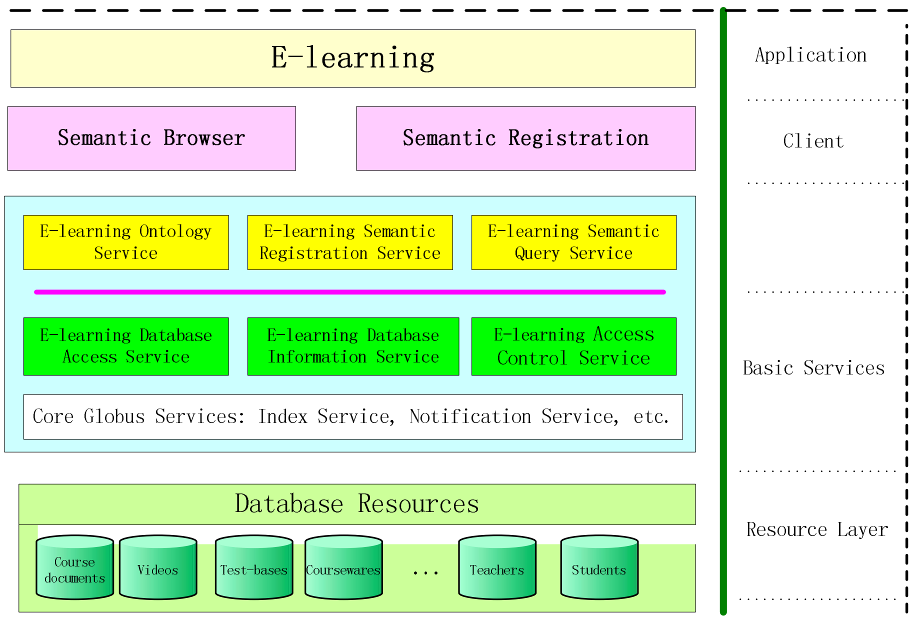

Fig. 1. Layered Architecture of E-learning Semantic Grid

At the basic service layer, three services are implemented.

1. E-learning Database Access Service. It supports the typical remote operations on educational resource contents, such as course documents, videos, test-bases, courseware, and teacher information etc. It also includes querying an education resources, insertion an education resources, deletion an education resources, and modification an education resources.

2. E-learning Database Information Service. It supports inquiring about meta information of the educational data resources such as DBMS descriptions, privilege information, statistics information that includes CPU utilization, available storage space, active session number etc..

3. E-learning Access Control Service. This service is developed for access control in E-learning Semantic Grid. For example, it provides the service of authorizing or authenticating students to access courseware resource.

We mainly contribute to the semantic service level. The services at this level are mainly designed for RDF-based relational schema mediation and semantic query processing.

1. E-learning Ontology Service. This service is used to expose the shared ontologies that are defined by RDF/OWL languages. The ontologies are used to mediate heterogeneous relational databases. For example, there are two parts in the courseware ontology. One part is defined based on CELTS or IMS. The other part is

defined as an extended set (Fig. 3.).The core set of CELTS has 11 elements as follows: Title, Subject, Keywords, Description, Identifier, Format, Date, Language, Type, Creator and Audience. The definition and determinant of these elements see also CELTS40 [18]. The extended set involves general architecture information class (FRAME) and page information class (PAGECONTENT).

2. E-learning Semantic Registration Service. Semantic registration establishes the mappings from source relational schema to sharing RDF ontologies. Semantic Registration Service maintains the mapping information and provides the service of registering and inquiring about this information. For example, it provides the service that enables teacher registering courseware and student inquiring about registration information of courseware.

3. E-learning Semantic Query Service. This service accepts RDF semantic queries, inquires of Semantic Registration Service to determine which databases are capable of providing the answer, then rewrites the RDF queries according to relational schema, namely, the RDF queries will be ultimately converted into a set of SQL queries. The results of SQL queries will be wrapped by RDF/OWL semantics and returned as RDF triples.

3 Working Scenario

3.1 Typical Use Cases of E-Learning Semantic Grid

Generally, there are two kinds of user roles in E-learning Semantic Grid, they are: Local Database Administrator (such as teachers), and Normal User (such as students). Fig.2. illustrates the relationship between these user roles and the core components of E-learning Semantic Grid.

Local Database Administrator (such as teachers). Education resources can be dynamically added into the sharing cycle of an e-learning semantic grid. E-learning Semantic Grid provides the education resource provider (such as a teacher) with a Semantic Mapping Tool. After a database grid service is setup, the teacher can use semantic mapping tool to register his database to the semantic grid. Typically, the mapping tool retrieves the e-learning ontologies from ontology service, and gets the relational schema from database grid service. Then the DBA can visually map the relational schema to e-learning ontologies. For example, the process that a teacher registers a courseware to the Semantic Registration Service is as follows:

1. Obtain the local database resource schema;
2. Obtain the domain ontologies on the ontology service;
3. Establish the semantic relational mapping between the local database resource schema and the sharing ontologies;
4. Submit the registering information to the semantic registration service.

Section 3.2.2 introduces the semantic mapping tool in more details.

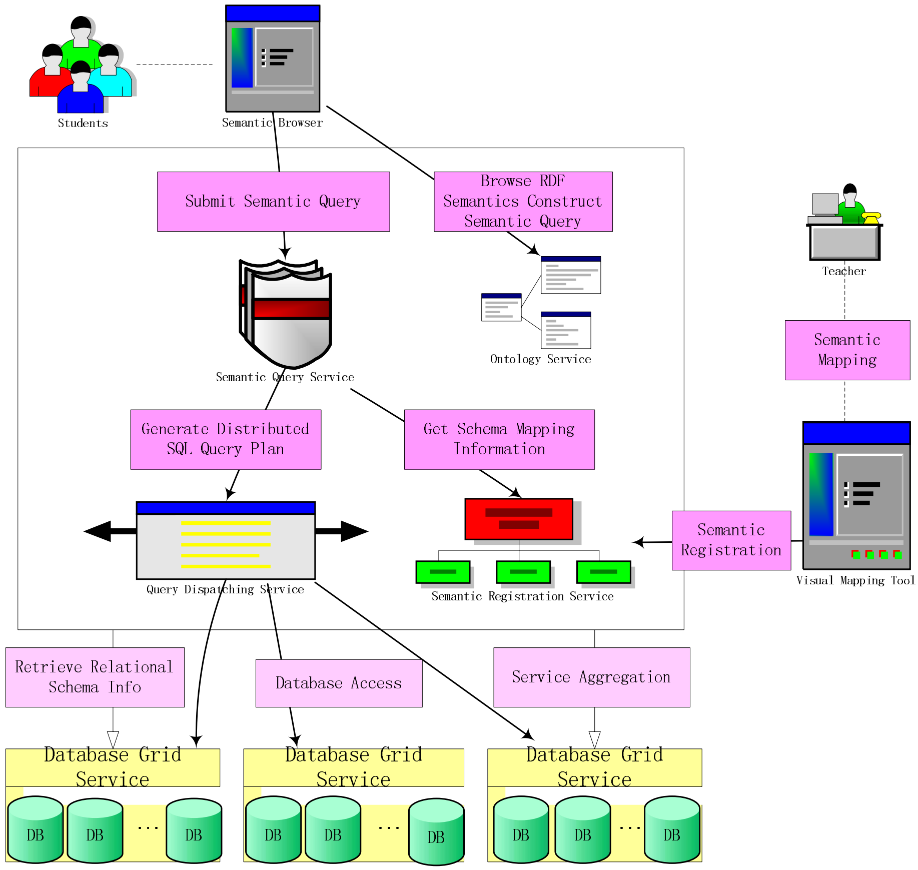

Fig. 2. Mapping of E-learning Semantic Grid

Normal User (such as a student) for normal users, E-learning Semantic Grid offers an intelligent user interface called Semantic Browser [8]. It is a visual interface enabling the user to graphically browse the RDF/OWL semantics and visually construct a RDF semantic query. For example, the process that a student inquires about a courseware is as follows:

1. User browser the e-learning ontology by using the semantic browser;
2. User visually construct a semantic query;
3. User submit the query to a semantic query service;
4. The semantic query service accesses the Semantic Registration Service to query the workable database resource and gets the schema mapping information;
5. The semantic query service generates the distributed SQL Query plan;
6. The semantic query service gets the data information from the idiographic database;
7. Return the result to the student.

Section 3.2.1 gives an example about how to construct a semantic query using this tool.

3.2 Semantic Tools for E-learning Semantic Grid

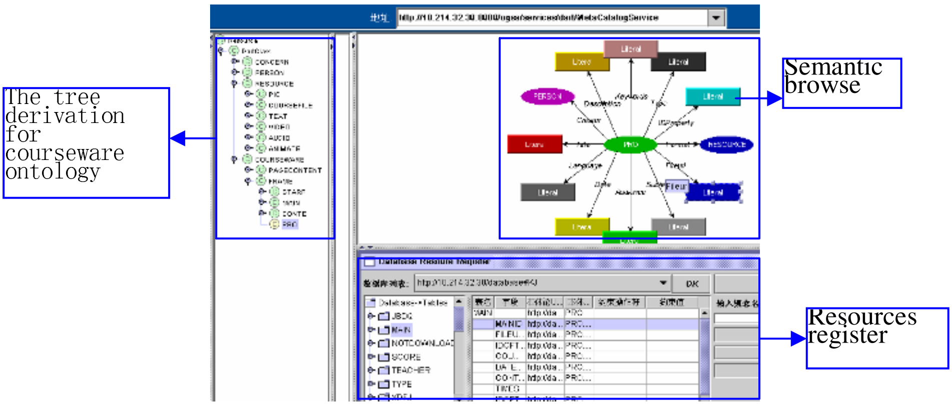

Fig. 3. Semantic Browser and Semantic Registration Tool For E-learning Semantic Grid

3.2.1 Semantic Browser

E-learning Semantic Grid offers a semantic browser [8] enabling user to interactively specify a semantic query. Users can search education information in E-learning semantic grid. Although the information come from all nodes, it is done transparently to users as if the user operates on the same database of the same computer.

For example, a user wants to search courseware with subject as “Java” from E-learning Semantic Grid. Fig 3 illustrates an example with our application. It showcases how users can step-by-step specify a semantic query to find out those needed courseware. In the first step (the left part of the Fig.3.), the user selects the courseware class and its subject property. In the second step, the user inputs a constraint which specifies that the name of the subject as "Java”. The Semantic Browser constructs the semantic query automatically. Here the query language is Q3 [15], a database query language like SQL semantically defined by Grid Computing Lab of ZheJiang University. It is a visual process to write q3 in Semantic Browser.

The Q3 language is as following:

```
[q3:context
   [q3:prefix(tcm:http://dart.zju.edu.cn/tcm)
    q3:variable{
      ?x1 a tcm:PRO
    }
   ]
 q3:pattern  ()
 q3:constraint{
   ?x1.tcm:Subject="Java"
 }]
```

At last, the result (the above of the right part of Fig.3.) records come from all databases of different nodes.

3.2.2 Semantic Registration

The task of defining semantic mapping from local relational schema to RDF ontologies is burdensome and erroneous. E-learning Semantic Grid offers a visual tool to facilitate the task of defining semantic mappings. As Fig.3 displays, the user can use the registration panel (the below of the right part in the fig.3) to view the table and column definition of the relational database, and use the semantic browsing panel (the above of the right part in the fig.3) to browse the RDF ontologies graphically. The user can then specify which RDF class one table should be mapped onto and which RDF property one table column should be mapped onto. After finishing the mapping, the tool automatically generates a registration entry in RDF/XML format, and submits it to the semantic registration service.

For example, a teacher wants to register courseware recourses. The Semantic Registration tool directly registers courseware resources to the Semantic Registration Center. It is a semantic mapping from the local courseware resources to the sharing semantic ontology.

During the registration, mapping information is written into a semantic registry. The courseware resource content itself is not uploaded to the registration center or any other centralized node. When the user searches resources, the Semantic Registration Center will look up the result from the resource registration table. User will download and browse the corresponding resources from the data node by linking it directly. The way it works is very similar to other P2P mode. The registration interface shows as Fig.3.

4 Related Work

There are a lot of relevant works. Within the domain of Grid research, there are many efforts about accessing and integrating e-learning database under the grid framework. Typical example is Realcourse [2]. Realcourse is a successful application of distributed computing [13] technology in a geographically wide area. Different from some traditional distributed fault-tolerant services like ISIS [14], Realcourse emphasizes giving clients access to the service with reasonable response time. For most cases, it means as much of the time as possible.

In [12], it is clear that standards like LOM, or Dublin Core are gaining importance. They provide more information on the learning material that is to be found in the web. However, their simple structure prevent them being used for modeling more complex knowledge. [10] Explains how Semantic Web technologies based on ontologies can improve different aspects of the management of E-Learning resources. Indeed, ontologies are a means of specifying the concepts and their relationships in a particular domain of interest. Web Ontology languages, like OWL, are specially designed to facilitate the sharing of knowledge between actors [17] in a distributed environment. We wish to emphasize here that Web Ontology languages have various advantages.

The significant difference, compared with others, is the RDF-based and semantic-web-oriented approach adopted in the Semantic Grid for E-learning. The Semantic Grid for E-learning complements those efforts with a semantic infrastructure for building database grid application and this infrastructure can provide information and knowledge services as other conventional portals. In addition, the use of multiple servers can semantically assist users in formulating their problem description, searching possible solutions on the Grid.

5 Summary and Future Work

The Semantic Grid will play a very important role for the wide acceptance of the Grid [9]. It will provide enhanced support for end users to access heterogeneous Grid services and resources by understanding their domain problems and providing solutions. We present a Semantic Grid for E-leaning based on DartGrid, and also put forward a dynamic, extensible Semantic-based distributed infrastructure for E-learning scenarios. We explore the essential and fundamental roles played by RDF semantics for e-learning resource sharing, and implement a set of semantically enabled tools and grid services for E-learning such as semantic browser, ontology service, semantic query service, and semantic registration service.

There are more works need to be done in this area. Semantic Grid for E-leaning, a DartGrid application, has many obvious attributes as a good test bed. As a typical DartGrid application by its nature, it stores various data classes that can be collected easily. Plus, the test result can be verified easily. The system needs to be further tested with more data classes and more grid nodes. More features are needed for the education resource management. In the meantime, DartGrid itself also needs to be continuously improved for perfection. Now we have a working prototype of an open education resource management system. The next step is to make it more powerful by fine-tuning its operability. As far as education is concerned, it is important to manage all education resources via the semantic grid for E-learning.

Acknowledgements

We gratefully acknowledge helpful discussions with other members in the Grid Computing Lab of Zhejiang University. This work is co-funded by subprogram of China 973 project (NO. 2003CB316906), a grant from Program for New Century Excellent Talents in University of Ministry of Education of China (NO. NCET-04-0545), China NSF program (NO. NSFC60503018), and Zhejiang Provincial Natural Science Foundation of China (NO. Y105463) .

References

1. I. Foster, C. Kesselman, The Grid, Blueprint for a New Computing Infrastructure, Morgan Kaufmann, San Francisco, USA, 1998
2. Jinyu Zhang, Xiaoming Li, The Model, Architecture and Mechanism Behind Realcourse, ISPA 2004, LNCS 3358, pp. 615–624, 2004.

3. http://news.11138.com/2001/sw/
4. H Zhuge, Semantic Grid: Scientific Issues, Infrastructure, and Methodology, Communication of the ACM Vol. 48, No. 4(2005) 197
5. T. Berners-Lee, J. Hendler, O. Lassila, Semantic Web, Sci. Am. 284 (5) (2001) 34–43.
6. J. Hendler, Agents and the semantic web, IEEE Intell. Syst. 16 (2) (2001) 30–37.
7. D. de Roure, N. R. Jennings and N. Shadbolt (2003) "The Semantic Grid: A future eScience infrastructure" Int. J. of Concurrency and Computation: Practice and Experience 15 (11)
8. Yuxin Mao, ZhaohuiWu, Huajun Chen. Semantic Browser: an Intelligent Client for Dart-Grid. Proceedings of International Conference on Computational Science (Lecture Notes in Computer Science, vol. 3036) , Springer:Berlin, 2004; 470-473.
9. M. Li, P. van Santen, D.W. Walker, O.F. Rana, M.A. Baker,SGrid: a service-oriented model for the Semantic Grid, Future Generation Computer Systems 20 (2004) 7–18
10. Ljiljana Stojanovic, Steffen Staab, and Rudi Studer. E-learning based on the semantic web. In WebNet2001 - World Conference on the WWW and Internet, Orlando, Florida, USA, 2001.
11. Zhang, L., Braden, B., Estrin, D., Herzog, S. and Jamin, S., RSVP: A new Resource ReSerVation Protocol. In IEEE Network, (1993), 8-18
12. Jan Brase and Wolfgang Nejdl. Ontologies and Metadata for eLearning, pages 579–598. Springer Verlag, 2003.
13. George Coulouris, Jean Dollimore, Tim Kindberg, Distributed System Concepts and Design (Third version), ISBN 7-111-11749-2, China Machine Press
14. Birman, K.P. (1993). The process group approach to reliable distributed computing. Comms. ACM, Vol. 36, No. 12, pp. 36-53
15. Huajun Chen, Zhaohui Wu, Guozhou Zheng, Yuxing Mao. DartGrid: a Semantic-based Approach for Data Integration Using Grid as the Platform
16. Christian Bizer, Andy Seaborne. D2RQ -Treating Non-RDF Databases as Virtual RDF Graphs. Presented at the 3rd International Semantic Web Conference (ISWC2004),November 2004.
17. Rudi Studer Steffen Staab, Hans-Peter Schnurr and York Sure. Knowledge processes and ontologies. IEEE Intelligent Systems, 16(1), 2001.
18. China ELearning Technology Standardization Committee, Education Informationize Technology Standard http://www.celtsc.edu.cn sub standard CELTS-40
19. Foster, I. and Kesselman, C. Globus: A Toolkit-Based Grid Architecture. In Foster, I. and Kesselman, C. eds. The Grid: Blueprint for a New Computing Infrastructure, Morgan Kaufmann, 1999, 259-278.
20. http://www.globus.org/toolkit/about.html
21. Czajkowski, K., Fitzgerald, S., Foster, I. and Kesselman, C., Grid Information Services for Distributed Resource Sharing. In 10th IEEE International Symposium on High Performance Distributed Computing, (2001), IEEE Press, 181-184
22. Czajkowski, K., Foster, I., Karonis, N., Kesselman, C., Martin, S., Smith, W. and Tuecke, S. A Resource Management Architecture for Metacomputing Systems. In 4th Workshop on Job Scheduling Strategies for Parallel Processing, Springer-Verlag, 1998, 62-82.
23. Raman, S. and McCanne, S. A Model, Analysis, and Protocol Framework for Soft State-based Communication. Computer Communication Review, 29 (4). 1999.