

**International Journal of Innovative Research in Science,
Engineering and Technology**

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 9, September 2013

A Grid-based Model for Integration of Geographically Distributed & Heterogeneous Educational Resources for Knowledge Extraction & Delivery

Dr. Sanskruti Patel¹, Dr. Priti Sajja²

Assistant Professor, Department of Computer Science and Applications, Charotar university of Science & Technology, Changa, Gujarat, India¹

Professor, Department of Computer Science, Sardar Patel university, V V Nagar, Gujarat, India²

Abstract: In a rising era of information and communication technology, data plays a crucial role in all types of cross-organizational research and business applications. Data grids rely on the coordinated sharing of and interaction across multiple autonomous database management systems to provide transparent access to heterogeneous and autonomous data resources stored in grid nodes. In this paper, we present a grid-based model which provides a uniform access interface and distributed query mechanism to access heterogeneous and geographically distributed educational digital resources. We first present the overview of the grid-based model and then discussed the architectural view and implementation details in regards of educational resources.

Keywords: Grid, Data Grid, knowledge-based System, Agent, Multi-agent System

I. INTRODUCTION

Emerging modern e-science and e-business applications require vast amounts of data which are being stored in many different formats by using different storage solutions. Currently, the increasing social demands of high quality data resources to process scientific and business applications cannot be fulfilled only by the available data resources which generally reside in the individual administrative domain. This has created a demand for solutions which offers the data access and integration. Thus data plays a fundamental role in such kinds of cross-organizational research and alliances, these organizations can be jointly deemed to form virtual organizations (VOs). Size of data, diversity of data sources and data freshness requirements makes data integration a complex task and this data may be geographically distributed. It is impossible to handle such large and diverse amounts of data within a single organization. Therefore, the need for the integration of database technologies with grid middleware is widely recognized.

Apart from this, organizations generally strive for the data for analysis and decision making purposes. Knowledge-based system is used to enhance the performance and reliability in the decision making process. The grid-based model presented here provides the different services like uniform query interface, metadata management, distributed query processing etc. thorough multiple collaborative agents. An agent is a piece of software program which acts proactively on behalf of the user [1]. Also, the knowledge-based component is incorporated on the top of the distributed database grid to provide knowledge extraction, presentation and delivery. It provides a great support to executives for decision making in any research and commercial applications.

II. AN OVERVIEW OF A GRID-BASED MODEL

While many modern academic applications (such as digital library, e-learning, etc.) involve accessing heterogeneous databases which are located in a geographically distributed environment, there is an increasing demand on integration of different data resources across different information systems or administrative domains to provide a uniform access interface for users. Emerging as a new paradigm for distributed computing, grid is defined as the secure, flexible, coordinated resource sharing mechanism among dynamic collections of individuals, institutions, and resources [2]. As

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 9, September 2013

a prospective solution for managing and collaborating distributed resources, data grid deals with the issue of heterogeneity by developing uniform and standardized interfaces for data access and integration of dissimilar data resources [3].

Universities often use information systems for operational and strategic tasks. These information systems may have been developed using different protocols, technologies, architectures and platforms. These systems containing a large number of digital objects that are stored in many different storage systems and data formats with differences in their schema, access rights, metadata attributes and so forth. In such scenario, it is a very difficult and challenging task to access these digital objects. For that, this case study proposes a grid-based model for sharing these academic resources in the heterogeneous environment. This model will improve the accessibility, integration, controlling and management of these academic resources.

The goal of the model is to provide a uniform query and browse mechanism for multiple distributed university databases. It shields their differences in data type, schema or site, and provide a new solution for the research on integration of university databases. There are good opportunities for exploiting synergies between grid and agents. The grid-based model provides a uniform and standardized interface which can be used to represent an abstract view of data resources, which can permit homogeneous access to heterogeneous university databases. Therefore, the said model is able to provide easy and transparent access to distributed heterogeneous resources span across different organizations and administrative domains.

III. AN ARCHITECTURAL VIEW OF A GRID-BASED MODEL

Fig 1 shows the grid-based model for integration of geographically distributed & heterogeneous educational resources for knowledge extraction & delivery.

The university is having a large number of digital objects stored in various distributed data repositories which may span across different organizations and administrative domains like Digital Library, R & D centre, Administrative section etc.

To establish a data grid environment, a data grid middleware technology named OGSA-DAI (Open Grid Services Architecture – Data Access and Integration) has been used. The OGSA-DAI has been studied and set of mechanisms have been identified to use data grid services. Open Grid Services Architecture - Data Access and Integration (OGSA-DAI) is a project that develops middleware solution to assist with access and integration of data from separate data resources via the data grid [4]. The project was funded by the UK e-Science core program and is working closely with the Global Grid Forum Data Access and Integration Services (GGF DAIS) work group [5]. OGSA-DAI is a database-centric middleware solution which mainly provides the data virtualization services.

Data grid services should be offered with minimum user intervention. To facilitate the access of data grid middleware services, we have implemented different collaborative agents those are working together to provide the core, user & knowledge level services to the users. Users are interacting with agents in order to access OGSA-DAI services. Agents are autonomous, adaptive and cooperative by nature and emerging as dynamic, flexible and extensible mediators for facilitating grid services in a data grid environment. Agents are working proactively with minimum user intervention and offer several data grid services.

We have used multiple agents which provide the user level services to the client applications and users. The users have to first interact with university portal to retrieve the information or knowledge stored in the form of digital objects. The grid-based model incorporates an agent-based framework which allows a university domain to be considered, designed and developed as an information environment being composed of different co-operating agents. Apart from this, it provides access to database nodes via the OGSA-DAI server. Thus, the databases need to be opened only for the OGSA-DAI server(s). In a grid environment, where an application can run on a large number of machines, this is the only way to provide better security for database providers.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 9, September 2013

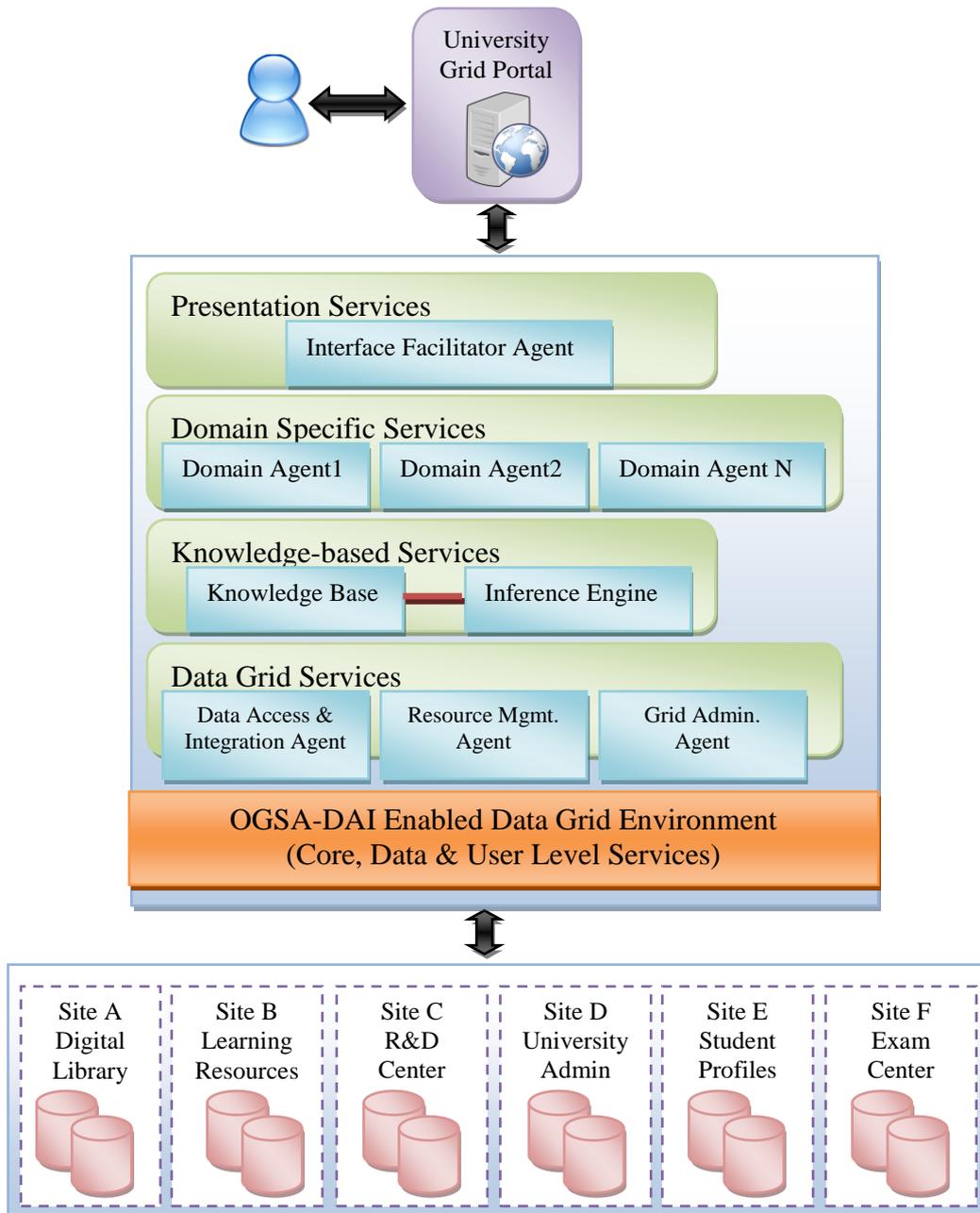


Fig. 1 A Grid-based Model for Integration of Geographically Distributed & Heterogeneous Educational Resources for Knowledge Extraction & Delivery

In multi-agent system environment, multiple collaborative agents are working together to accomplish the tasks. These agents are part of a multi-agent system and they are collaborating with each other in order to provide the services. These agents are triggered when a request is sent by a user or the state of the data grid is changed. As shown in the Fig 2, agents need to collaborate with each other in order to provide core, data and user level services.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 9, September 2013

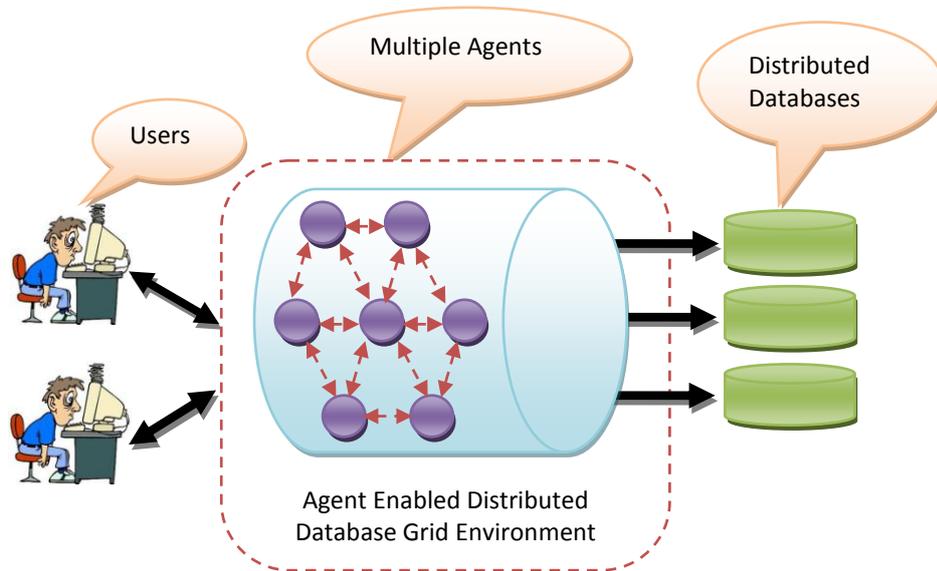


Fig. 2 Multiple Collaborative Agents in Distributed Database Grid Environment

Currently, the multi-agent system environment provides four types of typical services namely data grid services, knowledge-based services, domain services and presentation services.

A. DATA GRID SERVICES

The agent-enabled data grid model offered core, data and user level services. Through these services, the client applications and users may get access to digital objects of the data grid environment. The typical agents namely resource management agent, data access & integration agent and grid administrative agent provide the above said services.

The resource management agent is mainly responsible for the data grid node establishment through user-friendly and automated way. It hides the complexity of development and execution of the configuration script from the user and provides a standardized interface to manage data grid nodes.

Data access & integration agent Interacts with OGSA-DAI and allows data resources to be accessed, manipulated, integrated and federated across the network such as LAN or WAN. It gets the query execution request from the user and executes the workflow on the OGSA - DAI server. Also, it converts the data into most appropriate means. It also makes the union of query result if necessary and sends the response back to the user.

Grid administrative agent monitors the data grid nodes. For that, it executes a query to each master database node at every specific time interval. It shows the active data grid nodes and automatically switch between original and backup nodes as and when necessary to avoid single point of failure.

B. KNOWLEDGE-BASED SERVICES

The knowledge-based services are offered by integrating a typical agent. A knowledge-based system is a system that uses artificial intelligence techniques in problem solving processes to support human decision-making, learning, and action. The knowledge-based system is used to provide intelligent decisions with justification. Knowledge is acquired and represented using various knowledge representation techniques such as rules.

The knowledge-based system consists of mainly two components: knowledge base and inference engine. A knowledge base is a repository of domain knowledge and meta knowledge. The inference engine is a software program, which

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 9, September 2013

infers the knowledge available in the knowledge base. The domain agents of the model provide effective and seamless information retrieval and the typical agent provides knowledge-based services uses AI mechanism to turn this information into something more i.e. knowledge. Knowledge is derived from information in a similar way as information is derived from data. The knowledge base can be used as a repository of knowledge in various forms. The Inference Engine infers the knowledge that is available in the knowledge base. Fuzzy set theory and fuzzy logic are a highly suitable and applicable basis for developing knowledge-based systems in university for various tasks.

C. DOMAIN SERVICES

The domain services are application specific and customizable. With university environment, the typical domain agent could be user management agent, search agent, learning management agent, student profile agent, course management agent, advisory and help agent, e-communication agent, report generator agent and so forth.

The user management agent is responsible to create users in the system, assigns privileges to them, authenticates and authorizes the user by using metadata management service. Search agent retrieves specific information based on queries fired by users. It may use AI (Artificial Intelligence) mechanism for query retrieval. Learning management agent provides uniform interface for accessing learning resources to the users. Learning resources may store in heterogeneous form and on disparate location. This agent hides such differences from its underlying users and provides transparent access to them. Student profile agent maintains the students' profiles of entire university which may have different under graduate and post graduate departments. The various departments of the university may run different courses and the information about these courses are stored in digital form. The course management agent provides details about various courses to learners and visitors. It also provides an interface between instructor and system to manage and rearrange course content. Advisory and help agent provide support and help to instructors and learners. Also, they incorporate the decision-making processes by applying AI techniques. E-communication agent provides e-mail and chat facility among the users for communication. The report generator agent generates various reports by combining and using information stored in several distributed databases.

D. PRESENTATION SERVICES

The presentation services are provided by the interface facilitator agent. It provides an effective presentation of content and delivery mechanism. Also, it gets the request from users and sends a response back to them.

IV. IMPLEMENTATION DETAILS

We have used OGSA-DAI Axis Version 4.1 as a grid middleware service. The prerequisites to configure and use OGSA-DAI are as follows:

- OGSA-DAI require Java 1.6 or above
- OGSA-DAI Axis distributions required Jakarta Tomcat 5.5 or above
- OGSA-DAI needs ANT 1.6 or above as a build, installation and deployment tool

Along with this, the following environment variables need to be set to access OGSA-DAI services: CATALINA_HOME, OGSADAI_HOME and CLASSPATH.

To check whether the OGSA-DAI services are deployed on Jakarta Tomcat successfully or not, following are the steps to be performed.

1. Startup Tomcat.
2. Open a browser and write the following line in the address bar: <http://localhost:8080/dai/services>.
3. If OGSA-DAI is deployed properly, the following screen would be displayed which will contain a list of services and information on their operations and links to their WSDL service description documents.

We have realized agents and multi-agent system through realization of web services, we have used JAVAX-WS (also, known as JAX-WS) standard to implement web services. JAX-WS (Java API for XML Web Services) is a Java programming language APIs for creating web services. It is a part of the Java EE platform from Sun Microsystems. It is a technology for building web services and clients that communicate using XML. In JAX-WS, a web service operation invocation is represented by an XML-based protocol, such as SOAP. The SOAP protocol specification describes the encoding rules, the envelope structure, and conventions for representing web service invocations and responses. Here, calls and responses are transmitted as SOAP messages (XML files) over HTTP.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 9, September 2013

V. CONCLUSION

The work presented here provides an insight into the world of building a dynamic and scalable architecture for an agent-enabled distributed database grid which is composed of both data grid technology and agent oriented computing. The research goal of the grid-based model is to provide a uniform query and browse mechanism for multiple distributed university databases, shield their differences in data schema or location, provide a multi-agent system for management of the existing databases in a knowledge-based fashion, knowledge-based access of the domain information to enhance the utilization of scarce resources, and provide a new solution for the research on the integration of the domain databases.

REFERENCES

- [1] S. genesereth, M. Ketchpe, Software agents, Comm. of ACM. Volume 37, No. 7, Pages: 48-53, 1994.
- [2] Aaron Welch, Umut Onat, Carla Sotomayor, Service Oriented Architecture in Multi-Agent Systems retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.159.1588&rep=rep1&type=pdf> on February 2013.
- [3] Bepperling A., Lobov, A., Smit, H., Colombo, A.W., Lastra, J., Integration of Multi-Agent Systems and Service-Oriented Architecture for industrial automation, 6th IEEE International Conference on Industrial Informatics, INDIN 2008.
- [4] Karasavvas K, Antonioletti M, Atkinson M, Hong NC, Sugden T, Hume A, Jackson M, Krause A, Palansuriya C., Introduction to OGSA-DAI Services, Volume 3458, Pages:1-12, 2005.
- [5] Antonioletti M, Krause A, Paton NW, An outline of the Global Grid forum data access and integration service specifications. Data Management in Grids, 3836, doi: 10.1007/11611950_7, Pages: 71-84, 2005.