Research of Education Course Resource Sharing Technology Based on OGSA

YU Wanhai[1], GENG Peng[2], JI Qingshan[1], LIU Jianhua[2]

1. Xingtai Polytechnic College, Xingtai, China

2. The center of educational technology, Shijiazhuang Railway Institute, Shijiazhuang, China

e-mail: ywh76@163.com, sineagles@qq.com

Abstract: the OGSA-DAI Architecture and the OGSA-DAI principle were researched in the part 3**.** This paper researched a distributed education course resource data sharing system worked on Internet by OGSA-DAI technology in order to share and use the distributed and heterogeneous education course resource database. It is the Education Course Resource Sharing DataGrid (ECRSD). Finally, it researched the technology and principle of accessing the database by OGSA-DAI and the principle of realizing the sharing with the distributed heterogeneous database on the Internet.

Keywords: OGSA-DAI; education resource; GDSR; GDSF; GDS

1 Grid Technology

The grid technology is considered as the third wave of information technology after the Internet and web technology and is called as the third generation Internet[1]. At present it is in early period from the second generation Internet to the third generation Internet .The third generation Internet can be called as the knowledge grid, the education grid and the information grid. The grid technology makes use of the high-speed Internet to make all widely distributed compute resources (storage resources, data resources, information resources, database, software and other information capturing devices) into an integrity which is like a incomparably huge computer provides integrative services to users to obtain, publish, share and manage the resource. The OGSA-DAI[2] （Open Grid Services Architecture-Data Access and Integration） is a middleware run on the Globus Toolkit. The OGSA-DAI is designed to provide a simple method to access and integrate the data in the grid. Hence, a new Education Course Resource Sharing DataGrid (ECRSD) is designed in the paper by the OGSA-DAI middleware technology and the Globus Toolkit platform.

2 ECRSD Design

The ECRSD mainly is constituted by five parts: the client, the education course resource domain access service (ERDAS), the education course resource global access service (ERGAS), the register server (RS). The ERDAS and the ERGAS are a group of web services based on the Web Services standard and independently provide the supporting service for the system. The ECRSD is logically divided into three basic levels: the client layer, the service layer (ERDAS) and the resource layer (the distributed heterogeneous database). Figure 1 shows the system general structure.

The ERGAS is unique service to manage and access all the ERDAS in the system. The education course resource domain access Service provides the accessing data service and metadata service (MDS).Because the database grid is based on the data object attributes and not based on the physical location of the data, the ERDAS is designed to save all kinds of metadata which includes the meta-information, the configuration meta-information, the user meta-information and the data meta-information. The data access service is used to access and query the distributed and heterogeneous education course resources database in the domain.

The register server is used to certificate the users' privilege and returns the authority to the ERGAS.

The client can be used to query the data and the distributed heterogeneous data.

The access process is as follows: Firstly, the client accesses the ERGAS; the ERGAS accesses the register server which certificates the user's privilege and returns

Supported by Foundation of the eleventh five-year education science plan of Hebei province (06130061)

978-1-935068-06-8

the user's privilege to the ERGAS; The ERGAS according the users privilege sends the user's request to relative ERDAS; According the metadata of the user's accessing the data; The ERDAS establishes the connection with the data source by the data access service. The concrete operation on the data resource is executed by the basic database system and the database system returns the access result to the ERDAS; Finally the ERDAS sends the result to the client by the ERGAS. The client displays and processes the access results.

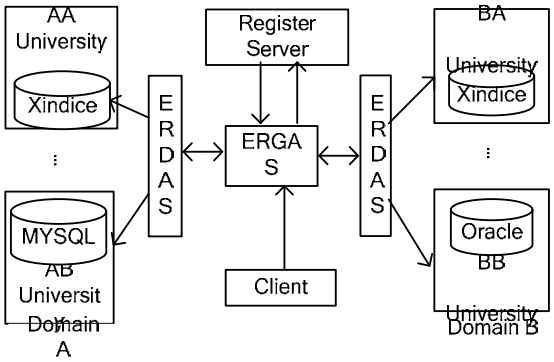

Figure 1. The principle of ECRSD system

3 OGSA-DAI Principle

The OGSA-DAI (Open Grid Services Architecture-Data Access and Integration) accords with the grid standard based on OGSA [3] and is developped based on the Globus Toolkits. The OGSA-DAI supports the some database systems such as the Oracle, Xindice, MySQL and DB2 etc. The OGSA-DAI project is devoting to build a middleware which accesses and integrates the isolated different data resources by grid idea and technology.

3.1 OGSA-DAI Architecture

The components in each layer and the interfaces between them are explained below, starting from the bottom layer and working upwards.

(1) Data Layer. The data layer consists of the data resources that can be exposed via OGSA-DAI. Currently these include:

- Relational databases such as MySQL, SQL Server, DB2, Oracle, PostgreSQL.

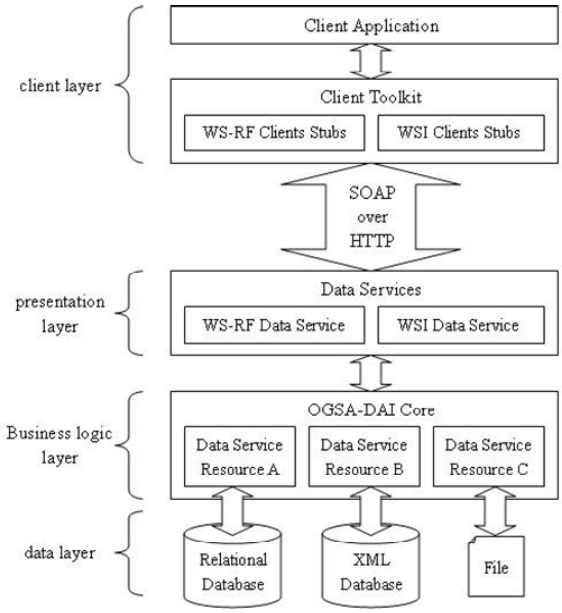

Figure 2. The high-level architecture of OGSA-DAI

- XML databases such as eXist, Xindice.
- Files and directories in formats such as OMIM, SWISSPROT and EMBL.

(2) Business Logic Layer. This layer encapsulates the core functionality of OGSA-DAI. It consists of components known as data service resources. As shown in Figure 1, multiple data service resources can be deployed to expose multiple data resources. There is a 1-1 relationship between data service resources and data resources. The responsibilities of a data service resource include:

- Execution of perform documents - a perform document describes the actions that a data service resource should to take on behalf of the client. Each action is known as an activity. OGSA-DAI already includes a large number of activities for performing common operations such as database queries, data transformations and data delivery.
- Generation of response documents - a response document describes the status of execution of a perform document and may contain result data, such as the results from a database query.
- Data resource access - interactions with data resources take place via the data resource accessor component.

- Data transport functionality - data can be streamed in and out of data service resources to and from clients and other data service resources.

(3) Presentation Layer. This layer encapsulates the functionality required to expose data service resources using web service interfaces. OGSA-DAI includes two realizations, one compliant with WSRF and the other compliant with WSI (that is a solution that only relies on the specifications mentioned in the WS-I basic profile, i.e. that do not use WSRF). These were schematically shown in Figure 1 as the WSRF and WSI data services. For each realization there is a WSDL document that describes the interface.

(4) Client Layer. A client can interact with a data service resource via a corresponding data service. Depending on whether a WSRF or WSI data service has been deployed, the client application must be compliant with the WSRF or WSI standards.

3.2 OGSA-DAI Workflow

The OGSA-DAI middleware must run on the Web container such as the Tomcat and includes three interfaces: the GDSR (the grid data service register), the GSDF (the grid data service factory) and the GDS (the grid data service) interface [4].

The GDSR (the grid data service register) is a special register service and can register the relative service compatible with the OGSI such as the service name, lifecycle, location information and the data resource connected with it. Hence, the client can know what service is there, the ability of the service and the data resource managed by it. The GDSR makes the application program find the database to be accessed by the user and shields the complicate query process.

The GDSF (the grid data service factory) is a special factory service which represents the data resource and creates the data resource interactively accessing with the client. If accessing the data resource, the GDSF must be properly configured and is used to connect with different data resources which makes the grid data service be able to access and update the data.

The GDS (the grid data service) makes the client not only integrate and find the data resource, but also transfer the data resource which includes the xml database, the relational database even some data saved in a usual file. An example GDS created by the GDSF can be received by the client in the form of the XML-based perform document. The GDS deals with the perform document and returns the results to the client it the form of the XML-based perform document.

3.3 The Distributed Query Principle

Figure 2 is the process that the client queries the data resource which shows the principle of sharing and accessing the data.

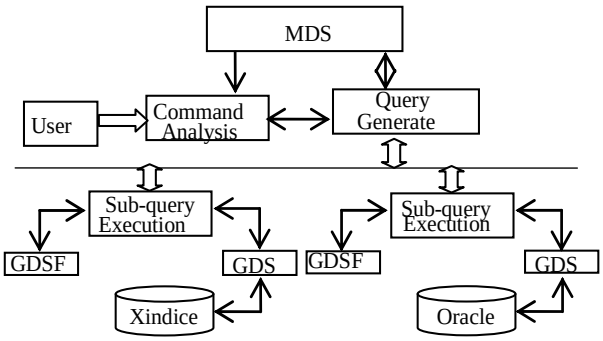

Figure 3. The principle of the distributed query of the database

Command analysis [5]: Receiving the SQL query command transferred from the client by the ERGAS is received in this step. Anglicizing the user query request and judging whether there are some grammatical errors and whether the data of the user query are available.Finally,if these steps are executived successfully, the internal expr- ession of the query statement will generate and be send to the next step.

Query Task Disassembly: obtaining the internal expression of the query statement in the command analysis. decomposing the internal expression into several subquery according some optimal algorithm.Finally, sending the sub-queries to Sub-query execution module.

Sub-query execution [6]: The GDS created by the GDSF accesses the database and obtains the data. Finally, the GDS returns the query results to the query user. The GDSF is a Web Service interface configured by the user

978-1-935068-06-8 © 2009 SciRes.

and access the local database.

Query process: The followings are the key codes about the distributing query from a XML database and a relational database.

4 The Algorithm and Some Codes

The algorithm executing process is as follows:

(1)Establishing the services handle;

(2)Creating GDSF (the grid data service factory) ;

(3)Creating the GDS example;

(4)Obtaining the SQL activity from perform document created by the SQL statement of the user input;

(5)Executing query request;

(6)Obtaining the results and data;

The main codes of its relevance are as follows:

```
// creating the service handle
String factoryHandle=
"http://202.206.43.123:8080/ogsa/services/ogsadai/GridDataServiceFactory";
// creating a temp table
String tableName "tempTable";
String createtable="createtable if not exists"+ tableName+"[ Property name 1 Property Type(Property length)…] ";
SQLUpdate create=new SQLUpdate(createtable);
factory=ServiceFetcher.getFactory(factoryHandle);
GridDataService
gds=factory.createGridDataService();
//creating a GDS example
SQLQuery query=new SQLQuery(create);
ActivityRequest request=new ActivityRequest();
request.addActivity(query);
//Constructs a simple activity
Response response=gds.perform(request);
//executing the request
ResultSet result=query.getResultSet();
//obtain the result set
ResultSetMetaData metadata= result.getMetaData();
//Anglicizing metadata
int columns=metadata.getColumnCount();
String[] labels=new String[columns];
int[] colwidths =new int[columns];
//get the line number label and width
for(int k=0; k<columns; k++) {
labels[k]=metadata.getColumnLabel(k+1);
colidths[k]=metadata.getColumnDisplaySize(k+1); }
while (result.next()){
for(int j=0; i<columns; j++){
String data= result.getString(j+1);
//get the data
} }
```

5 Conclusions

As the key next generation Internet technology, the grid can eliminate Knowledge Island and Information Island and make the information resources and education course resources be shared by more and more learners. So, the grid technology is a very good solution to the education course resources sharing. The grid technology makes the education course resources shared by the students from all universities can solve the problem of the inhomogeneous distribution of education course resources. Firstly, this paper researched the technology and principle of access the database by OGSA-DAI and the principle of realizing the query of the database. Finally, the ECRSD system which the Oracle and Xindice system database are used in the experiment is created and test. Used the ECRSD, the education course resources are able to be shared in the Internet with all the clients. The ECRSD system design concept can be beneficial to educational informatization and E-learning construction.

References

[1] Allcock Chervenak A,Foster, The data grid: Towards anarchitecture for the distributed management and analysis of large scientific datasets[J] ,Journal of network and Computer Applications, vol. 3, pp. 187-200, August 200.

[2] Allcock B,Fosrter I,Nefedova V.High-performance remote access to climate simulation data: A challenge problem for data grid tehnologies[C].USA: Conf of High-Performance Networking and Computing,Denver SC,vol. 4, pp. 283—297, january 2001.

[3] Alberto Sánchez, María S. Pérez, Konstantinos Karasavvas, Pilar Herrero, Antonio Pérez,MAPFS-DAI, an extension of OGSA-DAI based on a parallel file system[J],Future Generation Computer Systems, Vol. 23, pp 138-145, January 2007, 138-145.

[4] Yang Lin, Zang Yong-sheng, Xing Chang-ming, Information retrieval strategy based on education resource grid[J], Application Research of Computers,Vol. 26, pp 1484-1486, march 2009.

[5] Xing Chang-ming, Yang Lin, Liu Fang-ai,Replica Placement Strategy Based on Education Resource Grid[J],Computer Engineering,Vol. 34, pp 121-123, june 2008.

[6] Oscar Corcho, Pinar Alper, Ioannis Kotsiopoulos, Paolo Missier, Sean Bechhofer, Carole Goble.An overview of S-OGSA: A Reference Semantic Grid Architecture Web Semantics[J], Science, Services and Agents on the World Wide Web, vol. 4, pp. 102-115, June 2006.

978-1-935068-06-8 © 2009 SciRes.