

Data Extraction method for Heterogeneous Database Based on EJB

Zhenyou Zhang

Collage of Information Engineering
Hebei United University
Tangshan, China
e-mail: youzhenadd@163.com

Bochen Sun

Collage of Information Engineering
Hebei United University
Xi Qing Sub-branch
Tianjin Rural commercial Banks
Tiangjin, China
e-mail: 376055149@qq.com

Zhi Cao

Collage of Information Engineering
Hebei United University
Tangshan, China
e-mail: 526818718@qq.com

Abstract—Data island exists in most of the information system. And the system platform integration could eliminate the existence of the information island. This paper visited database information system through the EJB component technology, and using the metadata way to extracted data information from the heterogeneous database system, and it realized sharing information between the heterogeneous database system. In this system the B/S mode structure and EJB component technology, improve the system usability and expansibility.

Keywords- Data island; metadata; EJB component technology; B/S mode.

I. INTRODUCTION

The enterprise's normalization mostly lack long-term and overall planning from the reason analysis cause by itself. It caused the blindly imported the different application system only consider various local demand in different stage, such as there were superior department issued promotion, and it was also have to development or cooperative development. So most of the application system has no unified technology and data standard, the data could not be transfer automatically, and the data lack of effective connection and sharing, thus forming a isolated data island. Then, "data island", it was a computer application system which the function haven't mutual association, and the information haven't sharing of exchange and the information disjointed the business process application.

Integrated platform was a universal tool that can be adapted to information sharing in different system, it united enterprise business process, public data, application software, hardware and all kinds of standard through the enterprise application integration technology, and so it will realize seamless integration in different enterprise application system, and make them like a whole business processing and information sharing. When in a number of enterprises system for business transactions, the integrated platform could also realize system integration for different enterprise. The development with the information system integration technology, the system integration by driving information integration breadth and depth of the expanding, integrated

range by the enterprise internal gradually expanded to enterprise external.

II. THE THEORY ANALYSIS FOR THE DISTRIBUTED HETEROGENEOUS DATABASE SYSTEM

A. the system framework Analysis

The aim of system integration was to provide a unified interface for visit of heterogeneous data sources, and users need not consider the data model heterogeneous problem. It only need to designated concerns about data, without describes how to receive data. System integration platform store every data source for the node of heterogeneous database for the user. So the user could accessed to all heterogeneous databases. As shown in figure 1.

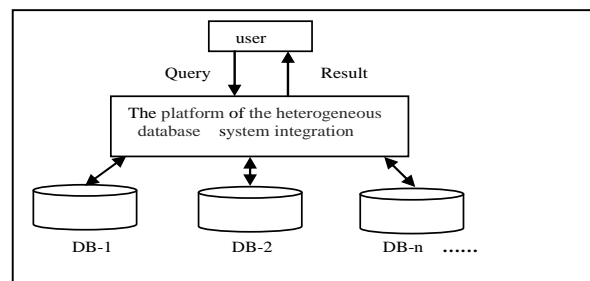


Figure 1. platform of the data integration

B. organizational structure design of the system

The way integration of heterogeneous database was normally has three: Federal Database, Data Warehouse and Intermediary Device. Federal database system was composed of a set of mutual cooperation, but keeps their self-governing member database system of composition; this member database system could be different degree of integration. Data warehouse was stored a single copies of the entire database from several data sources. Middleware technology was a kind of software architecture; it supports virtual view or view set.

It has good adaptability and expansibility. This study adopts such middleware technology to complete more heterogeneous database of integrated operation.

In order to make the system has good feasibility and expansibility, this study adopts B/S model .It has three-level structure and operation in the application server EJB components to solve the system with multiple heterogeneous database operation.

III. SYSTEM DESIGN BASED ON EJB COMPONENTS

EJB was a server component model designed by SUN, it was mainly use to the deployment of distributed application, but there were many ways to realize distributed application, it similar to Microsoft.net technology. With Java cross-platform superiority, the EJB technology deployment of distributed system can not limited to a particular platform. Enterprise JavaBean was part of the J2EE, it defines a standard for development the Enterprise multiple application based on the component. Its include network service support and core development tools (SDK).

In solving the problems that visit many heterogeneous databases, it adopting EJB components realizing database of various database operations, including data reading writing, inquires the etc. This system of concrete structure as figure2.

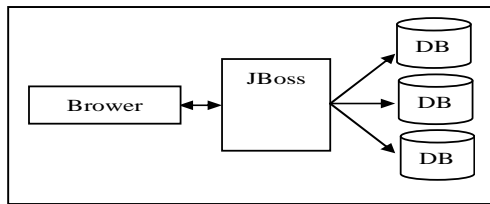


Figure 2. system structure

As is well known, the EJB have to run in the EJB container, and JBoss was a J2EE application server for run EJB, such as: database access JDBC, transaction (JTA/JTS), message mechanism (JMS), named mechanism (JNDI) and management support (JMX). It was open source project, follow the latest J2EE specification. JBoss and Web server run in the same Java virtual machine starting, and the Servlet invoke the EJB not through network, which greatly improve efficiency and safety performance.

So in figure 2 Browser was client, DB was the database server, web server was JBoss, it was a web server and EJB server.

A. development environment

This system was to develop a heterogeneous database data system integration platform. Users can add, modify and delete all kind of different database system in this platform. It lay a good foundation for various heterogeneous database operation in the future. The system development tool adopt Myeclipse, and using JBoss to deployment EJB servers and Web Server. The system could be integrated database Server have such as SQL Server 2000, Oracle10i, Mysql etc.

B. the module designed

Firstly, in order to integrate the various heterogeneous database systems to this operation platform, it must

complete the login information of the heterogeneous database system preservation, such as node name, type of database, driving class-name, connecting URL, user name and password .This information show in TABLE I. This study adopts EJB technology to implement for these information increases, modify and delete, and save the login information to the system of SQLServer2000 database.

TABLE I. DATABASE LOGIN INFORMATION

		Database Type	
	name	Mysql	SQLServer
1	Node name	department	center
2	Database Type	Mysql5.0	SQLServer2000
3	Driver name	org.gjt.mm.mysql.Driver	com.microsoft.jdbc.sqlserver.SQLServerDriver
4	URL	jdbc:mysql://127.0.0.1:3306/mysql	jdbc:microsoft:sqlserver://127.0.0.1:1433
5	User name	admin	sa
6	password	####	####

The more important aspect was the various heterogeneous database user data table on this system for complete database system of integration. This study adopts EJB component technology to achieve the user data sheets of integration to complete a variety of database connections. This process was divided into two steps, the first was extract the heterogeneous database user data table according to the user's require, the second was to choice the user concerned data table added to the system and management. The first step flowchart figure.3

User data table information was the main information of databases system integration. Including users table name, table creation time, table rows, etc., there were two key problems to solve in the inquiry in the process of user data table. The first was user data sheets of storage structure and position was different in different relational database.

The second was the user data table information item describing wasn't also at the same.

The following were SQLServer2000 and Mysql5.0 database query the user data table lookup routines.

User data table information stored in the corresponding database the sysobjects system tables and sysindexes system tables in SQLServer2000 database, so the SQL language for inquiring user information table for: select sysobjects.name,sysindexes.rows,sysobjects.crdate from sysobjects,sysindexes where sysobjects.id=sysindexes.id and sysobjects xtype ='U' and sysindexes.indid in(0,1). In here sysobjects. xtype = "U" said is the user table, sysindexes.indid in(0,1) says it is't index.

User data table information stored in the tables of system table of the database information_schema in Mysql5.0 database, so the SQL language for inquiring user information table for:Select table_name,create_time,table_rows from tables where

table_schema=' 'the name of the user database' ' order by table_name.

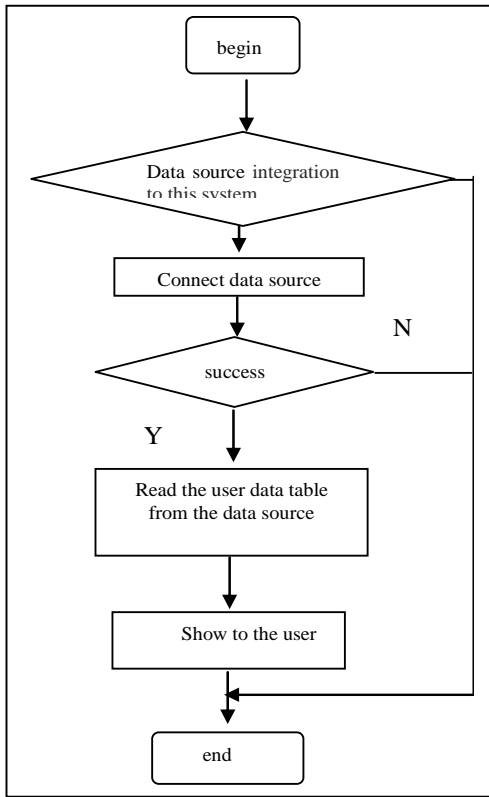


Figure 3. the query of the table information of user

It was must to displayed the message for the user, it must to put information undertake unity of the different relational database query for the user data table. Below give a simple example the data table name of the unified.

```

// SQLServer2000 users table name
private string name;
// Mysql5 users table name
private string table_name;
// create corresponding user variable get and set method
// reunified users table name
private string allname;
// unified get and set method
public String getAllname()
{
if(this.name!=null)
this.allname=this.name;
if(this.table_name!=null)this.allname=this.table_name;
return allname;}
public void setAllname(String allname)
{
this.allname =allname;
this.name=this.allname;
this.table_name=this.allname;
}
get consistent user data table information, so the data
show to the Web users, users according to the actual
  
```

situation of the selection and then add to the integrated system.

IV. SYSTEM DESIGN ON DATA EXACTRON

The system could extracted data from different heterogeneous database after completion the database information storage and user data table information storage. First it read the database information, and read the table structure information, and then creates the query sentence, last extraction out the data from the database. This was a simple single database table query. Flow show in figure 4. Multi-database and table query much more complicated.

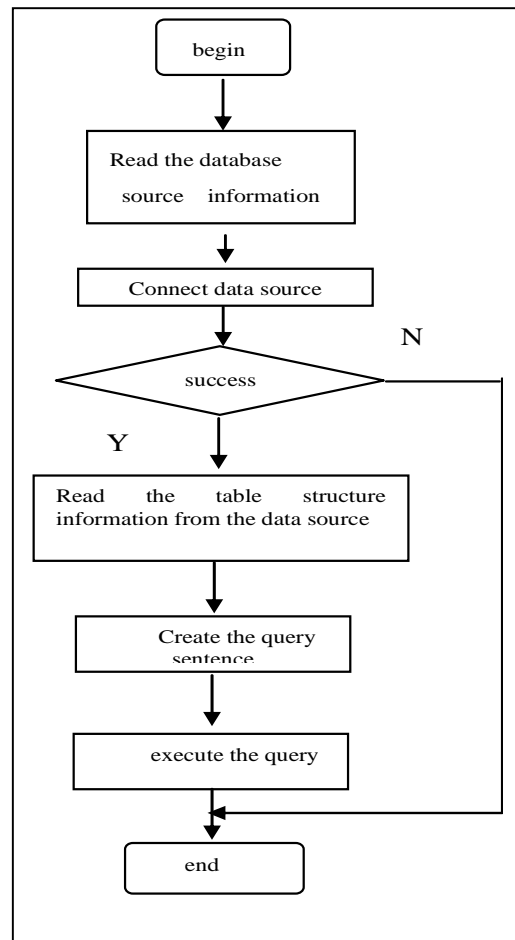


Figure 4. the query of the data information

V. SUMMARY

The heterogeneous relational database was to solve the key of the current variety of data and information management. This study adopts EJB component technology to dynamically linked more heterogeneous database system, and inquiring the different database user list information and put them integrated into the system in the center. It can realize the data sharing and improve efficiency in multiple node databases.

REFERENCES

- [1] Y.H.Wang etc.Distributed database systems. publishing house of electronics industry .2008.
- [2] H.Yan "distributed database data synchronization technique".naval engineering university journal, vol.6..pp11-15. 2007.
- [3] B.K.Ding.Distributed database system realization technology. Beijing: science press, 2006.
- [4] H.Shan etc. SQL Server data replicatio.Beijing: higher education press, 2006.
- [5] P.S Bernstein. Hipman D and Rothnie carol carroll oncurrency." In a Distributed Control System for Database (SDD - 1)". ACM Database System, Trans on.Apr 2007.
- [6] Ceri, s.d. istributed bear - flying and System . McGraw Hill, 2006.
- [7] Traiger I.L.T "ransactions and Consistency in Distributed bear". IEEE Transactions on Software Engineering,vol (9).pp175-179 2006.