Research the Synchronization Mechanism for Distributed Heterogeneous Database

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

LIU bo
College of Light Industry Hebei United University
Tangshan, Hebei, China
e-mail: liubo@heut.edu.cn

LIU shu
College of Information Science & Engineering Yanshan University
Qinhuangdao Hebei China
liushu@ysu.edu.cn

Abstract— **For the heterogeneous database of distribution in different physical location ,it was a key problem of how to guarantee the user data synchronization of heterogeneous database. This paper analyzed data synchronization system for the distributed heterogeneous database ,and the extracting module, data integration module, data update module and put forward a concrete realization of solutions.**

Keyword- synchronization mechanism of data; the minimal change sets; the minimal update sets

I. PREFACE

Research of the distributed heterogeneous database system was an application research for the integrated of the different kind of database system, and one of the important aspects was the data synchronization between distributed heterogeneous database system . The purpose of the data synchronization was designed to make each site located in the heterogeneous database system tables .It can maintain the consistency of the data according to the requirements of customers. This synchronization didn't change the original heterogeneous database system, therefore the autonomy in many nodes can be realized to realize data sharing between database.

II. THE THEORETICAL ANALYSIS OF SYNCHRONIZATION MECHANISM FOR DISTRIBUTED HETEROGENEOUS DATABASE

The fundamental purpose of the data synchronization of the heterogeneous database was made in different nodes of different types of databases to keep same of the data same in the table . Because of the autonomy of the heterogeneous database system, so the system makes each heterogeneous database data maintain consistency only by third party system to synchronize data transmission.

The essence for the heterogeneous database data synchronization was made other heterogeneous database also do the same change while a user data in the table in the node database at one time changed.

The collection of synchronous data is the foundation of the heterogeneous database synchronization mechanism . It established three trigger included INSERT, UPDATE and DELETE on the user's requirements for each synchronous data table. Each time, when a user operated the data table had trigger, trigger will be triggered. So it can be found those changes of data in the database called a single database changes sets.

The key of the heterogeneous database synchronization mechanism was to manage the synchronize data . The purpose was to get minimal change set according to the synchronous groups ,when the trigger get synchronous data and according to the synchronous groups. Again with each node database changes the difference operation, so getted various heterogeneous database of minimal update sets.

The data update for the heterogeneous database synchronization mechanism was put each of heterogeneous database of minimal update set to update the various heterogeneous database corresponding data table process.

III. THE DESIGN FOR THE HETEROGENEOUS DATABASE SYNCHRONIZATION MECHANISM

3.1 designed data extraction module

In this with SQL Server databases of test table as an example, use the trigger to track and extracted table of changes sets of the test table.

If the test table was to conduct data synchronization tables, created test_info_insert data table and test_info_delete tables,but it had an extra column that was the current time of the system with the test table. If write the SQL statement in "body of the trigger ",defining trigger,it can to write the data set into test_info_insert data table and test_info_delete table.

The SQL statements for written records into the test_info_insert data table are as follows:

1. Extracting insert record as follows:

```
//create a trigger named test_insert
CREATE test_insert TRIGGER ON test
//trigger type was insert trigger
FOR INSERT AS
```

978-1-61284-240-0/11/$26.00 ©2011 IEEE

//from inserted table,it's self-contained in the test_insert trigger reads record and place it into the test_info_insert table

INSERT INTO test_info_insert SELECT * FROM INSERTED

2. Extracting delete record as follows:

//create a trigger named test_delete

CREATE test_insert TRIGGER ON test

//trigger type was detele trigger

FOR DELETE AS

//from deleted table,it's self-contained in the test_delete trigger reads record and place it into the test_info_delete table

INSERT INTO test_info_delete SELECT * FROM DELETED

3. Extracting update record as follows:

// created a trigger named test_delete

CREATE test_update TRIGGER ON test

/ / trigger type was delete trigger

FOR the DELETE AS

//from inserted table,it's self-contained in the test_insert trigger reads record and place it into the test_info_insert table

//from deleteted table,it's self-contained in the test_delete trigger reads record and place it into the test_info_delete table

INSERT INTO test_info_insert SELECT * FROM INSERTED;

INSERT INTO test_info _delete SELECT * FROM DELETED

3.2 design of the data integration module

The purpose of the data integration is to get the minimal update sets of all of heterogeneous database . Therefore, it must first push the change sets in the all single database into system database. Then delete the same recordset according to the synchronous group, so it get the total minimal change set and have a minimal update sets of various heterogeneous database .

In this module, it has four sub module, it was the management for synchronous table, the centralize of the synchronous sets, generated the minimal change sets, generated the minimal update sets. The main function of the management for synchronous table was set up a table structure for the corresponding synchronous data table in the system's database ,it used to stored the changes sets from the data acquisition module. The main function of the centralize of the synchronous sets was replicated the data from the various heterogeneous database according to synchronous group into the system's database. The main function of the generated the minimal update sets was to processing the changes sets of the synchronous set ,and delete redundancy. The main function of the generated the minimal update sets was to use the generated the minimal update sets differencing the various heterogeneous database changes sets . the processing of the insert sets in the minimal change sets as figure1. the processing of the insert sets in the minimal update sets as figure2.

According to the same process could generate the delete sets and update sets, both of them was the minimal update sets of individual heterogeneous database. The minimal update sets in written back to the corresponding each node database can complete the heterogeneous database data synchronization.

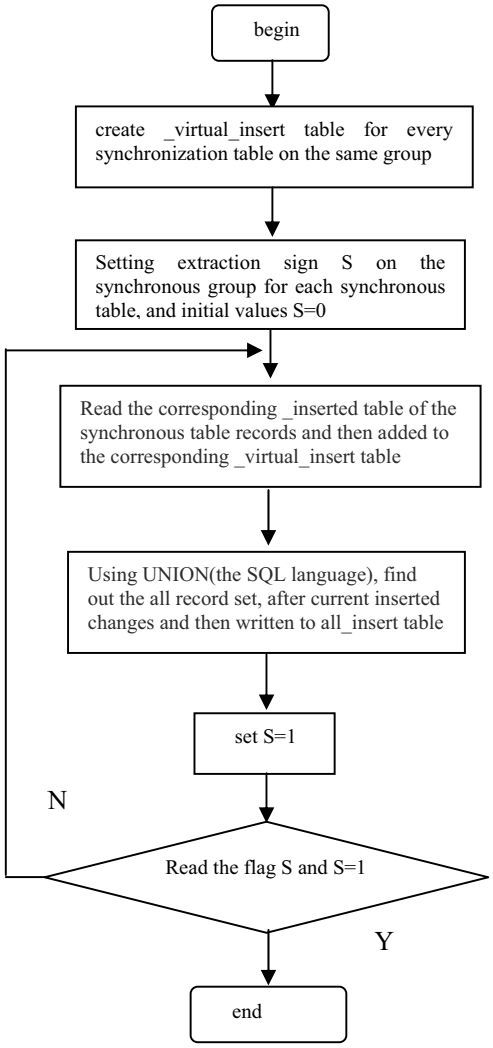

Figure 1 the insert sets in the minimal change sets

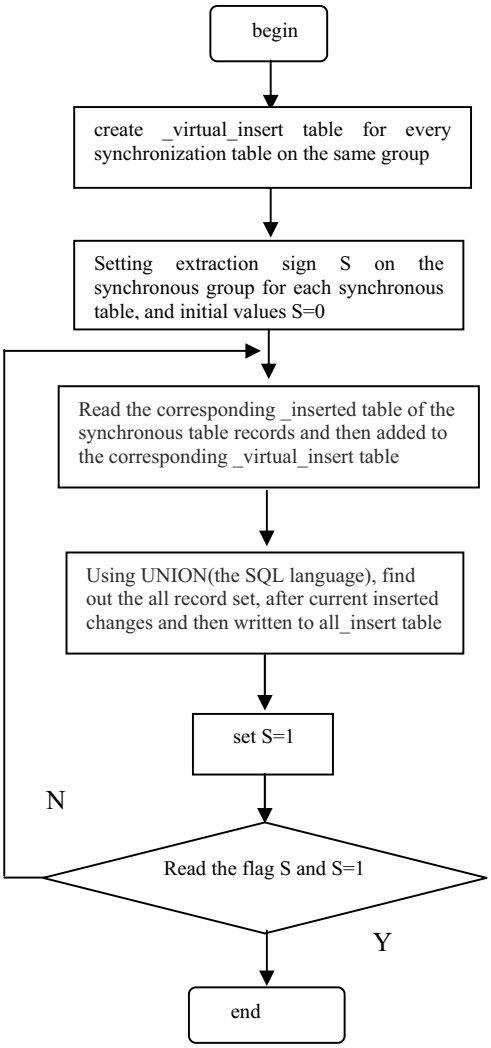

Figure 2 the insert sets in the minimal update sets

IV. SUMMARY

Data synchronization problem is currently an important problem of information management, if the traditional database replication technology has many disadvantages. In this paper the proposed synchronization mechanism solved data extraction and data update, and the minimal change sets can reduce synchronous data through put, and improves the efficiency of system. So it hoped some ideas for the heterogeneous database data synchronization study.

REFERENCES

[1] WangYiHe etc. Distributed database systems [M].beijing: publishing house of electronics industry, 2008,6.

[2] YanHao distributed database data synchronization technique [J].j naval engineering university journal, 2007,6 (3) : 11-15.

[3] DingBaoKang. Distributed database system realization technology [M]. Beijing: science press, 2006,8.

[4] huashan etc. SQL Server data replication [M].beijing: higher education press, 2006.

[5] P.S Bernstein. Hipman D and Rothnie j.carol carroll oncurrency. In a Distributed Control System for Database (SDD - 1) [J].j ACM Database System, Trans on 2007,4.

[6] Ceri, s.d. istributed bear - flying and System [J].j McGraw Hill, 2006,8. -

[7] Traiger I.L.T ransactions and Consistency in Distributed bear [J].j IEEE Transactions on Software Engineering, 2006,6.