Synchronous Optimization Method for Heterogeneous Database Based on the Minimal DataSet

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

Dongguang Wang
Collage of Information Engineering
Hebei United University
Xinhua Road Management Department
Tangshan Housing Fund Management Center
Tangshan, China
e-mail: wdg618@sohu.com

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

Abstract—The data synchronous of the distributed heterogeneous database was one of the important researches for the heterogeneous databases. This paper analyzed the data synchronization optimization mechanism of the distributed heterogeneous database and puts forward to get the minimum change data set from each node database and then get the system total change data set, and then find out each node database of minimal update data set to update each node database, so as to realize the synchronization of heterogeneous database.

Keywords- Synchronous optimization system; data changes se;, minimal changes se;, minimal changes update sets.

I. INTRODUCTION

Research the data synchronous in distributed heterogeneous databases was together multiple database system to realize the synchronization. The so-called synchronous refers to the node database in different between user data tables keep data in the agreement. Because the autonomy of heterogeneous database system, it cannot affect the original database of all kinds of user operation. It also realizes the share data in multiple node databases.

II. THE DATA SYNCHRONIZATION OPTIMIZATION THEORY ANALYSIS

Heterogeneous database data synchronization essence was the user table in a period of detains a node database, make other heterogeneous database also did the same change. There were three stages for realization the data synchronous in the heterogeneous database. The first was to collect the data of change in a user table on node database. The second was the concentration of these changes data. The third was updating the user table in the destination node database.

Assuming node database in the heterogeneous database system such as shown in table I.

TABLE I.   NODE DATABASE

	Local node Information		
	Local node name	*Table name*	*Field count*
1	Finance	workers	15
2	Personnel	employees	25
3	library	customer	12
4	workshop	employees	10

It used trigger mechanism to realize the data acquisition. According to the working principle, system created two triggers for each user data table, Insert action triggers (Insert) and Delete action triggers (Delete). Whenever the user data table that had two triggers when operating, trigger will be triggered. System can be found those changes, in which the data collection called single database changes. These data changes to store in table Ⅱ named data table. These data table and user data table have the same table structure, but plus a field named insert time field. When there was data.

Insert user data table, Insert trigger action. When there was data Delete a user data table, the Delete trigger action. When there is data update, Insert trigger action and Delete triggers actions. Execute actions, the changes data in user data table will be added to the data table shown in the table Ⅱ.

Change data concentrated stages was an important stage of synchronous data optimization. This phase is divided into three steps, the first step was to concentrate from a single database changes set to the system database. The second step was to change collation, and get tiny collection of change. In the process of change set finishing the collection was the process of optimizing synchronous data, the sorting process greatly reduces the amount of data synchronization transmission efficiency, and improved synchronization. The third was to use tiny changes the inclination and a single database changes the difference to get integrating minimal update set in a single node.

TABLE II.   CHANGES DATA CONCENTRATED IN NODE DATABASE

	Change Data in Node Database		
	Table name	Table name create by Insert trigger	Table name create by Delete trigger
1	workers	Finance_Workers _Insert	Finance_Workers _Delete
2	employees	Personnel_ Employees_Insert	Personnel_ Employees Delete
3	customer	Library_customer _Insert	Library_customer _Delete
4	employees	Workshop_Emplo yees_Insert	Workshop_Emplo yees_Delete

Updates the user data table in the destination node database was to used minimal update set to update the various heterogeneous databases in the process data table.

III.   IMPLEMENTATION THE SYNCHRONOUS OPTIMIZATION MECHANISM

A.   concentration from an single database changes sets

It was optimized foundation that the change of the user table had to integrating. Because of individual database changes sets in each node database, in order to focus them, first stage was established the corresponding user data table in the system's database, then copies the data in user data table to system database. Such data were all concentration changes to system database. But such concentration was a global user data table name for corresponding relationship, that is to say, with the same structure of the same global name in the user data table data were focused.

If it put the four use table in the four node database to the system database, the first stages will check the user table compatibility, namely three table fields type whether incompatible, if incompatible data table can proceed the concentration. In order to perform the synchronous optimization put these user data schedule uniformly to a user global table.

Its such as shown in table Ⅲ   after reunification information .

TABLE III.   TABLE NAME AFTER REUNIFICATION INFORMATION IN SYSTEM DATABASE

	Table Name In System Database		
	Global user table	Node name	User table name
1	user_Insert	Finance	User_Finance_Worke rs_Insert
2	user_Insert	Personnel	User_Personnel_ Employees_Insert
3	user_Insert	library	User_Library_custom er_Insert
4	user_Insert	workshop	User_Workshop_Em ployees_Insert
5	user_Delete	Finance	User_Finance_Worke rs_Delete
6	user_Delete	Personnel	User_Personnel_ Employees_Delete
7	user_Delete	library	User_Library_custom er_Insert_Delete
8	user_Delete	workshop	User_Workshop_Em ployees_Delete rt

The user table name in table Ⅲ was a node database user list the name in system database. While the global user table name was symbol that allowed synchronous in this distributed heterogeneous database.

The User Insert the data collection concentration process changes as shown in Figure.1.

B.   The minimal change set collection generated ions

The main function of the module of the minimal changes set collection was processing the generated by synchronous set each heterogeneous database data, delete redundancy. Insert collection set in the minimal changes set of the total concentration generative process as Figure.2.

According to the same process could generate the delete sets and update sets, both of them were 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.

C.   The minimal update set collection generated

Tiny update collection set was generated in the system's databases; it used to update the user data tables in node databases. The main idea was the tiny change in a single database changes collection and the differ set to get integrating a user data table of minimal update sets for node database. The processing of the insert sets in the minimal update sets as Figure.3.

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]   Zisman A, Kramer J. "Towards Interoperability in Heterogeneous Database Systems,". Network Management, vol. 29,pp.44-46,Mar.2007.

[2]   Amit P, James A. "Federated database systems for managing distributed, heterogeneous, and autonomous databases,"ACM Computing Surveys,vol.17.pp.12-14, 2008.

[3]   Ceri, S."Distributed Databases-Principle and System,".McGraw-Hill,Aug.2008.

[4]   Y. Zhou, W.Xu:heterogeneous database system structure. Beijing: science press, 2007.

[5]   S. Wilbur,B.Bacar."Building distributed system with remote procedure call,".Software Engineer Journal.Proceeding of SPIE,Bellingham,Washington,vol.(8).pp.251-155.2008.

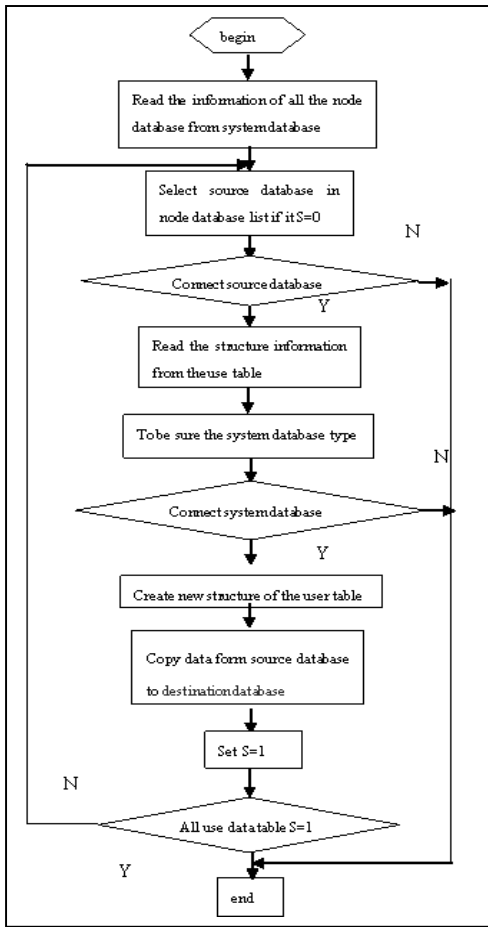[6]   P.Y. Shao:Distributed database systems and its application. Beijing: science press, 2008.

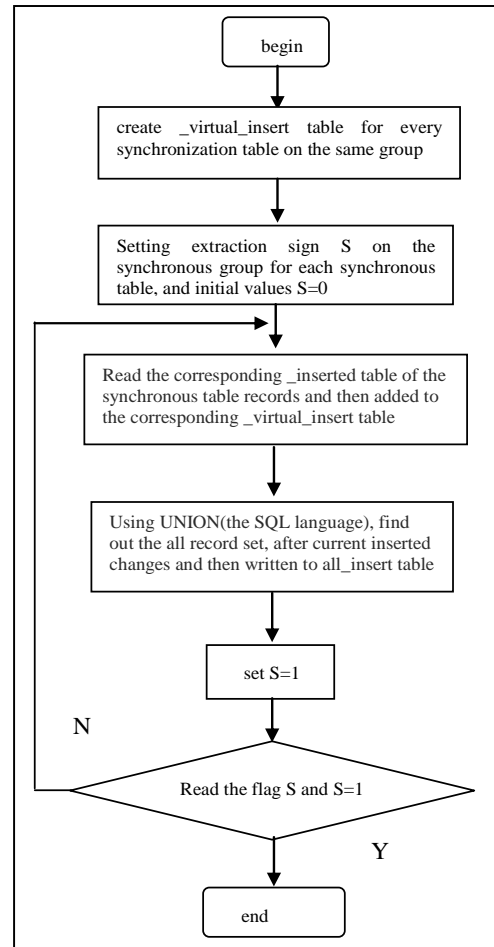Figure1. Uer_Insert user data collection concentration process

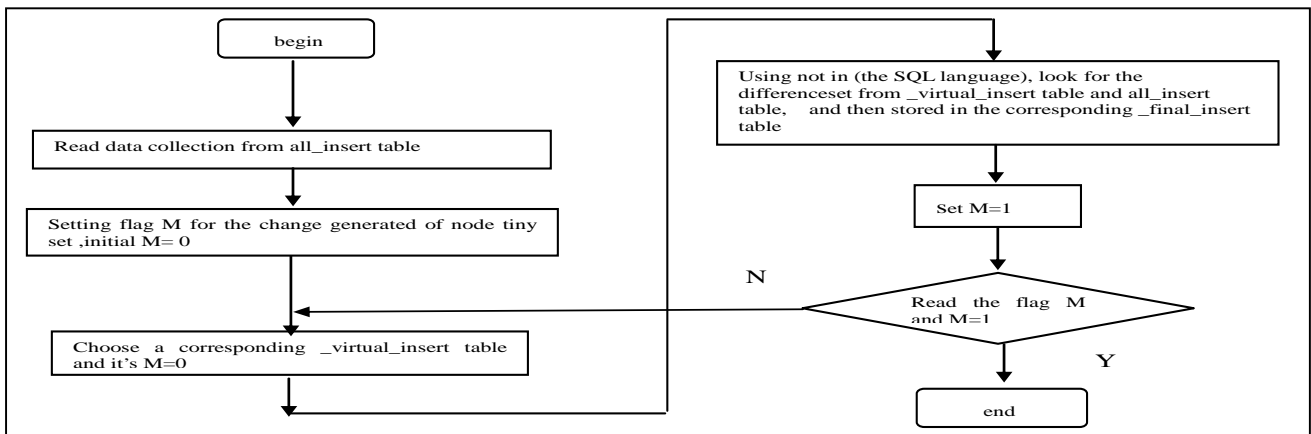Figure2. Insert collection set in the tiny changes set of the total concentration generative rocess

Figure3.The processing of the insert sets in the minimal update sets