

TITLE OF THE INVENTION

Data Reflecting Method and System

BACKGROUND OF THE INVENTION

5 1. Field of the Invention

The present invention relates to a technology for reflecting transmitted data to a database, and particularly to a technology for reflecting data to a database managed as a backup provided in case of disaster, for example.

10 2. Description of the Related Art

Nowadays, data has been made redundant under control. For example, there is a plan to realize business continuity through data to be used at a time of disaster. The data used in daily business is saved to another location. A
15 business continuity plan is formulated so that business still can be continued at an event of the disaster. Saving system data used during normal time to another location has been realized by use of a data transfer technology.

A technology for reducing the cost of ownership of a
20 backup system exclusive for the time of disaster has been disclosed in JP-2005-234749-A. Therefore, the backup system of JP-2005-234749-A is provided with a server resource management database which stores therein a management table including priority to assign a shared
25 server of a shared server resource group having a backup

database to back up a plurality of user systems to each user system, and a server allocating system for allocating the shared server to each user system. A plurality of shared servers are used in common by a plurality of user systems even at the time other than the backup. At the time of backup, the server allocating system selects a user system on the basis of the priority of the management table and assigns a shared server thereto.

10 SUMMARY OF THE INVENTION

At the above backup, data is transmitted (transferred) from the system that serves as the source. In the current data transfer technology, there is no established means for performing data storage processing in parallel when data is stored in a system used as a transfer destination, and improving efficiency in processing time required for the storage. In JP-2005-234749-A in particular, the allocation of a resource has been taken into consideration, but the transfer at that time is not taken account of. The present invention aims to improve efficiency in data transfer and reflection with the foregoing in view.

The present invention implements execution of data storage processing in parallel when storing data in a transfer destination system. As its implementing means,

the parallelization of data storage processing is performed with a system having realized a disaster recovery (disaster countermeasure) as an assumption.

More specifically, the present invention assures
5 ordering of elements constituting unit data like one transaction using a parallel key indicative of their ordering and executes parallelized reflection processing. This parallel key means that a column name that satisfies predetermined conditions is given the same parallel key, in
10 accordance with which its control is performed. As the predetermined conditions, the determination on whether the column names or the like are the same controls the reflection processing.

According to the embodiment of the present invention,
15 the reflection of data can be made efficient (speeded up).

The above, other objects and novel features of the present invention will become apparent from the description of the present specification and the accompanying drawings.

20 BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagram showing a parallel key definition file used in one embodiment of the present invention;

FIG. 2 is a diagram illustrating a conditions matching confirmation flow in the one embodiment of the present invention;

FIG. 3 is a diagram depicting a final determination flow for the propriety of parallelization in the one embodiment of the present invention;

FIG. 4 is a diagram showing a processing flow at data transfer in the one embodiment of the present invention; and

FIG. 5 is a diagram illustrating a system configuration example for executing the one embodiment of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

A preferred embodiment of the present invention will hereinafter be described by taking financial business for example. In the present embodiment, products for realizing data replication have heretofore been supplied from a software vender in terms of a disaster recovery, a data backup, or an aggregation of data from a plurality of systems to one database. These products are used particularly when the data are replicated between data centers.

In the related art technology, data (update information) created at a transfer source data center, at

the time of the data replication, have been serially transferred to each transfer destination data center in the order they have been created. Further, when the data sent to the transfer destination data center are persisted
5 (stored in database, for example) therein, the sent data have been persisted in order one by one. Since the data are persisted one by one even though they have been transmitted to the transfer destination data center in large quantities, this method takes time to complete data
10 transfer (to persist all transferred data at the transfer destination data center).

Therefore, a processing system for establishing means for processing the persistence of data in parallel and improving performance thereof is realized. With a view
15 toward realizing parallel processing of the data persistence, there has been devised a data parallel reflecting apparatus based on a processing system called a "parallel key." The present apparatus determines whether the data may be persisted in parallel when persisting the
20 transferred data. The data capable of being persisted in parallel are persisted in parallel, and the data not capable of being persisted in parallel are serially persisted in accordance with serial processing.

The below indicates how parallel propriety is
25 determined when two transactions {(1) and (2)} shown below

are processed by way of example. The transaction is a unit of processing when a system performs business processing. A plurality of queries (database language) are included in one transaction. Further, the unit of data to be persisted
5 by the present apparatus is a transaction unit.

An example shown below is of an example of processing of a data parallel reflecting apparatus used in a transfer destination system. Further, a system configuration example is shown in FIG. 5. In FIG. 5, the
10 data parallel reflecting apparatus is described as a data replication and integration framework server 521 of a transfer destination system 52. As shown in FIG. 5, the present system has a transfer source system 51 and the transfer destination system 52. These systems are
15 connected therebetween via network not shown. The transfer source system 51 has an application server 511 which executes business processing, a database server 512 which stores data used in the business processing, and a data replication and integration framework server 513 which
20 constructs a system base common to the transfer destination system 52. Further, the transfer destination system 52 has the data replication and integration framework server 521 which constructs a system base common to the transfer source system 51, and a database server 522 which stores
25 data similar to the data stored in the database server 512.

These servers are mutually connected therebetween by private line network such as a LAN.

The application server 511 has a business application processing unit 5111 which executes business processing. The database server 522 has a database 5121 which stores update information according to the business processing and a parallel key definition file to be described later. The data replication and integration framework server 513 has a data replication and integration framework processing unit 5131 which shares a system base with the transfer destination system 52 and cooperates or links data.

The data replication and integration framework server 521 has a data replication and integration framework processing unit 5211 which shares a system base with the transfer source system 51 and cooperates or links data. The database server 522 has a database 5221 similar to the data base that the above-described database server 522 includes. Each processing shown subsequently is carried out by the business application processing unit 5111, data replication and integration framework processing unit 5131, and data replication and integration framework processing unit 5211 mentioned above. The processing is actually implemented by allowing an arithmetic unit such as a CPU provided with each server to execute a program. (1) A

thousand yen, i.e., 1,000 yen is transferred from a bank account of Taro Hitachi (column ACCOUNT_ID=100) to a bank account of Hanako Hitachi (column ACCOUNT_ID=200). (2) An address is changed in personal information of Taro Hitachi
5 (column ACCOUNT_ID=100).

(1) Transaction A for transferring 1,000 Yen the bank account of Taro Hitachi to the bank account of Hanako Hitachi

When the transaction A of (1) is processed, two
10 queries are assumed to be executed by way of example. The queries to be executed are as follows:

(Query 1) 1,000 is subtracted from the balance of account information of Taro Hitachi in a table ACCOUNT_TABLE having account information of Taro Hitachi (column ACCOUNT_ID=100).
15 (Query 2) 1,000 is added to the balance of account information of Hanako Hitachi in a table ACCOUNT_TABLE having account information of Hanako Hitachi (column ACCOUNT_ID=200).

The above transaction A is applied to the conditions
20 matching confirmation flow of FIG. 2.

In a process at 202 in the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 included in the data replication and integration framework server 521 of
25 the transfer destination system 52 refers to a 1014 "table

name" of a parallel key definition file shown in FIG. 1 to be able to confirm that the "table name" corresponds to the table name ACCOUNT_TABLE included in the query 1.

Therefore, a determination result of the process at 202 of the conditions matching confirmation flow shown in FIG. 2 is "Yes". Incidentally, the parallel key definition file is stored in, for example, the database server of the transfer source system 51 or the transfer destination system 52. The description on each column in the parallel key definition file will be made later.

In a process at 203 in the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 refers to a 1013 "column name" of the parallel key definition file shown in FIG. 1 to be able to confirm that the "column name" corresponds to the column name ACCOUNT_ID included in the query 1. Therefore, a determination result of the process at 203 of the conditions matching confirmation flow shown in FIG. 2 is "Yes".

In a process at 204 of the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 refers to the values of the column ACCOUNT_IDs in the queries 1 and 2 included in the transaction A and determines whether both correspond to each other. Since in the transaction A the

value of ACCOUNT_ID is 100 in the query 1 and the value of ACCOUNT_ID is 200 in the query 2, the values of the column ACCOUNT_IDS in the transaction A are not the same. The data replication and integration framework processing unit
 5 5211, therefore, takes a determination result of the process at 204 in the conditions matching confirmation flow shown in FIG. 2 as "No." When the conditions matching confirmation flow shown in FIG. 2 is applied to the transaction A, a parallel key is accordingly determined to
 10 be "Gr_OTHER".

(2) Transaction B for changing the address in the personal information of Taro Hitachi (column ACCOUNT_ID=100)

When the transaction of (2) is processed, two
 15 queries are assumed to be executed by way of example. The queries to be executed are as follows:

(Query 1) The personal information is acquired from a table CUSTOMER_TABLE having personal information on Taro Hitachi (column ACCOUNT_ID=100).

20 (Query 2) The address of Taro Hitachi is changed with respect to address information of the table CUSTOMER_TABLE having the personal information on Taro Hitachi (column ACCOUNT_ID=100).

The above transaction B is applied to the conditions
 25 matching confirmation flow shown in FIG. 2. When the 1014

"table name" of the parallel key definition file shown in FIG. 1 is confirmed in the process at 202 of the conditions matching confirmation flow shown in FIG. 2, it is possible to confirm that the 1014 "table name" corresponds to the table name CUSTOMER_TABLE in the 1014 "table name" of No. 2 in the parallel key definition file shown in FIG. 1. Therefore, the determination result of the process at 202 of the conditions matching confirmation flow shown in FIG. 2 is "YES".

10 In the process at 203 of the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 refers to a 1013 "column name" of the parallel key definition file shown in FIG. 1 to be able to confirm that the "column name" corresponds to the column ACCOUNT_ID included in the query 2. Therefore, the determination result of the process at 203 of the conditions matching confirmation flow shown in FIG. 2 is "YES".

20 In the process at 204 of the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 refers to the values of the column ACCOUNT_IDS of the queries 1 and 2 included in the transaction B and determines whether both correspond to each other. Since in the transaction B the value of the ACCOUNT_ID is 100 in the query 1, and the

25

value of ACCOUNT_ID is 100 in the query 2 as well, the value of the column ACCOUNT_ID is the same in the transaction B. Therefore, the data replication and integration framework processing unit 5211 determines the determination result of the process at 204 of the conditions matching confirmation flow shown in FIG. 2 to be "Yes". When the conditions matching confirmation flow shown in FIG. 2 is applied to the transaction B, the parallel key accordingly assumes "Gr_B" at No2 of a 1012 "parallelization group name" in the parallel key definition file shown in FIG. 1.

As a result of application of the conditions matching confirmation flow of FIG. 2 to the above transactions A and B, the transaction A becomes "unparallelizable", and the transaction B becomes "parallelizable". The transactions A and B are subjected to the persisting process in parallel when a persisting process is performed.

The flow of the entire process in the present embodiment will be described below in accordance with the processing flow at the time of the data transfer in FIG. 4. (Determination of conditions for parallelization propriety by user (creation of parallel key definition file))

At 402 of FIG. 4, the user of this processing system defines conditions for parallelizable data and for non-

parallelizable data with respect to each column of the parallel key definition file of FIG. 1 from an input device of a computer connected to the application server 511 of the transfer source system 51 as shown in FIG. 5. The
5 columns to be confirmed that is used for determination as to the propriety of parallelization are the "table name", "column name", and "value of specified column".

The 1012 "parallelization group name" of the parallel key definition file shown in FIG. 1 defines the
10 name to uniquely specify the conditions for determining the parallelization propriety. Information (including member ID and stock name ID) to uniquely determine data is defined in the 1013 "column name" of the parallel key definition file shown in FIG. 1. It is possible to define only one
15 column name. A table name included in data targeted for conditions to be defined is defined in the 1014 "table name" of the parallel key definition file shown in FIG. 1. The table name can be defined in a plural form.

A flag for specifying whether a forced serial
20 process (unparallelizable) is taken is defined in a 1015 "forced serial process" of the parallel key definition file shown in FIG. 1. Upon receiving an input of such information from the input device, the computer sends the input to the application server 511, where the business
25 application processing unit 5111 thereof allows the

database server 5121 to store the information. The data replication and integration framework processing unit 5131 included in the data replication and integration framework server 513 of the transfer source system 51 acquires the
5 parallel key definition file stored in the database and transmits it to the transfer destination system 52. The data replication and integration framework processing unit 5211 of the transfer destination system 52 stores the parallel key definition file received from the transfer
10 source system 51 in the database 5221.

(Transfer of update information to transfer destination in transfer source system 51)

At 403 of FIG. 4, the data replication and integration framework processing unit 5131 of the data replication and integration framework server 513 included
15 in the transfer source system 51 transfers the update information of the database of the transfer source system 51, which has been updated by the business application processing unit 5111 of the application server 511, to the
20 transfer destination system 52 in a transaction unit. As the device for the data transfer, there can be utilized one that a bender provides.

(Determination as to data transferred to transfer destination)

At 404 of FIG. 4, the data replication and integration framework processing unit 5211 of the data replication and integration framework server 521 determines in accordance with the conditions matching confirmation flow shown in FIG. 2 whether the transferred data corresponds to any conditions of the parallel key definition file shown in FIG. 1. The result is eventually determined whether the group name should be "Gr_OTHER" or one other than that. Data in which the group name is determined to be "Gr_OTHER" becomes unparallelizable. Data in which the group name is determined to be the group name other than "Gr_OTHER" conforms to a forced serial process flag specified to its group. "Gr_OTHER" needs to always exist in the parallel key definition file shown in FIG. 1. Further, "Gr_OTHER" is a fixed value indicative of unparallelizable.

At 202 of the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 of the data replication and integration framework server 521 determines whether the table name defined in the 1014 "table name" of the parallel key definition file shown in FIG. 1 is included in the transferred data. At 203 of the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 determines

whether the column name defined in the 1013 "column name" of the parallel key definition file shown in FIG. 1 is included in the transferred data. At 204 of the conditions matching confirmation flow shown in FIG. 2, the data replication and integration framework processing unit 5211 determines whether the value of the column defined in the 1013 "column name" of the parallel key definition file of FIG. 1 takes the same value in the transferred data. There is a case where the same column name exists in a plural form within the transferred data (such as a case where a plurality of queries are included). When the same column name exists in a plural form and the column name is defined at 1013 of the parallel key definition file shown in FIG. 1, the value stored in each column is confirmed in the present processing system. The present apparatus confirms whether the values of the same column names are all the same in the transferred data. It is determined whether the transferred data can be parallelized in accordance with this determination.

20 (Determination as to parallelization propriety of persisting process at transfer destination)

A process at 405 of FIG. 4 will be described. At this step, it is determined to which conditions (group name) the data transferred up to 404 of FIG. 4 corresponds.

25 After the group name has been determined, the data

replication and integration framework processing unit 5211 finally determines the propriety of parallelization of the persisting process through the 1015 "forced serial process" flag of the parallel key definition file shown in FIG. 1 in the group name. The final determination as to the parallelization propriety of the persisting process confirms to the final determination flow for the propriety of parallelization shown in FIG. 3.

When the 1015 "forced serial process" of the parallel key definition file shown in FIG. 1 is "Yes", the data replication and integration framework processing unit 5211 determines the parallelization of the persisted process to be impossible in accordance with 302 in FIG. 3. That is, when "Yes" is taken, data having matched its conditions is treated to be same as the group name "Gr_OTHER". When the 1015 "forced serial process" of the parallel key definition file shown in FIG. 1 is "No", the data replication and integration framework processing unit 5211 determines the parallelization of the persisting process to be possible.

(Process after determination as to propriety of parallelization at transfer destination)

At 406 in FIG. 4, the data replication and integration framework processing unit 5211 performs the persisting process, based on the parallelization propriety

of the persisting process of the transferred data, which has been determined up to 405 in FIG. 4.

When the transferred data corresponds to the group name "Gr_OTHER", for example, the data replication and
5 integration framework processing unit 5211 does not parallelize the data but serially persists all data corresponding to "Gr_OTHER" one by one.

At the same time, the data replication and integration framework processing unit 5211 persists
10 corresponding data other than the group name "Gr_OTHER" in parallel with the data corresponding to "Gr_OTHER".

What is claimed is:

1. A data reflecting method for reflecting data of a transfer source system to a transfer destination system,
5 the method comprising:

applying a parallel key indicative of an ordering of each element constituting data targeted for reflection to the each element in the transfer source system;

transferring the data from the transfer source
10 system to the transfer destination system; and

parallelizing the element in accordance with the parallel key in the transfer destination system and reflecting the each element to a database in the transfer destination system.

15 2. The data reflecting method according to claim 1, wherein:

the data is transaction data for financial business;
and

the element is processing content that constitutes
20 the transaction.

3. The data reflecting method according to claim 1 or 2, wherein the parallel key satisfies a predetermined condition for a column name included in the data.

4. A data reflecting system for reflecting data of a transfer source system to a transfer destination system, comprising:

5 a storage unit for storing a parallel key definition file in which a parallel key for determining whether a plurality of data are processed in parallel is defined, the key including a column name and a table name included in the data; and

a data replication and integration framework
10 processing unit for determining, on a basis of a column name and a table name included in the data to be processed and a column name and a table name stored in the parallel key definition file, whether each of the data to be processed is executable in parallel, and for reflecting the
15 data to the transfer destination system in parallel when the data are determined to be executable in parallel.

5. The data reflecting system according to claim 4, wherein the data replication and integration framework processing unit is forced to serially process the data when
20 the data is determined not to be executable in parallel.

6. The data reflecting system according to claim 4, wherein:

the data processed by the data reflecting system is transaction data; and

the data replication and integration framework processing unit determines the data to be processable in parallel when the column name and the table name included in the transaction data and the column name and the table name stored in the storage unit correspond to each other and a value of these column names is same.

7. The data reflecting system according to claim 4, wherein the each data determined as to whether to be processable in parallel are reflected to the transfer destination system as a result of updating to the database included in the transfer destination system.

8. A data reflecting method for reflecting data of a transfer source system to a transfer destination system, the method comprising:

15 determining, on the basis of a column name and a table name included in data to be processed and a column name and a table name stored in a parallel key definition file in which a parallel key for determining whether a plurality of data are processed in parallel, and includes

20 the column name and the table name included in the respective data is defined, whether the respective data among the plurality of the data to be processed are executable in parallel; and

reflecting the each data to the transfer destination system in parallel when the data are determined to be executable in parallel.

9. The data reflecting method according to claim 8,
5 wherein the respective data are forced to be serially-processed when the respective data are determined not to be executable in parallel.

10. The data reflecting method according to claim 8, wherein:

10 the data to be processed are transaction data; and
the data are determined to be processable in parallel when the column and table names included in the transaction data the column name and the table name stored in the storage unit correspond to each other and a value of
15 these column names is same.

11. The data reflecting method according to claim 8, wherein the respective data determined as to whether to be processable in parallel are reflected to the transfer destination system as a result of updating to a database
20 included in the transfer destination system.