

2014

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(11), 2014 [5682-5689]

Study on application of availability data acquisition based on database mirroring

Ming Yang, Yong Yang

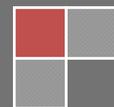
Jilin Technology College of Electronic Information, Jilin 132021, (CHINA)

ABSTRACT

Many businesses in the securities industry need to rely on a lot of information systems to maintain. Nowadays, with the increasing task amount in securities market, the demand of data information during operation is growing, and in order to fulfill the need for data, securities industry must analyze and process the data information. Therefore, securities enterprises' demand of data system is getting higher and higher, requiring a higher availability of data acquisition system. This article focuses on the application of securities industry commonly applied data warehouse, analyzing the defect of data acquisition by data warehouse during the operation of securities businesses. This paper makes a deep analysis of the securities company's data acquisition application, designing a new data acquisition application solution in two aspects of data acquisition task mechanism and system failover about the fault tolerance and availability research. So that the problems related to data acquisition in securities business can be solved by corresponding securities information system, and the securities business data can be analyzed and processed reasonably.

KEYWORDS

Data Acquisition, Data warehouse; Database mirroring; Failover.



INTRODUCTION

In era of information, the operation and management of an enterprise is conducted with the support of information system, especially large and medium-sized enterprises. As large enterprises, Securities companies have a certain particularity in their businesses. In business operations, for example, the securities companies are heavily dependent on information system, and a lot of information and data are produced in the operation process of information system, these data can be used by enterprise managers to acquire the company's application situation, the economic situation, and risk index conveniently. In this paper, the operation of securities companies is supervised and controlled through the information platform, of which the information platform is the data warehouse system. In specific business, this system supports the whole business operation of an enterprise by data information acquisition, including a star-schema data organization. During this process, each user in the securities company can obtain the information they need in various ways. The data warehouse system is divided into three main components in functional structure, including data acquisition, data storage and data access.

Securities companies are large enterprises, during operation process of their businesses, the data sources are enormously needed in data warehouse, with the amount of data reaching dozens of GB or even hundreds of GB. Synthesizing the above analysis of huge data volume and long collection time, unmanned automated database system is needed in data warehouse to correctly operate data acquisition within the specified time. Any error occurring in this detailed acquisition process will generate negative effects to the entire demand and acquisition of data. And any error in data system will bring tremendous pressure on corporate networks, with employees delayed in the working process, network congestion and lower operation speed. Therefore, in such serious situation, a set of data acquisition system with high availability is heavily needed for the enterprise. This system can switch automatically through redundancy backup to complete acquisition work when any default occurs.

THE PRESENT SITUATION AND CORE TECHNOLOGIES OF DATA ACQUISITION SYSTEM IN SECURITIES INDUSTRY

Structure of data acquisition application system in securities industry

Data warehouse mainly adopts three-layer architecture in gathering information, including the collection layer, load layer and transfer layer, of which the first two mainly apply the connection mode of communication transmission. Data acquisition system with three-layer architecture is as shown in Figure 1:

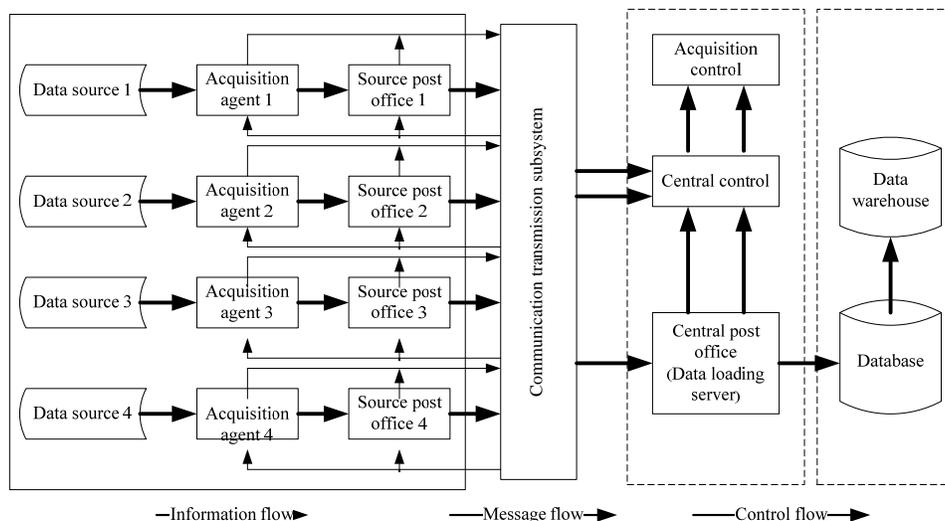


Figure 1 : Data acquisition system structure in securities industry**Application of redundancy backup mechanism in data system of securities**

IBM server is applied as database server for data information storage, and the data storage function which adopts asynchronous PPRC technology is used for data synchronization. The cold standby^[1] is used for information acquisition during this process. Because PPRC asynchronous technology is used for storage and the backup database has been in stop mode, that is to say, if any error occurs in the entire data acquisition process, the database needs to terminate the data synchronization task, and change the server from read-only status to read-write status. Then the database service in auxiliary machine will be enabled to switch database. After this, manually modify the data information in the database by the application configuration. The entire switch procedure will take 30 minutes or so. Acquisition agent facility doesn't realize backup throughout the data bank switching process, if any problem occurs, relevant re-configuration will be needed, which takes a lot of manpower and resources. Any default of information acquisition facility will not result in systemic breakdown of entire database system, but will cause great negative effect^[2] on information acquisition.

In the entire process of data management of securities businesses, some securities information need to be recorded in database table, such as stock rise and fall of different enterprises. For example, some investor has bought three different kinds of stock, A, B, C, of which the compensation is shown in TABLE 1.

TABLE 1 : Compensation of three stocks

Month	Compensation amount of stock A	Compensation amount of stock B	Compensation amount of stock C	Total compensation amount
January	1000	4000	-4000	1000
February	-1500	3000	-1500	0
March	2000	-500	500	2000

Among them, the calculation formulas of single stock return rate and standard deviation are as follows:

$$E(R_i) = R_i = \sum R_{ij} P_{ij}$$

The probability calculation formula is as follows:

$$\sigma_i = \sqrt{\sum (R_{ij} - P_i)^2 P_{ij}}$$

R_{ij} in the last formula represents the j -th return rate of the stock i ;

P_{ij} represents the j -th return rate of stock i .

The data table of the database also needs to record the risk measurement information of the securities industry, such as a portfolio investment of security A and security B, the economic situation, probability and investment return rate of the two securities are shown in TABLE 2:

TABLE 2 : The economic situation, probability and investment return rate of the two securities

The economic situation	Probability P	Rate of return on investment%	
		P (A)	P (B)
prosperous	0.5	30	10

general	0.3	10	-5
depressed	0.2	-25	15
grand total	1.0	-	-

The expected return rate and standard deviation of each security are as shown in TABLE 3:

TABLE 3 : Expected return rate, standard deviation and correlation formulas of securities

The economic situation	Probability P	Rate of return on investment%		E (A)	E (B)	$\sum (R_{Aj} - R_A)P_j$	$\sum (R_{Bj} - R_B)P_j$
		P (A)	P (A)				
prosperous	0.5	30	10	0.15	0.10	0.01445	0.0036125
general	0.3	10	-5	0.03	0.015	0.00027	0.0081675
depressed	0.2	-25	15	-0.05	0.03	0.02888	0.000245
grand total	1.0	-	-	0.13	0.115	0.436	0.005125

Among them, the formula of expected return rate of securities portfolio is:

$$E(R_p) = \sum X_i P_i = 0.5 \times 13\% + 0.5 \times 11.5\% = 12.25\%$$

standard deviation of securities portfolio is:

$$\begin{aligned} \sigma_p &= \sqrt{X_1^2 \sigma_1^2 + X_2^2 \sigma_2^2 + 2X_1 X_2 \sigma_{12}} \\ \therefore \sigma_{12} &= E[(R_{1j} - R_1)(R_{2j} - R_2)] \\ &= 0.5(0.3 - 0.13)(0.2 - 0.115) + 0.3 \\ &= 0.605\% \end{aligned}$$

$$\therefore \sigma_p = \sqrt{0.5^2 \times 0.2088^2 + 0.5^2 \times 0.0716^2 + 2 \times 0.5 \times 0.5 \times 0.00605}$$

Mirroring technology of database

Mirroring technology of database is referring to the database real-time backup mechanism based on software technology, mainly including software facility of Microsoft SQL Server provided by Microsoft. Microsoft SQL Server is a C/S mode Server, its main language is Transact-SQL language which plays a managerial role of data information in transmission of data warehouse information. Microsoft SQL Server software plays a strong promotive role in data warehouse, especially a supportive role to the securities business management when dealing with defaults of database system.

The mirroring technology of Microsoft SQL Server 2005 or higher versions is an effective solution of database defaults with high availability and compatibility, requiring no support of any hardware or software. The main function of this technology is data redundancy, finally realizing hot backup of database^[4].

Microsoft provides log stream output technology for data mirroring technology by Microsoft SQL Server 2008 to roundly optimize the system performance and minimize the network bandwidth finally. Database mirroring technology needs two management devices, namely the principal database and mirror database, installed on different servers and connected by TCP network technology. Database transaction processing system is more powerful than synchronous and asynchronous function of database log in specific function implementation, with a automatic failover function^[5].

In the securities business management system based on data mirroring, the main database plays a supportive role for external server, while the data mirroring facility has the same function with every

database in the entire process. Mirroring database has constantly been in reparation status for database system, information inquiry can be realized by creating database snapshots. In the whole database system, three work modes support database mirroring technology, namely automatic failover mode, manual failover mode, high performance manual failover mode^[6].

High availability mode

High availability mode gives support to the availability of database, with the primary mirror system sending the transaction log to the mirror server and waiting for the mirror server to confirm the entire process for the purpose to make sure the data synchronization in wrong process. This is replication of data synchronization, which enable the high performance business dealing primary server to play a part on function of mirror server. Data synchronization way is as shown in Figure 2. Among them, the data replication needs four request processes, namely:

- (1) request for data updates in database;
- (2) request for confirmation of transaction log sent to mirror database after data updates of primary database;
- (3) updates request of the mirror database after data updates of the database;
- (4) the primary database returns to application system successfully.

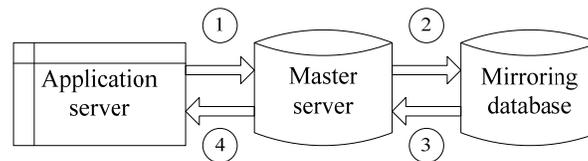


Figure 2 : Schematic diagram of database mirror synchronous replication

High protection mode

High availability mode is referring to the database performance and data integrity getting high-performance maintenance, and system keeping high transaction safety. Any error occurring in database system in this model needs to be handled by manual method, enabling database by the primary database^[7]. Operation of this mode needs support of high performance data, require no automatic failover configuration described above.

Master server, in high performance protection system mode, sends the transaction log to database mirror server, then sends a request to the database client. No wait is needed to confirm database mirror server by system in this process which calls data asynchronous replication during which the database mirror server doesn't have any effect on the master server. Asynchronous replication performs a minimum transaction latency on data through database mirror server. Any error in this process will be switched to database mirror by manual operation, causing loss of some data in database. The above mode requires support of high performance data system, and the entire error transfer form of database needs support of high protection mode.

Comparing the high performance failover mechanism with the traditional failover technology, the database mirror technology of the former has certain advantages of two relatively independent supporting database servers and two database files. In this process, only database level operation errors have negative effect on database mirror system; During the whole operation, it is convenient to execute the system setup and easy to operate, and no special requirements are needed^[8]. In addition, the database mirror technology itself has some limitations, for example, the process of providing solutions of software facility in database operation system may have influence on database performance; In addition, some contents in the database can't be synchronized in user login; Database failover will change IP address of the database, and require support of database client.

DEMAND ANALYSIS OF THE HIGH AVAILABILITY DATA ACQUISITION APPLICATION SYSTEM

System target analysis

This paper has deeply researched the application of data acquisition system in business procedures of securities enterprises, especially the receipt data acquisition of the risk assessment for securities enterprises, providing effective basis for risk index assessment. Data information acquisition system of securities companies has its own characteristics, including broad range of data information, diversified data format and abundant information gathering volume.

In the information acquisition of data availability based on database mirroring, several subsystems are applied in data source to support the system. The database mirror technology itself has some limitations, for example, the process of providing solutions of software facility in database operation system may have influence on database performance; In addition, some contents in the database can't be synchronized in user login; Database failover will change IP address of the database, and require support of database client.

Data of data mirror technology consists of two supporting parts of data, namely internal and external data sources and information subsystems. External information subsystem and data information mainly includes specific operation of business data information acquisition or so in securities firms. In addition, the internal data source of data mirroring is mainly composed of the subsystem used by securities company for business transaction, business transaction information, evaluation system of stock or so. Subsystem in this aspect provides effective basis and direct data security for risk assessment of the securities company.

The information acquired in the information gathering process by securities company can be divided into two main categories of data source, that is file data source and relationship data source. The external data source during the whole securities company business operation process is mainly the file data, adopting such data formats as Excel; The internal information during the whole securities company business operation process is mainly the relationship data. Among them, lots of data information is gathered in transaction of business system, mainly including basic transaction information of users and account balances in user system which requires daily collection and management by enterprise managers. Typically, the customer's basic information account contains many information collection resources. The data sources of the enterprise management contains file data. Use the availability resources of data information based on the data of mirroring to make certain evaluation about existing risk in transaction of securities company business.

Function requirement analysis of database system

In the whole operation of database system, comprehensive analysis of the data acquisition process should be based on requirement of data format and target data in actual practice. In the whole data acquisition process, the system needs to refine this task, so as to greatly improve the performance of the database and finally regard the data collected in the database system as a task. Due to the certain automatic function of the data acquisition in database system, corresponding record of data dates, which is used to confirm whether the acquisition has been done successfully, is needed in every transaction operation.

Acquisition date needs to be calculated when the data acquisition in the database system is started. On the basis of acquisition date set after midnight, the date can be set to the former trading day. Complete tasks in the task list after the specific acquisition date is figured out. Throughout the whole operation process, the system needs to restrict the tasks in waiting status. If there are no task in waiting status, manually set all the data date to waiting status. Then start the data acquisition thread, the system will automatically read data acquisition tasks and then return to the task executing status, showing

success or error message. The system will show completion status after the whole task is completed. Data acquisition work flow diagram of database system is as shown in Figure 3.

Performance demand analysis of database system

The emphatic analysis of database work shows that the database system needs the ability of high performance data service, which requires the data acquisition work to be done within the time limited by the system. In this case, the system can make comprehensive evaluation and analysis of the daily data information collected by risk supervisory control. High performance data acquisition application system of the database has strong effect on database maintenance, including data source performance, network bandwidth, performance acquisition of client and target database qualification, etc.

The data sources performance in the whole system operation process is free of control by any external factor. Moreover, network bandwidth and the database server performance have certain supportive effect on data acquisition. Improving the performance of database can make the database server write data information during the system connection process.

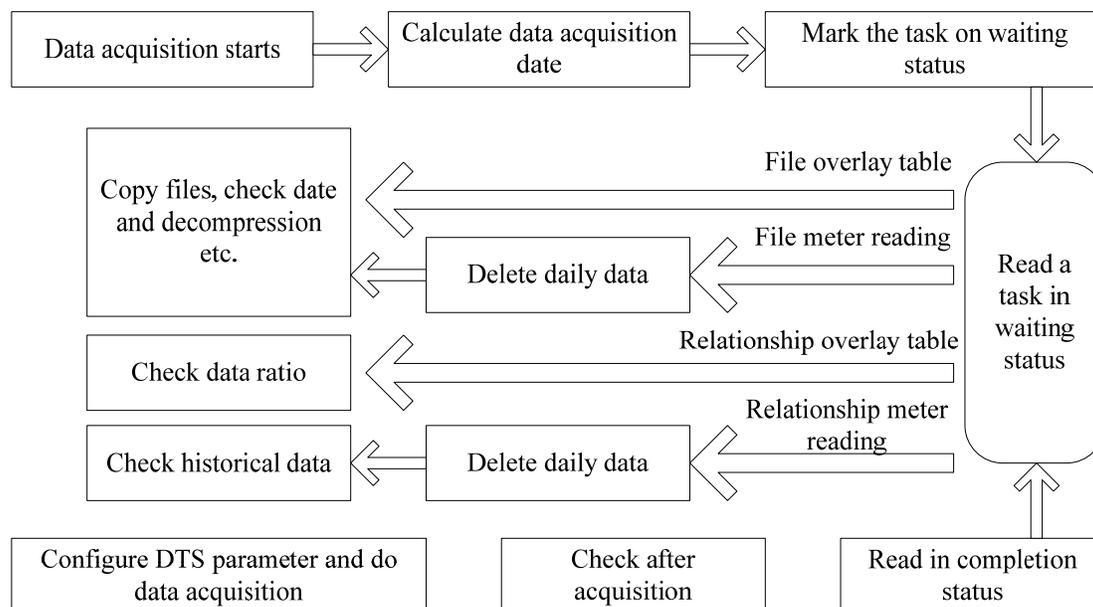


Figure 3 : Specific data acquisition work flow diagram

CONCLUSION

This paper has analyzed the specific demand of data acquisition in securities enterprises transaction operation by data mirroring principle. In this process, data acquisition system is a subsystem to locate the risk system in securities enterprises and a transaction management system based on data mirroring. Its primary database has certain supportive effect on external server, and the data mirroring facility in the whole process has the same function of each database. Mirroring database has constantly been in reparation status for database system, information inquiry can be realized by creating database snapshots. High performance data acquisition application system of the database has strong effect on database maintenance, including data source performance, network bandwidth, performance acquisition of client and target database qualification, etc. This paper has analyzed specific requirement according to the three correlation characteristics in securities business, namely concurrency, availability, fault tolerance of data information acquisition. The core of this article is the detailed analysis of the securities risk information, and judge risk level through correlation formula.

REFERENCES

- [1] J.Kang; Development and Application of Real-time Bridge Scour Monitoring System, *Engineering*, **3(10)**, 978-985 (2011).
- [2] Eroglu, B.Westrick; General Purpose Wireless Communication System Using Data Acquisition, *Communications and Network*, **4(1)**, 48-53 (2012).
- [3] R.Chebie, S.Wachanga, J.Kiboss; Effects of Science Process Skills Mastery Learning Approach on Students' Acquisition of Selected Chemistry Practical Skills in School, *Creative Education*, **3**, 1291-1296 (2012).
- [4] G.Sandoval-Romero, E.Pinzoń-Escobar; A Simple Detection System for Two Classical Sagnac Interferometer Configurations, *Engineering*, **4(12)**, 903-907 (2012).
- [5] R.Dallasheh-Khatib, R.Ibrahim, A.Karni; Longitudinal Data on the Relations of Morphological and Phonological Training to Reading Acquisition in First Grade: The Case of Arabic Language, *Psychology*, **5**, 918-940 (2014).
- [6] K.Mollaiyan, R.Santerre, R.Landry; Acquisition of Weak Signals in Multi-Constellation Frequency Domain Receivers, *Positioning*, **4(2)**, 144-152 (2013).
- [7] J.Murungi, G.Muthaa, L.Micheni, E.Mwenda; Strategies for Improving Parental Involvement in Acquisition of Literacy Skills among Pupils in Primary Schools in Akithi Division, Tigania West District, Kenya, *Open Journal of Social Sciences*, **2**, 250-255 (2014).
- [8] R.Pallugna, A.Cultura, C.Gozon, N.Estoperez; Low-cost Remote Rain and Stream Data Acquisition System for Mapping of Potential Micro-Hydro Sites, *Energy and Power Engineering*, **5(4B)**, 56-62 (2013).