

Design and Implementation of Grid Data Asset Management Platform Based on SOA

P. LOU & J. Chen & L.M. Lu

State Grid Huzhou Power Supply Company, Huzhou 313000, Zhejiang Province, China.

S.P. Ji & X.H. Li & Y. He & Y. Lin

China Realtime Database Co., Ltd, Nanjing 21012, Jiangsu Province, China

ABSTRACT: Power grid data has become an important asset of the enterprise, but power grid enterprises lack effective technical means to solve the grid data assets life cycle monitoring and management. From the design principle of component, loose coupling and standardization, the power grid data assert management platform based on SOA is designed and realized. A captured through the database log method to real-time acquisition of data assets change information, and account management, management of the panoramic view, daily monitoring and management, traceability management, valuation management and process management data asset management the core function is designed, which realizes the standardized management of grid data assets, and shows the status of data assets from the business, technology, management and other perspectives.

KEYWORD: SOA; Power Grid Data Assert; Database Log; Data Transaction

1 INTRODUCTION

With the advances in information communication technology, power companies have already gone through the stage of process-centric, and entered the stage of data-center. Building digital enterprise, enhancing operational efficiency is becoming the urgent requirement of grid development. With the development of business and level of information in enterprise, it needs to efficiently manage, analysis and utilize the all kinds of massive business data, and to gradually transform the data into core assets, fully exploit the value of data assets, then change enterprise management from "business-driven "to" data-driven "(Li, J.H. 2013). Therefore, power companies come to realize that data is immutable, and it is the most valuable assets of the enterprise.

Grid enterprises clearly put forward that data needs to be managed as a strategic asset, and establish the concept of data asset management, eliminate professional barriers, maximize asset utilization efficiency of enterprise data.

Centralized administration of operational data assets needs to strengthen the integration management, analysis and utilization of operational data, and mine the value of data assets. However, current management means and mechanisms are far unable to meet the requirements for the management of massive data assets, mainly include:

1) It lacks of global physical view. Such as it wants to know the number of data table, index and views in a particular database.

2) It lacks of different dimensions dictionary view and corresponding IT support. For the system already existed, it wants to know what are the key business data, who owns this data asset and who did modify or change?

3) It lacks of traceability of problem. How to calculate index A of statement and where is the original data source.

4) It lacks of change management. Why is there an increase of 100 tables in database? Why a field of this table is deleted, and when deleted?

5) It lacks of management of data running state. How much data for a certain sheets underlying table, and how the data growth?

6) It lacks of audit auditing mechanisms. In the design documents one table has 10 fields, then why there are 12 fields in the database?

The main target of data asset management is data. Faced with massive business data, data ledger management, data monitoring, data model, data traceability and other work will be difficult to carry out relying on traditional offline methods, and it can't guarantee effectiveness of data asset management data. Thus it urgent needs to develop data assets management platform, to achieve monitoring and management of data life cycle relying on the support platform (Xiong, X. 2014).

2 PLATFORM ARCHITECTURE

Service Oriented Architecture (SOA) is an architectural model, it advocates component-based, loosely coupled, isolated concern, standardization and so on design principles of architecture. The actual application can package business functions, information or process as the basic components- standardized services, these services can serve other applications through a loosely coupled SOA architecture (Li, M. & Wei, W. 2015). SOA includes service providers, service requesters and service agents, which are responsible for the service releases, service discovery and service bindings. Service operator releases service to service agents, the service requestor finds the required service by service agents and bind to these services, and the service provider and service requester can interact. For applications using a service, they do not care about the technology used by service providers, the position of service and what hardware platform provides the service, the process of using service is completely loose and transparent. On the other hand, the business system providing services can be reused, and it will not be invaded by other system, and the evolution of business systems platform will not affect the business system using its services. Grid data asset management platform based on SOA architecture can reduce coupling of the platform, and it also can provide ideal solution for the integration and extension of platform (Dong, Z.Y. 2015).

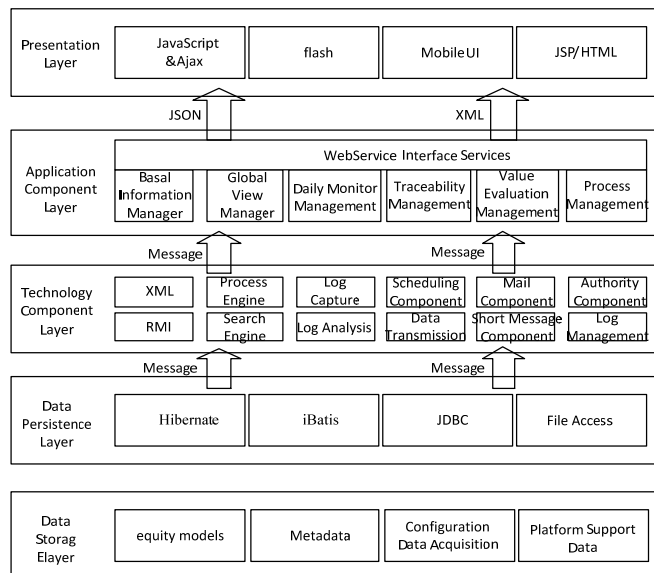


Figure 1. The general architecture grid data asset management platform

Press SOA architecture ideas, platform is divided into data storage layer, data persistence layer, technology component layer, application service layer, presentation layer, as shown in Figure 1.

Data storage: it mainly storages platform assets model, metadata, collected configuration data and platform support data.

Data persistence layer: the method is a combination of Hibernate, iBatis and JDBC, and Hibernate, iBatis is a more practical access plan for small amount of data with precise structure. Due to the inability to protect performance of Hibernate and iBatis for large and high requirements of data access, it uses JDBC directly read and write data. In addition, it provides separate file access components for the data of file type.

Technology component Layer: It packages and deploys many basic services alone on the platform so the services can be called and deployed easily, and it also uses Spring transaction management mechanisms and Ioc mechanisms based on Spring Framework. Business logic layer provides the interface for the control layer, and WebService interfaces and RMI Remote Method Invocation Interface for analysis and collection and other applications.

Application service layer: The platform provides a variety of data asset management applications, it achieves services by encapsulating logic components, and it can call services from other business systems to implement some of the features. These services can be called by various applications from interface layer.

Presentation Layer: It uses a variety of UI framework, includes Flash, Ajax, mobile UI and so on. For the analysis class graphics show, it can achieve good interactivity through Flash.

3 KEY TECHNOLOGIES OF PLATFORM

Grid data asset management platform needs to manage a wide variety of enterprise data assets real-time and effectively, so it involves automated data collection from business system. It requires real-time acquisition and minimal impact on existing business systems on the acquisition approach, and most of the core business system data is stored in relational database like Oracle, SqlServer, DB2 and other. Therefore, it is crucial to collect real-time data trading information in this class of databases (Qu, Z.Y. 2012). It presents a log capture technology to collect data trading information real-time. The working principle of the technology is shown in Figure 2:

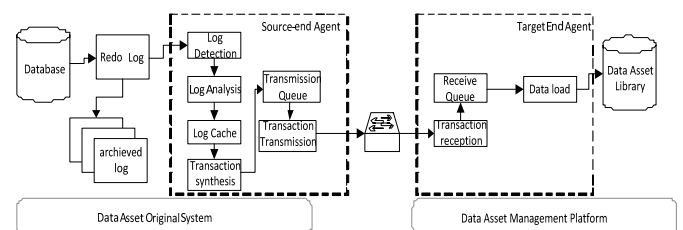


Figure 2. Automated data collection assets schematics assets

Agent processes are installed at the source and destination end, the agent at source monitor logs in Ora-

cle, SqlServer, DB2 and other, when find changes update target database timely. When the application system operates database at the source end, the information will be stored in the RedoLog. Agent obtains the orders and transactions data in this operation through analyzing the real-time log, and transforms these data into a compact format, then transmits these data to target system over the network. Agent in target system receives database package, then checks the check code, when make sure it is the correct database package, calls Oracle, SqlServer, DB2 and other functions to execute the transaction in the target system in accordance with the order of the transaction.

3.1 Log Detection

Agent determines whether there are new trading in Oracle, SqlServer and DB2 at source through checking the current SCN number recorded in Controlfile periodic. When confirms there is a new trading in Controlfile, Agent gets the current position of Redo Log (or Transaction Log) group and the latest log file. Agent module analyzes the Log between positions of the last crawled record and the lasted according to these information, and saves the data in the Log Cache file and waits for the next step synthesis processing transactions.

3.2 Log Analysis

All changes to Oracle, SqlServer, DB2 and other databases are recorded in the log. When we need to understand the transactions made in the database, we can achieve it through log analysis capabilities integrated in Agent. Through log analysis, it can get all SQL commands in database, and these commands will generate in a compact format expression. It can directly convert the compact format expression into Oracle, SqlServer, DB2 and other internal data representation format through a proprietary algorithm. When analysis and convert, it needs only a minimum format conversion, hereby improves performance and efficiency.

3.3 Trading Synthesis

There are two characteristics of the orders analyzed by Redo Log (or Transaction Log) in Oracle, SqlServer, DB2 and other:

These instructions are cross arise. A plurality of SQL commands of a transaction (Transaction) are stored non-contiguous; SQL commands in a plurality of transactions are interpenetrating.

All transactions are recorded in Redo log, including the already Commit transaction and not Commit transaction. In order to improve the system controllability, ensure the logical integrity and prevent data loss, it is best to copy the smallest unit as a transaction (Transaction), so that the load on the target side of the transaction is more easily controlled.

While for the replication of data, only those Commit data is significant in the target, Rollback data is unnecessary copied into the target system. So not each SQL statement is necessary copied, it copies the data integrated after transaction (Transaction), and only copies Commit transaction.

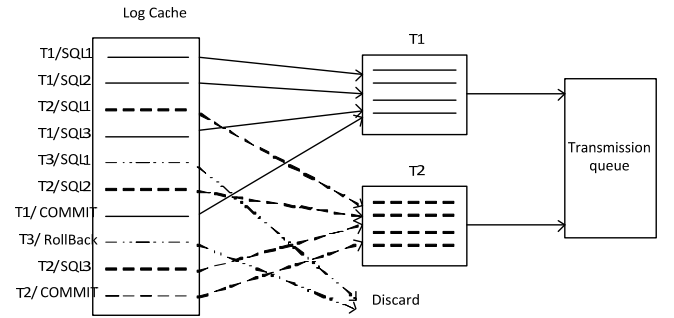


Figure 3. Database transactions synthesis scheme

Shown as Figure 3, Online Log Cache file contains the Commit transaction, the not Commit transaction and Rollback transaction. Trading synthesis module first divides SQL statement in accordance with the transaction number, then passes the Commit transaction to the transmission queue at the unit of transaction, stores the not Commit transaction locally and discards the Rollback transaction. When the transaction synthesis module find that the not Commit transaction has been submitted, the transaction will be immediately sent to the transmission queue.

3.4 Trading Transmission

In order to ensure the safety and reliability of data transmission, transaction data is firstly stored in transmission queue in the Agent at source end before transmission. The transmission module reads transactions from the queue in accordance with the FIFO principle and packages it to TCP / IP data packets, then transmits data to the receiver module on the target side. On the target side, after receiving module receives transaction data packets, it firstly checks the legality of the transmission according to the packet size described by header to determine whether transmission complete. If transmission is not completed, feedback Agent at source end and then retransmit. If complete, it will store the transaction data in the receive queue, then load module load transaction information in strict accordance with the order from the queue.

3.5 Data Load

Data loading module will accept transactions data in queue, and write data into asset database directly by the underlying function through calling Oracle, SqlServer, DB2 and other database I / O layer interfaces. It makes the performance and processing power of the loading end improved significantly.

4 FUNCTIONAL DESIGN OF PLATFORM

Grid data asset management is entire process and comprehensive method for the transformation, asset operation and maintenance as well as asset value applications of enterprise data assets. The goal is to carry out intensive, unified, standardized asset management for operational data, and to conduct data quality control and assessment, to promote effective data analysis and utilization, finally do data mining to maximize data assets value.

Based on demand of data asset management, grid data asset management includes the following features:

4.1 Basic Information Management of Data Assets

Basic information management of data assets primarily manage the categorical, accounting, coding, business objects, data models, rules and relationships of data.

1) Data asset classification data management: It defines data asset classification, forms data assets classification criteria, forms a clear range of data assets.

2) Ledger archiving of data assets: It establishes a detailed accounting of data assets, provides data assets ledger maintenance functions, maintain the technical attributes business property, property management and basic properties of data assets.

3) Uniform code maintenance: It unified manages the encoding involved with data assets, defines data encoding specification.

4) Key business object management: It manages critical business object involved with data assets.

5) Data model management: It establishes data model, unified manages and maintain business data entry, reference data.

6) Rules management: It manages and maintain quality validation rules, daily monitoring and early warning valuation rules involved with data asset.

7) Data association management: It manages relationship with data, maintain data related audit rules, sets verify relationship between across business field data.

4.2 Panoramic View Management of Data Assets

It displays enterprise-wide data assets with a panoramic view, and reflects the full range of data assets from the business, technology and management perspective.

1) Panoramic view shows: It displays a panoramic view of the data assets graphically.

2) Panoramic view maintenance: It supports for new panoramic view, modifying and deleting functions, supports visualization, supports for dynamic configuration techniques property, business property and management property, supports for defining relationship models.

3) Multidimensional perspective display: It allows to view data from a technical perspective, business perspective, management perspective and asset perspective.

4) Completion rate of data assets show: It display data assets completion rate graphical, reflects the data assets completion levels and provide support for the data assets completion rate evaluation.

4.3 Daily Monitoring and Management of Data Assets

According to panoramic view of data assets, it sets key monitoring point to provide routine monitoring of critical control point changes, data timeliness, completeness, accuracy, verification, data tampering and the like.

1) Automatic monitoring of data assets: It develops automated data monitoring program, maintain content based on each node evaluation rules, and carry out the automatic evaluation to alarm the data exceeds the set threshold.

2) Key monitoring of data asset: It sets data asset monitoring points in the business and technical aspects based on data assets traceable graph, monitors and early warns to the monitoring points automatically according to monitoring rules.

3) Monitor of data association: It completes follow-up analysis through association validation rules between data across the business.

4) Data tamper monitoring: It reviews access to the database, develops data tamper-proof rules, monitors data tampering real-time according to the rules.

5) Manual inspection of data assets: It supports for artificial checking detailed data of each node, automatic calculates results in accordance with the definition, supports site audit work.

4.4 Traceability Management of Data Assets

It combs the data assets thoroughly from the source, manages the data assets traceable graph based on the combing results. The data assets traceable graph can be dynamically maintained and drill query.

1) Data traceability node configuration: It dynamically defines and configures the data traceability nodes.

2) Data traceability graph visual presentation: It has visual representation of data traceability diagram, and supports properties view.

3) Data traceability graph maintenance: It can new add, modify and delete visual data traceability graphs.

4) Traceability graph version management: It records change history of data traceability graph, provide version control function.

5) Origin node data details check: It supports for data detail query of each storage node in traceable graph.

4.5 Valuation Management of Data Assets

Data assets valuation model maintenance: It provides new add, delete and modify features, manages data assets graphically based on valuation model.

Data asset valuation: It values each node in data traceability graph based on data assets valuation model. It evaluates data reliability, timeliness and completeness based on valuation rules, and calculates the value of data assets in accordance with valuation model.

4.6 Process Management of Data Assets

Through the data asset management approach, it manages the formation, changes and deactivation process of data assets in the means of information, and establish a standardization process of data asset management, it is shown in Figure 4.

1) Data assets add, change or deactivation application: It provides new adding and adjusting demand application function of data assets, and the application is proposed by business sector.

2) Data asset preliminary: It provides trial function to data assets. Data assets management sector carried out trial to the demand proposed by business sector, if the trial is not pass then return it to business sector.

3) Access verification of data assets: Data assets management department carefully verifies the data assets after preliminary approval, then carry out the access verification. It needs to determine whether the data consistency and data completeness meet the requirements, if not, the business sector needs to rectify data assets.

4) Data assets rectification: Business sector rectifies the problems of data assets, records the rectification of data assets for all types of problems exit and control data assets version changes.

5) Audit data assets: For the data compliance with access rules, the asset management department is responsible for auditing and confirming.

6) Archiving data assets: After the audit of data assets completed, the business sector improves business property of data assets, information sector improves its technology properties then a completed data assets information is formed. Asset management department is responsible for data archiving.

5 CONCLUSION

The grid data asset management platform in this paper has been successfully applied in China Southern Power Grid Corporation. It manages the enterprise data as a core asset. The platform continue to monitor data value through the whole process of management and rational use of operational data then to achieve the transformation of data to information, followed by the formation of business value and achieve core business capacity building, information value chain expansion and optimization of enterprise business model. Meanwhile, it deepens the data governance, improves the quality of the data and provide an effective way for company to monitor the health of the business promptly and quickly.

REFERENCE

- Cui, J.F. 2014. Data Asset Management Research and Exploration. China Power Enterprise Management 12: 94-95.
- Dong, Z.Y. 2015. On the enterprise data quality management. China Management Informationization 3: 78-79.
- Li, J.H. 2013. Study on East China Power Grid Calculation Unified Data Management Platform. East China Electric Power 1: 69-72.
- Li, M. & Wei, W. 2015. Design and Practice of Master Data Management Architecture Based on SOA. Ordnance Industry Automation 8: 49-51.
- Qu, Z.Y. 2012. Real-time Duplicate Software in DB2 Large Database. Computer Systems & Applications 3: 198-201.
- Xiong, X. 2014. Design and Implementation for the Master Data Management Platform of Iron and Steel Enterprise Application Integration. Information Technology and Informatization 1: 21-24.

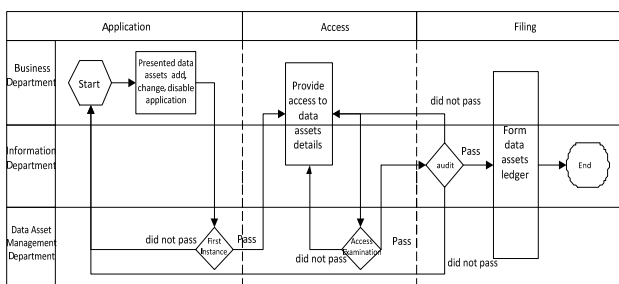


Figure 4. Data asset management flowchart