

## Application of Data Mining Theory in Power Grid Equipment Defect

Wang Hongbin<sup>1,a</sup>, Zhang Hu<sup>1,a</sup>, Li Feng<sup>1,a</sup>

<sup>1</sup>Electric Power Research Institute of Guangdong Power Grid Co., Ltd., Guangzhou  
510080, Guangdong Province, China

<sup>a</sup>wanghongbin\_gd@163.com

**Keywords:** Power Grid Planning; Analogue Simulation; Application Data; System Design

**Abstract.** In order to meet the structural and functional demands of power grid planning research simulation laboratory for application data processing, a set of application data subsystem serving the power grid planning research simulation laboratory is designed according to the actual demand of power grid planning simulation work. Not only the overall framework but also the specific functions are described in this article in order to present the design thought for the data subsystem serving the power grid planning research simulation laboratory, wherein the data bus based integration method is adopted for the application data subsystem, and CORBA is adopted to solve the problem regarding the interoperability of the distributed objects of different heterogeneous systems, and open structure is adopted to provide the network architecture with key function redundancy and able to support distributed processing environment in order to meet the demands for expandability, safe reliability, openness and easy maintenance. Specifically, the main functions realized thereby include: data exchange platform, data center, graphic center, data query, information service, WEB, uniform information maintenance platform, uniform platform support for accessed/embedded third-party application functions and other basic functions, and these functions aim at improving power grid planning simulation level and verifying more reasonable planning scheme in order to create greater benefits.

### Introduction

Point-to-point integration framework is adopted in traditional integration mode, and due to the non-reusable and non-upgradable interfaces, etc., the interface type data integration cannot adapt to the long-term integration objective of an enterprise. In this article, according to the achievements of information technology development, we select the newest data bus based integration method to establish the data integration platform in the data center, the complete interface system and the data exchange platform on the basis of CIM/CIS and data standardization research in order to maximally realize data access, sharing and exchange.

### Overall Architecture Design for Application Data Subsystem of Power Grid Planning Research Simulation Laboratory

In allusion to the informatization characteristics of the application data subsystem of the power grid planning research simulation laboratory, we strive to construct a set of practical, advanced, mature and reliable application data subsystem for the power grid planning research simulation laboratory according to the design objective and the functional requirements of the system and meanwhile combine the present advanced and mature information technology and idea for system design, thus to provide reliable and comprehensive basic data for power grid planning simulation, improve power grid planning simulation level, verify and find reasonable planning scheme and create greater economic benefit and social benefit.

Basic data are sourced from ERP system, PMS system and SCADA system, and these data are collected to the data warehouse through different data adapter interfaces and uniform data service bus, wherein these data are divided into real-time data, historical data, graphic data and GIS information data. Data service is corresponding to application service, and the application data subsystem provides such relevant application services as algorithm call to the power grid planning simulation platform through the application service bus.

## **Software Architecture of Application Data Subsystem of Power Grid Planning Research Simulation Laboratory**

CORBA, as one of the ideal schemes for realizing the distributed computing technology of different heterogeneous systems, is mainly used to solve the problem regarding the interoperability of the distributed objects in a heterogeneous environment. Among lots of middleware such COM and DCOM, CORBA has the best timeliness. Meanwhile, CORBA not only allows static call, but also allows dynamic call, and has such advantages as high-level language binding and system self-description. CORBA distributed object model which is combined with J2EE environment through JavaIDL represents the most interoperable distributed standard model. Therefore, the application data subsystem of the power grid planning research simulation laboratory shall be located in the software system of CORBA/J2EE based multilayer C/S and B/S hybrid structure.

Software configuration and design of the system shall follow the basic principles below:

The system shall be able to provide the real-time, multitask and multiuser operating environment and the mechanism which allows the high-priority process to interrupt the low-priority process, and shall be able to monitor the high-resolution clock and timely awaken corresponding process, respond to and process various hardware and software interrupt requests, and automatically arrange the priority of these requests.

UNIX operating system must be adopted for database server, application server, WEB server, etc., and all workstations shall not only support UNIX/Linux, but also support WINDOWS system. The system shall provide corresponding mechanism to prevent data and files from being destroyed, and only the allowable privilege process can write and read the designated data area. Additionally, various hardware drive programs shall be included in the operating system.

Database is the core of the information system, so all applications such as WEB, information exchange and data processing shall be focused on the database. Specifically, the central information system database must meet the requirements for the concurrent users to safely access the data in the database, and the data stored therein must have large expansion capacity, high openness and good interoperability, thus to meet the high reliability requirement of key businesses. Additionally, the database adopted thereby shall be enterprise-level Oracle.

## **Functional Design for Application Data Subsystem of Power Grid Planning Research Simulation Laboratory**

The general functions of the application data subsystem of the power grid planning research simulation laboratory include: data exchange platform, data center, model center, graphic center, data display, data query, information service, WEB, uniform information maintenance platform, uniform platform support for accessed/embedded third-party application functions and other basic functions.

### **Functional design for application middleware layer**

#### **1) System management**

System management as the core function of the system is responsible for the construction and the maintenance of the system core. Specifically, the system management includes the following basic functions: basic system configuration management; system node management; system server management; system process management; basic system operating environment configuration.

#### **2) System security**

The system security management mainly aims at providing system security solutions to ensure safe computer operation, and includes operating system security strategy, network security (prevention of hacker invasion), virus detection, vulnerability scanning, password management, etc.

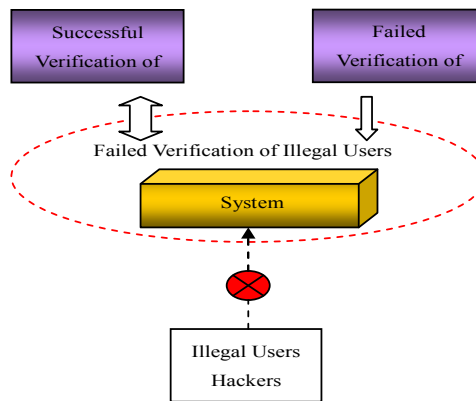


Fig.1. System security management diagram

### 3) System diagnosis

The system is provided with many monitoring and diagnostic tools to prevent the occurrence of faults or rapidly recover the system when the system is about to have faults or has already suffered from faults, thus to ensure the safe and reliable operation of the system. Specifically, these monitoring and diagnostic tools include the following functions: monitoring system operating condition; monitoring the local shared information of the communication gateway and receiving relevant information; monitoring the connection management of the communication gateway and the operating condition; fault diagnosis; inquiring and managing the monitoring and diagnostic results; diagnosing the system according to system logs.

### 4) Network management

The application data subsystem of the power grid planning research simulation laboratory is a distributed system and has complex network structure and many network nodes, and in order to ensure the reliable network operation, the system is provided with a series of network management functions as follows: configuring system network node; configuring server network node; monitoring network operating condition; automatically switching dual network; automatically balancing dual network loads.

### 5) Process management

Due to complex system functions and many applications, each node needs to operate a lot of processes to realize different functions. In order to ensure normal process operation, the system is provided with the perfect and reliable process management functions as follows: configuring the processes running at different nodes; configuring the processes running at different servers; configuring process starting mode including daemon starting, periodic starting, timed starting, etc.; monitoring the operating conditions of the processes at different nodes; sending alarm for process exception and automatically restarting relevant process.

### 6) Redundancy management

In order to ensure system operation security, such redundancy modes as network redundancy and service redundancy are adopted for the system.

### 7) Service location

The system is provided with uniform service location mechanism, so when an application needs to interact with a certain service, this application only needs to provide service name, without the need to care about the node location, IP, etc. of the service, and the service location module at the middle layer of the system platform will immediately associate the application with the corresponding service.

### 8) Real-time database

As the real-time data core of the application data subsystem of the power grid planning research simulation laboratory, the real-time database adopts client/server structure and takes high-speed RAM memorizer as the memory medium to provide rapid data access to the clients through the standard, uniform, open and transparent data access. Such real-time database not only has good timeliness, but also can support concurrent multiuser access.

The real-time database has following functional characteristics: distributed client/server system

structure; rich access modes; flexible and open access interfaces; online real-time database construction; automatic consistency; data dump and recovery; real-time database display tool; real-time database maintenance tool.

#### 9) Log record

The system is provided with general log function, and each application can define different log formats according to the actual needs thereof. Specifically, the main log functions are as follows: operation log; running log; log query.

#### 10) CORBA based soft bus

CORBA provides basic facilities to solve the problem regarding the heterogeneity of the database system, and meanwhile provides excellent method and middleware function to design and realize the distributed multi-data source integration system, thus to significantly reduce development workload and make the system have good openness, expandability and scalability.

### Functional design for data support layer

#### 1) CIM/CIS based data bus

CIM/CIS data bus is the data access bus based on CORBA standard technology and following IEC international norm CIM/CIS. Such bus supports multi-language access, e.g. C++, Java and PHP, and can provide open service for development, thus to be favorable for the realization in the application layer. Additionally, such bus is based on CIM/CIS and can ensure that other relevant application systems following CIM/CIS can friendly access the data in the application data subsystem.

#### 2) Log service

The system has rich log information for convenient system debugging and troubleshooting. All log information is uniformly managed in the system, and meanwhile the shared buffer is also provided in the system, and all application programs and system log information are put in this buffer. Additionally, special tools are also provided not only for convenient log information search, but also for remote inquiry. Actually, the detailed log information is crucially important for system debugging and troubleshooting, and is favorable for timely solving relevant problems to ensure the continuous and reliable system operation.

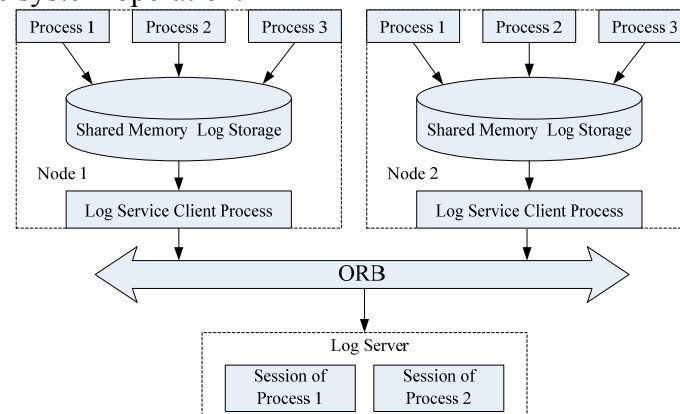


Fig.2. System log function diagram

#### 3) Global data dictionary

Metadata refer to the “data related to data” and are used to describe the position, source, content, attribute and state of the data in the existing application system. Metadata mapping refers to the mapping between the class in CIM and the table in the existing database. CIM is object model and most of the existing databases are relational databases, so some classes and attributes cannot be mapped to the tables in the existing database one by one. In such case, metadata mapping can be used to define certain operating methods for these classes or attributes in order to effectively realize model transformation.

Different data models in the existing system can be uniformly mapped into the models in accordance with IEC61970 standard through metadata mapping, thus to effectively eliminate the barrier brought by inconsistent data model to enterprise integration.

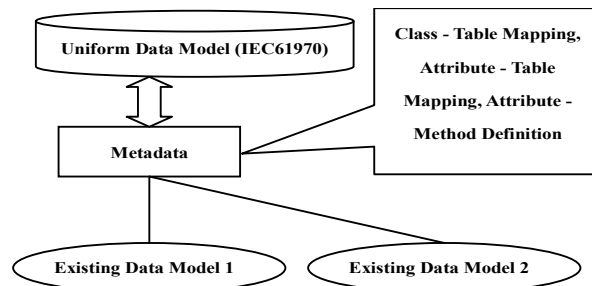


Fig.3. Metadata mapping diagram

## Conclusion

The complexity of power grid planning work has determined the difficulty in designing the application data subsystem of the power grid research simulation laboratory. Not only the overall framework but also the specific functions are described in this article in order to present the design thought for the data subsystem serving the power grid planning research simulation laboratory. The newest integration method based on data bus and service bus is adopted for the system based on CIM/CIS and data standardization. Specifically, multiple application service interfaces are provided to the upper layer to realize the uniform and multi-platform function for the upper layer; multiple data adapter interfaces are provided to the lower layer to shield the difference in the data structures of different data systems. The data integration exchange platform can maximally realize data access, sharing and exchange. Structurally, the open structure is adopted to provide the network architecture with key function redundancy and able to support distributed processing environment in order to meet the demands for expandability, safe reliability, openness and easy maintenance.

At present, the system has been preliminarily applied in Zhejiang Economic Research Institute. The application data subsystem of the power grid planning research simulation laboratory designed in this article has provided comprehensive and reliable data service for the power grid planning work in Zhejiang Province, and such system can not only improve power grid planning efficiency, but also improve the accuracy of the power grid planning scheme, thus to have high promotion value.

## Reference

- [1] Zhang, Mengxin, Zhihan Lv, Xiaolei Zhang, Ge Chen, and Ke Zhang. "Research and Application of the 3D Virtual Community Based on WEBVR and RIA." *Computer and Information Science* 2, no. 1 (2009): p84.
- [2] Su, Tianyun, Zhihan Lv, Shan Gao, Xiaolong Li, and Haibin Lv. "3D seabed: 3D modeling and visualization platform for the seabed." In *Multimedia and Expo Workshops (ICMEW)*, 2014 IEEE International Conference on, pp. 1-6. IEEE, 2014.
- [3] Jiang, Dingde, Zhengzheng Xu, Peng Zhang, and Ting Zhu. "A transform domain-based anomaly detection approach to network-wide traffic." *Journal of Network and Computer Applications* 40 (2014): 292-306.
- [4] J. He, Y. Geng and K. Pahlavan, Modeling Indoor TOA Ranging Error for Body Mounted Sensors, 2012 IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Sydney, Australia Sep. 2012 (page 682-686)