Research on Technical Architecture and Application of Big Data Cloud Platform for Electric Power Measurement

Shang Huaiying1,a, Liu Yan1, Zheng Angang1, Qin Songfeng2, Wang Yong2 and Zhang Wulei2

1China Electric Power Research Institute, Beijing 100192, China
2State Grid Henan Electric Power Co., Ltd., Zhengzhou 450052, China
a Corresponding author: shanghuaiyingip@163.com

Abstract. This paper studies the technical architecture and composition of the cloud computing platform. The data types and data source related to the big data of the power energy measurement are discussed. The design of the large data application service scene of the power energy measurement is discussed in combination with the technical targets and business objectives. Finally, the development of cloud platform and its application in business field are prospected.

1. Introduction
At present, the world is in a new round of technological revolution and industrial transformation led by the Internet. The relevant national documents point out that the energy Internet is the key strategic support to promote China's energy revolution. It is the related technology and thought of comprehensive collection, processing, analysis and application of data in the energy field such as electricity, oil and gas, and data in other fields such as population, geography and meteorology[1]. In the application of new technologies to form an intelligent, multi-energy cooperation, symmetrical information, decentralized supply and demand, flat system and open trading power supply internet, large data of power mainly come from power generation, transmission, substation, distribution, power consumption and dispatch, which can be roughly divided into three categories: first, power grid operation and equipment detection or monitoring data, such as power grid operation and equipment monitoring data, transformer insulation status, arrester current data, circuit breaker contact temperature and so on; second, marketing data of power enterprises, such as transaction price, electricity sales, electricity customers and other data; third, management data of power enterprises, such as metering point, electricity consumption, electricity consumption and so on[2].

At present, the work of power metering, meter reading and charging has realized the mode of time-sharing metering and collecting, basically getting rid of the backward mode of manual meter reading and manual settlement, but there is still a big gap between the real-time metering, information exchange and active control under the mode of energy internet.[3] The relevant data acquisition methods can not meet the requirements of timeliness, security, multi-dimensionality and multi-data sources of large energy data. At the same time, government agencies, power grid enterprises, measurement equipment manufacturers, power users and so on have increasingly strong demand for the application and service of power energy measurement data[4,5].

This paper mainly discusses the technical framework of big data of power metering under the background of energy Internet to complete the process from data acquisition to storage, calculation
and business scenario analysis. At the same time, combined with the current situation of measurement data of the State Grid, this paper discusses the subject of quality of measuring instruments and puts forward some specific application scenarios.

2. Cloud Platform Technology Architecture Research

2.1. Objectives of Cloud Platform Technical Architecture

The goal of the technical framework of big data cloud platform for electric power measurement is to provide a set of technical framework to realize the collection of complex data in the field of electric power measurement, the storage of structured and unstructured data, the large data operation of intensive data and the development of various application scenarios. Through the layered framework and technology, the collection, storage, analysis and content output of various business data and related business scenarios are completed one by one.

2.2. Data Acquisition Mode

From a technical point of view, the characteristics of data used to show business status (such as whether it is digital, discrete or continuous), the storage form of source data and the frequency of data generation are different. In order to ensure the integrity and effectiveness of data acquisition process, different acquisition modes should be adopted. For big data of electric power metering, the data covers a wide range from the design and development of instruments, production and manufacturing, logistics transportation to verification, inspection, installation and operation. Data classes include structured information such as type, voltage and current of measuring instruments, and unstructured information such as technical standards, contract accessories, etc. The data type is complex. The data can be generated periodically, such as watt-hour meter operation detection, or real-time, such as watt-hour meter faults in the process of operation. The generation period of data is different. In view of the above characteristics of big data of measurement, its acquisition mode is shown in Figure 1.

![Data acquisition mode of power metering data.](image)

Data collection and acquisition is an important basis for the follow-up technical links. The process of large data acquisition for power metering is shown in the figure above. Data sources can be divided into five categories. One is data related to enterprise operation and production, such as data generated by enterprise planning (ERP), product life cycle management (PLM), production execution system (MES), supply chain management (SCM), and data generated by power application system such as marketing business application system, metering production scheduling system, etc. IOT data is used to generate real-time data for collecting and transferring equipment operation status, working conditions and environmental parameters. Network and social media data, which are related to the production and operation activities of enterprises, and the data comes from the external Internet. Machine generated data, such as terminal positioning data, network signaling, machine operation log and so on. In the end, other data.

These data are input and collected by database interface program, stream data acquisition tool, human-computer interaction, intelligent terminal PDA, IOT sensor, RFID identification, bar code scanner, network crawler, etc.
2.3. Data Storage and Computing

Among the four V's of big data, Value feature is the most important, that is, no matter how complex the size and structure of data are, the data that can bring value is the most important data, and the value of data is reflected by analysis and calculation. For the analysis and processing of large data, the current typical scheme is OldSQL (old relational database) + NewSQL (new relational data) + NoSQL (non-relational database) + other new technologies (stream, real-time, memory, etc.). Through the combination of technology application, the problem of inadequate cross-analysis ability, deep mining ability and iterative analysis ability of traditional database is solved, while taking into account non-knotting, feature extraction of structured data, content retrieval and value mining of semi-structured data. The storage and calculation framework of large measurement data is shown in Figure 2.

Figure 2. Storage and computing architecture of large data cloud platform.

For storage, relational database, distributed database and distributed file system are used for storage according to different data types. Distributed file system is used to store unstructured data, such as various technical standards and specifications related to measuring instruments. Data is stored on physically dispersed nodes, providing users with high data access capacity through file system access interface. Distributed database is used to store semi-structured data and structured data, such as the producers of measuring instruments manufacturing enterprises. Distributed time series database and distributed memory database are used to store data, such as automatic verification and detection pipeline test result data of measuring instruments. And relational database is used to store structured data, such as fault data during the operation of electric energy meters. After that, all kinds of data preprocessing are carried out.

After data pretreatment, data calculation is carried out. According to the requirement of data processing time in business scenarios, data calculation can be divided into three types: on-line, near-line and off-line. The former can complete data processing in seconds, near-line and processing in minutes. Usually, memory computing is used. The latter takes days as the unit of processing time and uses off-line computing method.

Data analysis is the embodiment of the capability of big data platform, and it is also the key to determine the output of the final service scenario. Data analysis methods include statistical analysis, data mining and machine learning. Different methods are adopted according to the needs of scenarios. Statistical analysis is based on the analysis of data to describe the nature of the object, such as describing the relationship between the power meter fault and the region. Data mining is based on the analysis of data to automatically find useful information, such as classification decision tree algorithm to early warn and judge the abnormalities in the production and verification process of the power meter. Machine learning is obtained by automatic analysis from the data through "self-learning" of the
system. The rule is used to predict the unknown data. For example, based on the abnormal clearing fault of the existing watt-hour meter, the relationship between the fault and the moisture-proof ability of the non-components such as filter capacitor is analyzed, and whether the watt-hour meter with a certain type of components can be used in the high-voltage and high-wetland area is automatically predicted.

3. Business and Application Research

3.1. Business and Application Objectives
Based on the technology platform, through opening up the business flow, data flow and information flow among the relevant parties in the measurement field of smart grid, the scenario applications of quality monitoring, quality analysis and early warning, and whole process quality evaluation are realized, which provides valuable information and services for the manufacturers, power companies, power users and governments of measuring instruments, and helps the producers of measuring instruments improve their design. The quality of production will help power companies to improve the inspection quality and operation level, and promote the overall level of development of the industry.

3.2. Data Source and Data Acquisition
According to the orientation of business and business objectives, there are four types of data sources involved in the large data cloud platform of electric power measurement, namely, enterprise operation and production data, equipment material link data, network and social media data, machine generated data. Among the data listed, the data of enterprise operation and production include the data of measuring instrument manufacturer and the data of power company inspection, operation and maintenance. According to the relationship between upstream and downstream industries or the execution order of business, these data mainly include eight links: research and design, material procurement, production and manufacturing, ex-factory supply, acceptance and inspection, warehousing and distribution, installation and operation, dismantling and discarding. The data generated in the first four links are from the production enterprises, and the data generated in the last four links are from the power companies. Specifically, the data of the first four links are mainly generated in the production enterprise’s product life cycle management (PLM), production execution system (MES), supply chain management (SCM), customer relationship system (CRM) and other systems. The data of the latter four links are mainly produced in the marketing business application system, the metering production dispatching platform and the electric energy acquisition system of the electric power company.

In the four types of data, the process of collecting data related to enterprise operation and production is shown in Figure 3. Data acquisition methods can be divided into three types according to their sources: first, data interface services can be directly collected from various business systems of enterprises, such data can be directly collected through web services, intermediate databases or files; second, through various types of intelligent business terminals. The third is the data that cannot be collected directly and need to be added manually by PC or terminal. All data are produced through the
data collector unified access and cloud data access management services to complete the work of data into the cloud. The other three types of data are collected by means of data interface and network crawler.

3.3. Service and Application Scenarios

Cloud platform services are mainly composed of application services and data services. In figure 4, application services are mainly provided to users in a direct visual way. Users complete access through browsers or intelligent terminals. Data services are provided in a non-visual way. Users complete access through data interfaces to support subsequent personalized processing.

![Figure 4. Cloud service type.](image)

Application services (service scenarios) can be classified into four categories according to different service objects, as shown in Figure 5, which are oriented to production enterprises, grid companies, government, third parties and other users.

![Figure 5. Service scenario for different users.](image)

In the service scenario for manufacturing enterprises in the figure above, it includes all kinds of service scenarios for manufacturing enterprises, such as the analysis of the comprehensive research and design capability level of enterprise capability domain. The main content of this service scenario is to analyze the comprehensive research and design capability level of manufacturers in the industry, and to analyze the factory from three dimensions of research and design tool optimization, research and design management capability and product development capability. To evaluate the research and design capability of a family, three steps are used to calculate and analyze the data for each dimension. One is the refinement of the indicators, the second is the calculation of the optimal value and average value, and the third is the vertical and horizontal comparison.

In the application of scenarios for power companies, it includes various service scenarios for users of power companies, such as operation fault analysis and early warning in operation quality domain. The main content of this service scenario is to analyze and measure on-site operation faults, discover fault rules, and carry out horizontal early warning for typical batch faults. Scene application analyses the collection and measurement faults of measurement field operation, counts out the production enterprises and batches of multiple typical faults, and gives early warning to the power grid companies of the same production enterprises and different regions of production batches.

In the application of government-oriented scenarios, it includes various service scenarios for the government, such as the formulation of the power meter rotation policy in the management decision-making assistance. The content of this service scenario is to show the quality level of the industry by collecting and analyzing the failure situation, life span and operation situation of the national power
meter, to provide data support for the formulation of the government power meter rotation policy, and to assist in setting up the service scenario. The policy of electricity meter rotation which conforms to the use of electricity meter and saves cost is worked out.

The cloud platform mainly includes basic support, big data platform, cloud portal, cloud portal operation and management, etc. The basic support mainly provides basic resources for the development, operation, monitoring, operation and maintenance, security and other aspects of the entire cloud platform. Big data platform is used to provide data access, data cleaning, data storage, data integration, data calculation, data analysis, data exchange and other services for various applications. It is convenient for upper applications and services to use large data storage, computing and other resources quickly. Cloud portal provides application services and data services for various user groups, such as producers, users, regulators, developers and so on. Services for each group are nested in the cloud portal and jump to specific service function pages through menus or links. Operations management system provides services for the daily operation and management of cloud platform.

4. Conclusion
In the era of big data, cloud computing and big data technology are a new mode to solve business problems. For metrology big data cloud platform, with the deepening of related business in metrology field and the upgrading of various management requirements, the related application requirements will also become more and more complex. The problem characteristics in the business field must be accurately grasped. Through the improvement of service types and the expansion of service content, the service scenarios should be enriched. Through the comprehensive application of technology, the ability of data acquisition should be continuously improved and the data analysis methods should be optimized, so as to continuously improve and optimize the effect of various application scenarios and continuously meet the future development.

Acknowledgements
This study is supported by Science and Technology Project of SGCC. (Research on measurement technology framework supporting "Internet +" smart energy, No. 52110418001P)

References
[1] China Institute of Standardization of Electronic Technology, White Paper on Industrial Big Data (2017 edition) [R], 22-25(2017)
[2] Wang Jiye, Smart Grid Big Data [M], China Power Press, 36-37(2017)
[3] Cao Junwei et al. Southern Power Grid Technology, Overview of Big Data Analysis Technology of Energy Internet [J], 9 (11), 3-7(2015)
[4] Wang Jiye, Cheng Zhihua, PENG Lin, et al. Electric Power, Summary on Cloud Computing and Its Application Prospect in Power Sector[J], 47(7): 108–112(2014)
[5] Wu Kaifeng, Liu Wantao, Li Yanhu, et al. Electric Power, Cloud-Computing Based Power Big Data Analysis Technology and Its Application [J], 48(2): 111–116(2015)