

# Technical Research on the Electric Power Big Data Platform of Smart Grid

MA Ruiguang<sup>1</sup>, WANG Haiyan<sup>1</sup>, ZHANG Quanming<sup>1</sup>, LIANG Yuan<sup>2</sup>

<sup>1</sup>State Grid Sichuan Economic Research Institute SCSERI Chengdu, China

<sup>2</sup>College of Management Science, Chengdu University of Technology, Chengdu, China

**Abstract.** Through elaborating on the associated relationship among electric power big data, cloud computing and smart grid, this paper put forward general framework of electric power big data platform based on the smart grid. The general framework of the platform is divided into five layers, namely data source layer, data integration and storage layer, data processing and scheduling layer, data analysis layer and application layer. This paper makes in-depth exploration and studies the integrated management technology based on big data, the index technology based on electric power big data, the analysis technology based on electric power big data and the excavation technology based on electric power big data, to achieve the breakthrough of key technologies of electric power big data. Finally, the typical case of big data applications shows that the key technology application of electric power big data platform will bring changes for the power enterprises, which has demonstration meaning and important reference meaning.

## 1 Introduction

With the advent of the Age of Big Data, countries around the world attach great importance to the development of big data. In 2012, the United States proposed the "Big data Research and Development Initiative", with the investment of more than 200 million funds of scientific research for the acquisition, organization and management of massive data, and authorized the government's six federal agencies to join, in to promote the application and development of big data related technologies. The Chinese government proposed the plan of "The Implementation of the National Big Data Strategy" in 2013 and issued *The Outline of The Action of Promoting the Development of Big Data*, to comprehensively promote the development of big data and accelerate the construction of great power of data. Big data has become a national basic strategic resources, and the "second petroleum". The rise of big data has whipped up huge waves in many fields of the country, and also to brought changes to a lot of industries, such as the telecommunications industry, power industry, consuming industry, which are based on big data. Of course, big data has also provided opportunities of development for many enterprises [1].

The energy industries, including electricity, petroleum, gas, new energy and other energy fields, generate massive data due to related technology applications of collection, processing, analysis and application, which form the big data of energy. At present, the main sources of big data are electricity, petroleum, gas, new energy, population, geography, meteorology and other data. The energy big data makes the energy production, consumption and

relevant technical revolution deeply integrated into the big data philosophy, and promotes the thorough application of big data technology in the energy field, and also accelerates the development of energy industry and business model innovation.

In 2013, the issuance of White Paper for the Development of Chinese Electric Power Big Data aroused a hot trend of research on the electric power big data, which brought vitality to the traditional power industry, promoting the transformation of the power industry. Electricity mainly includes power generation, transmission, power transformation, power distribution, power consumption and scheduling, each producing a lot of data. The data are divided into several categories: the first one is the management data of the power enterprise; the second one is the power grid operation data; the third one is the electric power marketing data and the fourth one is the power grid monitoring data. Just like other big data, electric power big data is featured by 4V: (1) Volume: The automatic system of power enterprises collect, store and compute huge amount of data everyday; (2) Velocity: the data circulation and processing speed of power enterprises are very fast, which realizes real-time processing and analysis of a great deal of data in a fraction of a second; (3) Variety: The electric power big data have various types including structured data and unstructured data, and the sources are diversified, which may be relational data or text, pictures, audio and other data; (4) Very low value: There are not so much valuable data in the massive data, because the majority of data are normal. For the power monitoring data, the abnormalities of the data are the important data, because the abnormal data is the most important basis for condition based maintenance [2].

\* Corresponding author: luckfunny\_wang@163.com

With the in-depth advancing of big data and the intelligent transformation of the power grid, the electrical power system of each link has achieved digitization, informatization and intellectualization. The use of a variety of intelligent electric meters, sensors and information systems has injected a powerful driving force into the construction of electrical power system, which penetrate into the power generation, transmission, power transformation, power distribution, and power consumption and scheduling, and other links of production and management. The use of these devices brings more data sources for the electric power big data, and also provides continuous massive data for the power sector. Therefore, strengthening the relevant technical research on the electric power big data help promote the sustainable development of the power industry, and also promote the use of electric power big data.

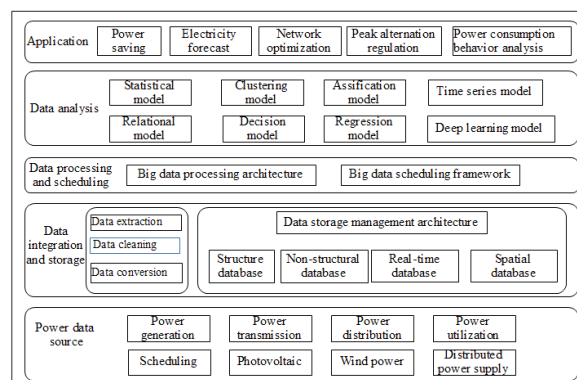
## 2 Relation among the Electric Power Big Data, Cloud Computing and Smart Grid

The smart grid integrates information, computer, Internet and communication, and forms a new type of smart grid combined with devices of power transmission and distribution, which has high energy efficiency and power supply security. A mass of data generated by the smart grid carry out information interaction through the modern network and communication technology, and make information exchange with the power grid device, to ensure the exchange of information. The power grid completes the automated management of smart grid through automatic collection, automatic control, automatic monitoring and other functions. Automated management is based on the analysis of big data. It is just because of the existence of big data analysis that massive data analysis of smart grid is applied. The relationship between the cloud computing and big data is based on the powerful computing power of cloud computing, which can deal with massive data quickly, and also provide demanded service conveniently; the big data storage, management and analysis technologies are built on the basis of cloud computing, rooted in the cloud computing, which are also the basis of big data application [3-4]. The development of big data can promote the development of cloud computing, and the cloud computing can promote the realization of cloud computing application, both of which supplement each other.

The smart grid, big data and cloud computing are complementary to each other. The smart grid is the usage scenario of big data platform, and the cloud computing provides handling ability to the massive data storage and calculation of electrical power system. The platform which integrates the electric power big data is helpful for improving the on-line analysis ability and real-time control ability of the smart grid, and improving the reliability of the power supply and reducing the power loss of the electric energy of transmission grid, and alleviating the environmental impact and reducing the cost of the basic equipment [5].

## 3 Ease of Use General Framework of the Electric Power Big Data Platform

The general framework of electric power big data has good versatility, which can regulate and define the electric power big data platform. The general framework defines the functional procedure of electric power big data, which plays an exemplary role [6] for the power enterprise in the platform technology selection. As shown in Figure 1, based on other architecture platforms, with the integration of relevant characteristics of power enterprises, the five-layer framework is built on the basis of electric power big data, which is composed of the data sources, integration and storage, data processing and scheduling, data analysis, application. This framework integrates the data of different fields in the power sector, to form unified data resource pool; the data integration and storage layer and the data processing and scheduling layer are open and expandable, which can provide interface support to the data from other sources and data processing; the data analysis layer can provide mature and advanced big data analysis services; the application layer can provide personalized, secure and reliable application services, which provides public service to the society [7].

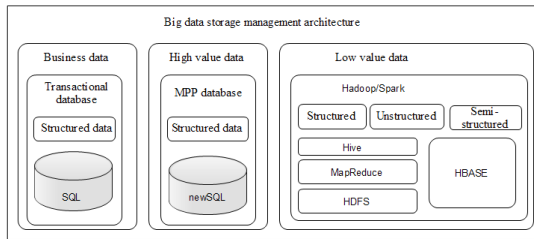


**Figure 1.** The General Framework Based on Electric Power Big Data

The power data source layer integrates the data sources of each links of power generation and management, such as power generation, transmission, power transformation, power distribution, and power consumption and scheduling, and also includes the photovoltaic and wind power, distributed power supply and other new energies, which provide ceaseless massive data for the power industry, laying the data basis for electric power big data [8].

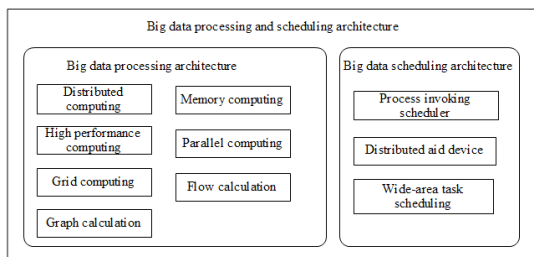
At the data integration and storage layer, with the rise of cloud computing, the traditional centralized storage is no longer the mainstream storage architecture of the data center, but replaced by distributed storage. Distributed storage architecture is featured by retractility and scalability, which provides solutions to big data storage. As shown in Figure 2, the general business data uses relational database storage, and high-value, high-density data uses MPP database, low-value low-density structured, unstructured, semi-structured data use distributed file system. The architecture of sorted storing is conducive to improving the storage efficiency of data [9]. Distributed

file system has become the core module of the big data platform, which is used for storing all types of electric power big data.



**Figure 2.** Big Data Storage and Management Architecture

The big data processing and scheduling layer, as shown in Figure 3, includes big data processing and big data scheduling and other related content. The big data processing includes distributed computing, high-performance computing, parallel computing, memory computing, stream computing, grid computing, etc. The big data scheduling, includes streamline scheduling, distributed assist, device, task scheduling in WAN, which achieves big data organization and scheduling, thus providing the necessary conditions for analysis.



**Figure 3.** Big Data Processing and Scheduling Framework

Big data analysis layer achieves big data related analysis methods and models, such as statistical analysis, clustering analysis, classification analysis, time series analysis, correlation analysis, regression analysis, decision analysis, neural network analysis, etc. The big data platform can be used to carry out data analysis, and provide mature and advanced big data analysis service.

The application layer provides application service based on the analysis of big data, including energy saving, power consumption predication, grid structure optimization, tagging power consumption, power consumption behavior analysis and other application services.

## 4 Key Technology of Electric Power Big Data

### A. Research on Integrated Management Technology Based on Electric Power Big Data

Power data synthesizes multiple links of data, involving different formats, features and properties of data that should be stored in the database organically and orderly. Therefore the big data platform should be able to handle multi-source, complex, and diverse data [10]. Therefore, the technical research on the integrated management based on the electric power big data aims to solve the problem of massive multi-source heterogeneous data

storage in the power sector. Through data extraction, cleaning and transition, it extracts the entities and relationships and then achieves high-efficiency storage to make the data for big data analysis stable and reliable. In addition, it supplies good expansibility and efficient data access ability.

### B. Research on the Index Technology Based on Electric Power Big Data

In order to effectively manage massive and diverse power data and provide "demand-obtaining" services, the electric power big data platform must be extensible, usable, and capable of running in heterogeneous environment and provide efficient query [11]. Therefore, it is also one of the keys to strengthen the index technology research based on electric power big data and provide good index technology solutions [12]. Because the query performance will seriously affect the service quality, and good index technology can improve the query processing performance effectively, to reduce the CPU using time and disk reading frequency. It is the core technology of the big data platform and a research topic of great importance to build effective index structure and increase the query processing performance.

### C. Research on the Analysis Technology Based on Electric Power Big Data

With the use of electric power big data analysis technology, potential mode and laws can be found in the massive data of electric power, which provides decision support to decision-makers after analysis and processing. The big data analysis technology includes statistical analysis, clustering analysis, classification analysis, time series analysis, correlation analysis, regression analysis, decision analysis, neural network analysis and other methods, which is mainly used in online analysis of power grid security, line state analysis, user power consumption behavior analysis, etc., which provides reliable data analysis [13-14] for the safe and stable operation of the electrical power system.

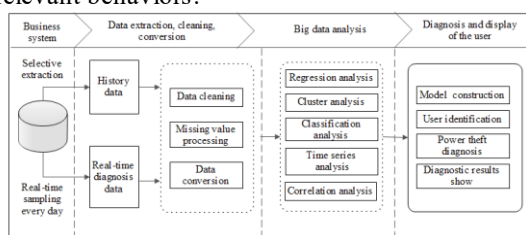
### D. Research on the Mining Technology Based on Electric Power Big Data

The mining of electric power big data refers to extracting knowledge from complex data set to assist with decision-making mainly through using the association rules, classification, clustering and other mining methods. Electric power data mining includes data preparation, law searching and law representation. Data preparation refers to the data set used for excavation; the law searching refers to finding the law of data using different methods; the law representation refers to showing the laws found in a certain way or visual means.

### E. Stealing Electricity and Electric Leakage Automatic Diagnosis Platform Based on Electric Power Big Data

With the popularization of SGCC smart grid and the improvement of data collection means, various relevant information about the user can be collected, which lays a data basis for the big data analysis. There are many applications based on electric power big data, so the stealing electricity and electric leakage automatic diagnosis platform is taken for example here to illustrate the application prospect of big data in the smart grid. The intelligent acquisition system collects power load data and power consumption abnormal data of current, voltage,

power of each phase and other factors from the electric energy meter, voltage transformer, secondary metering circuit, electric power measurement bin and other relevant device, and stores it in the database system of big data platform. As shown in Figure 4, the stealing electricity and electric leakage automatic diagnosis platform based on electric power big data extracts relevant data from the database selectively through the data extraction module and read relevant data in real time, for cleaning and conversion treatment, and then provide effective, high-quality data to big data analysis module for analysis and processing, and the big data analysis module will use big data analysis method for analysis and processing, before sending the results to the user diagnosis and display module. Through building user identification model based on big data, automatic detection and automatic diagnosis, it identifies stealing electricity and electric leakage and relevant behaviors.



**Figure 4.** Stealing Electricity and Electric Leakage Automatic Diagnosis Platform Based on Electric Power Big Data

Through elaborating on the associated relationship among electric power big data, cloud computing and smart grid, this paper put forward general framework of electric power big data platform based on the smart grid. The general framework of the platform is divided into five layers, namely data source layer, data integration and storage layer, data processing and scheduling layer, data analysis layer and application layer. The framework points out the direction for the power enterprise informatization, which is of reference value to the promotion of the electric power big data platform. This paper also focuses on the integration of big data in the framework and storage architecture, and the big data processing and scheduling architecture, and then solves the problem of structured data and unstructured data storage in the way of distributed storage; and solves the calculation and processing problem through big data organization and scheduling. Among various key technologies, this paper focuses on the integrated management technology based on big data, the index technology based on big data, the analysis technology based on electric power big data and the excavation technology based on electric power big data, to achieve the breakthrough of key technologies of electric power big data. Finally, in terms of the typical case, it uses stealing electricity and electric leakage automatic diagnosis platform based on electric power big data as the study case, which has demonstration effect on power enterprises in the development of big data related researches.

## References

1. Yao Wei, Cheng Shijie, et al. *Application of HVDC technology in grid - connected offshore wind farm* CEP, **40** (2007)
2. Yan Long-chuan, Li Ya-xi, Li Bin-chen, Zhao Zi-yan, Opportunity and Challenge of Big Data for the Power Industry, J. Electric Power Information and Communication Technology. 11(2013) 1-4.
3. Deng Chuan-hua, The Application Prospects and Challenges Research of Power Big Data, J. CA. Journal of the Hebei Academy of Sciences. 3(2015) 66-69.
4. Yan Long-chuan, Chen Liang, Zhao Zi-yan, Cloud computing data center architecture supporting power big data research, J. Distribution & Utilization. 8(2014) 37- 40.
5. Wu Kai-feng, Liu Wan-tao, Li Yan-hu, Su Yi-peng, Xiao Zheng, Pei Xu-bin, Hu Song lin, Study on the Impacts of Short-Term Demand Response on Generation Expansion Planning with Wind Power, J. CSCD. Electric Power. 2(2015) 111-116 + 127.
6. Zhang Ming, Wang wei, Shi Jian-hua, Zhao De-wei, Optimization of Dispatching Cloud Based on Power Big Data, J. CSCD. Computer Simulation. 11(2014) 123-126 + 137.
7. Duan jun-hong, Zhang Nai-dan Zhao Bo, Yan Xiao-bin, Research on Basic System Architecture and Application of Electric Power Big Data, J. Electric Power Information and Communication Technology. 2(2015) 92-95.
8. Zhu Chao-yang, Wang ji-ye, Deng Chun-yu, Platform Research and Design of Electrical Big Data ,J. Electric Power Information and Communication Technology. 6(2015) 1-7.
9. Peng Xiao-sheng, Deng De-yuan, Cheng Shi-jie, Wen Jin-yu, Li Chao-yang, Niu Lin: Key Technologies of Electric Power Big Data and Its Application Prospects in Smart Grid , J. CSCD. Proceedings of the CSEE, 3(2015) 503-511.
10. Dan Tao, The key technology analysis of the Power Big Data for Smart Grid Applications [J]. ublic Communication of Science & Technology. 20(2015) 31 + 29.
11. Qiao Mao-bin, The Key Technology of the Power Big Data for Smart Grid Applications [J]. Resources Economization & Environmental Protection. 8(2015) 2-4.
12. Luan Kai-ning, Zheng Hai-yan, Ding Chen, Li Kun-ming, Dynamic Indexing Technology of Fast Combination Query for Power Large Data ,J. Electric Engineering, 1(2015) 113-116 + 122.
13. Yang De-sheng, Chen Jiang-jiang, Zhang Ming, Power Big Data High speed Data Storage And Retrieval And Application Of Key Technologies ,J. Electronic Test. 3(2014) 62-63 + 61.
14. Shen Hao-bin, The Application Status of Power Big Data and Multi-Source Heterogeneous Data Analysis Technology J. Heilongjiang Science and Technology Information. 24(2015) 181-183.