Research About Solutions to the Bottleneck of Big Data Processing in Power System

Ning Chen, Chuanyong Wang, Peng Han, Jian Zhang, Kun Wang, Ergang Dai, Wenwen Kang, Fengwen Yang, Baofeng Sun, and Guang Guo(✉)

State Grid Zaozhuang Power Supply Company, Zaozhuang 277100, China
43383047@qq.com

Abstract. The big data technology provides a new opportunity to the electric power system in various fields. In view of the shortage in traditional power system computing platform in computing, storage, information integration and analysis, this article puts forward a platform of power system based on cloud computing. Firstly, this paper outlines the development of power system can produce a large amount of data, and cloud computing technology is widely used in big data processing as a kind of new model, we can take chance of its application in power system. This paper discusses the relationship between cloud computing, big data and power system, and makes the conclusion that cloud computing technology can meet the demand of mass data storage and computing power system. Then we do research about the architecture, software technology of cloud platform system in electric power system. Finally, the feasibility of the application and development trend were discussed.

Keywords: Power system · Big data · Cloud computing

1 Introduction

With the development of long distance transmission system and the strength of interconnected power system, large scale power systems have covered a number of countries continuously. The continuous expansion of power system and the complex structure is becoming increasingly difficult in safety evaluation, economic operation and system control. Electric power system will produce a large amount of data, the power system dispatching center should have powerful computation ability and information collection, integration and analysis capabilities in the future. The existing centralized power system computing platform is hard to meet the requirements, and it has become one of the main bottleneck of the realization of the smart grid.

Cloud computing is a new computing mode, and has gained rapid development in recent years, it includes several new computing technologies. So far, for cloud computing, there is no standard definition from authority. China's cloud Computing network cloud is defined as: cloud Computing is Distributed Computing, Parallel Computing and the development of Grid Computing, Grid Computing, or the commercial implementation of the concept of science. It is generally believed that cloud computing represents a large scale of distributed computing model based on the Internet.

J. Wan et al. (Eds.): Industrial IoT 2016, LNICST 173, pp. 44–51, 2016.
DOI: 10.1007/978-3-319-44350-8_5

Cloud computing platform firstly using the Internet to a variety of wide-area and heterogeneous computing resources integration, in order to form an abstract and dynamically extension of virtual computing resources pool; And then through the Internet to users on demand provides computing power, storage capacity, software platform and applications, etc.

Through the establishment of power system, big data cloud computing platform can effectively integrate existing computing resources system, providing powerful computing and storage capacity to support a variety of analytical computing tasks. Cloud computing can support a variety of heterogeneous computing resources, compared with a centralized supercomputer, which can be highly scalable and can be easily upgraded in the existing computing power is insufficient. In addition, compared with the traditional computing model, cloud computing also has the ease of information integration and analysis, to facilitate the development of software systems, maintenance. In short, the establishment of the core system is calculated based on the power of cloud computing platform that can effectively address some of the key challenges in computing power system and information processing encountered in the future. This article will be defined for the cloud, features, technology, architecture, in power system research and application prospects and other issues were discussed in detail.

2 Big Data of Power System

Data generated by computing power system belongs to structured data, the main features are

(1) Multi-type isomers
Different calculation software, different types of elements, different models of the same elements, and the results of different types of data structures are quite different;
(2) The homogeneity of calculating data online and offline
Power System Simulation online data is usually calculated by putting the measurement information together with offline data, similar to the data structure, which has research data analysis on two kinds of data may be common;
(3) Diverse storage
Online data is typically stored centrally, updated regularly, and most of the offline data are scattered in the staff's personal computer;
(4) Huge volume
With the extensive use of smart grid scheduling technology support system (D5000), online data to calculate the rapid accumulation of volume will be achieved PB level [1]. While off-line calculation of the raw data is smaller, but huge volume of data analysis and calculation results are produced. For example, the entire network of 10,000 nodes grid bus get on short circuit current scanning, and calculate the whole network voltage of each bus fault under way will produce nine months results bus voltage 10, the data file reaches GB level [2].

As big data generated by power system has the main features above, we need a more efficient big data processing.

3 Power System Needs Big Data Cloud Computing Platform

Figure 1 briefly describes the relationship between power system, cloud computing, big data: Electric power system produces a large amount of data and needs to use cloud computing to save, processing and management the data; Cloud computing technology provides tools and methods to the development of big data technology, meanwhile, the latter provides application scenarios to the former, and both of them can share services for electric power system.

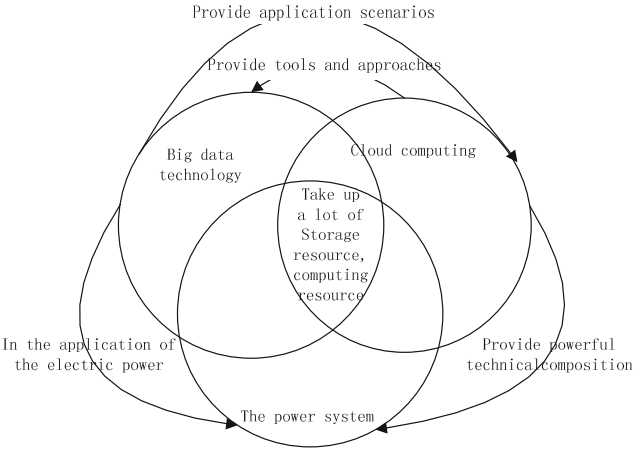

Fig. 1. Relationships among power system, big data and cloud computing technologies

Cloud computing can integrate intelligent power computing and storage resources, and improve power system and interactive processing capacity, becoming a strong technical composition of grid; Big Data technology and services based on business needs, rooted in the cloud computing; Cloud computing platform in power system can be considered to be big data applications in abstract, so relationship is between the three and they interact with each other [3].

The relationship between power systems, big data and cloud computing is a power system products of different stages of development from the more deep-seated, and it has heritage features.

Intelligent Power System is the result of information technology, computer technology, artificial intelligence technology in the traditional power system applications precipitation, meeting the power system information, intelligence, clean high-level operational and management needs, not only on the traditional grid inheritance, but also to carry forward the traditional power grid, so its development must be synchronized with the new technology. Forefront of cloud computing and big data technologies from most computer and information technology is their stage of development and technical application level that has a landmark new technology [4].

Cloud computing technology in the distributed storage technology and parallel computing technology, satisfying the mass data storage and computing power system

requirements, so power cloud is proposed soon after the cloud computing technology, the concept of cloud computing technology application in power system is also shows a tendency of flowers gradually, of which promoting the development of the power system. Big data technology is not only the continuation of traditional data analysis and mining technology, but also an inevitable product of the knowledge mining and business application demand, so most of the applications of big data technology based on distributed storage and processing technology with the key technology of cloud computing. In a sense, the development of the power of big data technology can be regarded as cloud computing technology in power system and the implementation process of high-level business requirements.

4 Cloud Platform for Processing Big Data in Power System

The power systems cloud computing platform is achieved by architecture and technologies, and is a complex entity composed by a variety of devices and users that connected to each other through Internet component (see Fig. 2). Generally speaking, cloud computing platform can be divided into two main parts, namely the control center and a variety of computing resources integrated by cloud computing platform. The main function of the control center in cloud computing is based on the user's request, dividing user's computing tasks into several sub-tasks, and then dynamically assigning sub-tasks by the Internet to the computing devices that integrated by cloud computing platform [5]. After each sub-task is completed, its results will be re-aggregated to the control center via the Internet, and finally back to the user.

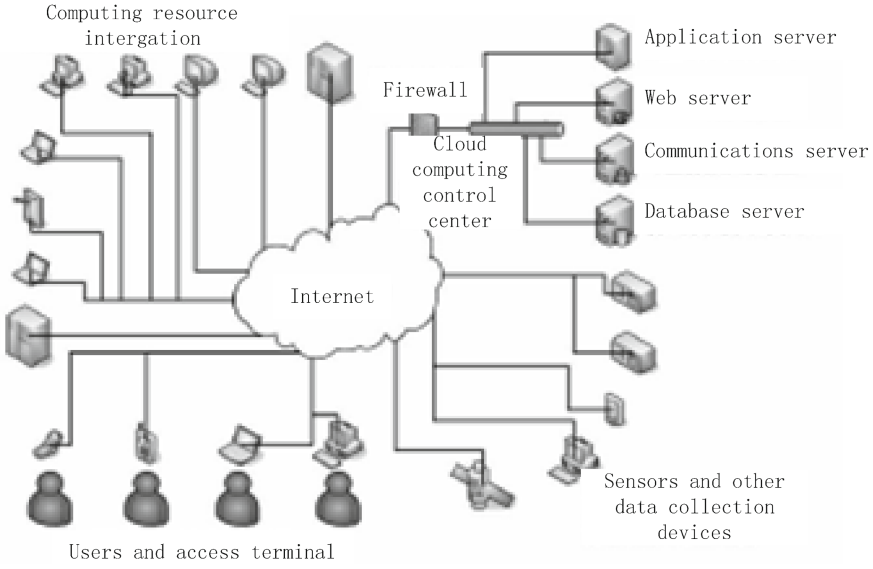

Fig. 2. Structure of cloud computing

Cloud computing platforms and data acquisition are connected via Internet network and a large number of sensors and other data collection devices. For power systems, future data collection network not only includes traditional SCADA system sensor, but also PMU and placement of smart meters in the end-user at home, or even a variety of smart appliances embedded systems. These devices can provide a full range of system information. In addition, the power system can also connect cloud computing platform and other data sources, such as: regional meteorological database connected to each other, to get the temperature, humidity, wind speed, sunshine and other data. The amount of data generated by such a large-scale network collected would be amazing, it can be stored and analyzed only by virtue of cloud computing platform's computing power. Taking into account a lot of power system analysis tasks on time-critical, you should consider creating a dedicated high-performance network to connect to cloud computing platforms and data collection network, to improve the reliability of data transmission [6]. Power Systems cloud computing system architecture platform as shown in Fig. 3.

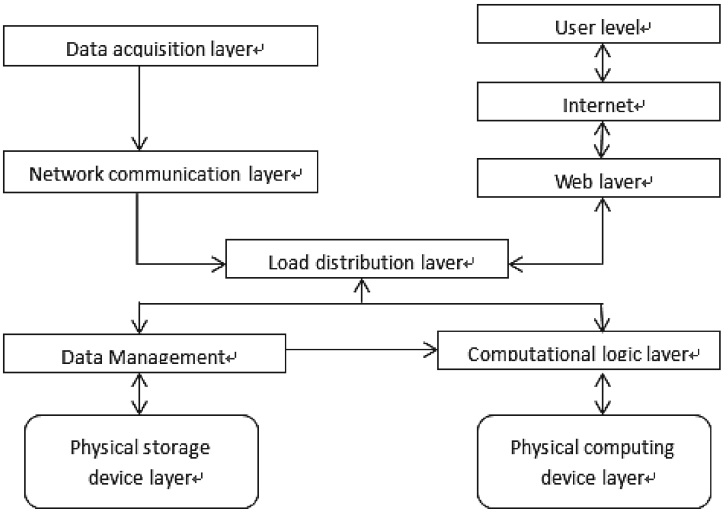

Fig. 3. System architecture of the cloud computing platform for power systems

From the system architecture point of view, cloud computing platform consists of Web layer, load distribution layer, data management, calculation logic layer, physical computing device layer and the physical storage device layer [7]. Wherein, Web layer is responsible for implementing cloud computing platform, Web site that is the only interface to the user to access the cloud computing platform.

Load distribution layer is the core component of cloud computing platforms. This layer has four main functions: computing tasks of user will be divided into several parts, and decided to enable what device for each task; the data to be stored is divided into several parts, and determined the appropriate storage device; computing and integrating results the logic layer returned, and then feedback it to the user; according to the read data request, the data management layer reads data and then outputs it after integration.

The four layers above consist of the software part of the cloud computing platform. Physical computing device layer and the physical storage device layer represents all the physical devices that make up the hardware part of the cloud computing platform [8].

The following discusses several important software technology can be used to achieve cloud computing power system.

(1) An important feature of the service-oriented architecture of cloud computing is online expansion and upgrades, which puts forward higher requirements of flexibility for the software [9]. Service-oriented architecture is a rapidly growing software design method in recent years. In the traditional design methods and software to function or class as a basic function module to the application programming interface (API) as a means of communication between different programs, SOA is to serve as the basic functional modules. Compared with the function, service representative higher-level application requirements (for example: to print the data from the database and to read the report of the entire process can be abstracted into a service, but to read the database is only a function). According to user needs, each one of the main functions is packaged in the form of services by SOA, and each service is independent, only by Extensible Markup Language to realize communication [10]. When any kind of function needs to be updated, you just need to replace the corresponding services. In addition, SOA-based architecture can freely combine a number of services to the rapid formation of a new system. For example: it can package flow calculation as a service, so there is no need to do the task flow calculation, what should be done is to tide computing services and other related services to by online combination. Visibly, SOA applications can greatly improve the flexibility and software systems to upgrade the speed of development. Of course, raise the level of abstraction of software generally at the expense of the efficiency of communication. Summarize the history of software development methods, it develops from process-oriented to object-oriented, and then to service-oriented. Software development has trend to improve the level of abstraction overall, and this is compatible with the software itself increasing complexity. With the continuous development of power system, the function of power information system growing, and structure of itself is increasingly complexity, which will bring difficulties to developers who use traditional development methods. Therefore, raising the level of abstraction of software is an inevitable trend. For communication efficiency, you can define a service level of abstraction appropriate to strike a balance.

(2) Dynamic load distribution

Load distribution algorithm is the core of cloud computing. Dynamic load distribution has proven to be a class of more effective allocation algorithm for computing tasks. The basic principle is to allocate computing tasks based on the calculation speed of each device dynamically; computing device of greater speed is assigned the bigger task, of which to ensure all computing devices simultaneously return results. For large-scale cloud computing platform, you can also consider using the job replication methods to improve reliability, it is about to copy a number of copies of each sub-task, and distribute them to multiple computing devices to synchronously processing, thus avoid slow the overall efficiency down by the occurrence of a device failure that results in redistributing sub-tasks.

In addition to the discussion above, there are still a variety of commercial or open source software technologies can be used to cloud computing, such as GoogleMapReduce, Google File System (GFS), Microsoft's Dryad/DryadLINQ, and other open-source distributed computing framework [12]. Considering the power system is an important national infrastructure, we think open source software technology should be the foundation to realize cloud computing platform in power system.

5 Prospects

With the development of technology, electricity big data platform will be similar to EMS, ERP and other systems, and becoming an "infrastructure" of planning, design, operation and management for the power system in the near future. Perspective in all areas of power system will change from current information and results to trends against data traceability to the source of the problem. At the same time, it will comply with the variation of the data, explore new trends, and find hidden in one of the next scene.

Cloud platform for big data processing power system, as a guarantee for validity of the data processing, all the research we do need to have a solid theoretical foundation. Thus, the research should from the discovery of phenomenon to the analysis of phenomenon to discovery of reasons behind it, and rise to the level of theory as far as possible, which brings a new challenge for applying cloud computing technology to the applications.

In summary, the combination of power Big Data technologies and Cloud Computing will be able to provide a new means of system analysis, perspective, and even methods.

References

1. Kwon, O., Lee, N., Shin, B.: Data quality management, data usage experience and acquisition intention of big data analytics. Int. J. Inf. Manag. **18**(1), 156–157 (2014)
2. Yaqi, S., Guoliang, Z., Yongli, Z.: Present status and challenges of big data processing in smart grid. Power Syst. Technol. **37**(4), 927–935 (2013)
3. Xiaofeng, M., Xiang, C.: Big data management: concepts, techniques and challenges. J. Comput. Res. Dev. **50**(1), 146–169 (2013)
4. Divyakant, A., Philip, B., Elisa, B., et al.: Challenges and opportunities with big data. Proc. VLDB Endowment **5**(12), 2032–2033 (2012)
5. Meng, L., Xiaodong, C., Wen, Z., et al.: Design of cloud computing architecture for distributed load control. Power Syst. Technol. **36**(8), 140–144 (2012)
6. Xiaoyang, T., Shengyong, Y.: A survey on application of data mining in transient stability assessment of power system. Power Syst. Technol. **33**(20), 88–93 (2009)
7. Chen, C.L.P., Zhang, C.-Y.: Data-intensive applications, challenges, techniques and technologies: a survey on big data. Inf. Sci. **14**(9), 118–119 (2014)
8. Ni Zhang, Yu., Yan, S.X., Wencong, S.: A distributed data storage and processing framework for next-generation residential distribution systems. Electr. Power Syst. Res. **19**(26), 77–80 (2014)

9. Prasad, A., Green, P., Heales, J.: On governance structures for the cloud computing services and assessing their effectiveness. Int. J. Account. Inf. Syst. **15**(4), 335–356 (2014)
10. Zhao, X., Zhou, C., Zhao, L.: Condition evaluation model of fluid power system in gradual failure based on data envelopment analysis. Comput. Fluids **27**(31), 199–201 (2014)
11. Pereira, L.E.S., da Costa, V.M.: Interval analysis applied to the maximum loading point of electric power systems considering load data uncertainties. Int. J. Electr. Power Energ. Syst. **9**(14), 97–99 (2014)
12. Shin, S.-J., Woo, J., Rachuri, S.: Predictive analytics model for power consumption in manufacturing. Procedia CIRP **29**(30), 191–194 (2014)