Study of the key technologies of electric power big data and its application prospects in smart grid

Jie Zhan, Jinxin Huang, Lin Niu
State Grid of China Technology Colleague
State Grid Corporation of China
Jinan, China

Xiaosheng Peng, Diyuan Deng, Shijie Cheng
School of Electrical and Electronic Engineering
Huazhong University of Science and Technology
Wuhan, China
Email: XiaoshengPeng@hust.edu.cn

***Abstract* — Application of big data techniques in power system will contribute to the sustainable development of power industry companies and the establishment of strong smart grid. This article introduces a universal framework of electric power big data platform, based on the analysis of the relationships among the big data, cloud computing and smart grid. Then key techniques of electric power big data is discussed in four aspects, including big data management techniques, big data analysing techniques, big data processing techniques and big data visualization techniques. Finally, the article presents three typical application examples of electric power big data techniques which are new and renewable energy integration, wind turbine condition monitoring and assessment and data base integrative backup for electric power enterprises.**

***Index Terms*—Big data, cloud computing, smart grid, data integration, electric power big data, visualization.**

I. INTRODUCTION

The earliest reference of Big Data can be traced back to Nutch which is an open source project launched by Apache Software Foundation. In Nutch project, Big Data is used to describe the large volume of data sets in web searching application [1]. According to reference [2] big data has the "3 Vs" characteristics, which are Volume, Variety and Velocity. Because of the "3 Vs" characteristics of the big data, the traditional data processing techniques are not always applicable for big data application. The characteristics of electric power big data can be summarized as "3 Vs" which are Volume, Variety and Velocity and "3 Es" which are Energy, Exchange and Empathy [3]. Electric power big data can be found in all sections of electric power generation and management. It is the key technology of electric power to overcome the challenges of limited resources and harsh environment, renewable energy integration, load fluctuation and sustainable development.

A. Smart grid and big data

Smart grid is a modernized electrical grid which integrates digital information technology, communication technology, computer technology and power facilities [4]. Smart grid has several advantages which can improve energy efficiency, improve the safety and reliability of power supply [4]. When large numbers of renewable energy integrated into power systems, more variables are brought into the system and more data need to be processed. The growing number of electric vehicles also makes the situation to be more and more severe. Big data technologies are with great advantages of prediction and analyzing of these random factors. Essentially, Smart Grid is the electrical grid with the application of the big data technologies [5].

B. Big data and cloud computing technologies

From a technical point of view, big data technologies are rooted in the cloud computing technologies. Cloud computing technologies are the foundations of some of the big data technologies. Cloud computing technologies provide powerful tools for big data. Big data provide practical application for cloud computing technologies. Overall, big data and cloud computing are complementary [7].

C. The relationships of smart grid, big data technologies, cloud computing technologies

The relationship of smart grid, big data and cloud computing technologies are shown in Figure 1.

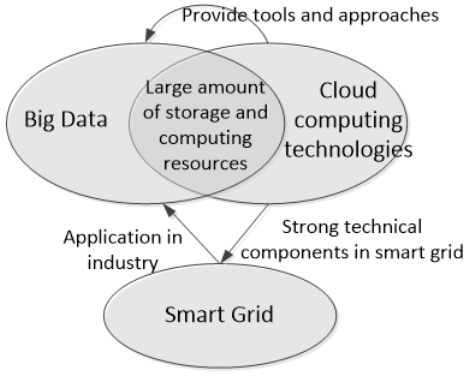

Figure 1. Relationships of smart grid, big data and cloud computing technologies

978-1-4799-7537-2/14/$31.00 ©2014 IEEE

Cloud computing technologies can integrate computational processing resources and storage resources within the smart grid, and improve data processing and information interactive capabilities, which are powerful technical components in smart grid [8]. Big Data technologies are rooted in cloud computing technologies. Smart Grid is achievable when the big data technologies applied in electric power industry.

II. THE UNIVERSAL FRAMEWORK OF ELECTRIC POWER BIG DATA PLATFORM

Figure 2 shows a universal framework of big data platform which is an open source technology proposed by the Apache Foundation [8]. The key points of the platform are HDFS distributed files system based data storage framework, and Map-Reduce computing technology based data processing framework. Both distributed files system and Map-Reduce computing technology could deal with PB, ZB size of data. The frame also includes another modules such as business intelligence applications, traditional data warehouses, big data access frameworks, big data scheduling framework, the network layer, operating systems, servers, backup and recovery and data management.

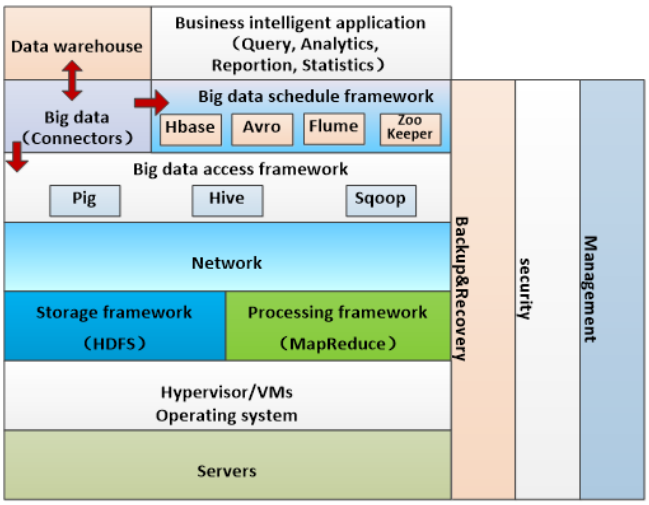

Figure 2. Universal framework of big data platform [8]

Big data storage framework and big data processing framework are based on ordinary computer servers, operating systems or virtual machine which means the system is low-cost and with high scalability.

The big data access framework connected with the big data storage and processing framework by the network layer. The big data access framework consists of several sub-modules, which are parallel computer programming language Pig, data warehousing tool Hive, open source data transfer tool Sqoop. Operator could access to both the distributed file storage system and traditional data warehouse by big data connectors and open source data transfer tool called Sqoop, which is in the big data access framework.

The data process flowchart of electric power big data system is not totally different from the data process flowchart of traditional data system. The most significant difference is in big data system there are both large volume of data and unstructured data, in which situation the parallel processing techniques, such as MapReduce, are essential. The key technologies of electric power big data system can be divided into data integration and management techniques, data analysis techniques, data processing techniques and data visualization techniques, which are presented in the next chapter.

III. THE KEY TECHNOLOGIES OF ELECTRIC POWER BIG DATA SYSTEM

A. *Data integration and management techniques of electric power big data system*

The integration and management technologies of electric power big data, which could create new application functionalities for enterprises, combine data from two or more information systems. In another words, the technologies integrate the data which are from different sources, of different formats, with different features, in logical storage media.

The integration and management technologies of electric power big data include data fusion and data integration, database management technology, Extract-Transform-Load (ETL) technology, which is shown in Figure 3.

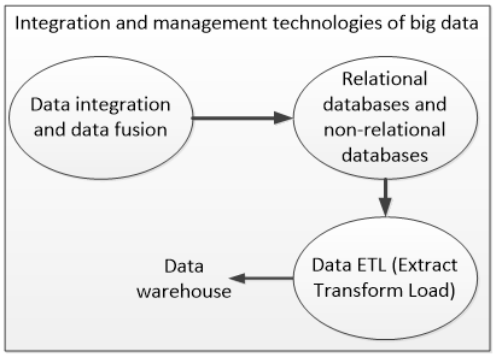

Figure 3. Integration and management technologies of electric power big data

B. *Data analysis techniques of electric power big data*

Data analysis techniques of electric power big data include correlation analysis techniques, machine learning techniques and data mining techniques. Valuable information can be extracted from the mass data by data analysis techniques of electric power big data.

Correlation analysis techniques could identify correlation or causal relationship, which are widely applied in investment decisions of power system and the electric power load forecasting.

Machine learning techniques belong to artificial intelligence technology, which are widely applied in different

areas in power system, such as power equipment condition monitoring, power system transient stability assessment [9].

Data mining is currently one of the most popular data analysis techniques. The main application areas of data mining in power system are for predicting or presentation.

C. *Data processing techniques of electric power big data*

Data processing techniques of electric power big data includes distributed computing technology, memory computing technology, stream processing technology, which is shown in Figure 4.

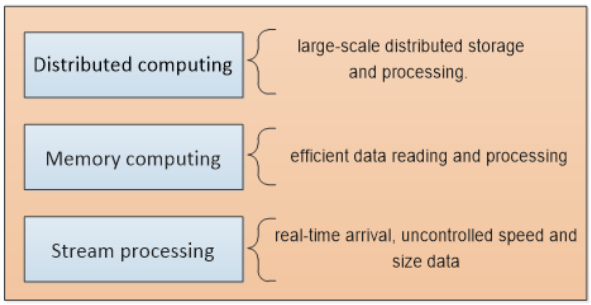

Figure 4. Data processing techniques of electric power big data

Memory computing technology is an innovative tool which handles vast amounts of historical and real-time information. Distributed computing technology is a breakthrough and a major revolution in database technology. Stream processing is mainly used to solve real-time processing problems.

D. *Data visualization techniques of electric power big data*

Data visualization of electric power big data includes 3 dimensional visualization, geographical information visualization, and historical information visualization, which could express electric power information intuitively and accurately. These three technologies are all widely applied in the power system. Visualization techniques, which are applied for power system status real-time monitoring, improve the automation level of power system. Visualization can also be combined with theories of complex system network to identify potential patterns and relationships in power system.

IV. APPLICATIONS EXAMPLES OF ELECTRIC POWER BIG DATA TECHNOLOGIES IN SMART GRID

Smart Grid is the electrical grid with the application of the big data technologies. With the development of smart grid, electric power big data technologies will play more and more important role in the smart grid. Three application examples of electric power big data are demonstrated as following.

A. *Case one: Big Data technologies application in new and renewable energy integration*

In electric power generation process, the integration of new and renewable energy changes the traditional power generation mode. The measurement and management of electric power generation data has become increasingly complex. Big Data technologies can make better predictions for electric power generation and electric grid companies.

Vestas Wind Technology Group of Denmark optimizes the wind turbine geographical layout and improves the efficiency of electric power generation by the analysis of the big data including weather reports, tidal information, geographical information and satellite images. IBM proposed a micro-siting solution for wind farm location selection based on high-precision weather forecast, which maximizes the efficiency of wind resources and minimizes the cost.

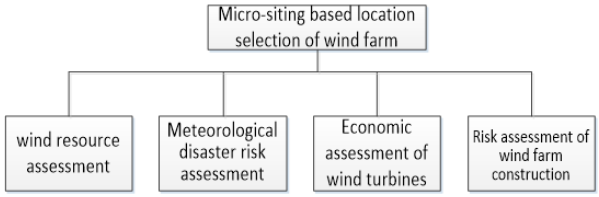

Figure 5. Micro-siting solution of wind farm [10]

The principle of micro-siting solution is shown in Figure 5. Based on atmospheric dynamical processes analysis, physical processes analysis and high-precision numerical weather model, the accuracy and reliability of the model are greatly improved. The micro-siting solution can comprehensively predict the distribution of the resources throughout the wind farm in longer-term, which can avoid the biases created when the predicting time is too short.

B. *Case two: Big Data technologies application in condition monitoring of wind turbines*

The operating environment of wind turbines is very harsh, which are affected by thunderstorms, salt fog, snow and other weather factors. Big data technologies based comprehensive condition assessment system of wind turbines can indicate potential catastrophic faults caused by parts wear or fatigue. Wind turbine equipment failure or parts wear can be identified by the diagnostic system. Simultaneously, operational safety of wind turbine is improved. Catastrophic faults are avoided so the operation and maintenance costs are reduced.

The flowchart of the condition monitoring and assessment of wind turbine is shown in Figure 6. The major part of the assessment system is an online ultra-sphere model, which is based on ultra-sphere model of monitored device and on-line monitoring data. The ultra-sphere model of monitored device is based on device properties and corresponding historical data of monitoring data. The output of the online ultra-sphere model is the input of the similarity curve monitoring, which will indicate the status of the device in terms of state warning, predictive value of measuring point, etc.

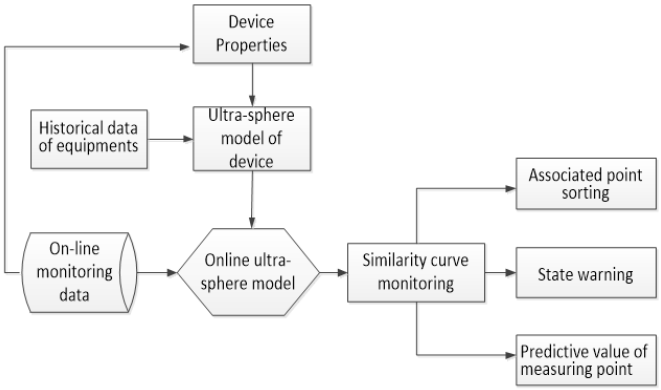

Figure 6. Flowchart of the condition monitoring and assessment of wind turbine [11]

C. *Case three :Big data technologyies application in backup and recovery of electric power generation enterprises*

With the rapid development of electric power enterprises, large amount of data management centers are established, in which the size of data grows exponentially. These data are the strategic resources for electric power enterprises, once damaged, catastrophic losses will be brought to the power utilities. Electric power companies need an integrated data protection solution which can achieve enterprise-class fast backup and recovery, provide comprehensive backup and recovery strategy and ensure security and business continuity of enterprise data.

IBM Tivoli Storage Manager (TSM) is a big data backup solution designed for large institutional data. TSM system has become the first choice in the field of data backup for different application fields, including financial services, telecommunications, etc.

System diagram of the TMS solution is shown in Figure 7.

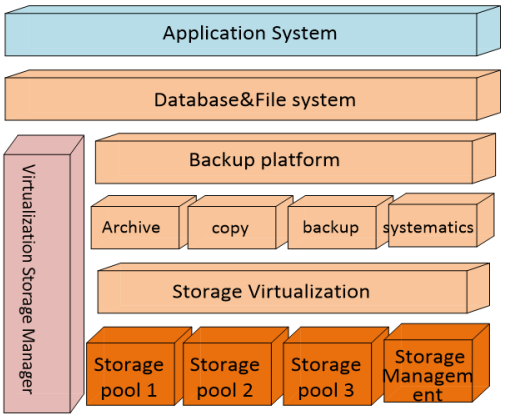

Figure 7. System diagram of TSM backup and recovery platform [12]

TSM backup platform is based on storage pools, which are managed by storage management system and virtualization storage manager. The data base and file system are built on backup platform, which has the functionalities of archiving，copying， backup, etc. The TSM backup platform is applicable for the data back and recovery for electric power generation companies. The system is fully protected by simple management tools, in which effective data recovery can be carried out when faults occur and the reliability and security of data are fully protected.

V. CONCLUSIONS AND FUTURE WORK

Big data technologies are of significant importance for the development of smart grid. The paper introduced a universal framework of electric power big data platform, which can be adopted for different applicant areas in smart grid. The key technologies of electric big data are presented in the paper in details and suggested for further research and industrial applications. Three typical application examples of electric power big data techniques are presented in the paper, including new and renewable energy integration, wind turbine condition assessment and data base integrative backup for electric power enterprises, which prove that big data technologies are with brilliant future in power system and will make significant contribution to the development of smart grid.

The most significant challenge of the application of electric power big data in smart grid is not the technical challenge but the application cases, which means how to use electric power big data is the biggest challenge. Future work should be focused on how to use electric power big data. Once the special applicant cases are identified, big value behind big data can be extracted by the big data technologies presented in the paper.

REFERENCES

[1] C. Chen, S. S. Zhang, and S. W. Shang, F. Sun, "Research for electric power big data quality evaluation model and dynamic exploration technology," Modern Electronics Technique, Vol.24, pp. 8-14, 2013.

[2] G. B. Li, H. Luo, "Intelligent data analysis under the background of big data," Science & Technology Information, Vol.30, pp.11-12, 2013.

[3] Electric power informatization special committee of CSEE, *China electric power big data development white paper*, Beijing: Proceedings of the CSEE, 2013, pp.10-15.

[4] W. L. Zhang, Z. Z. Liu, M. J. Wang, and X. S. Yang, "Research status and development trend of smart grid," Power System Technology, Vol. 13, pp.1-11, 2009.

[5] B. L. Sun, "Big data technology and its application in power industry," Electric Age, Vol 8, pp.18-23, 2013.

[6] J. Z. Luo, J. H. Jin, and A. B. Song, F. Dong, "Cloud computing: architecture and key technologies," Journal on Communications, Vol7, pp.3-21, 2011.

[7] H. Y. Yao, "Big data and cloud computing," Information Technology & Standardization, Vol5, pp.21-22, 2013.

[8] G. Zhao, *Big Data technology and application practice*, Vol3. Beijing ： Publishing House of Electronics Industry, 2013, pp.56-58.

[9] S. Y. Ye, "Study on power system transient stability assessment based on machine learning method," Ph.D. Thesis, Dept. EE. Southwest Jiaotong Univ. China, 2010.

[10] IBM., "power wind site down time," *IBM's solution in power industry*, IBM,2012.[Online]. Available: http://www31.ibm.com /solutions/cn/industries/energy/thankyou/energy_wp.shtml. [Accessed: Aug. 17, 2014]

[11] T. Y. Lu, Y. D. Li, W. Jiang, W. B. Zhou, H. Yang, Y. T. Jia, etc, "Comprehensive assessment system of safe and economic operation state for wind turbines," in *IT&OT convergence summit seminar of china power industry*, 2013, p. 10.

[12] IBM., "inter backup regeneration," *IBM's solution in power industry*, IBM,2012.[Online]. Available: http://www31.ibm.com /solutions/cn/industries/energy/thankyou/energy_wp.shtml. [Accessed: Aug. 17, 2014]