Research on Power System Dispatching Automation Technology Based on Big Data

Peishuang Zhang¹ and Rui Geng²,*
¹Tianjin University of Technology, Tianjin 300384, China
²State Grid Corporation of China, Beijing 100031, China

*Corresponding author: rui-geng@sgcc.com.cn

Abstract. The perfect combination of power dispatch automation system and the analysis and processing of large amounts of data in big data mining technology can meet people's needs for knowledge; at the same time, the application of mining technology in the power dispatch automation control system realizes the accurate collection and sorting of power data, Has a broad application space and practical application significance. This article introduces the source of grid data in the dispatch automation system, discusses how to extract and analyse the big data in the dispatch automation system efficiently and safely, and provide scientific data support for the economic and stable operation of the power grid and regional planning and development.

Keywords: Big data, power system, power dispatch, power automation technology.

1. Introduction

Today's power system is moving towards automation, and the degree of automation is getting higher and higher. In the development process of the power system, the automation of the power system has played an important role in the safe and stable operation of the power system, and the level of automation has also made a qualitative leap along with the changes in demand and the development of technology. From the initial power dispatch centre to understand and dispatch the equipment of various power plants and substations through the telephone, it has developed into the current comprehensive automation equipment based on computer technology to monitor power equipment, which is truly automated and has produced huge economic benefits.

The use of modern information technology can effectively allocate power dispatching, thereby promoting the rapid development of the power industry. With the advancement of science and technology, my country's modern information technology has made great progress, and the difficulty of power dispatching is gradually increasing. The traditional automatic power dispatching system has been unable to keep up with the pace of social development and the development needs of the power industry. Regarding this phenomenon, my country's power industry is undergoing a brand-new reform [1]. The traditional power dispatching system has slowly developed in the direction of informationization, and the power dispatching system is constantly innovating and improving. However, in the process of power reform, the information and network of the power dispatch automation system have made the use of the power dispatch system a certain hidden danger, and even
some data loss, system damage, etc., have hindered the power industry. The development of the power industry has gradually become the focus of attention in the future development of the power industry. This article re-examines the technological development direction of the power dispatch automation system industry in accordance with the Internet + thinking mode, and discusses the feasibility of using the big data cloud computing architecture for the dispatch automation system platform [2]. On this basis, a new generation dispatch automation system architecture scheme based on big data cloud technology is proposed, and technical ideas for grid operation data pool and big data analysis application are proposed.

2. A new generation of dispatch automation system solution based on big data cloud computing architecture

2.1. Hardware architecture
The hardware architecture of the new generation of dispatch automation system adopts cloud computing cluster technology. As shown in Figure 1, a cluster is composed of tens, hundreds, and thousands of servers connected through a local area network to provide a service, such as ECS (Elastic Computing Service). The cluster provides virtual servers, the storage and analysis computing cluster provide distributed file system, real-time library and database services, the cloud desktop cluster provides user terminal services, and the control cluster provides automatic deployment and monitoring services of resources.

![Cloud computing platform hardware architecture of dispatch automation system.](image)

Figure 1. Cloud computing platform hardware architecture of dispatch automation system.

2.2. Platform architecture
The cloud computing platform architecture of the new generation dispatch automation system is shown in Figure 2. The platform fully adopts a distributed computing architecture, which requires that it can process hundreds of millions or even hundreds of billions of data points, and the data computing
efficiency is within seconds [3]. The data types processed include: unstructured data; second-level/streaming data and millisecond-level data.

The modules of the platform include: 1) Unstructured data distributed architecture, such as MapReduce, ODPS (Alibaba Open Data Processing Service); 2) Real-time data distributed computing architecture, such as Storm, Spark, Analytic DB, ADS (Alibaba’s Real-time high-concurrency online analysis database); 3) Distributed relational database, such as RDS (Alibaba Relational Database Service); 4) Distributed real-time database, such as HBASE (Distributed column-oriented open source database on Hadoop); 5) Distributed File systems, such as HDFS and Panga; 6) Automatic cluster deployment and monitoring; 7) Elastic computing service (ECS) and provision of virtual host services; 8) Cloud security monitoring and protection; 9) Load balancing.

The scheduling cloud computing platform completely abandons the classic IOE architecture and will change the existing system construction model. It will no longer purchase hardware, operating systems, and databases separately to build business systems, but use the virtual resources provided by the cloud computing platform to build subsystems or replace the traditional the architecture professional system is migrated to the cloud platform [4]. With the help of cloud clusters and distributed computing technology, the functions on the platform are completely decoupled from the hardware, and the business system will gradually be integrated into functions to solve the real-time library capacity and performance bottlenecks, the technical barriers of the message bus, and the system reliability. At the same time, drawing on the ideas of the Internet, the construction of grid operation data pool and APP pool in the cloud platform will provide a new model of big data application development.

3. Task automation scheduling model
Based on the existing resources of the power system, this paper establishes a task scheduling model from provincial dispatch to local dispatch; in this model, provincial dispatch is responsible for decomposing cloud tasks into multiple cloud task streams of appropriate size, and local dispatch Responsible for processing the corresponding subtasks of the calculation [5]. The function of this model is mainly to solve the sudden tasks in the power system that need to be processed quickly. This article mainly introduces the task scheduling strategy of the model from provincial to local, so that cloud tasks can be processed quickly and efficiently, and the maximum completion time of tasks is
minimized; the maximum completion time of tasks includes task processing time and task transmission time. The task scheduling model is shown in Figure 3.

Figure 3. Task scheduling model.

Several elements of the task scheduling model are as follows:
(1) Cloud task set \( T = \{ t_1, t_2, ..., t_m \} \) represents a task set with \( m \) subtasks, \( t_i = \{ Num, Length \} \). \( Num \) represents the task label, \( Length \) represents the task length.
(2) Resource node set \( V = \{ v_1, v_2, ..., v_n \} \) represents a set of \( n \) resource nodes (ground modulation), where each resource node is represented by \( E \), where \( Num, Mips, Bw \) represents the resource node number, processing speed and bandwidth.
(3) Task processing time \( EEC = \{ dTime_{ij} \} ; i = 1,2,...,m; j = 1,2,...,n \). \( dTime_{ij} \) represents the processing time of task \( i \) on resource node \( j \). Among them \( dTime_{ij} = Length_i / Mips_j \).
(4) Task transmission time \( ETC = \{ tTime_{ij} \} ; i = 1,2,...,m; j = 1,2,...,n \). \( tTime_{ij} \) represents the transmission time of task \( i \) to resource node \( j \). Where \( tTime_{ij} = Length_i / Bw_j \).
(5) Task allocation vector \( R = \{ r_1, r_2, ..., r_n \} \). \( n \) is the number of tasks, and \( r_j \) is the number of the resource node. The vector \( R \) represents a scheme of task allocation. Due to the parallelism of resource nodes when processing cloud tasks, the total completion time of the task should be the total time for all the resource nodes to complete the task [6]. It can be expressed as:

\[
CT = \max_{j=1}^{n} (TT_j)
\]

4. Experimental research
The thesis uses a parallel network composed of 10 resource nodes to represent the geostationary centre. The 10 resource node parameters are shown in Table 1. Select 13 task sets with tasks ranging from 20 to 140 in length. Through the simulation experiment, the assignment of the 13 groups of task combinations to the parallel network composed of 10 resource nodes is obtained, and the task completion time is shown in Figure 4. Cloudsim simulation platform, its core categories include cloud computing platforms and simulation resources such as Data Centre, Data Centre Broker, and Host. In
this paper, by extending the Data Centre Broker class, the bind Cloudlet ToVm( ) method is overloaded, and the bind Cloudlet ToGA( ), bind-Cloudlet ToGS( ), and bind Cloudlet ToSQ( ) methods are added to the cloud computing resource scheduling.

| Numbering | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|-----------|----|----|----|----|----|----|----|----|----|----|
| Mips      | 278| 132| 209| 250| 500| 100| 300| 320| 261| 415|
| Bw        | 100| 150| 180| 260| 500| 320| 440| 155| 410| 320|

5. Conclusion
In the power dispatch automation system, the application of data mining techniques such as fuzzy analysis, gray analysis and neural network methods can allow the power dispatch automation system to effectively collect data, which can improve the quality of power dispatching work and make external power supply more efficient. Reliability improves the safety and stability of grid operation. With the in-depth application of big data technology in the construction of smart grids, there will inevitably be a large number of new requirements, new technologies, and new applications. The author will closely follow the latest application trends and research hotspots of related technologies, and carry out in-depth work. Enhance the level of intelligent management of the power system.

References
[1] Song, Y., Zhou, G., & Zhu, Y. Present status and challenges of big data processing in smart grid. Power System Technology, 37 (4) (2013) 927-935.
[2] Yu, D., Song, S., Zhang, B., & Han, X. Synergistic dispatch of PEVs charging and wind power in Chinese regional power grids. Dianli Xitong Zidonghua (Automation of Electric Power Systems), 35 (14) (2011) 24-29.
[3] Zhao, J., Liu, S., Zhou, M., Guo, X., & Qi, L. Modified cuckoo search algorithm to solve economic power dispatch optimization problems. IEEE/CAA Journal of Automatica Sinica, 5 (4) (2018) 794-806.
[4] Wang, B., Fang, B., Wang, Y., Liu, H., & Liu, Y. Power system transient stability assessment based on big data and the core vector machine. IEEE Transactions on Smart Grid, 7 (5) (2016) 2561-2570.
[5] Brayne, S. Big data surveillance: The case of policing. American sociological review, 82 (5) (2017) 977-1008.
[6] Diamantoulakis, P. D., Kapinas, V. M., & Karagiannidis, G. K. Big data analytics for dynamic energy management in smart grids. Big Data Research, 2 (3) (2015) 94-101.

[7] Li, J., Tao, F., Cheng, Y., & Zhao, L. Big data in product lifecycle management. The International Journal of Advanced Manufacturing Technology, 81 (4) (2015) 667-684.