

# Study on Data Mining Platform based on Cloud Computing and Its Key Techniques

Li Li

Yunnan Vocational College of Mechanical and Electrical Technology, Kunming, Yunnan, 650203

**Keywords:** Cloud Computing; Data Mining Platform; Key Technologies

**Abstract:** With the continuous development of science and technology in our country, the Internet has become an indispensable part in people's life. The cloud computing with Internet as the core has also been widely used in social production. Due to its high efficiency, large capacity and dynamic processing, cloud computing is favored by the commercial field and scientific research field. Cloud computing can be effectively used in the data mining platform architecture by organically applying cloud computing. Improve data mining efficiency and create good conditions for the collection of relevant data. Based on the actual situation, this paper analyzes the framework of cloud computing-based data mining platform and its key technologies to promote the continuous improvement of cloud computing data mining platform architecture.

## 1. Introduction

In the course of development, various advanced technologies have been gradually derived from the era of development. With the rapid development of social economy and science and technology, advanced technologies have been increasingly improved. In the era of the Internet, cloud computing has become a more common technical subjects, is widely used in various fields, because of its high efficiency, large capacity, dynamic processing and other characteristics, is used in data mining platform architecture to effectively improve Data mining efficiency. The application of cloud computing in data mining platform architecture can facilitate the collection of user data and help users quickly filter out the data information related to user requirements from a large amount of data and information and provide the stability of the data mining platform architecture Strong protection. Based on the actual situation, this article analyzes the framework of cloud computing-based data mining platform and its key technologies, with a view to providing valuable experience for future related work.

## 2. Overview of cloud computing

With the continuous development of science and technology in our country, cloud computing has been widely applied in all fields of society and has become an important tool to promote the continuous development of society. Cloud computing mainly refers to the distribution of multiple computing tasks on a resource pool composed of computers, so that all the systems in the computer can obtain the corresponding computing capabilities, information services and storage space according to their own needs so as to effectively improve the computer work efficiency. As cloud computing technology is a combination of a variety of mature technologies, together with the strong support from the state made cloud computing in recent years has been rapid development and lay a good foundation for future development, and in the promotion of cloud computing under the part of the soft Hardware IT companies, Internet companies and local governments all join in the promotion of cloud computing applications, so that the market potential of cloud computing is further increased, accelerating the promotion and development of cloud computing to a certain extent. In the information society, the reason why cloud computing will be widely promoted, mainly due to the following characteristics: First, cost-effective. By using the cloud computing model, the computational ability of computer information data can be effectively improved, the storage space is effectively broadened, and the computer management is more convenient. The user only needs to

input some simple instructions at the client terminal to complete the complicated calculation of information data and effectively save The user's time, for the user's life, work, study to facilitate. Second, high scalability. In the course of the development of the times, users are constantly demanding computers. Therefore, effectively increasing the computing scalability has become an inevitable trend in computing and development. In the cloud computing model, the computing power can be increased or decreased by adding or removing servers in the "cloud" To achieve, so that computer expansion capabilities have been effectively promoted, prompting users to work and learn more efficiently. Third, high reliability. Through the use of cloud computing can effectively improve the reliability of data and information, and mainly reflected in two aspects: ① data is reliable, data information in the cloud computing mode will store information in the cloud data space, so that data will not be damaged due to computer information Loss of damage; ② reliable data. As users in the application process by the service provider through a unified deployment in a "cloud" environment to operate, so will not be subject to personal computer problems affecting the accuracy of data and information. Fourth, high flexibility. In cloud computing mode, calculators can share information anytime, anywhere and compatible users can apply for cloud computing in different states and under different conditions in different ways, and at the same time, users can be provided personalized service according to their own characteristics and needs, For the convenience of users.

### **3. Cloud computing data mining platform architecture analysis**

Data mining belongs to a comprehensive technology. It mainly combines a series of knowledge such as statistics, information retrieval, intelligent acquisition and artificial recognition to form data mining technology. In actual use, this technology can extract the data information needed by people in the vast data information database to facilitate people's life and work. Data mining has been widely used in the fields of Internet, mobile Internet, telecommunication, finance and scientific research. For example, Facebook's friends recommend the application of products recommended by Amazon and Taobao, and the anti-fraud analysis of banks. By organically applying the cloud computing to the data mining platform architecture, the efficiency of data mining can be effectively improved, and the data collection work can be facilitated. With the help of cloud computing platform, various network service functions can be realized to store and analyze network data, and data mining platform can excavate data needs of users. Since these two technologies have common features in some aspects, all can promote artificial intelligence technology development and implementation. The data mining platform architecture based on cloud computing combines the key points of data information collection, processing, processing and storage and performs functions such as data calculation, mining and service on the cloud platform respectively. Because data mining capabilities are mainly based on data computing, mining and service functions, it is possible to schedule and process information data on a cloud computing platform to create favorable conditions for efficient operation of data mining platform architecture.

In order to cloud computing applications in the data mining platform architecture, researchers in the actual development process should be strictly controlled cloud computing data mining platform architecture development environment, providing automatic expansion and dynamic services, in order to improve the use of this platform architecture Safety and stability in the process. Under normal circumstances cloud computing data mining platform architecture is composed of user interaction, application, data mining and other subsystems, mainly cloud computing as the core technology, through the use of the Internet to connect the terminal device and data platform for the platform Data mining to create good conditions, in the process of building user interaction, applications, data mining and many other subsystems for organic combination, better to provide users with the demand for information. Among them, the user interaction subsystem as a user and cloud computing data mining platform interactive window, including registration, cancellation, user authorization application, user authentication, etc., for the user's property, account security to provide a strong guarantee; application Subsystem is the interface between user management and

data mining. The user inputs the demand data by using the interactive subsystems, and the application subsystem can accurately describe the relationship between the data and algorithm according to the user's needs. And the main task is to achieve the relevant data and algorithms to describe, and registration and cancellation, user applications and other functions are driven by the application subsystem; data mining subsystem in the relevant data to dig deeper, but also both data computing, information processing and other functions, including data mining, pretreatment, data evaluation, special services and other functional modules. After using the data subsystem to process and analyze the data information, the data can be automatically stored in the database according to a fixed format to meet the information needs of different users and facilitate the use of the data.

In addition, researchers in the process of cloud computing data mining platform architecture should be based on rapid analysis of this model based on the actual needs of users to develop the platform architecture, and do a good job of related investigations, so that the cloud Computing data mining platform architecture in the use of a certain degree of system security. In order to fully guarantee the reliability of supporting technologies, the development of cloud computing data mining platform architecture, should be based on the actual situation and use to adapt to the development of software tools. This requires researchers to combine cloud computing data mining platform architecture specific characteristics of the actual needs of users of the platform architecture to evaluate the effectiveness of platform architecture to ensure the rationality.

#### **4. The key technologies of cloud computing-based data mining platform architecture**

Cloud computing technology applications. Currently, cloud computing applications are mainly divided into distributed storage technologies, virtualization technologies, and parallel cloud computing technologies. Among them, distributed storage technologies mainly use distributed storage to store a large amount of data and fully guarantee the economic performance and reliability of cloud computing data processing Sexual, stability. Under normal circumstances, in order to make the cloud computing data to achieve the best effect, data mining platform architecture is mainly used to have better reliability of software to make up for the lack of hardware, and cloud computing in the distributed storage technology, a certain To some extent, the current status quo has been changed to provide strong technical support for the data mining platform architecture and to reduce the cost of the cloud computing data mining platform architecture with greater economic advantages. By applying the cloud computing technology to the data mining platform architecture, the stability of the platform operation can be fully ensured, the reliability and security of the platform can be improved, and the favorable conditions for the long-term development of the cloud computing data mining platform architecture can be created. Virtualization technology is mainly to make the data in the full virtual environment for application and integration, the platform can be achieved in cloud computing data mining capabilities, in the dual role of virtualization and cloud computing, cloud computing data mining platform for resource scheduling It can integrate data across systems, use IT resources to collect a large amount of data, and then present it as a dynamic form to users in front of users to provide users with virtualized resources. Only in this way can we fully guarantee services collected, processed and integrated through the cloud computing data mining platform The tasks are highly accurate, improve the operating efficiency of the platform and make the application of platform data more in line with the development needs of the current era. As the parallel cloud computing technology can effectively accomplish the related service tasks in actual application process, it is one of the key technologies in cloud computing data mining platform architecture, fully guarantee the platform operation efficiency, and have the encapsulation function at the same time, and minimize the clustering performance index As the main idea to carry out related work, and how to parallel update clustering center as a parallel development ideas can be part of the technical details of the package for the user to provide a strong guarantee of security.

Data Collection and Dispatch Center Under the promotion of science and technology and strong support from the state, the technologies used in cloud computing data mining platform architecture are gradually improved. By utilizing the cloud computing data mining platform architecture, a large amount of data and information in the network can be collated, analyzed and combined with users

Demand to provide users with data and information, for people's lives has brought great convenience. In the actual operation process of the cloud computing data mining platform, the data collection and dispatch center is a more important and crucial foundation to help the cloud computing data mining platform architecture to collect and organize data. This technology can collect different types of information that exist in the network, meanwhile, it can also collect the internal business data that is accessed to the cloud computing data mining platform and support multiple different source data formats in actual operation, and can provide different users The same service effectively solves the problem of statutory rules caused by the differences between the data and makes the cloud computing data mining platform run more efficiently.

Service scheduling and management techniques. Based on cloud computing to establish data mining platform architecture, expanding the platform application, so that all sectors of social production can be combined with their own needs and reasonable use of cloud computing data mining platform to effectively improve work efficiency, which requires the platform in the operation process needs to have Management and scheduling functions. Therefore, in the construction of cloud computing data mining platform should take full account of management, service scheduling issues, and service scheduling and management techniques applied to cloud computing data mining platform to strengthen platform management. By applying the service scheduling and management technology to the data mining platform, each system in the platform can be fully guaranteed to be mutually exclusive and mutually isolated during operation so as to ensure the safe operation of the cloud computing data mining platform in an efficient and effective manner Sex, to provide users with better service. In addition, the application of this technology to meet the unified registration of services to support third-party data mining capabilities, expanding the development of cloud computing data mining platform to promote research and development.

## 5. Conclusion

In summary, the emergence of cloud computing technology is mainly based on the parallel computing technology, network technology, software technology is established, is the inevitable product of the development of the times, because of its own high reliability, expanding features and is widely used Social production in all areas. Through the application of cloud computing, the further development of data mining platform has provided powerful technical support for the development of data mining technology and realized parallel mining of data in different formats, laying a good foundation for the stable operation of data mining, analysis and processing.

## References

- [1] Ding Yan, Yang Qingping, Qian Yuming, et al. Study on the architecture of data mining platform based on cloud computing and its key technologies [J]. ZTE Technologies, 2013,19 (1): 53-56,60
- [2] Liu Jian Dong. Cloud computing platform for data mining and exploration of key technologies [J]. Technology and Innovation, 2017 (6): 128,132.
- [3] Wang Dewen. Editorial office of Beijing University of Posts and Telecommunications, 2012,36 (11): 67-71,107. Centralized power data center based on cloud computing and its key technologies [J]. Automation of Electric Power Systems, 2012,36 (11): 67-71,107.
- [4] Ge Xiaogu, Liu Jie. Cloud computing-based data mining platform architecture and its key technologies [J] Jingdezhen Institute of Education, 2017,32 (3): 26-29.
- [5] Chen Lei. Cloud computing-based data mining platform architecture and its key technologies [J]. Computer programming skills and maintenance, 2017 (6): 64-65.
- [6] Zhang Lizhong. Discussion on cloud computing-based data mining platform architecture and its key technologies [J]. World of Communications, 2015 (16): 193-193.