

Research and Implementation of Secure Storage of Big Data Based on Distributed Database

Tingting Zhang^{1*}, Ze Liu²

¹Sichuan Aerospace Vocational College, Guanghan Sichuan, 618300, China

²Sichuan Aerospace Vocational College, Guanghan Sichuan, 618300, China

*Corresponding Author

Keywords: Distributed, Database, Safe storage, Study

Abstract: With the continuous development and progress of the society, the information technology industry has achieved unprecedented development. Data statistics of information systems also show explosive growth, gradually increasing from the original TB level (TB = 1024 GB) to the PB level (1PB=1024TB). So we can see from the data that today's data is growing very fast. In addition, due to the continuous growth of data, which has increased people's demand for data storage, the big data era is no longer far away from people. It can be said that the big data era has come. With the development and progress of various industries in the society, the amount of data used in each industry will grow rapidly no matter which field. And big data will also be an important standard for information and industrial development in the future. Therefore, this paper will mainly discuss the research and implementation of big data secure storage based on distributed database.

1. Introduction

Under the current situation, big data storage has gradually become the main development trend of today's storage technology. Big data storage is a new mode of storage today. In the process of storage, people not only need to transfer data storage from traditional personal storage devices to the network to realize online storage. At the same time, it is necessary to consider the reaction time of the data in the process of access, the inquiry method, the data security and the storage performance in the storage process. With the continuous development of big data, the research on big data storage technology has become an inevitable requirement for today's big data applications. The application of distributed database ensures the security of data storage and the realization of data storage on the network. In recent years, information security has been widely valued by various countries. However, compared with western developed countries, China still has a late start and is still in a passive situation. The security storage of distributed database can ensure the security of data in the storage process. Therefore, it is necessary to apply the distributed database technology to the process of data storage.

2. Significance of Distributed Database

With the rapid development of today's era, traditional information storage methods are far from meeting the needs of today's people. Traditional centralized database collation has not kept pace with the times. Therefore, it is necessary to find new storage methods in the storage process to improve this situation. Then distributed database arises spontaneously and is developed and applied under the new background of data storage. The application of distributed database has many advantages that traditional database does not have in the process of application: the first point is that distributed database has very large expansion capability compared with traditional centralized database, and the expansion capability is also an important embodiment of the development of today's database. The traditional centralized database does not have the advantage of expansion, which leads to the elimination of traditional database storage methods. The second point is that

distributed databases are lower in cost than traditional centralized databases. If the traditional database is still used in today's large databases, it will affect the cost of capacity expansion and greatly increase the cost. And after spending a lot of money, the data information obtained is still very limited. However, in the application process of distributed database, due to its capacity expansion, its cost is greatly reduced. Finally, the most important point is that the application of distributed database in big data can well ensure the storage security of users in the application process. Distributed database also makes storage more convenient and easy, so distributed database technology needs to be applied to the process of data storage in a more perfect and comprehensive way.

3. Analysis of Distributed Database Technology

Nowadays, the information content of big data is continuously expanding, which makes distributed database technology widely used in big data. Distributed database is mainly completed by various processing units combined, and each unit is also a very complete system. In this system, whether it is data storage or data analysis, for a certain unit, their work and status are equal, and each unit is also very independent. Distributed database, which contains mixed storage technology, breaks through the traditional technology and realizes the mixed storage mode of rows and columns. From this we can see that the application of this technology in big data can effectively improve the performance of big databases. Therefore, the distributed database has made great progress in speed and flexibility. In addition, another technology is intelligent guidance technology. In the application process of distributed database, the application of intelligent guidance technology gradually reduces the original space occupation and solves the problem that the database runs slowly. In addition, distributed database technology also includes many advanced technologies to support the development and application of distributed database. For example, the parallel processing technology, efficient and transparent compression technology, and various technologies that do not exist in the traditional database storage process, etc. included in the distributed database have higher performance than the traditional centralized data storage mode in all aspects. Therefore, this is one of the reasons why distributed databases can be widely used in today's big data.

4. Application of Distributed Database in Big Data

4.1 Distributed Database Ensures Data Security in Mixed Applications

With the rapid development of big data information today, more and more businesses are needed in all walks of life. In the process of data storage, such as data loading, data processing and analysis, more and more places need to use distributed databases in data. In the era of big data, the use of data is not limited to a single database, but distributed database can provide users with a variety of databases to choose from in the process of application. Distributed database technology can realize the mixed application of data in the application process, so it can embody the important supporting function for data. Distributed database technology can not only make reasonable screening of data, but also effectively screen the information needed by users. At the same time, in the process of screening, it can also process information for users, thus ensuring that the database application is carried out according to a certain program and meeting the information needs of all walks of life in the world today. In addition, due to the fact that the distributed database technology processes the user during the information retrieval process, the security problem of the user during the data application is better guaranteed, and the realization of data storage security is improved.

4.2 Distributed Database Applied in Various Services

The foundation of data and information is the development of the Internet. With the continuous progress of science and technology in the world today, the Internet technology is also continuously developing. The economy of various industries is developing towards a virtual economy. This makes the business volume among various industries continuously increase, and brings great problems to the business management among various industries. However, the application of

distributed database technology can enable various industries to meet the actual needs in the process of business processing, so that the use in various industries can realize the common use of multiple users. Distributed database not only has better security performance, but also has super load management capability. In front of a large number of businesses and data, the application of distributed database technology can solve the problems existing in the current operation stage well, and can also classify the data, thus reflecting its higher level of business management. In addition, distributed database technology can also realize priority management in the process of business handling. For example, users with VIP can occupy more resources in the application process. For data, users can extract hot data from large databases and set it as priority data. Through this way of data processing, the performance of distributed database in the process of business processing can be successfully improved. Therefore, it can be seen that the security of distributed database can be applied in various industries.

4.3 Application of Distributed Database in Query Process

The application of distributed database technology not only plays an important role in the above aspects, but also has certain advantages in the process of query, and the usage in the process of query is also very large. In the process of application, there will be a platform that can provide self-service for people. This platform is mainly to help people analyze data effectively. The number of users faced by this platform during the application process is very large, which greatly increases the records of simultaneous access between users. And in the process of accessing, there will also be the phenomenon of multiple people accessing together, which requires the application of concurrency control in distributed database technology. The investigation of the actual situation shows that the distributed database can accommodate more than 200 people to enter and query information at the same time. In addition, if there are less than 100 million pieces of data in the query process, the distributed database can strictly limit the processing time to two hours. If there are more than 200 million pieces, distributed data can also put the time within 40 minutes. It can be seen that in the process of data query, the application of distributed database occupies a leading position today, which can effectively control the time in the query process. In this way, the safety performance in the process of data storage is well guaranteed.

4.4 Application of Distributed Database Technology in Big Data Loading Process

Distributed database has very strong loading capability even in the actual application process. Nowadays, with the rapid development of big data, the ability of data processing and statistical loading is very important, and distributed database can solve this problem well. It can also be seen from this that one of the prerequisites for the widespread application of distributed database in the application process is its very strong loading capability. According to relevant scientific investigations, the data statistics of the distributed database can be guaranteed at about 3 billion pieces per day. This ensures that each loading node can ensure a loading speed of 105,000 bars per second during the loading process. Then if monthly statistics are carried out, it can reach about 30 billion articles per month. This is a very frightening set of numbers, which is one of the outstanding performance of distributed database technology today. Therefore, it can be seen that the distributed database technology can well display its unique characteristics in the process of big data loading and drive the safety of the whole big data storage process.

5. Conclusion

In a word, the application of distributed database under big data can well guarantee the security problems in the process of information storage. And can also ensure the performance problems of users in the application process, improve the efficiency of solving data loading, reduce the cost of data storage and expansion, thus obtaining greater benefits. In addition, the application of distributed database technology can also improve the shortcomings of traditional storage technology, so that it can follow the pace of the development of the times. Keeping pace with the times and developing with the trend of the times are the necessary strategic means to achieve long-term

development. Although distributed database ensures the security of data storage process, there are still some problems in the application process of distributed database technology. This requires relevant researchers to be able to continuously discover and overcome suffering. Therefore, the development of the motherland in information will be more prosperous, and China's international influence will continue to rise in the international arena.

References

- [1] Shen Yang. Research on the implementation of spatial big data distributed storage in NoSQL database [D]. Science and technology wind, 2018 (23): 224
- [2] Cai Manyi. Research on PostgreSQL data analysis extension middleware based on spark [D]. University of Chinese Academy of Sciences (Shenzhen Institute of advanced technology, Chinese Academy of Sciences), 2017
- [3] Yan jianzhuo, Gao Kaili, Xu Hongxia, Yu Yongchuan. Design of distributed big data storage platform for water affairs based on Virtualization [D]. Water conservancy informatization, 2019 (03): 17-24