

Research on Adaptive Migration and Fusion of Stored Data Based on Block Chain Technology

Gao Fei

School of information science and technology, Tibet University, Lhasa Tibet, 850000, China

Science and Research Office of Tibet University, Lhasa Tibet, 850000, China

Keywords: Block chain technology; Storage data; Adaptation; Migration and fusion;

Abstract: In recent years, with the rapid development of storage technology, the performance of storage devices has been continuously improved, and many new storage technologies have emerged. Based on this, an adaptive migration and fusion method of storage data based on block chain technology is proposed. Firstly, the adaptive modal weight is adjusted, the storage metadata is designed, and finally the migration and fusion of data blocks is completed. By comparing the experimental results, the adaptive migration and fusion method of storage data based on block chain technology designed in this paper has a positive proportion relationship between storage data and data migration and fusion ability, that is to say, with the increase of the amount of storage data, its ability also increases. Experiments show that the method designed in this paper has great advantages over traditional methods.

1. Introduction

Internet technology is gradually affecting every aspect of people's daily life, and many social applications are also emerging with huge amounts of data. There are more and more applications that need to access links and operate huge amounts of data [1]. These applications must be popularized on a safe, reliable, scalable and practical platform, so it is necessary that the underlying database can perform better in function and cost-effective use, and also have the ability to gradually influence and serve data in various regions of the world and access links. Since the 1930s, with the emergence of many Internet companies, the amount of data that need to be operated has exceeded the amount of data that traditional companies need to operate at any time in human history. According to a survey of a statistical report, the amount of data worldwide is increasing rapidly. Thirty percent of the world's data come from recent years, which can be said to be an explosive growth. Take Facebook, the largest social networking application, for example. Since its inception, the company has collected more than 300 PB of user data, which is more than 150 times the total data collected by American libraries since 2000 [2]. In today's big data application, data must be constantly collected and analyzed, so as to effectively promote the continuous development and progress of national economy and social innovation. Enterprises or organizations can use the data they collect to provide personalized services, optimize the decision-making ability of the company, and evaluate the future development trend. In other words, today's data are an asset of great potential and economic value to the human economy.

2. Adaptive migration and fusion method of storage data based on block chain technology

Since mass media technology broke through the previous means of information exchange, the rapid development of social informationization, especially the emergence of new technologies such as computer, cloud computing and big data, has greatly improved the efficiency of information transmission to a certain extent, at the same time, it has also brought blowout bursts of social and resident information such as Internet [3]. Block chain technology, as a new technology of de-centralization, opens a gap for the means of information transmission on the Internet from the side, which makes the main body and carrier of information dissemination fission to a certain extent.

From the past, information media mastered information resources, controlled information channels, and now online users in the information application platform receive information, transmit information, transmit information interactive participation, so that users can speak boldly; in the future, the extensive promotion and use of block chains, making human beings as the main body of information transmission, so as to continue to voice in the decentralized information dissemination network. From the standpoint of data location and computational power, block chain is also called distributed accounting technology. It mainly takes a distributed computing network as the precondition, and does not need the examination of other central networks. It thoroughly smashes every data file. At the same time, it allows network users to repeatedly encrypt their login keys and disperse them in the computer network [4]. On the other hand, the concept of "work proof" introduced in block chain technology is to make use of the competition of computing power to ensure that the recorders can make as much effort as possible in data writing, thus effectively realizing the storage of data, computing and network virtualization, and corresponding to the distributed storage of block chain and the demand for computing power.

2.1 Adaptive Modal Weight Adjustment

In the process of data storage and aggregation, parameters specify the basic weight of data similarity for each stored data. The more effective the feature modes of clustering fusion can be distinguished, the higher the basic weight ratio [5]. In the process of incremental co-clustering algorithm for multi-modal data storage, with the gradual increase of data instances, the data clustering structure is gradually updated and developed, and the corresponding matching data modal weights are constantly changing. Therefore, it is more important to design an adaptive update mode for storing data modal weights than to set a fixed value of modal basic weights by experiment. Generally, a clustering cluster with matched feature modes for storing data has a fairly dense distribution of data instances. The dense distribution performance of this cluster can be measured in detail by using the difference between the actual case of storing data in the cluster and the center point. In the clustering process, since the initial modes have some similarity discriminant features, the basic weight of each mode needs to be set at the same value. Data information is distributed evenly on each virtual node by hashing algorithm to complete the centralized migration of stored data. When data is updated, Redis Live monitors the data cluster in real time and collects important information. Once the data carrying capacity of the whole system is below the standard threshold specified by the equilibrium degree, it will trigger the data migration operation immediately, otherwise, it will remain in the current state; the selection of migration scheme is particularly important. The key of data migration and migration lies in the selection of nodes and the size of migration cost, which are the key points to measure its integration. Taking these factors into account, the storage data migration is formulated. The best choice to move.

2.2 Design of Storage Metadata

In order to effectively solve the problems of related management and access links in storage data and complete the task of scientific and efficient data migration and fusion, it is necessary to design metadata for each storage data. Metadata design in storage data includes file location, size, access time, read and write rate, access record and specific time of input file in storage database [6]. This paper adopts metadata server, namely Meta Data Server, abbreviated as M D S, which refers to the management of all metadata in the database on behalf of others. Metadata servers are generally divided into centralized and distributed services. Centralized metadata servers mainly use special MDS to manage metadata, so they have a high degree of sharing and can achieve the requirements of simple operation. Distributed metadata servers use master/slave backup to manage metadata, but their security and reliability are poor, for example, UCS C and Lustre belong to this way. It should be noted that distributed MD S mainly refers to the absence of a special MDS manager in the database, so that metadata can be stored in devices in a decentralized manner. All storage devices can store both metadata and ordinary data. This means can greatly improve its performance and data scalability to a certain extent.

2.3 Data Block Migration and Fusion

In order not to adversely affect the hybrid storage database, the data migration and fusion link must be completed when the database is not fully operational, and it is necessary to predict whether SSD is in idle state at intervals. Once in idle state, data migration and fusion must be completed. It is necessary to distribute the storage database scientifically in SSD for the storage database that meets the operating conditions effectively. According to the detailed information recorded in the actual observation record, the maximum values of all data blocks are calculated. In order to maximize the use of SSD, storage databases must be distributed as centrally as possible throughout the SSD, and it is necessary to keep the sum of the maximum values distributed in the SSD to the maximum. The migration and fusion of storage data is carried out from top to bottom. When the remaining amount of storage devices only accounts for 20% of the total capacity, the migration and fusion conditions can be triggered immediately, which can effectively activate the progress of migration and fusion. The data with the lowest value accounting for 20% of the total capacity can be formed into a migration and fusion queue in turn and begin to perform tasks.

3. Experiments and Effect Analysis

In order to see the effect of this method more clearly and concretely, this paper compares it with the traditional method of adaptive migration and fusion of stored data, and compares its data migration and storage capacity.

3.1 Experimental preparation

In order to ensure the accuracy of the experiment, two methods of adaptive migration and fusion of stored data are put in the same experimental environment to carry out the related experiments of data migration and storage capacity. The setting of test environment is shown in the table below.

Table 1 Laboratory Environment Settings

Project configuration	Servers and workstations
Memory	3GB and intel pentium 4 4
CPU / GHz	AMD athlon 2. 42 and 500 MB
Hard disk / GB	WD SATA 40
Network card	8270KM and 8050BM

3.2 Analysis of experimental results

During the experiment, two different methods of adaptive migration and fusion of stored data are used in the same environment at the same time, and the change of data migration and fusion ability is analyzed. The actual effect comparison is shown in the figure below.

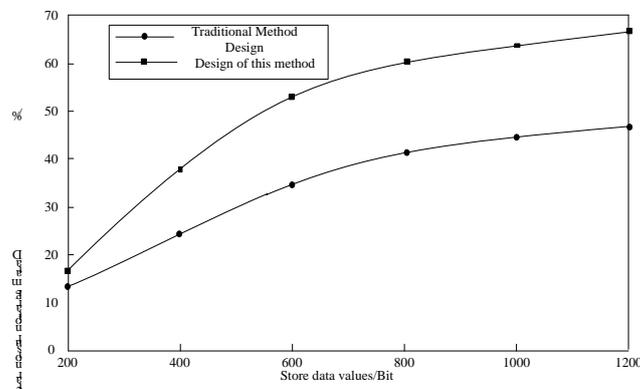


Fig. 1 Experimental comparison chart

By comparing the experimental results, we can clearly see that the adaptive migration and fusion method of storage data based on block chain technology designed in this paper has a positive proportion relationship between storage data and data migration and fusion ability, that is to say,

with the increase of the amount of storage data, its ability also increases. Experiments show that the method designed in this paper has great advantages over traditional methods.

4. Conclusion

In this paper, the method of adaptive migration and fusion of storage data based on block chain technology is analyzed. According to the advantages of high speed, convenience, simplicity and effectiveness of block chain technology, the method of adaptive migration and fusion of storage data is adjusted according to the technical points of a series of processes of adaptive migration and fusion of storage data, and the design of this paper is realized. The experimental demonstration shows that the method designed in this paper is highly effective. It is hoped that this study can provide a theoretical basis for the method of adaptive migration and fusion of stored data based on block chain technology.

Acknowledgement

Project name: Tibet Autonomous Region's 2017 Key R&D and Transformation Plan"Key Technology Development and Solution Design of E-Commerce Application Platform Based on Blockchain"

Project No.:xz201703-gc-10

References

- [1] Liu Nan, Liu Li. Integration of Block Chain and Cloud Computing to Develop BaaS[J]. Communications World, 2017, 22(17):61-62.
- [2] Wang Kang, Li Dongjing, Chen Haiguang. An Improved Consistent Hashing Algorithm in Distributed Storage System[J]. Computer Technology and Development, 2016, 26(7):24-29.
- [3] Dong Hongbin, Teng Xuyang, Yang Xue. Feature Selection Based on the Measurement of Correlation Information Entropy[J]. Journal of Computer Research and Development, 2016, 53(8):1684-1695.
- [4] Xu Chuanfu, Che Yonggan, WangZhenghua, et al. Scalable and Evenly Distributed Simulation Method for Computer Architecture[J]. Journal of Software, 2014, 25(8):1844-1857.
- [5] Qin Xiulei, Zhang Wenbo, Wang Wei, et al. An Overhead-Sensitive Data Migration Method for Cloud Key/Value Storage Systems[J]. Journal of Software, 2018, 24(6):1403-417.
- [6] Gao Changyuan, Zhang Yunhui, Zhang Shuchen, et al. Cloud database dynamic path optimization based on immune polymorphic ant colony algorithm[J]. Application Research of Computers, 2015, 32(10):2955-2959.