Research and Design of Education Management System Based on Cloud Storage

Qingyan Cheng(✉)

Nanchang Institute of Technology, Jiangxi 330044, China
ck123456m@163.com

Abstract. Information construction in Colleges and universities is the key factor to realize modern education. In the process of information construction in Colleges and universities, due to the similarity of teaching data, there are a lot of duplicate data between the education management systems independently built by colleges and universities, resulting in a great waste of storage resources, and the teaching data is growing exponentially, which brings great storage pressure to the traditional data center. All users share a resource pool, resulting in low resource search efficiency. Moreover, due to the unclear role and authority of users, teaching resources can not be safely shared among college users.

Keywords: Cloud storage · Management system · Resource sharing · Role

1 Introduction

With the rapid development of information technology and the popularity of personal computers and intelligent machines, more and more users begin to communicate with the outside world through the Internet, understand and obtain all kinds of information resources they need, and manage their own data through network storage. In the environment of colleges and universities, realizing the scientization and automation of daily education management is an important symbol of realizing education modernization. Campus network is a key factor in the modernization and information construction of colleges and universities. Relying on the convenient campus network, using advanced computer technology, implementing more scientific, effective and convenient online management of teaching work, and establishing a series of education management systems matching with college education can further accelerate the work efficiency of colleges and universities and improve the teaching level.

Analyzing the current process of University modernization, we can find that there are many problems in its education management system. For example, the education management system used by each university is generally built, operated and maintained independently. Due to the existence of a large number of similar resources or the use of the same teaching data among colleges and universities, the direct consequence is that the data between colleges and universities is highly repeated, resulting in a waste of a large number of storage resources and increasing the capital investment in network teaching

M. A. Jan and F. Khan (Eds.): BigIoT-EDU 2022, LNICST 465, pp. 537–542, 2023.
https://doi.org/10.1007/978-3-031-23950-2_58

of major colleges and universities. Moreover, most colleges and universities may have multiple education management systems at the same time. A variety of teaching systems lead to complex campus system maintenance and a great waste of human and financial resources. At the same time, this paper analyzes the current resource sharing strategies of colleges and universities. In order to obtain relevant shared resources, users often need to consume a lot of search time, resulting in the slow progress of campus digitization, so that the rapidly developing computer network technology can not timely serve today's teaching work, which seriously hinders the process of college education modernization [1].

In the past two years, cloud computing technology has developed rapidly. Its core idea is to provide software, platform and it equipment resources in the form of services. The goal is to integrate a large number of cheap physical computing resources with low performance into a powerful computer system with high computing power, large storage capacity, and provide a unified service interface. Users can pay for software, platform or underlying hardware equipment according to their actual needs.

2 Overview of Key Technologies

2.1 Cloud Storage Technology

With the emergence of cloud computing technology, cloud storage technology provides a new solution for massive data storage. The cloud storage platform has the advantages of storage space Virtualization (by aggregating the space of multiple storage devices, flexibly allocating and deploying storage space, avoiding unnecessary hardware device costs and improving the utilization of storage space), and dynamic scalability of storage capacity (on the premise that the original storage system is not affected, storage nodes are dynamically added or deleted to meet the needs of dynamic changes of data) and data security (ensure the continuity and integrity of data and ensure that users can obtain the data they want anytime, anywhere) Based on the above characteristics, cloud storage platform can provide enterprises with cheap, dynamically scalable storage capacity and safe and complete data, and provide a method to reduce the pressure of massive data storage for traditional data centers.

2.2 Data Layout Algorithm

Cloud storage platform is composed of a large number of heterogeneous storage devices. How data is stored and arranged in these storage nodes not only affects whether the storage devices can be fully utilized, but also affects the access efficiency of cloud storage platform. At the same time, cloud storage platform has dynamic scalability, and storage nodes can be dynamically added and removed. When the storage capacity of cloud storage platform changes, it is better Xiu's data layout algorithm can reduce the amount of migrated data and improve the performance of the cloud storage platform. Therefore, it is necessary to study the data layout algorithm to build the cloud storage platform.

The data layout algorithm mainly solves the two major problems of fairness and adaptability. The actual amount of data obtained by each device in the system is related

to its performance, and the amount of data obtained by different devices is directly proportional to its performance. If - data layout algorithms meet the above relationship, it is considered that this layout strategy meets the fairness requirements. The adaptability of the layout strategy can be achieved by adding or removing device nodes It can be measured by the ratio of the actual amount of data migrated to the optimal amount of data that needs to be migrated in theory, or if it can be ensured that data migration only occurs between old equipment and new equipment, it can also be said that the corresponding strategy can well meet the requirements of self adaptability.

1. Assign a weight value to all devices according to their performance (storage location, storage capacity, network bandwidth, etc.):

$$\sum_{i=1}^{n} w_l = 1 \tag{1}$$

2. Assume that the minimum weight value of all devices is w_m. Allocate *klogn* virtual storage nodes to the storage node, where k is a constant and N is the number of actual storage nodes. For other storage nodes, the weight is assumed to be w_i. The algorithm will assign n to the device node N_i virtual storage nodes:

$$N_i = (w_i/w_m) \cdot k \log n \tag{2}$$

Therefore, a total of N storage locations are required in the ring to map all storage nodes (including physical storage nodes and virtual storage nodes) [2].

The algorithm can solve the fairness problem of the data layout algorithm. Its disadvantage is that it introduces a large number of virtual nodes, increases the storage overhead and increases the complexity of the algorithm. Especially in the case of large performance differences between devices, the number of virtual storage nodes introduced by the algorithm will be huge, resulting in serious storage overhead. At the same time, the algorithm still does not consider redundancy The remaining problem is that only one piece of data is stored in the storage space. When a storage node fails, the data stored in this node will not be recoverable.

3 Design of Education Management System Based on Cloud Storage

With the deepening of the process of educational modernization, the amount of data in the educational management system has increased exponentially, which has brought severe challenges to the traditional data center. The most direct way to store massive data is to increase the number of storage devices in the data center. However, relying solely on the addition of storage devices to solve the storage problems caused by the explosive growth of data will make college storage a success In the face of the explosive growth of data, how to realize the efficient storage of this part of data is an urgent problem to be solved in Colleges and universities.

Through the analysis and research of various data layout algorithms in cloud storage technology, an improved data layout algorithm is proposed to meet the requirements of

fairness and adaptability at the same time. In addition, according to the characteristics of college teaching data, a local duplicate data deletion algorithm based on the characteristics of teaching data is proposed to minimize data redundancy and reduce the pressure of network bandwidth, so as to reduce the cost The storage pressure of massive data on the traditional data center and the complexity of system upgrading and maintenance in the future.

The construction of digital campus has achieved initial results in Colleges and universities. The use of network, multimedia and other high-tech means for auxiliary teaching or online teaching has been gradually carried out By analyzing the current daily teaching process in Colleges and universities, it can be found that although the campus network has been built, the actual utilization rate of each teaching system is not high, and the auxiliary role for daily teaching work is not complete. For example, there are two common teaching activities: teaching courseware sharing between teachers and students and online submission of students' classroom homework. At present, there are three main ways: one is Copy through USB flash disk, that is, students use USB flash disk to copy teachers' teaching courseware in spare time such as recess; second, teachers share teaching resources by sending e-mail, that is, teachers uniformly set and publish a shared mailbox, and student users in a specific class use password to share teaching resources and submit classroom homework by logging in to the shared mailbox; third, FTP server is used.

There are two main backup strategies for information data backup using backup system: real-time backup and regular backup. Both strategies have their own advantages and disadvantages. In comparison, real-time backup has the advantages of high data integrity and good security. The disadvantage is that for the large storage system such as teaching system, real-time storage consumes too much resources and there is a great waste of resources At present, it is the most widely used. Its advantage is that it can make more full use of system resources. Its disadvantage is that when the timing policy is not set properly, it may cause the loss of data information.

Based on the frequency characteristics of data resources in the education system, the strategy of combining real-time backup, regular backup and manual backup is designed to back up the education system. In addition, the backup mode is dynamically and flexibly adjusted according to the semester characteristics of school teaching activities [3]. For example, the frequency of courseware resources in different chapters of curriculum courseware is different in the semester of the beginning of the course With the increase of class time, the use frequency of new courseware resources is high, and the use frequency of courseware resources in historical chapters is reduced. Therefore, different backup strategies need to be adopted for courseware resources in different chapters. The real-time backup strategy is adopted for courseware resources in new chapters. When the use frequency is reduced, the backup mode is dynamically changed to regular backup mode, so as to realize dynamic adjustment of storage Storage strategy to increase storage efficiency and resource utilization.

4 Cloud Storage Platform Implementation

The cloud storage platform is further modified and implemented on the basis of Apache Cassandra program, which is mainly composed of three parts: external interface, HTTP

server and back-end storage device group. External interface: the system provides storage, reading and deletion interfaces for the upper layer developers. Developers can call the interfaces provided by the platform to complete the storage, reading, modification and deletion of data The main implementation interfaces include:

1. Performput (string nativepath) / L storage interface. The parameter is the local storage path;
2. Performget (string nativepath, string remoteobjectname) / / read interface. The parameters are the data storage path and the ID of the read data;
3. Performmodify (string objectname) / / modify the data interface. The parameter is the ID of the modified data;
4. Performdelete (string objectname) / / delete the data interface. The parameter is the ID of the deleted data.

HTTP server: this server completes the parsing of data packets transmitted between the upper layer developers and the back-end storage node. The interfaces involved include get, put, post and delete. The get interface encapsulates the back-end data and returns it to the upper layer developers. The put interface stores the front-end data in the back-end storage node. The post interface is based on the needs of the front-end developers It is required to modify the data stored in the back-end. The delete interface is responsible for data deletion.

The system realizes the classified storage of all data and stores the structured data in the database. The design of each data table is shown in Fig. 1.

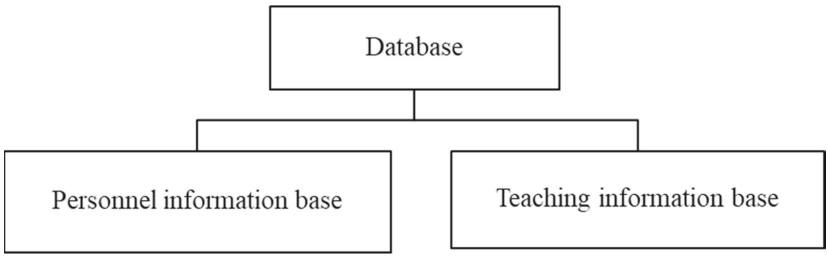

Fig. 1. Database implementation

The administrator first imports the question bank into the system and establishes the index of the question bank [4]. When a new question is added to the question bank, only one index record needs to be added; According to the item bank index, each item tester needs to save a mapping table. Mapping the item to the electronic item bank can complete the recording of an item; For students' answers, judge the correctness of the answers based on the question bank index and the test paper mapping table. For correct answers, they are mapped to the standard answer location of the question record by default. Only the wrong answers are stored, and the mapping between the storage location of the wrong answers and the students' test paper can be completed.

5 Conclusion

This paper deeply studies the problems of massive data storage and resource security sharing, and puts forward the corresponding solutions combined with the current popular technologies. On this basis, an education management system based on cloud storage is developed and implemented. Research and propose solutions to the problem of massive data storage. The cloud storage technology and data deduplication technology are deeply studied respectively. According to the advantages and disadvantages of each data layout algorithm in cloud storage technology, an improved data layout algorithm is proposed to better meet the performance requirements of fairness and adaptability in data layout algorithm at the same time. Research and propose solutions to the problem of resource security sharing. Based on the study of various resource sharing methods in Colleges and universities, a resource sharing strategy based on college resource space and role is proposed and implemented.

References

1. Yijie, W., Weidong, S., Song, Z., Xiaoqiang, P., Yong, L.: Key technologies of distributed storage in cloud computing environment. J. Softw. **23**(4), 926–986 (2012)
2. Yanxia, G., Jun, Y.: Research on massive data storage mode. Comput. Digital Eng. **36** (11), 153–169 (2008)
3. Junzhou, L., Jiahui, J., Aibo, S., Dongfang.: Cloud Computing: architecture and key technologies. J. Commun. **32**(7), 3–21 (2011)
4. Dean, J., Ghemawat, S.: MapReduce: a flexible data processing tool. Commun. ACM **53**(1), 72–77 (2010)