

The Research on the Implementation of a Cloud Computing System Architecture

Wuguang Han ^a, Haoyou Peng ^b, Jian Yao ^c

¹Hainan College of Software Technology, Qionghai 571400, China.

^a77910772@qq.com, ^b53845545@qq.com, ^c2424735736@qq.com

Keywords: cloud computing, partition mapping, parallel storage.

Abstract. In this paper, through the study of cloud computing system, it designs the cloud computing system architecture, and establishes a relational cloud data model. It uses partition mapping to improve the efficiency of cloud data query. At the same time, the parallel storage and data virtualization technology is used to improve the file transfer rate, overcomes the load imbalance problem of the server in the cloud computing system. It realizes the load balancing, and improves the reliability and usability of the cloud computing system.

1. Introduction

Since the concept of cloud computing had put forward formally by IBM in 2007, the definitions of cloud computing have hundreds. The definition of the cloud technology that put forward in 2011 is a dynamic and easy expansion of virtualization software and data resources that provides the user to calculate on the Internet. Like the power grid, the user do not need to understand the details of the cloud technology, and do not need to manage the infrastructure that supports the cloud computing. Therefore, the cloud computing is the applications of various forms of service and data centers that include the software and hardware facilities on the Internet. Cloud computing is the improvement system of the Distributed Computing, Parallel Computing, Grid Computing and distributed database. It is a cloud including the hardware cloud, software cloud and service cloud, and also an overall structure that is full virtual and transparent from the top to the bottom in the cloud [1-3]. In a word, the cloud computing is a convenient way and a service mode to use, and provides a resource pool model that includes networks, servers, storage, applications and services over the Internet. It uses the smallest management work to provide the fastest service to users. This paper mainly studies the cloud computing system architecture and the implementation of the cloud data access methods, and achieves the prototype of cloud computing.

2. The Design of Cloud Computing System

The cloud computing system architecture is mainly composed of the backstage of the cloud computing system and the cloud client. The backstage of the cloud computing system mainly includes the primary server, the monitoring manager, the cluster of the storage node and the cluster of the computing node. The primary server is responsible for responding the client requests, validating legitimate users, load balancing, and categorizing requests. The monitoring manager main monitors the primary server and schedules the storage nodes. The primary service selects the idle storage node according to the monitoring manager, and then achieves a balance of responsibility. The cluster of the storage node is consisted of a mass storage system or a large disk array system that is responsible for the storage of cloud data and the virtualization of cloud data [4]. The cluster of the computing node is mainly responsible for the calculation of a large number of requirements, similar to the functions of the operator. The cloud client mainly provides a virtualized, transparent desktop similarly to the network operating system, logs in at anywhere from the cloud desktop and customizes dynamically and flexibly personalized service on the network connecting to the Internet [5]. The cloud computing system architecture is shown in Figure 1:

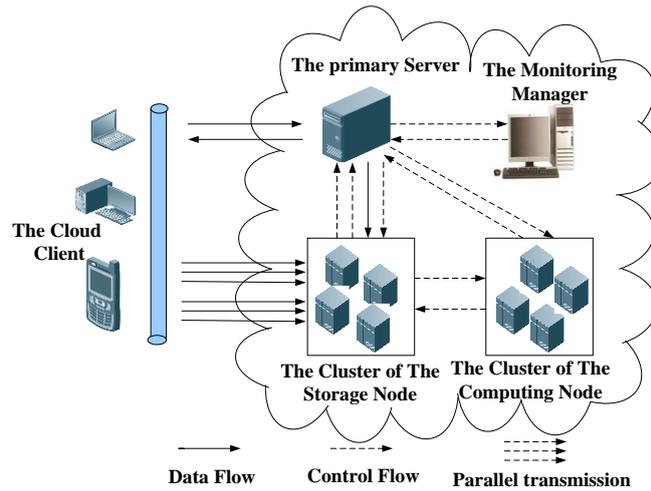


Fig. 1 The architecture of the cloud system

In the cloud computing system architecture, it uses three computers to simulate the cluster of the storage node, uses a computer to simulate the primary server and the monitoring manager, uses a computer to simulate the cluster of the computing node and a computer to simulate cloud client.

3. The Implementation of Cloud System

3.1 Partitioned Search.

In this paper, the model of cloud data is established by using the relationship model, and the relational model is established according to the logical relationship. The data model is mainly composed of the table group and the row group. A table is a logical relationship that contains a partitioning key for partitioning. A set of multiple tables with the same partition key is called the table group. A set of rows with the same keys of the partition is called row group, the rows in a row group are always assigned to a data node. Each table group may contain multiple row groups, which are assigned to different data nodes. A data partition contains multiple rows, so each node is stored in a partition that contains all rows.

As shown in Figure 2, a table group contains two related tables, which called Table 1 and table 2 that contain the same partition key. Each partition key column is composed of multiple partition keys, and the partitioned columns in the two tables have the same data type, both of them are the ID type. The ID columns in Table 1 and Table 2 have the primary foreign key association. As a result, the 2 row column value ID of 6 in table 2 and the 1 row column value ID of 6 in Table 1 are a row group, it can be seen from the figure 2, the table group is divided into multiple partitions, and the data in the same row group are located in a partition.

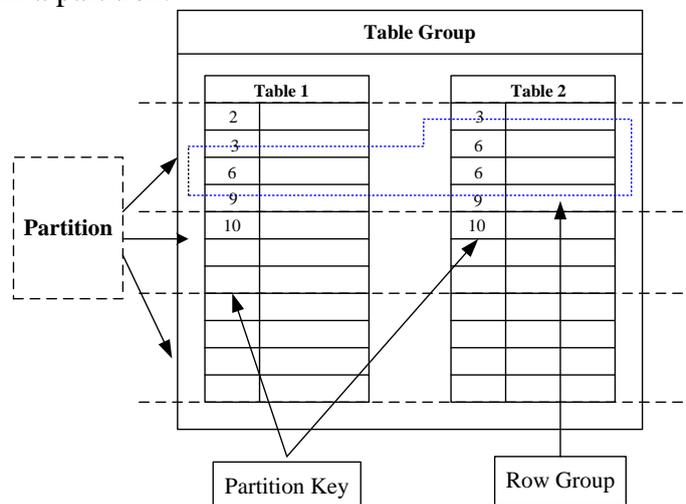


Fig. 2 Data model of relational cloud database

A table group usually contains a number of tables that will be used by an application at the same time. Therefore, it is needed to store these tables together, so as to increase the query efficiency. For example, there are two tables of the Student Info and Score info. The former records the information that contains the name, gender, age, telephone, etc. The latter records the information about the scores of each subject. When teachers inquire the score of students, it is usually need to inquire the student name, grade, class, subjects, grades, test time and other information. This requires a join between the two tables of the student info and Score info. If the two tables are stored in the same data node, as shown in Figure 2, the data stored in the row group which the partition key's ID is six. The query will be faster through the connection operation, and it improves the efficiency of data query. Therefore, the multiple rows in the same row group are stored in a data partition, which is associated with the primary key and foreign key, and it improves the efficiency of query.

As shown in Figure 3, when the client requests data by using the partition mapping search, it first requests a partition map from the cloud data manager, and the manager sends a partition map to the client. After that the client receives, it searches in the graph according to the partition key to find their desired data storage location. And then the client requests data from the specified the node of data storage. In the end, the data is returned to the client from the node of the data storage. In fact, in order to improve the performance and avoid the bottleneck of the performance in the cloud data manager, the client will usually use the common partition map in its cache, so that the client does not have to interact with the cloud data manager to access the corresponding data nodes directly in many cases.

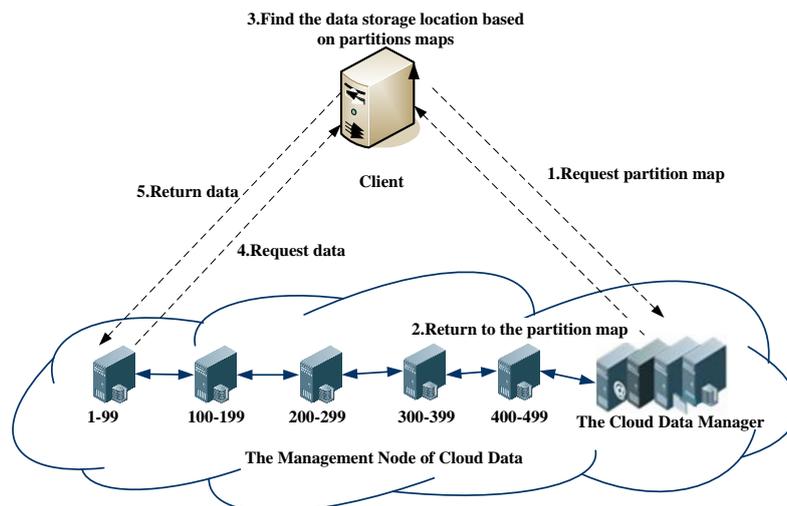


Fig. 3 Partition mapping search

3.2 The Implementation of Parallel Storage.

In order to solve the problem of load imbalance and transmission bottleneck of the data storage, the access method uses parallel uploading and parallel downloading of the cloud storage technology, it makes full use of the cluster of the backstage service and each node. The server accords to the file information for dicing calculation, allocates the file block to each available node and writes file information and dicing information to the database. After that the server will put block information (file number, block size, etc.) and the corresponding node address to the client. When the client receives a return message, it will cut the file and parallel upload the file block to the corresponding storage node. When download a file, the server first finds the block of information from the database, then the file blocks are downloaded to the computer from each storage node in parallel. In the end, the file blocks are integrated into a complete file and delete the file blocks. Because each storage node is a parallel transmission, the load balance of the server is realized [6], and the transmission rate is increased.

Parallel upload process:

(1)The client to send upload request: the client upload the file information (such as file size, file name and other information) to the main server.

(2)The server dicing calculation and the monitoring system allocation node: The primary server divides the received files into N blocks. In addition, the monitoring system allocates the storage blocks

according to the state of the storage nodes, allocates the data blocks in the database according to the partition mapping mode.

(3) The server response request: The server sends the file dicing information and the allocation node information to the client;

(4) Parallel upload file: The client uploads the file blocks to the corresponding storage node in parallel according to the allocation storage node;

(5) Upload Confirmation: After the storage node successfully receives the file blocks, it sends a confirmation message to the server. When all the file blocks are uploaded successfully, the server database changes the file status to 1;

(6) Single node upload failed: When a file block upload fails in a node, it immediately retransmitted.

Parallel download process:

(1) Download request: select the download file and send a download request.

(2) Find file blocks: Find the storage node and find file blocks information according to the partition mapping.

(3) Response to download request: The server sends the file blocks information to the client.

(4) Parallel download file blocks: The client selects the storage node according to the file blocks information, establishes the process, and downloads them to the local temporary file.

(5) Download success: The client integrates the download file blocks into a complete file, and deletes the file blocks.

(6) Single node download failed: The monitoring information found that the download failed, it immediately reassigned a new node and reloaded;

3.3 Data Virtualization.

The cloud data is deployed in the cluster of the storage node and the data is virtualized [6-8]. A database will be scattered in different disks, the cloud data management node is responsible for the work that each disk data is rewritten and read. Cloud data management node accords the load situation to reduce or increase the number of disks to adapt to the user's access, which can protect the efficient access the data. Different storage node cluster send their own use of information regularly to the server for ensuring the effective use of the cluster of the storage node and Load balancing. Therefore, through data virtualization technology, it greatly improves the storage capacity of the database, it makes the using and upgrades of the software and hardware becomes easier [4]. As shown in Figure 4, the customers do not need to understand the underlying details of cloud data in cloud data applications. All the bottom of the cloud data is transparent to the customer who uses these databases as easy and simple as using a single database server, while gaining almost unlimited storage and processing capacity.

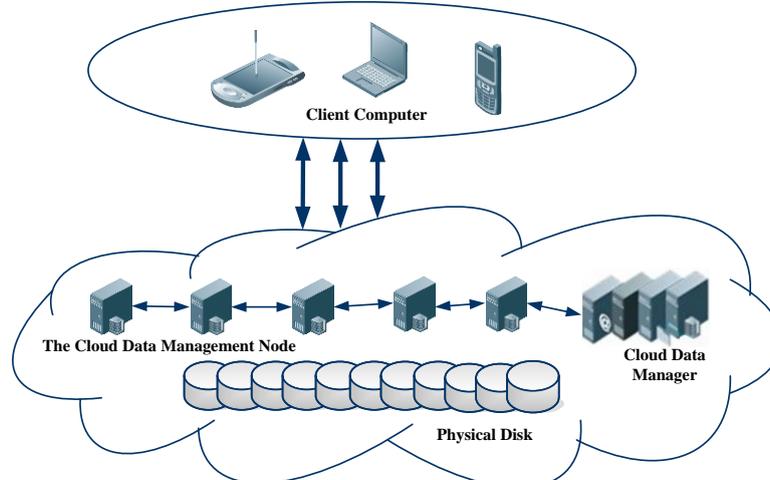


Fig. 4 Data virtualization

4. Summary

This paper mainly studies the cloud computing system, designs the architecture model of the cloud computing system, and builds the simulation platform of the cloud computing system. This paper focuses on the realization of cloud data in the cloud computing system, it uses the relational data and partition mapping search method to establish cloud storage data model in the implementation of the cloud data, it improves the query efficiency of cloud data by partition mapping search. At the same time, it uses the parallel storage methods and data virtualization technology to increase the cloud transmission rate and achieve a load balancing so that it improves the reliability and effectiveness of cloud computing systems.

This work was supported by Hainan Province Natural Science Fund Project (617166).

References

- [1]. Lu qian. Constructing Smart Campus Based on the Cloud Computing and the Internet of Things [J]. Computer Science, 2011.10.p.18.
- [2]. chen kang, zheng weimin. Cloud Computing: System Instances and Current Research [J]. Journal of Software, 2009, 20(5), p.1337-1348.
- [3]. Youseff L,Butrico M,Silva D D. Toward a Unified Ontology of Cloud Computing[C]. Grid Computing Environments Workshop. IEEE, USA, 2008, p.1-10.
- [4]. Li Qiao, zheng xiao. Research Survey of Cloud Computing [J]. Computer Science, 2011, 4, p.34.
- [5]. Liu zhengwei,Wen Zhongling,Zhang Haitao. Cloud Computing and Cloud Data Management Technology [J]. Journal of Computer Research and Development.2012, (49), p.26-31.
- [6]. YE Ke-Jiang, Wu Zhaohui, Jiang Xiaohong,He QinMing. Power management of virtualized cloud computing platform[J].Chinese journal of computers.2012,7,P.1263-1275
- [7]. Yoon JP. Access control and trustiness for resource management in cloud databases [M]. Grid and Cloud Database Management. Springer Berlin Heidelberg, 2011:109-131.
- [8]. Barham P, Dragovic B,Fraser K. Xen and the art of virtualization[C].proceedings of the 19th ACM Symposium on Operating Systems Principles(SOSP'03).New York,USA,2003,p.164-177.