

Analysis of Data Pooling and Optimization of Storage Management in Cloud

SANDHYA B.L¹, MAMATHA G.S²,

¹ M.Tech, Dept. Of Information Science, R.V College of Engineering, Bangalore, India-560059

² Assistant Professor, Dept. Of Information Science, R.V College of Engineering, Bangalore, India-560059

Abstract:- This paper focus on analysing and explaining structure of data pooling, online data pooling and partition Storage model, enterprise Cloud Storage System structure proposes the design for data control and Storage in Cloud. Storage Management Control is an effective method that will reduce the working time to large-scale in data Storage management. Combining both Storage devices and control management software will provide system data sharing and system high applicable. By using Cloud Storage Management control mechanism, most of those business enterprises could be benefited.

Keywords:- Data pooling, Storage Management, partition

I. INTRODUCTION

Nowadays every business enterprise facing the problem of growing endless business data which requires more Storage space to support the activity that can be done continuously, so, the cost of the data storing and managing is also high. For example, if storing a communication record needs 60B Storage spaces, then to store data of one hour needs 60 MB Storage at least. Besides, large-scale data and the conventional maintenance task can increase the operation cost with respect to space and time. This will lead to slow down the speed of system responding to query request of the user and user's incapability to regularly visit the data, So there is a need to build up the network storage management control specifically for large scale data and to realize high-performance administration and high applicability.

In GSM system, data management of VLR is an important factor that affects network load and performance of system. The data management of VLR consists of how data is organized and part data. As users are added, Mobility of the users is improved unceasingly; the visiting load of network database will be increased. Facing with more and more data, system must optimize Storage devices configuration, extend Storage devices, provide the history data applicability, and realize that physical Storage resource corresponds with data logical resource source, improve data query performance, and enhance trouble recovery ability.

II. ENTERPRISES CLOUD STORAGE SYSTEM

A. Cloud Storage

Cloud Storage is an aggregation that has two parts: Storage devices and servers. Using Cloud Storage System, the user will gain data visiting service. The Cloud Storage consists of application software and Storage devices. Creating application software can make you to realize that the changeover from Storage devices to Storage service. Cloud Storage is depending on Storage devices. It works with cluster function and distributed file system, and provides certain Storage service and visiting service for users, running certain Application software or using Application interface. Any users with permission can link with cloud Storage.

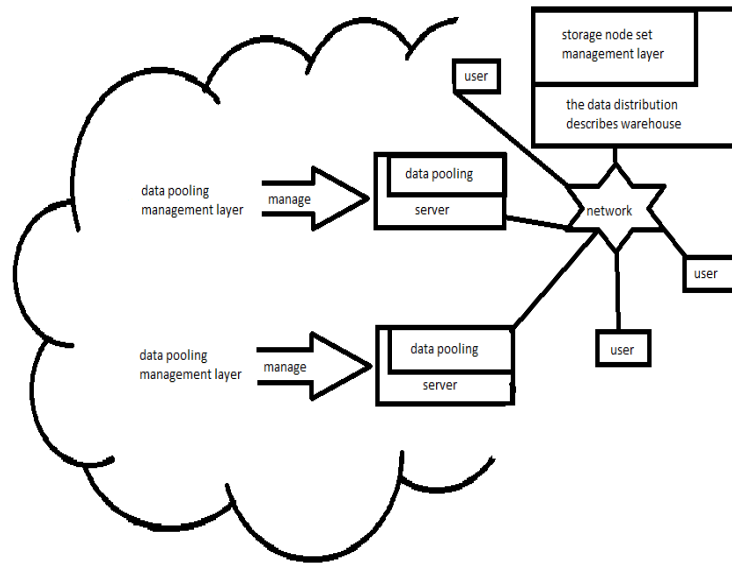


Figure 1. Structure model of enterprise Cloud Storage System

Cloud Storage correlate with data Storage nodes spread over a wide geographic area. Storage node consists of Storage devices and data pooling. Cloud Storage System manages and handles co-ordinately dispersed Storage node.

B. Structure model of enterprise Cloud Storage System

Structure model of Enterprise Cloud Storage System Structure model of an Enterprise Cloud Storage System consists of Storage node (Storage devices, data pooling, data pooling management layer), Storage node set, Storage nodes set management layer, application interface layer. Please see Fig. 1.

Storage node is the basic part of Cloud Storage, it has three parts: Storage devices, data pooling, data pooling management layer. Storage Devices can select IP Storage Devices, FC Storage Devices, DAS Storage Devices, and so on. Data pooling determines management layer monitor, analysis and the automation carrying out the tasks that is distributed data Storage and the call configuration.

Data pooling is a data set that consists of backup pooling, history data pooling, data warehouse, online data pooling. With various data pooling managed by data pooling management layer, the data saved in data pooling can be migrated dynamically, monitored and controlled, in order to prevent the data in pool from running off, realize the more accessibility, integrity and privacy of the data.

Storage nodes spreading over a wide area are connected via network, which is a Storage node set. Storage node management layer is a software platform that manages Storage node set. Via cluster, distributed file system and grid computing technology, we must develop management software to manage Cloud Storage nodes, to realize that many Storage nodes work co-ordinately in Cloud Storage, to observe and maintain the state of Storage node, to realize trouble warning and the automatic maintenance, to make many Storage nodes provide stronger data visiting performance for users. Via various data backup technologies, we shall prevent the data stored in Cloud Storage from losing, in order to guarantee security and stable-stability of Cloud Storage. The data distribution describes warehouse records the position of data saved in Storage node.

Application interface layer is an application service interface; any user can gain data from Cloud Storage System, via public application interface. According to Service demands of different enterprise, we configure different visiting type of Cloud Storage and visiting method. In application interface layer, we must define the process of authorization management, in order to assure the security of data stored in Cloud Storage System.

III. DATA POOLING STRUCTURE

In enterprise Cloud Storage System, data Storage node consists of data pooling and data pooling Storage management layer. Data pooling has five parts: online data pooling, backup pooling, history data pooling, data warehouse, data table model. The main task of data pooling Storage management layer is to manage various data , for example, auto configuration of Storage capacity, auto-update of data restored in data pooling, adding database copy, to realize auto maintenance of data Storage node. Data pooling flow diagram see Fig. 2. If we can optimize Storage source management technology, dynamically balance databases visiting load, reduce the cost of enterprise Storage management, then enterprise will gain higher commerce benefit.

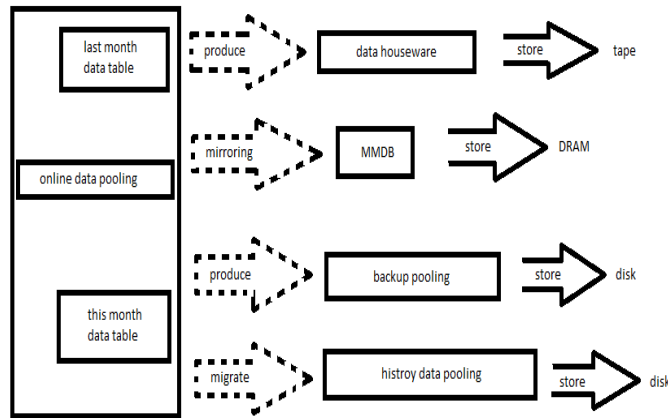


Figure 2. Data pooling structure

A. Online data pooling

Mobile user’s information (business data, site data etc.) is restored in Central database HLR of mobile Communication system, VLR is a copy of HLR, and database capacity is less, therefore, we can restore VLR in the DRAM, which forms MMDB. MMDB has removed the I/O bottleneck. Online data pooling that has all data that are restored in MMDB is the mirroring of MMDB. Online data pooling consists of online this month data table and last month data table, is the full copy of MMDB. When MMDB is lost, we can recover quickly it from online data pooling and guarantee the system high applicability.

Besides, system must control the data capacity of online data pooling, via migrating history data. We must also control the data inputting and outputting pool.

In the process of data migration, the history data pooling is formed. In order to maintain automatically the migration of history data, data must be parted by time, in order to sufficiently assure that logical partition bring into relation with physical Storage partition.

B. Online data pooling partition storage model

Data Storage model is an important factor that affects network load and system performance, in which the important questions are: How should data be organized for control? How should data be parted? Relational database (Oracle, SQL Server etc.) can be used at the lowest layer database platform of system. By using partition technology of database, partition table store in physical Storage. According to date time type, data table is parted logically and levelly, every physical Storage area stores data logically, according to Storage scheme. The Partition Storage model. See Fig. 3

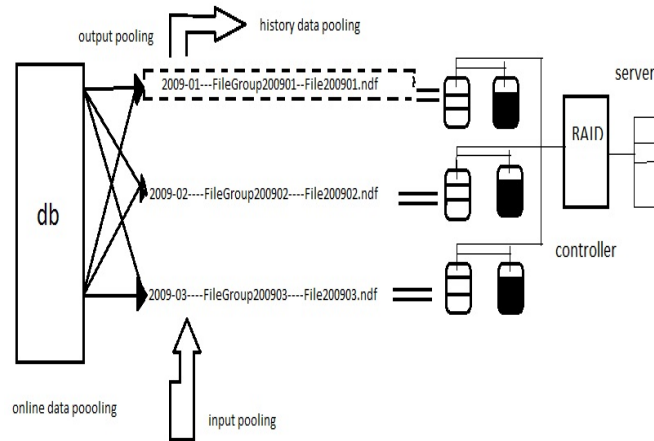


Figure 3. The Partition Storage model

For example, according to communication time, the data related by communication record is parted levelly. Every logical partition has a corresponding data table, with every data table stores in different physical Storage site. The physical Storage site of logical partition data binds data source, so every monthly data stored in the single-handed different disk.

At the beginning of every month, previous month data, is migrated to history data pooling according to data output pooling maintenance plan, and prepare Storage spaces for input pooling last month data. The users

have no need to know about output and input maintenance task. According to the time of query data, system can judge data that the user will query store in online data pooling or history data pooling. Data Input pooling algorithm description:

If (day (get date ()) = 1)

(To prepare Storage spaces for next month; System automatically produces file group of the next month and secondary data file that belong to next month file group; To add partition boundary values, bind partition plan with next month file group; Via data table model, automatically produces next month data table set).

Data Output pooling algorithm description:

If (day (get date ()) = 1)

(In history database, to prepare Storage spaces; previous month data is migrated to history database; Delete previous month data in online data pooling; Merge partition boundary values).

C. History data pooling

History data pooling stores history data. In order to enhance the usage of the Storage devices, reduce the cost of Storage, data often is migrated to history database, which will reduce the history data to affect system performance.

D. Data warehouse and backup pooling

At present, database system all provide the function to create data snapshot, data snapshot store data static copy set read only, According to certain time gap, to create data snapshot set, and form the data warehouse. Collecting the history data from data warehouse, providing decision-making for future Storage development of enterprise. It is easier to decide when the number Storage capacity is increased.

Backup pooling store log backup set of online data, full back up set and differential backup set.

E. The strategy of monitoring and controlling performance

With time fading away, the Storage capacity of online Data, backup and history data pooling, data warehouse began the snowball effect, the Storage space of data pooling is insufficient will lead that the data are lost, or preserved nowhere. By means of monitoring and controlling the Storage capacity of data pooling, when Storage capacity approaches threshold value of Storage capacity, carrying out the suitable strategy of monitoring and controlling performance can timely automatically deploy Storage capacity.

The first strategy: Define the age of history data and backup file, migrate the data which age exceeds the time to tape, and accomplish a data saved offline.

Second strategy: Define threshold value of Storage capacity, set up performance alert when Storage capacity approaches threshold value of Storage capacity, automatically migrate backup file and history data; call Storage Administrator, add new Storage devices in order to provide more Storage space for new data and ensure the system applicability.

Third strategy: Define Storage performance standard, monitor and control server performance when server performance reduces obviously, call Administrator, analyse the system configurations and re-configure system Storage space.

IV. CONCLUSIONS

The paper mainly concentrates on analysing and discussing about the data pooling structure, online data pooling partition Storage model and the structure model of enterprise Cloud Storage System, it proposes the design plan of data control and Cloud Storage. The Storage management control optimizing will be an effective method that can reduce the working time to large-scale data Storage management. The Combination of Storage devices and control management software will be able to provide system data sharing and system high applicability. And with Cloud Storage Management control mechanism being applied, so these business enterprises could be benefited from this.

REFERENCES

- [1] Han J, Kamber M., Data Mining: Concepts and Techniques [D]. Simon Fraser University, 2000.
- [2] KOU Wei-hua, ZHENG Feng-bin, Baton Model for Distributed Database Based on Distributed Application Techniques [J]. Sichuan: Computer Applications, 2003 (12)
- [3] YI Fa-ling, File System Design on Volume Holographic Storage [J].Sichuan: Computer Applications, 2006(08)
- [4] Microsoft, SQL Server 2005 Database Development [M]. Higher Education Press, 2007(9)
- [5] Zuidweg J., Next Generation Intelligent Networks (1st edition) [M]. Artech House, 2002