Research on Load Balance Strategy Based on Grey Prediction Theory in Cloud Storage

Yu Mao[a], Jie Ling[b*]

Faculty of Computer, Guangdong University of Technology, Guangzhou, China

[a]maodiablo@163.com , [b]jsjgdut1628@21cn.com

Keywords: cloud storage; replica strategy; load balance; grey prediction theory

Abstract: In the HDFS cloud storage environment, the existing replica strategy is not able to adopt the optimal replica according to the current system environment dynamically, which affects the performance of the Cloud Storage system. This paper aims at the replica selection strategies in the load balancing problem, proposes a Replica Strategy based on grey prediction theory in Cloud Storage and predicts the node load rate of different replicas at a moment, selecting the minimum load node dynamically, eventually to load balance.

Introduction

In recent years, along with the increasing popularity and high speed development of the Internet, together with the wide application of multimedia, the data run on different system platforms exponentially grows. According to IDC survey statistics, in 2011, the creation and duplication data quantity in global is 1.8ZB (1ZB=10^{21}B), 11 times that of global total data quantity in 2006, 161EB (1EB=10^{18}B), and the speed of data growth continues increase. Cloud Storage is a system that groups a wide variety of different types of storage equipment in network to work together collaboratively through application software, through functions such as the cluster application, the grid technology or distributed file system, commonly providing data storage and business access functions externally. The advantages of Cloud Storage is mass storage, high performance, low cost, expansibility, not affected by specific geographic position limitation, based on commercial components, charging according to use, crossing different applications, and so on. Therefore, Cloud Storage system becomes the effective solution of EB level mass data storage problem. Some current world IT giant all has launched cloud storage services, such as Google's GFS [1], Apache's HDFS and IBM's Blue Cloud [2], etc.

In order to realizing the high expansibility of system, to ensuring high availability and reliability of data, the data of Cloud Storage are generally stored and managed in distributed way. However, the reasons such as unpredictable hardware equipment damage, power supply interrupt, hacker attacks, virus invasion, fire, earthquake, and terrorist attacks, are all likely to cause the damage and loss of data, and thus affect the normal data access. So Cloud Storage system use redundant storage to achieve system fault tolerance purpose, and improve system reliability. Redundant storage refers to preparing a lot of data replicas on the same data block, reading the optimal replica from lots of data backups when reading the data block, so as to ensure efficient and stable operation of the system. Among them, replica selection strategy is one of the core technologies in cloud storage system of replica management. A good replica selection strategy can give full play to best performance of cloud storage system, rising consumers' user experience, reducing the operation cost of service providers, and achieve a win-win situation.

HDFS Cloud Storage Replica Strategy

The number and place of replica

HDFS is a master-slave structure system. HDFS cluster is organized by NameNode and

* Corresponding author

DataNode. NameNode is namespace of management files and the main server of adjusting client access files, while DataNode is used to store data. The data in HDFS can only have one write in and more assessors at any moment, it is a file system of write-one-read-many access model. Each file is segmented for Blocks which have fixed size, 64MB, and then record by numbers. Each data blocks are scattered in different DataNode, and at the same time, HDFS make a redundant backup for each block, and copy several replicas [3]. The set of replica number is directly related to system environment and the access heat of file itself. When the file access users are too much, increasing the number of replicas can alleviate system pressure and improve the data global efficiency, on the other hand, when file access users are too less, too many replicas will increase system load and reduce storage equipment utilization rate. By default, the number of replicas of each data block is 3(including the source data block). HDFS replica placement strategy refers to storing replicas in DataNode on local frame, another DataNode on local frame and DataNode on different frames. This strategy improves system reliability and balance system load. Fig.1 shows the distribution in DataNode of 3 replicas of one data block.

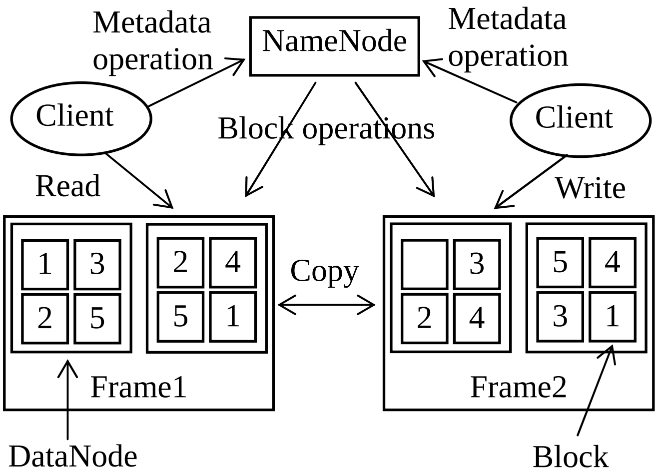

Fig.1. The distribution in DataNode of 3 replicas

The Selection of Replica

The selection of replicas is the core of HDFS cloud storage replica strategy, which directly affects the bandwidth utilization and the user access rate of the entire system. In current HDFS, by default, the client selects the strategy that which replica has the shortest topological distance to it [4]. The bandwidth between the nodes decides the data transfer rate. We can know from the network characteristics of HDFS that the bandwidth between nodes can be approximately expressed as the topology distance between nodes. The farther the distance is, the lower the bandwidth is, oppositely, the closer the distance is, the higher the bandwidth is. However, this selection strategy has obvious flaws. It can cause problems such as the high load of nodes, leading to network congestion and performance reduce, when a lot of optimal replicas are placed on the same frame or DataNode.

The Adjustment of Replica

Cloud storage system runs on ordinary hardware equipment, so it is easy to appear unbalanced disk utilization situation when the storage capacities of different nodes are quite different, like node which has big storage capacity has low storage rate, whereas node which has small storage capacity is already saturated. HDFS will adjust replicas dynamically after finishing creating them. If the available space of a node below specific preset value, system will transfer part of the data of this node to relative free node, in order to achieving a balanced load purpose. If the access numbers of a file increased dramatically, system will copy more replicas to scatter in the cluster to meet the requirements.

Research on Load Balance Strategy Based on Grey Prediction Theory in Cloud Storage

In order to finding suitable replica selection strategy, the main factors which affect replica selection should be analyzed firstly[5]: (1) Read speed of disk I/O, faster read speed can reduce access duration; (2) Load situation of storage nodes, load problem is highlighted when multiple tasks access the replica on the same data node at the same time; (3) replica response time, the uncertainty of this factor is very strong, so it is hard to determine the optimal solution, can only approximately select according to the real-time situation; (4) Bandwidth between request node and data node, the transfer rate of data depends on bandwidth. The higher bandwidth will reduce data transmission time. The shorter the time is, the less the risk is that happened in the process of transmission, such as line reliability and delay. In this paper, the research of data replica selection strategy will mainly consider the load factors of storage nodes.

The number of replicas is marked with k, any node is N_i and all the optional replicas of one block can be expressed as a node set $N=\{N_1,N_2,N_3,...,N_k\}$, for any node N_i at the t moment, the number of total data blocks the node has is $T_{N_i}(t)$; the number of data blocks in ongoing tasks is $P_{N_i}(t)$; bandwidth of node bus is $BW_{N_i}(t)$; the average access speed of ongoing tasks is $AS_{N_i}(t)$;used disk space is $US_{N_i}(t)$;total disk space is $TS_{N_i}(t)$.For nodes load, it is divided into task load, bandwidth load and storage load, whose corresponding weights are P_1, P_2, P_3 ;and the load rate of N_i at t moment can be expressed as:

$$L_{N_i}(t)=P_1\times\frac{P_{N_i}(t)}{T_{N_i}(t)}+P_2\times\frac{AS_{N_i}(t)\times P_{N_i}(t)}{BW_{N_i}(t)}+P_3\times\frac{US_{N_i}(t)}{TS_{N_i}(t)} \tag{1}$$

Where: $P_1+P_2+P_3=1$;

Since cloud storage system runs on ordinary hardware, the newly added machine has strong processing ability and large capacity, so the value of P1/P2 is smaller than the older node when determine the weights.

The nodes locations of the entire optional data replicas the users need can be found through the data replica positioning of NameNode. The load rate of node at t+1 moment can be predicted through the record of history information, select the node which has lowest load rate as the one, that is optimal replica.

The following is the prediction process of load rate based on GM (1, 1) model in Grey Prediction Theory [6]:

1) Each node read previous t moments, number of total data blocks, number of data blocks in ongoing tasks, bandwidth of node bus, the average access speed of ongoing tasks ,used disk space and total disk space, according to system logs, then it can work out the load rate of previous t moments according to formula (1). The generated data sequence is as the original sequence of prediction model:

$$L_{N_i}^{(0)}=\{L_{N_i}^{(0)}(1),L_{N_i}^{(0)}(2),...,L_{N_i}^{(0)}(t)\} \tag{2}$$

2) Accumulate the random data sequence of formula (2) successively, weaken its randomness, to get a regular generation number:

$$L_{N_i}^{(1)}(1)=L_{N_i}^{(0)}(1)$$

$$L_{N_i}^{(1)}(2)=L_{N_i}^{(0)}(1)+L_{N_i}^{(0)}(2)$$

......

$$L_{N_i}^{(1)}(t)=\sum_{i=1}^{t}L_{N_i}^{(0)}(i)$$

$$L_{N_i}^{(1)}=\{L_{N_i}^{(1)}(1),L_{N_i}^{(1)}(2),...,L_{N_i}^{(1)}(t)\} \tag{3}$$

3) After accumulating process, the new generated $L_{N_i}^{(1)}$ has characteristics of stability enhancement and volatility reducing, compared to $L_{N_i}^{(0)}$. Build first order differential equation based on GM (1, 1) to formula (3):

$$\frac{dL_{N_i}^{(1)}(t)}{dt}+aL_{N_i}^{(1)}(t)=b \tag{4}$$

Where: a is development coefficient, b is for grey action quantity

Use the least square method to fitted getting the estimated parameters:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (A^T A)^{-1} A^T B_t$$

In the formula: $$A = \begin{bmatrix} -\frac{1}{2}[L_{N_i}^{(1)}(1) + L_{N_i}^{(1)}(2)] & 1 \\ -\frac{1}{2}[L_{N_i}^{(1)}(2) + L_{N_i}^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[L_{N_i}^{(1)}(t-1) + L_{N_i}^{(1)}(t)] & 1 \end{bmatrix} \tag{5}$$

$$B_t = [L_{N_i}^{(0)}(2), L_{N_i}^{(0)}(3), \ldots, L_{N_i}^{(0)}(t)]^T \tag{6}$$

Work out: $$\hat{L}_{N_i}^{(1)}(t) = (L_{N_i}^{(0)}(1) - \frac{\hat{b}}{\hat{a}}) \times e^{-\hat{a}(t-1)} + \frac{\hat{b}}{\hat{a}} \tag{7}$$

The load rate prediction sequence can be obtained through formula (2):

$$\hat{L}_{N_i}^{(0)}(1) = L_{N_i}^{(0)}(1)$$

$$\hat{L}_{N_i}^{(0)}(2) = \hat{L}_{N_i}^{(1)}(2) - \hat{L}_{N_i}^{(1)}(1)$$

......

$$\hat{L}_{N_i}^{(0)}(t) = \hat{L}_{N_i}^{(1)}(t) - \hat{L}_{N_i}^{(1)}(t-1)$$

$$\hat{L}_{N_i}^{(0)} = \{\hat{L}_{N_i}^{(0)}(1), \hat{L}_{N_i}^{(0)}(2), \ldots, \hat{L}_{N_i}^{(0)}(t)\} \tag{8}$$

4) Accumulate $\hat{L}_{N_i}^{(0)}$ successively can obtain:

$$\hat{L}_{N_i}^{(0)}(t) = (L_{N_i}^{(0)}(1) - \frac{\hat{b}}{\hat{a}}) \times (1 - e^{\hat{a}}) \times e^{-\hat{a}(t-1)} \tag{9}$$

According to formula (9), the load rate of node N_i that contains optional replica is:

$$\hat{L}_{N_i}^{(0)}(t+1) = (L_{N_i}^{(0)}(1) - \frac{\hat{b}}{\hat{a}}) \times (1 - e^{\hat{a}}) \times e^{-\hat{a}t} \tag{10}$$

According to formula (10), we can work out the load rate prediction sequence that k nodes correspond to at the t+1 moment:

$$\{L_{N_1}(t+1), L_{N_2}(t+1), L_{N_3}(t+1), \ldots, L_{N_k}(t+1)\} \tag{11}$$

The clients can select the nodes with lowest load rates to read replicas. If access fails, select the node with load rate that only second to the previous one until success. As time goes on, the load rates of nodes contain different replicas of the same data block constantly change, always select the one with lowest load rate dynamically, which make the possibility of the large access to the same node on the same moment is greatly reduced, realizing the dynamic balance of load efficiently.

Summary

Cloud Storage cluster has characteristics of big differences between hardware, therefore, how to reasonably select the optimal replica has big importance on load of each storage node. This paper proposes to use grey prediction theory to do the real-time prediction for load rate of each node ,realize load balance of each node dynamically, so as to ensure the cloud storage system maintain efficient and stable operation.

Acknowledgement

This work is supported by the Industry-Education-Research Cooperation Project of Guangdong Province and the Ministry of Education (No.2011A090200068), and the Project of National Science and Technology Support Program (No.2012BAH27F05).

References

[1] GFS(Google File System)[EB/OL].http://www.codechina.org/doc/Google/gfs-paper/2008. 9.

[2] Blue Cloud, IBM. http://www.ibm.com/cloud-computing/us/en/.

[3] XU Xiao-yong, PAN Yu, DING Yan-yan: *Strategy of replica selection based on grey Markov chain prediction model in HDFS*, Journal of Computer Applications, 2010, s2:39-42.

[4] WHITE T. *Hadoop: The Definitive Guide*. *Version 2[M]*. Beijing: Tsinghua University Press, 2011.

[5] CHEN Lei, YANG Peng: *Application research of ant algorithm to replica selection in data grid*, Computer Engineering and Design, 2008, 23(29):6157-6160.

[6] Julong Deng: *Grey Prediction and Grey Decision [M]*. Wuhan: Huazhong University of Science and Technology Press, 2002.