Applied Mathematics & Information Sciences 5 (2) (2011), 235S-242S
– An International Journal

Study on Redundant Strategies in Peer to Peer Cloud Storage Systems

Wu Ji-yi[1], Zhang Jian-lin[1], Wang Tong[2] and Shen Qian-li[1]

[1]Key Lab of E-Business and Information Security, Hangzhou Normal University,Hangzhou 310036,PR China
[2]Information and Communication Engineering College, Harbin Engineering University, Harbin 150001,PR China
Email Address: wangtong@hrbeu.edu.cn

Received June 22, 2010; Revised March 5, 2011

Abstract: Relative to the current Master/Slave computing model cloud storage file systems, including GFS, HDFS, Sector, we proposed a general model of Peer-to-Peer based Cloud Storage file system, and constructed a prototype system named MingCloud based on Kademlia algorithm. Focused on redundant strategy selection and design of the storage system, our experiment used the Cauchy Codes as coding algorithm to assess the system from many aspects. Compared with the redundant mode of complete copy, Erasure Code model provided more satisfactory system availability, and more suitable to apply to our P2P Cloud Storage System.

Keywords: cloud computing, cloud storage, redundancy, peer-to-peer.

1 Introduction

As the latest hotspot in distributed storage nearly 3 years, there are many valuable research on cloud storage from academic circles. Robert L. Grossman from University of Illinois put forward and implemented a wide area network based high performance computing and storage cloud, Sector/Sphere [1,2], and experimental results show it's better performance than Hadoop. MetaCDN were designed and proposed by Australia University of Melbourne's James Broberg [3], to integrate cloud storage service from different service providers, and provide a unified high-performance and low-cost content distributed storage and distribution services for content creators. Kevin D. Bowers et al of RSA Laboratories [4] proposed a high reliability, completeness cloud storage model HAIL, and carried out experiments on security and efficiency. Knowledge storage model to dynamic fuse local storage and cloud storage were put forward by David Tarrant, etc [5] from University of Southampton, UK.

Relative to the current Master/Slave model cloud storage file systems, including

GFS, HDFS, Sector, we proposed a general model of Peer-to-Peer based Cloud Storage file system, and constructed a prototype system named MingCloud based on Kademlia algorithm. The results from a simulation experiment show that the system has high availability and performance, which can be actually applied to the Internet dynamic and open environment after improving and optimizing, and to provide higher quality of cloud storage service.

2 Redundant strategies in P2P Cloud Storage System

2.1. P2P Cloud Storage System

Highlighting the system security and reliability of data, our system can provide basic data storage, read, delete, search and other storage services. To adopt highly scalable P2P system architecture, the basic idea of the MingCloud is to build a high performance, high reliability, low-cost distributed cloud storage system by integrating these idle storage resources in Internet connected desktop computers. According to the scale of storage nodes number, the system need to be divided into different Domains or Clusters, configured one Master Server to store user's directory information, file index information and other meta-data, and complete domain user authentication.

With a hierarchical architecture, as shown in Fig. (2.1), MingCloud can be divided into five layers: Physical Layer, Router Layer, Data Layer, Session Layer, Application Layer, from a functional point of view.

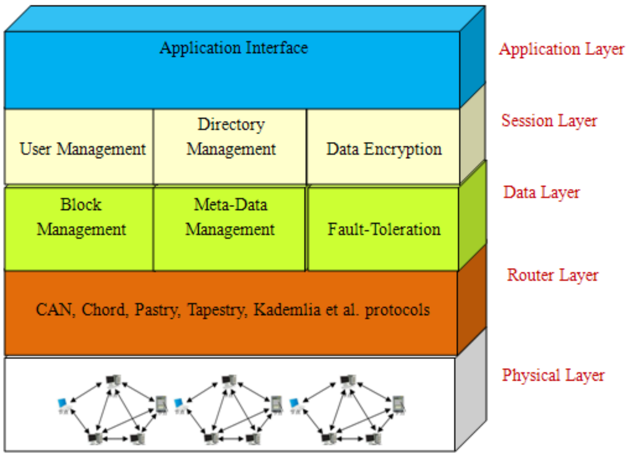

Figure 2.1: Layered architecture in MingCloud P2P cloud storage system

2.2. Redundant strategies selection

At present, there are many coding algorithms, such as Streaming Codes, Shamir Threshold Scheme, Forward Error Correction, Reed-Solomon, Vandermonde, Cauchy Codes [6] and Tornado Codes [7]. Are these algorithms suitable for P2P cloud storage system? How to choose and realize an adaptive algorithm for our MingCloud P2P cloud system, is critical.

After experimental analysis, combined with related research on encoding algorithm in paper [8], Vandermonde, Cauchy Codes and Tornado codes algorithms are more suitable for our P2P cloud storage systems.

The paper [8] also experimental proved that Tornado Codes have advantages and high efficiency compared with Cauchy Codes more early. More than 100 times in execution speed of Cauchy Codes; Tornado Codes operate and process through pair of graphs, as simple, fast and practical algorithms.

Taking into account algorithm protection by the Digital Fountain Company, our system temporarily adopt the relatively good performance Cauchy Codes as the core algorithm. In future study, our team also hopes to design a coding algorithm close to Tornado in performance.

3 Experiments and performance analysis

3.1. Experimental Environment and Purpose

Our experiments will compare the influence on system availability by using different redundancy strategy. A Java based P2P simulator PeerSim is adopted, which is run in JDK1.6 and eclipse on a PC with dual-core CPU 1.8 GHZ and 2G memory whose operation system is Windows XP.

The purpose of our experiments is to verify the correctness of Eq. (3.1) and Eq. (3.2). First, we compare the system availabilities under different k-bucket sizes, in order to choose an appropriate k value; Second, we compare the system availabilities under different block number of file for each redundant strategy, in order to show the applicable scene of each redundant strategy; At last, we compare the system availabilities under different redundant strategies with the same redundant rate, in order to compare these two redundant strategies directly.

$$A = (1-(1-p)^r)^n \tag{3.1}$$

$$A = \sum_{i=m}^{mr} \binom{mr}{i} p^i (1-p)^{mr-i} \tag{3.2}$$

3.2. Experimental results and analysis

As shown in Fig. (3.1), when the parameters value is assigned as followed: p=0.6, r=2.0, n=1, m=1, system availability increases as the size of k-buckets increase, whichever the adopted redundant strategy is. It is because some data may be much closer to a new node, so other nodes may tend to request data from this new node, which leads to reading failure. When the size of k-bucket k is little, other nodes request data from few nods, which make it more likely to fail. When $k \geq 4$, system availability tends to be stable, because there tends to be at least one effective node among these k nodes when k is large enough.

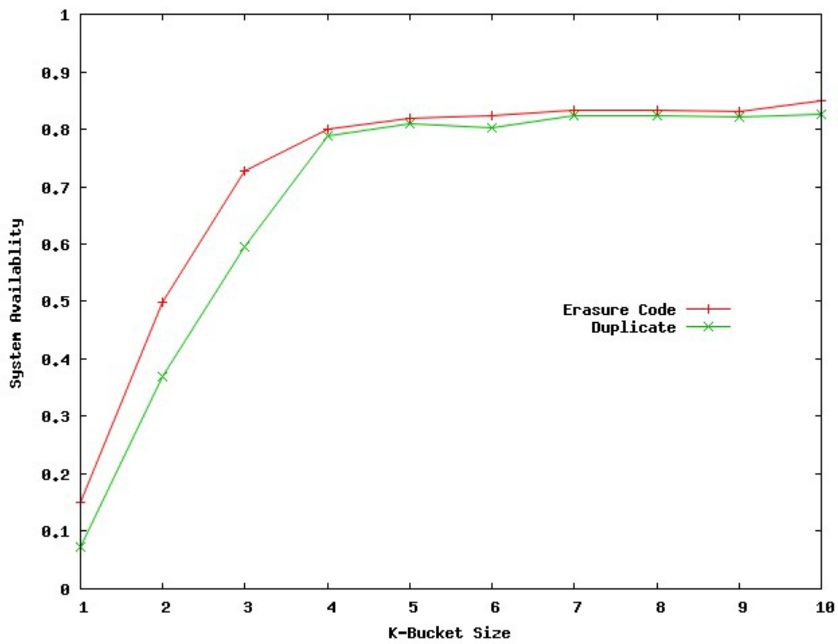

Figure 3.1: The influence of different k-bucket sizes on system availability

According to the Eq. (3.1), we can see that the theoretical value of system availability is 0.84 when full replica redundant strategy is adopted. According to the Eq. (3.2), we can see that the theoretical value of system availability is 0.84 when error correction code redundant strategy is adopted. As shown in Fig. (3.1), the experiment result is close to the theoretical value when $k \geq 4$.

When file is large, it is wise to split the file into small blocks, which can reduce the extra spending when transmission failure happens. This experiment will measure the influence of different block number on system availability, where p=0.6, r=2.0 and k=8. When error code redundant strategy is adopted, file is first divided into m blocks, and then the m blocks are encoded into mr new blocks. The experiment result is shown in Fig. (3.2).

As the result shows, the system availability decreases rapidly as the block number increases when full replica redundant strategy is adopted, it may because the probability that all of the blocks exists in the network decreases as the block number increase. So it is wise not to split the file when full replica redundant strategy is adopted in order to keep high system availability. The experiment result is close to the theoretical value, for example, the theoretical system availability is 0.21 when block number equals to 9, and the system availability is 0.175 when block number is 10.

System availability remains stable as the block number increases when error code redundancy is adopted, it is because the error code redundancy strategy does not require all of the blocks exist in the network. The error code redundancy strategy can restore the original file even if only some of the file blocks exist. Furthermore, system availability increases slowly as the block number increases according to Eq. (3.2), and it remains stable when $m \geq 5$.

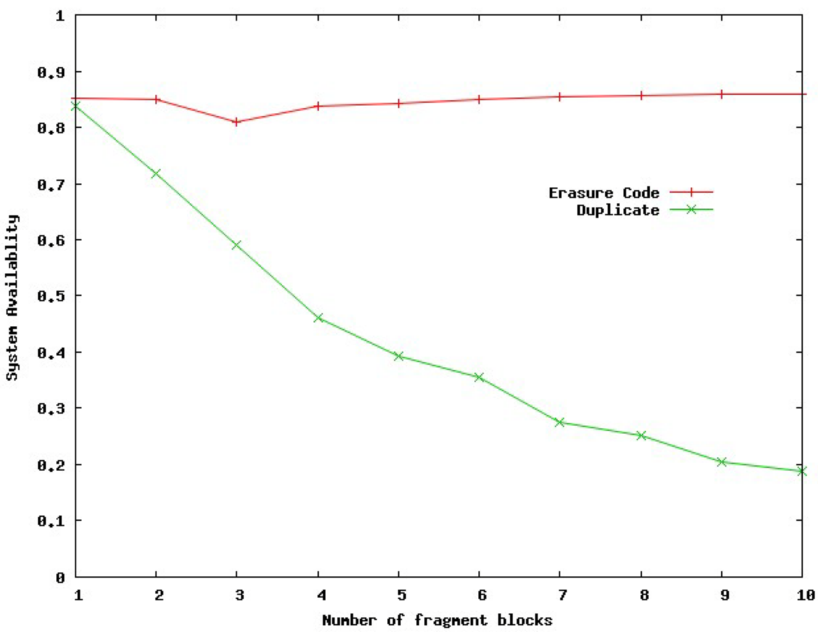

Figure 3.2: The influence of different block number on system availability

Comparing these two redundancy strategies, we knows that full replica redundant strategy is only suitable to save small files which don't need to be split, but error replica redundant strategy is suitable to save any size of files, which make it more suitable to P2P file storage system.

As shown is Fig. (3.3), the system availability increases as the redundant rate increases, whichever the adopted redundant strategy is. When the redundant rate is small ($r < 2$), full replica redundant strategy obtains a higher system availability, which can be derived from Eq. (3.1) and Eq. (3.2) too. When the redundant rate is large ($r \geq 2$), error code redundant strategy obtains a higher system availability, which can also be derived from Eq. (3.1) and Eq. (3.2).

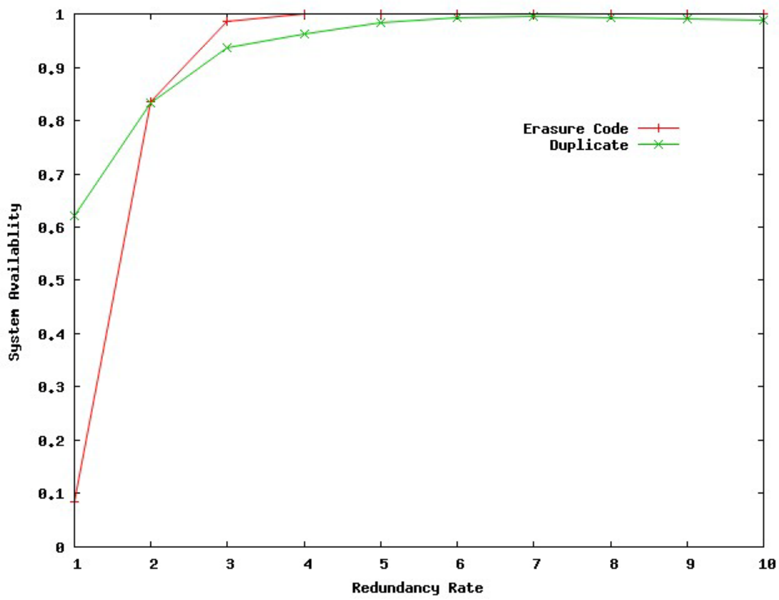

Figure 3.3: The influence of different redundant strategies on system availability

When redundant rate is 1, system availability is poor, because in this case, error code redundant strategy has degenerated to full replica redundant strategy, and what's more important, the block number is 5, so the system availability becomes poor according to Eq. (3.1).

In summary, erasure codes redundant strategy will obtains higher system availability than full replica redundant strategy when redundant rate is high.

4 Conclusions

Compared with the redundant mode of complete copy, the Erasure Code model provided more satisfactory system availability, and more suitable to apply to our P2P Cloud Storage System. Therefore, the availability of our systems needs to be further validated and tested. Based on the work in this paper, we will have the further in-depth study in routing algorithm optimization and design, redundancy algorithm optimization and design, data security, trust and reputation mechanism, incentives mechanism.

Acknowledgements

Funding for this research was provided in part by the Scientific Research Program of Zhejiang Educational Department under Grant No.20071371, National Natural Science Foundation of China (Grant No.61070153). We like to thank anonymous reviewers for their valuable comments.

References

[1] Yunhong Gu and Robert L.Grossman. Sector and Sphere: the design and implementation of a high-performance data cloud. Philosophical Trans. of the Royal Society. A367: 2429 ,2009.

[2] Robert L Grossman,Yunhong Gu. Data Mining Using High Performance Data Clouds: Experimental Studies Using Sector and Sphere. In Proc. of the 14th ACM SIGKDD,920,2008.

[3] James Broberg and Zahir Tari. MetaCDN: Harnessing Storage Clouds for High Performance Content Delivery. In: Proc. of ICSOC 2008,LNCS 5364, 730,2008.

[4] Kevin D.Bowers,Ari Juels, and Alina Oprea. HAIL: A High-Availability and Integrity Layer for Cloud Storage. Cryptology ePrint Archive, Report 2008/489, Retrieved from http://eprint.iacr.org/,2011.

[5] David Tarrant,Tim Brody, and Leslie Carr. From the Desktop to the Cloud: Leveraging Hybrid Storage Architectures in your Repository,2011.

[6] James.S.Plank.Optimizing Cauchy Reed-Solomon Codes for Fault-Tolerant Storage Applications.Technical Report CS-OS-569 Deparment of CoScience University ofTenessee.December,2005.

[7] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman.Efficient erasure correcting codes [J]. IEEE Transactions on Information Theory, 47(2):569,2001.

[8] John W.Byers, Michael Luby, Michael Mitzenmacher. A Digital Fountain Approach to Reliable Distribution of Bulk Data. ACM SIGCOMM Computer Communication Review, 28(4):56,1998.

Wu Jiyi received M.Eng. degree in 2005 from Zhejiang University, Hangzhou China, in computer science. He is currently a PhD candidate in the School of Computer Science and Technology, Zhejiang University, and associate professor of the Key Lab of E-Business and Information Security, Hangzhou Normal University. His research interests include services computing, trust and reputation.

Zhang Jianlin is a professor at Alibaba Business School, Hangzhou Normal University. He received Master's degree in Computer Application from Zhejiang University in 1998. His research interests include Cloud Computing, SaaS and information security. Prof. Zhang was born in 1966. He received B.S degree from Zhejiang University of Technology in 1989.

Wang Tong is an associate professor at Information and Communication Engineering College, Harbin Engineering University. He received Doctor's degree in Computer Application from Harbin Engineering University in 2006. His research interests include Cloud Computing, SaaS and information security. He was born in 1977. He received Master's degree from Harbin Engineering University in 2003.

Shen Qianli is a lecturer at Alibaba Business School, Hangzhou Normal University. His research interests include SaaS and information security.