# Modeling a Load-adaptive Data Replication in Cloud Environments

Julia Myint and Axel Hunger
*Department of Computer Engineering, University of Duisburg Essen, Essen, Germany*

Keywords: Replication, Cloud Storage, Reliability, Data Popularity.

Abstract: Replication is an essential cornerstone of cloud storage where 24x7 availability is needed. Failures are normal rather than exceptional in the cloud computing environments. Aiming to provide high reliability and cost effective storage, replicating based on data popularity is an advisable choice. Before committing a service level agreement (SLA) to the customers of a cloud, the service provider needs to carry out analysis of the system on which cloud storage is hosted. Hadoop Distributed File System (HDFS) is an open source storage platform and designed to be deployed in low-cost hardware. PC Cluster based Cloud Storage System is implemented with HDFS by enhancing replication management scheme. Data objects are distributed and replicated in a cluster of commodity nodes located in the cloud. In this paper, we propose a Markov chain model for replication system which is able to adapt the load changes of cloud storage. According to the performance evaluation, the system can be able to adapt the different workloads (i.e data access rates) while maintaining the high reliability and long mean time to absorption.

## 1 INTRODUCTION

Cloud computing is a new computing paradigm that is gaining increased popularity. It enables enterprise and individual users to enjoy flexible, on demand and high quality services such as high volume data storage and processing without the need to invest on expensive infrastructure, platform or maintenance.

Although high performance storage servers are the ultimate solution for the challenges, the implementation of inexpensive storage system remains an open issue. Moreover, the economic situation and the advent of new technologies have sparked strong interest in the cloud storage provider model. Cloud storage providers deliver economics of scale by using the same storage capacity to meet the needs of storage user, passing the cost saving to their storage. Data replication is a well-known technique from distributed systems and the main mechanism used in the cloud for reducing user waiting time, increasing data availability and minimizing cloud system bandwidth consumption by offering the user different replicas with a coherent state of the same service (Sun et al., 2012). This paper proposes a modelling approach of cloud storage system in replication aspect.

The rest of this paper is organized as follows. In the next section, we discuss the related papers with our paper. In section 3, we propose the cloud storage architecture and model the system by using Markov model. Then the model is analyzed in section 4. Finally, we conclude our paper.

## 2 RELATED WORK

Among a large amount of researches in storage system for cloud computing, Google File storage system for cloud computing, Google File System (GFS) (Ghemawat et al., 2003) and Hadoop distributed file system (HDFS) (Borthakur, 2007) are widely used and most popular. Other cloud storage systems that use key-values mechanisms are Dynamo (Decandia et al., 2007), Pnuts (Cooper et al., 2008) and Cassandra (Lakshman and April, 2010).

Replication management has been active research issue in Cloud storage system proposed by (Vo et al., 2010), (Jagadish et al., 2005), (Wei et al., 2010) and (Ye et al., 2010). Modeling and analysis of cloud computing has been an active research for availability, reliability, scalability and security issues. (longo et al., 2011), (Chuob et al., 2011), (Sun et al., 2012) proposed modelling approach for this purpose.

The above cloud storage systems and models

Myint J. and Hunger A..
Modeling a Load-adaptive Data Replication in Cloud Environments.
DOI: 10.5220/0004374805110514
In *Proceedings of the 3rd International Conference on Cloud Computing and Services Science* (CLOSER-2013), pages 511-514
ISBN: 978-989-8565-52-5
Copyright © 2013 SCITEPRESS (Science and Technology Publications, Lda.)

apply different strategies for effective storage. However they do not consider the rapid changes of data popularity in modelling cloud storage replication system with Markov chain approach. In this paper, therefore, a model for an efficient replication scheme which is able to adapt the data popularity is proposed with the analysis of data reliability and mean time until absorption (failure).

# 3 PROPOSED SYSTEM OVERVIEW

The proposed cloud storage system is implemented with Hadoop storage cluster (HDFS). HDFS applies tri-replication and configurable per file. However, HDFS does not provide policy to determine the replication factor. In this section, the architecture of proposed Cloud storage system and modeling of this architecture will be presented.

## 3.1 Cloud Storage Architecture

As a **cloud user**, applications are browsed through browsing interfaces. The applications may involve data storage and file retrieval applications.

**Cloud storage client** is an interface between user interface and storage servers. The storage client is a code library that exports HDFS client module. Replication Manager considers how many replicas should be replicated to cope data popularity. After that, it updates the replication number in HDFS configuration file, contacts the Name node together with the configuration setting and requests for the list of Data nodes that host replicas of the blocks of the file. It then contacts a Data node directly and requests the transfer of the desired block. When a client writes, it first asks the Name node to choose Data nodes to host replicas. The designed cloud storage is based on PC cluster. The PC cluster consists of a single Name node and a number of Data nodes.

## 3.2 Modeling

The modeling of Cloud storage system is based on proposed system architecture shown in figure 1. The cloud storage system *S* consists of some number of nodes (i.e DataNodes) *n* on which replicas of data *D* can be created. A node participates some duration of time $t_p$ i.e an exponentially distributed random variable with mean $1/\lambda$ where $\lambda$ is the rate of departure. We assume that $t_p$ is independent and identically distributed for all nodes in the system. Over a period of time, node departures decrease the number of replicas of *D* present in the system. In the system *S*, we assume that the cloud user access to data *D* in cloud storage system with mean $1/\alpha$. Likewise, we also assume that the user request to the file does not arrive with mean $1/\beta$. In these two cases, $\alpha$ and $\beta$ are data access rate and data cold rate respectively.

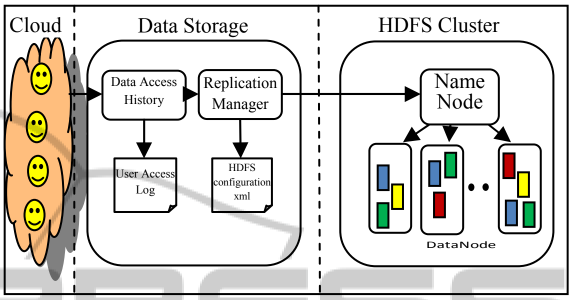



Figure 1: Cloud Storage Architecture.

To be more reliable and maintainable in cloud storage system, the system must also use a repair mechanism that creates new replicas to account for lost ones. The repair mechanism must first detect the loss of a replica, and then create a new one by copying *D* to another node from an existing replica. The whole process may take the system some duration of time $t_r$ i.e an exponentially distributed random variable with mean $1/\mu$ where $\mu$ is the rate of repair.

In any state of system, the system can replicate data *D* for the duration of its participation in the system or for the popularity of data access rates. When a node leaves the system, we assume that its state is lost. We note that the number of replica *n* is a parameter that the system can choose depending on the storage limit which may impose the upper bound on *n*.

### 3.2.1 Markov Chain

To analyse the system, it is reduced to a Markov chain. The proposed system is adapted from the problem of Gambler's ruin found in (Epstein, 1977) and (Feller, 1968). In Markov chain model, the system has k functioning replicas where $k$ ($0 \leq k \leq n$ ) and $l$ data access level where $l$ (min $\leq l \leq$ max). The remaining $n$-$k$ are being repaired. Thus the system has ($n$+1)*(max-min+1) possible states in Markov chain. If the system is in state ($k$,$l$), there are $k$ functioning replicas and $l$ data access level. In state ($k$,$l$), any one of the k functioning replicas can fail or access frequency can be lower than $l$ data access level. In the former case, the system goes to state ($k$-

1,*l*) and in the later case, the system goes to state (*k*-1,*l*-1). There is another possibility in state *k*. One of the *n*-*k* non-functioning replicas is repaired in which case the system goes to state (*k*+1,*l*)or if data access frequency exceeds the level (*k*+1,*l*) with rate $\mu$ , to (*k*+1,*l*+1) with rate α, to (*k*-1,*l*) with rate $\lambda$, to (*k*-1,*l*-1) with rate $\beta$. Note that state 0 is an absorbing state and the system no longer recover *D* when there are no more functioning replicas left. Figure 2 shows the Markov chain model of the system and the definitions of the symbols are described in table 1.

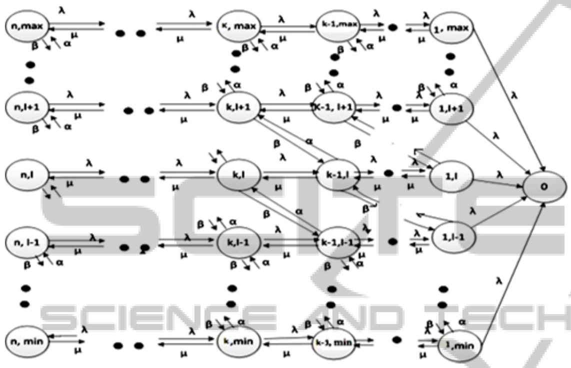


Figure 2: Markov Chain moDel for a Load Adaptive Replication.

Table 1: Symbols used in Figure 2

| Symbols | Descriptions |
|---|---|
| $n$ | Number of replica |
| $\lambda$ | Departure rate |
| $\mu$ | Repair rate |
| $\alpha$ | Access rate due to increase popularity |
| $\beta$ | Data cold rate due to decrease popularity |
| $k$ | Number of active replica |
| $l$ | Data access level |
| max, min | Maximum data access level and minimum data access level |

To analyze the Markov chain model of the proposed system shown in figure 1, we reduce the system with *n*=3 (*k*=[1,2,3]), max=2, min=0, and *l*=[0,1,2]. Then the simplified model is shown in figure 3.

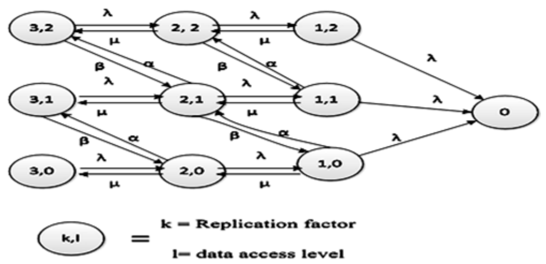



Figure 3: The State Diagram of Maximum Replication Factor=3 and 3 Data Access Levels.

# 5 NUMERICAL RESULT

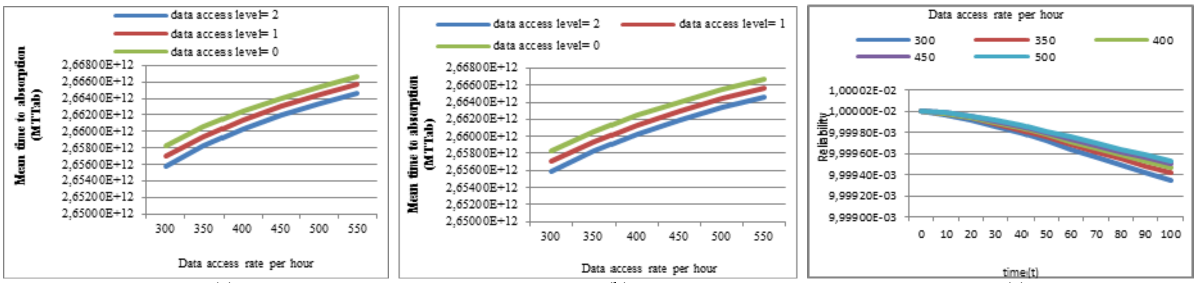
The proposed replication model is evaluated by using SHARPE software package (Trivedi et al., 2009). We evaluated the system with two solutions- (1) Mean time to absorption (failure) and (2) Reliability. In this system model, the effect of changing data access frequency, initial configuration of workload and initial replication factor on the system are analysed. As the cloud storage system is implemented with commodity computing nodes, the assumption of mean time to failure (MTTF) is 1 month and we calculate the other parameters are listed in Table 2.

According to the parameter values in table 2, we quantified how long a replicated system can maintain some states before it is lost permanently due to the dynamics of the data access rates. Figure 4(a) shows the mean time to absorption by varying data access rates The effect of changing initial replication factor on mean time to absorption is illustrated in figure 4(b).

Table 2: Operating Parameter Values for the Model.

| Parameters | Values |
|---|---|
| The failure rate is $\lambda = \frac{1}{MTTF} = \frac{1}{1*30*24\text{hours}}$ | $1.389*10^{-3}$ per hour |
| The failed machine needs 1 day to recover normal condition. The repair rate is $\mu = \frac{1}{MTTR} = \frac{1}{24\ hours}$ | $4.167*10^{-2}$ per hour |
| Data access rate is $\alpha = \frac{1}{Mean\ time\ to\ Data\ Access\ (MTDA)}$ | 300 to 550 per hour |
| We assume the data access is slow with mean 7 days. Data cold rate is $\beta = \frac{1}{Mean\ time\ to\ Data\ Cold(MTDC)} = \frac{1}{7*24\ hours}$ | $5.952*10^{-3}$ per hour |

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS




(a) (b) (c)

Figure 4: (a) Comparison of Mean Time to Absorption with the Starting State of k=2and l=[0,1,2] (b) Mean Time to Absorption Starting with k=[1,2,3] and l=2 (c) Reliability Comparison of Various Data Access Rate.

In order to evaluate how much reliable in the proposed system, we compared the reliability values by varying the data access rate from 300 to 550 per hour. The results are shown in figure 4(c). In the figure, it can be said that the system is more reliable in accordance with higher data access rate.

## 6 CONCLUSIONS

Data replication is an essential technique to reduce user waiting time, speeding up data access by providing users with different replicas of the same service. To take advantage of these, we propose an effective replication model to manage replication degree in which it takes failure rate and data access popularity into account. In this paper, we quantify the effects of variations in workload (i.e data access rate) and initial system configuration (setting up the replica number and data access level) on cloud storage quality in terms of reliability and mean time to failure. The experimental results demonstrate that the proposed model is able to adapt the varying data access load and therefore it can be more efficient in cloud data storage.

## REFERENCES


Borthakur, D., 2007, "The Hadoop Distributed File System: Architecture and Design". *The Apache Software Foundation*.

Chuob, S. and et.al., 2011, "Modeling and Analysis of Cloud Computing Availability based on Eucalyptus Platform for E-Government Data Center", *2011 Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*.

Cooper, B. F., Ramakrishnan, R., Srivastava, U., Silberstein , A., Bohannon, P., and et.al., 2008, "Pnuts: Yahoo!'s Hosted Data Serving Platform",*In VLDB*.

Cristina, L. and et. al., 2012 "A strong-Centric Analysis of MapReduce Workloads: File Popularity, Temporal Locality and Arrival Patterns", *In Proc. IEEE IISWC*.

Decandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., Sivasubramanian, S., Vosshall, P. and Vogels, W., 2007,"Dynamo: Amazon's Highly Available Key-value Store", *In SOSP*.

Epstein, R., 1977, *The Theory of Gambling and Statistical Logic*. Academic Press.

Feller, W., 1968, *An Introduction to Probability Theory and Its Applications*. John Wiley and Sons.

Ghemawat, S., Gobioff, H., Leung, S. T., October, 2003, "The Google File System", *Proceedings of 19th ACM Symposium on Operating Systems Principles(SOSP 2003)*, New York USA.

Jagadish, H. V., Ooi, B. C., and Vu, Q. H., 2005, "BATON: A Balanced Tree Structure for Peer-to-Peer Networks". *In VLDB*.

Lakshman, Malik, P., April 2010, "Cassandra - A Decentralized Structured Storage System", *ACM SIGOPS Operating Systems Review, Volume 44 Issue 2*.

Longo, F., Ghosh, R., Naik, V. K. and Trivedi, K. S., 2011 "A Scalable Availability Model for Infrastructure-as-a-Service Cloud", *DSN*.

Ramabhadran, S. and Pasquale, J., 2006, "*Analysis of Long-Running Replicated Systems*", In the Proceedings of 25th IEEE International Conference on Computer Communications, pp. 1--9.

Sun, D. W. and et.al., Mar. 2012,"Modeling a Dyanmic Data Replicatin Strategy to Increase System Availability in Cloud Computing Environments", *Journal of Computer Science and Technology* 27(2):256-272. DOI 10.1007/s11390-012-1221-4.

Trivedi, K. S. and Sahner R., March 2009, "SHARPE at the age of twenty two," *ACM Sigmetrics Performance Evaluation Review*, vol . 36, no. 4, pp. 52–57.

Vo, H. T., Chen, C., Oo, B. C., September 13-17 2010, "Towards Elastic Transactional Cloud Storage with Range Query Support", *International Conference on Very Large Data Bases* , Singapore.

Wei, Q. and et. al., 2010 "CDRM: A Cost-effective Dynamic Replication Management Scheme for Cloud Storage Cluster", *IEEE International Conference on Cluster Computing*.

Ye, Y. and et. al., 2010 "Cloud Storage Design Based on Hybrid of Replication and Data Partitioning", *16th International Conference on Parallel and Distributed Systems*.


SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS