

# Improving Consistency in Cloud by using Audit Cloud Service

Sudeshna  
Shinde  
Department of  
Computer  
Engineering  
Savitribai  
Phule Pune  
University, Pune,  
India,

Payal Tatiya  
Department of  
Computer  
Engineering  
Savitribai  
Phule Pune  
University, Pune,  
India,

Priyanka Vidhate  
Department of  
Computer  
Engineering  
Savitribai  
Phule Pune  
University, Pune,  
India,

Priyanka Yadav  
Department of  
Computer  
Engineering  
Savitribai  
Phule Pune  
University, Pune,  
India,

Kirti Deshpande  
Department of  
Computer  
Engineering  
Savitribai  
Phule Pune  
University, Pune,  
India

## ABSTRACT

Cloud storage has become more popular due to their many advantages. To provide access everywhere, a cloud service provider maintains many replicas for each piece of data on geographically distributed servers. Using the replication technique in cloud it has one problem that is it is very costly to maintain strong consistency on a large scale. Here a Consistency as a Service (CaaS) model is presented, which consists of a large data cloud and audit service. CaaS model consists of a data cloud which is maintained by a CSP (Cloud Service Providers) and audit service to verify whether the data cloud provides the level of consistency or not. There is a two-level auditing architecture, which requires a secured synchronized clock in the cloud. Algorithms are designed to express the problems. Finally, a heuristic auditing strategy (HAS) is created to find as many violations as possible.

## Keywords

Cloud storage, consistency as a service (CaaS), two-level auditing and heuristic auditing strategy.

## 1. INTRODUCTION

Cloud refers to the Internet or is more precisely connected to Internet through datacenter full of servers. Cloud is individual IT environment that is designed for the purpose of remotely accessing different services such as storage, servers.

Cloud Computing is network of remote servers hosted on Internet for manipulating, arranging and accessing the hardware and software resources, storing data. It provides online data storage, infrastructure and application. Cloud Computing is delivery of on-demand computing resources means storing and accessing the data and programs over the internet instead of computer hard drive. Cloud Computing is the dreamed vision of computing as a public utility. Cloud Computing provides services such as data storage, virtualized infrastructure etc. e.g. Amazon, SimpleDB etc. For continuous access, cloud service stores data replicas on multiple geographically distributed servers. The use of the replication technique is to achieve strong consistency and user is ensured to see the latest updates.

Service models are as follow:-

- 1) SaaS defines any cloud service where users can access software applications over the internet.
- 2) PaaS that provides a platform and environment to which developers use to build applications and

services over the internet.

- 3) IaaS that provide access, monitoring and managing remote infrastructures, such as compute, storage, networking services.
- 4) CaaS consists of a large data cloud formed by CSP and audit services that can check whether the data cloud provide consistency or not.

## 2. RELATED WORK

A cloud is a large-scale distributed system where each piece of data is replicated geographically by distributed servers to accomplish high availability and high performance. Thus, first review is done. Ref. [1] proposed that how to assess and check the consistency actually observed by the clients when using these key-value stores rather than assuming the worst-case consistency that is promised. Ref. [2] proposed cases where the platform offers consumers a choice between stronger and weaker consistency and used only one user to read data in the trials, and showed that few inconsistencies exist in S3. Ref. [3] proposed Consistency Rationing to enhance the runtime cost of a database system in the cloud when discrepancies incur a penalty cost. The consistency requirements differ over time depending on actual availability of the data, and provide techniques that make the system enthusiastically adapt to the consistency level by monitoring the state of the data. Ref. [4] proposed a fresh consistency model for replicated data across data centers with framework and allows it to adjust the consistency levels for different semantic data. Ref. [5] used multiple geographically-distributed users to read data, and it is seen that it violates monotonic-read consistency. Ref. [6] presents a client-centric method for knowing eventual consistency in distributed key-value storage systems. Ref. [7] proposed simple figures to quantify differences between cloud and conventional Computing and finding the top technical and non-technical obstacles and opportunities of Cloud Computing. Ref. [8] proposed an online verification algorithm and used different techniques to quantify the severity of violations. Ref. [9] proposed the design of a cloud storage system with a range of consistency options that lie between strong and eventual consistency. Ref. [10] proposed the types of scalability, choosing the correct scalability and other issues.

### 3. SYSTEM ARCHITECTURE

- The CaaS (Consistency as a Service) model consists of a large data cloud and multiple users.
- The Data Cloud is key-value data storage system where each data is identified by a unique key.
- The Audit Service manages the large data cloud and replicates all of the data on cloud.
- The consistencies are verified locally and globally then after that the data is transferred to the large data cloud.
- Locally checking is done of the monotonic read consistency and read your write consistency. And globally we are checking the casual consistency.

Basically system architecture is divided into four modules, they are as follows:-

#### Module 1: User Module

Input: User information & file to be uploaded.

Output: Registered user, uploaded file, and also the files which are present in the database.

#### Module 2: Auditor Module

Input: A file which also called revision report.

Output: Accepting or rejecting the file.

#### Module 3: Admin Module

Input: Registered user details, uploaded file details.

Output: View of user information & uploaded file details.

#### Module 4: Data Upload Module

Input: File to be uploaded.

Output: Storage of uploaded files.



Fig 1: System Architecture

### 4. TECHNOLOGY USED

The system has heuristic auditing strategy (HAS) which adds appropriate reads to know as many violations as possible. Key contributions are as follows:

- 1) A novel consistency as a service (CaaS) model is presented in which a group of users that constitute an audit service can check whether the data cloud provides the level of consistency or not.
- 2) A two-level auditing structure is proposed, which only needs a loosely synchronized clock for ordering operations in a cloud. The two level auditing structure is called local level auditing and global level auditing. In local level auditing, each users will perform the auditing individually with their own UOT (User Operation Table). In the local level auditing, more focus is on monotonic read consistency and read your write consistency. In the global level auditing, audit service will automatically perform the auditing. In the global level auditing focus is

on causal consistency.

- 3) UOT (User Operation Table): One of the main methods is generating the UOT. UOT is maintained by each user for storing the local operations. The consistencies are verified using the User Operation Table.
- 4) Algorithms are designed to quantify the severity of violations with different metrics.
- 5) Heuristic Auditing Strategy (HAS) is developed to remove as many violations as possible.

#### Building Blocks:

##### 1. Local consistency auditing:

1. Initial UOT with  $\emptyset$
2. while issue an operation op do
3. if op = W(a) then  
record W(a) in UOT

4. if  $op = r(a)$  then

$W(b) \in UOT$  is the last write

5. if  $W(a) \rightarrow W(b)$  then

Read-your-write consistency is violated

6.  $R(c) \in UOT$  is the last read

7. if  $W(a) \rightarrow W(c)$  then

Monotonic-read consistency is violated

8. record  $r(a)$  in UOT

## 2. Global consistency auditing:

Each operation is denoted by a vertex

1. for any two operations  $op1$  and  $op2$  do

2. if  $op1 \rightarrow op2$  then

Time edge is added from  $op1$  to  $op2$

3. if  $op1 = W(a)$ ,  $op2 = R(a)$ , and two operations come from various users then data edge is added from  $op1$  to  $op2$

4. if  $op1 = W(a)$ ,  $op2 = W(b)$ , two operations come from various users, and  $W(a)$  is on the route from  $W(b)$  to  $R(b)$  then

A causal edge is added from  $op1$  to  $op2$

5. Check whether the graph is a DAG by topological sorting

### Monotonic-read consistency

If any user reads the value of data, then other reads on data will always return that same value or updated value.

### Read-your-write consistency

The every write by a user on data will always be seen by a successive read on data by the same user.

## 5. CONCLUSION

Cloud Computing is an rising commercial infrastructure concept that removes the need for maintaining expensive computing hardware. As market is growing the threat of maintaining and losing data is also growing. So, consistency in cloud computing is very much needed as data in the cloud storage is not updated regularly and it also require 24/7 access to cloud. In this paper, a Consistency as a Service model (CaaS) is presented in which audit cloud service is implemented that can automatically check whether the data cloud provides the level of consistency or not & also various replicas are created to maintain the consistency. For future work, a thorough theoretical study of consistency models in cloud computing is to be conducted.

## 6. ACKNOWLEDGMENTS

Our thanks to project guide Ms.Kirti Deshpande who has contributed towards research of this work.

## 7. REFERENCES

- [1] E. Anderson, X. Li, M. Shah, J. Tucek, and J. Wylie, "What consistency does your key-value store actually provide," in Proc. 2010 USENIX HotDep.
- [2] H. Wada, A. Fekete, L. Zhao, K. Lee, and A. Liu, "Data consistency properties and the trade-offs in commercial cloud storages: the consumers' perspective," in Proc. 2011 CIDR.
- [3] T. Kraska, M. Hentschel, G. Alonso, and D. Kossmann, "Consistency rationing in the cloud: pay only when it matters," in Proc. 2009 VLDB.
- [4] S. Esteves, J. Silva, and L. Veiga, "Quality-of-service for consistency of data geo-replication in cloud computing," Euro-Par 2012 Parallel Processing, vol. 7484, 2012.
- [5] D. Bermbach and S. Tai, "Eventual consistency: how soon is eventual?" in Proc. 2011 MW4SOC.
- [6] M. Rahman, W. Golab, A. AuYoung, K. Keeton, and J. Wylie, "Toward a principled framework for benchmarking consistency," in Proc. 2012 Workshop on HotDep.
- [7] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, et al., "A view of cloud computing," Commun. ACM, vol. 53, no. 4, 2010.
- [8] W. Golab, X. Li, and M. Shah, "Analyzing consistency properties for fun and profit," in Proc. 2011 ACM PODC.
- [9] Douglas B. Terry, Vijayan Prabhakaran, Ramakrishna Kotla, Mahesh Balakrishnan, Marcos K. Aguilera, Hussam Abu-Libdeh, "Consistency-Based Service Level Agreements for Cloud Storage".
- [10] R. Anandhi, K. Chitra, "A Challenge in Improving the Consistency of Transactions in Cloud Databases – Scalability".
- [11] M. Rahman, W. Golab, A. AuYoung, K. Keeton, and J. Wylie, "Toward a principled framework for benchmarking consistency," in Proc. 2012 Workshop on HotDep.
- [12] A. Tanenbaum and M. Van Steen, *Distributed Systems: Principles and Paradigms*. Prentice Hall PTR, 2002.
- [13] W. Vogels, "Data access patterns in the Amazon.com technology platform," in Proc. 2007 VLDB.
- [14] "Pushing the CAP: strategies for consistency and availability," *Computer*, vol. 45, no. 2, 2012.
- [15] G. DeCandia, D. Hastorun, M. Jampani, G. Kakulapati, A. Lakshman, A. Pilchin, S. Sivasubramanian, P. Vosshall, and W. Vogels, "Dynamo: Amazon's highly available key-value store," in Proc. 2007 ACM SOSP.