iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** IV **Month of publication:** April 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Heuristic Auditing Strategy Using Big Data

Thirunavukkarasu.K[1], Appandairaj.A[2]
[1] M.E Student, Ganadipathy Tulsi's Jain Engineering College, Vellore Dist, T.N, India
[2] Assistant Professor, Ganadipathy Tulsi's Jain Engineering College, Vellore Dist, T.N, India

Abstract- Cloud storage services have turn into commercially famous now a day. To give unique and updated access of user file on cloud, a cloud service provider (CSP) keeps up multiple replicas for each every bit of information or data on geographically conveyed servers. The primary issue of using the replication technique as a part of clouds is very costly to maintain consistency among large data on clouds. In this paper, we first present a consistency as a service (CaaS) model, which have large data access as well conduct multiple small audits in clouds .Auditing use to ensure the consistency of the data played in cloud globally. Here I we study many existing papers and develop auditing model which uses heuristic auditing strategy that have two level of auditing among cloud data. These auditing uses a loosely synchronized clock and we develop algorithms to evaluate the severity of violation with two metrics such us commonality of violation and mustiness of the reading the data content and heuristic auditing strategy (HAS) to disclose as many violations as possible.

I. INTRODUCTION

Cloud computing is widely used IT technology which provide services like network access, resources, platform, infrastructure, rapid resource elasticity as per user require. The data of user is most important asset in cloud, the user can use the cloud service anywhere and anytime through internet. The NIST definition of cloud computing "Cloud computing is a model for allowing convenient and on- demand network access to a shared pool of configurable computing resources which can be rapidly provisioned and released with minimal management effort or service provider interaction"[1]. Storage system in the cloud is kind of network prototype for online storage in which the user data stored in virtualized pools of storage system generally maintain by cloud service provider. There are two type of cloud storage, such as store in remote system and temporarily cached on mobile phones, desktop computers or other internet devices. The benefit of storing data on cloud is avoidance of capital expenditure on personal maintenances, software, Hardware, relief of online burden of data storage [2].Many users from remote location use cloud services continuously so there may arise many issues such as security, privacy, dynamic updates, data integrity. It is difficult to check data consistency every time by each user. So user wants to maintain data consistency, privacy and data integrity forever. On other hand they do not know that the CSP can misuse their data, and correctness of data [4]. The cloud server stores large amount of data which does not offer guarantee on data integrity and consistency. This problem is taking in hand and solve by public auditing for secure cloud.

Our model contain large amount of cloud data and small audit on cloud data. The data floating in clout maintain by their own clouds service provider. Cloud audit can be done by many people involved in cloud data manipulation, may be user, data owner, cloud service provider they have to cooperate on this job, such as prepare document, maintain service level agreement(SLA).the data cloud and audit cloud should be engaged by service level agreement(SLA).This SLA use to ensure what level of consistency the data should provide by cloud service provider and how much monetary paid if the data cloud break or violate the service level agreement between them. Due to virtualization technique used in cloud computing it is very difficult to implement the data cloud to all users. So it is very hard for the end user to verify whether each data in data cloud comes from latest replication or not. Thus in our model we propose audit cloud which allow user to verify data consistency by analyzing a trace of interactive operations. This model not have global clock for all user for all operation, here we use synchronized clock used, and it is more suitable for audit cloud. In our project the data consistency checked by upload and downloads a file from cloud. The following goal should be meet to ensure cloud consistency.

A. The cloud service provider should give basic consistency as service.
B. Service Availability
C. Understand the importance of CaaS model.
D. Maintain synchronized clock at audit clouds which has responsible for checking weather at least minimum level consistency provide by CSP

II. RELATED WORK

As per our survey there are many previously study done in the field of cloud computing. Because cloud have provide many service like online marketing, healthcare ,banking ,social media, thus maintain consistency of the data in cloud is hot research topic today. This highlights the concern on consistency of cloud data, for these growing consistencies as a service many technology and method have been proposed.
First we take the consistency models in cloud system ref[10] propose two consistency model ,

A. Data-centric consistency model
B. client-centric consistency model

Data-centric consistency model consider the in-house state of a storage system, i.e., how guarantees the system can provide with respect to updates flow in storage system. However, by customer perspective, really they does not matter whether or not a storage system internally have any stale copies of data.
Client-centric consistency model give attention to on what specific customers need and what they need to get, i.e., how the customers monitor data updates against their need. Their works also illustrate different levels of consistency in cloud computing systems, from strict to weak consistency. High consistency means high cost and less availability of data. Ref. [11] states that strict consistency in cloud cannot be practice in real environment. Many cloud systems sacrifice strict consistency for high availability and reduce access cost.
In Ref. [14], the consistency necessities vary over time depending on availability of the data, and the data owner should provide techniques that make the system dynamically adapt to the consistency level by monitoring the state of the data. We Ref. [15] proposed a novel consistency model that allows it to automatically adjust the consistency levels for different semantic data.
We analysis the effort on make sure the levels of consistency provided by the cloud service provider from the end users' point of view. Existing solutions can be classified into benchmark-based verifications and trace-based verifications. The trace-based verifications concentrates on three consistency semantics: atomicity , safety and regularity, which are proposed by Lamport. A registration is safe if a read that is not concurrent with any write returns the value of the most recent write, and a read that is concurrent with a write can return any value.
We consider that cloud computing, networking and storage must all focus on horizontal scalability of virtualized resources rather than on single nodule performance. In paper [2] they deal with the way to store user file on cloud storage and key contribution of COPS is its scalability, which can implement causal dependencies between keys stored across an entire cluster, rather than a single server. But storing a file on a cluster is making enormous problem for providing consistency. Measuring consistency is a very essential task in this system because monitoring and controlling consistency is major purpose of proposed system.
In [3] they are provided new perspective to show the need of consistency as a service (CaaS) and cloud computing storage services, every request has an associated cost according to service type. It is possible to assign a monetary cost to consistency protocols. Therefore, in cloud storage services, consistency not only influences the availability and performance of the systems. Dengyong Zhou and Thomas Navin Lal propose local and global consistency model . The key to semi-supervised learning problems is the consistency assumption, which essentially requires a classifying function to be sufficiently smooth with respect to the intrinsic structure revealed by a huge amount of labeled and unlabeled points. We proposed a simple algorithm to obtain such a solution, which demonstrated effective use of unlabeled data in experiments including toy data, digit recognition and text categorization.

III. PROPOSED SYSTEM DESIGN

A. Design and Implementation Constraints

Main purpose of this paper is to upload and download a user file from cloud. While downloading document or user file from cloud, we are checking whether the record is upgraded version form or not. Likewise discovering end time requires updating the document on multiple servers.

1) Giving the upload and download tab with temples in browser.
2) Button to submit the user file on cloud.
3) Label to show warnings messages and alert messages to the enduser while error occur.

B. Assumptions and Dependencies

1) Assumptions

a) First assumption stands that the user should complete knowledge of the system .

b) The software tool should be user friendly as possible and other hand it should meet user requirements.

c) The user PCs must have Windows OS and .NET framework should be configured to form a cluster.

2) Dependencies: A dependency means any one going to use this system should follow international standard on every process within environment. The user should fill the related information in the proper format.

C. User Classes and Characteristics

There are various type of users for the product. Who are:-

1) Admin who have whole control on entire project.

2) Customer or end user who going to use this system.

Both users play vital role in cloud service and make use of cloud services provided by CSP.

IV. PRELIMINARIES

In this part, we demonstrate the consistency as a service (CaaS) model. Then, we explain the structure of *the user operation table* (UOT). Finally, we present an overview of the two-level auditing structure and related definitions.

A. Consistency as a Service (CaaS) Model

The CaaS model consists of a *data cloud* and multiple *audit clouds shows in below figure*.

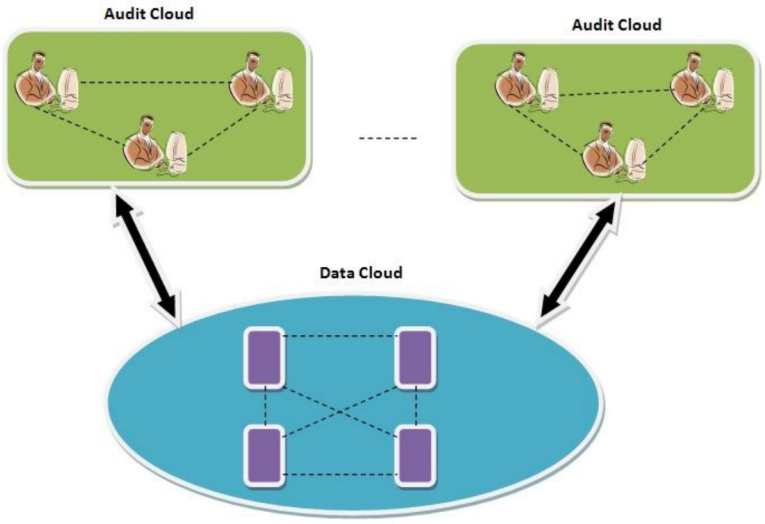


Figure 1. Consistency as a service model

The data cloud, maintained and accessed by the cloud service provider, where each part of data is identified by its unique key value. To provide greater service availability, the CSP replicates data on multiple geographically cloud servers. An audit cloud consists of a group of users that cooperate on verifying consistency off the data. Each and every user in the audit cloud is identified by a unique ID and it should be fully confidential. Service level agreement (SLA), which stipulates the assure level of consistency that should be given by the data cloud. The audit cloud be real to verify whether the data cloud violates the SLA and to quantify its severity.

B. User Operation Table (UOT)

Each user maintains a *User Operation Table (*UOT) for recording local operations. Every record in the UOT is described by three elements: *operation, physical vector and logical vector*. While issuing an operation, a user will record this operation, as well as his current logical vector and physical vector, in his UOT.

Volume 3 Issue IV, April 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

V. ALGORITHM

According to [8] we use algorithm which describe below. This algorithm is for Local Consistency Auditing. The procedure in algorithm is as follows:

Initialize User Operation Table (UOT) with Ø
While issue an operation op do
If op = W (a) then
Record W (a) in User Operation Table (UOT)
If op = r (a) then
W (b) ϵ User Operation Table (UOT) is the last write
If W (a) → W (b) then
Read-your-write consistency is violated
R(c) ϵ UOT is the last read
If W (a) → W(c) then
Monotonic-read consistency is violated
Record r (a) in User Operation Table (UOT)
Where,
W (a) - Write operation
R (a) - Read operation

VI. DETAILED DESIGN

A. *State-Transition Diagram*

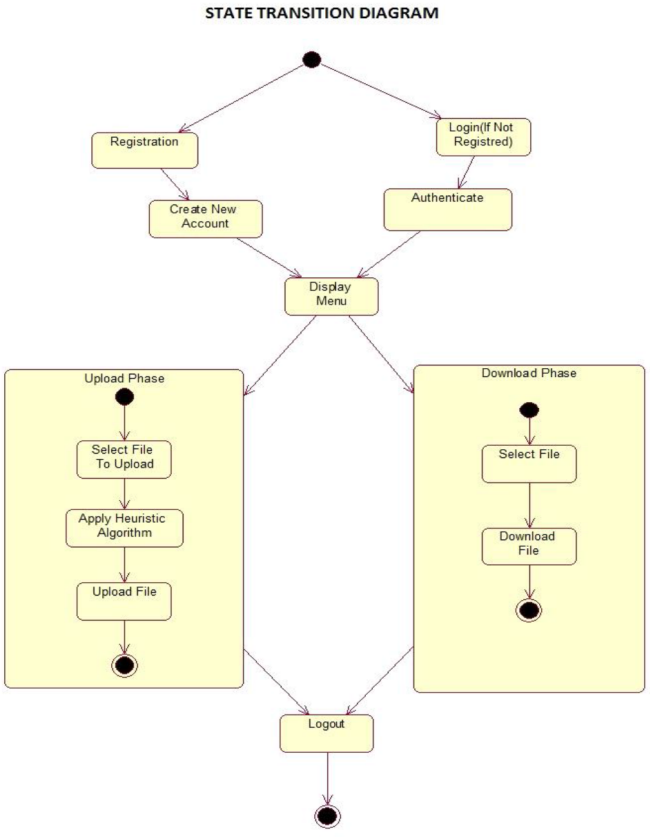


Figure 2. State Transition Diagram

Volume 3 Issue IV, April 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

B. Data Flow Diagram

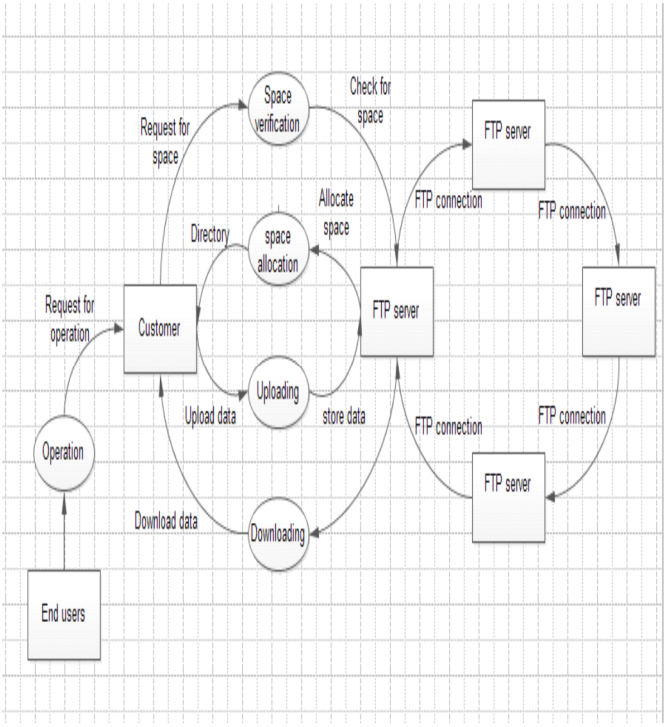


Figure 3. Data Flow Diagram

Data flow diagram shows flow of data and result in each stage .The process started by request by end user.

C. System Architecture Diagram

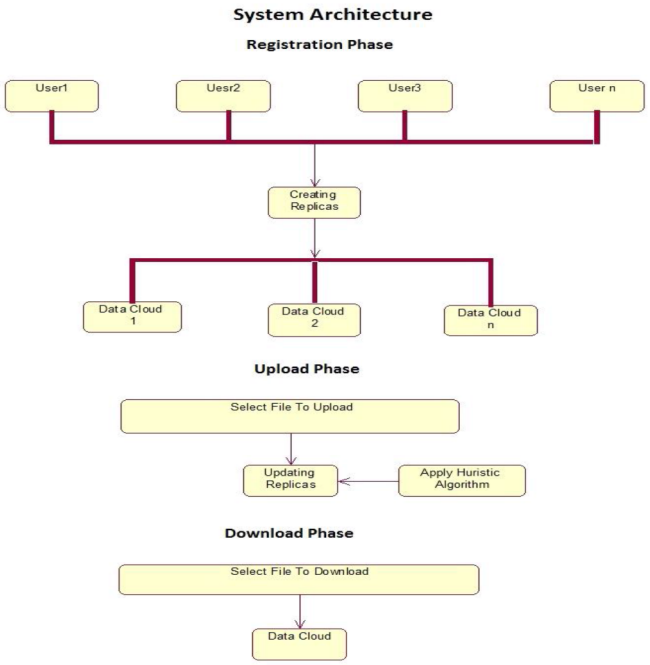


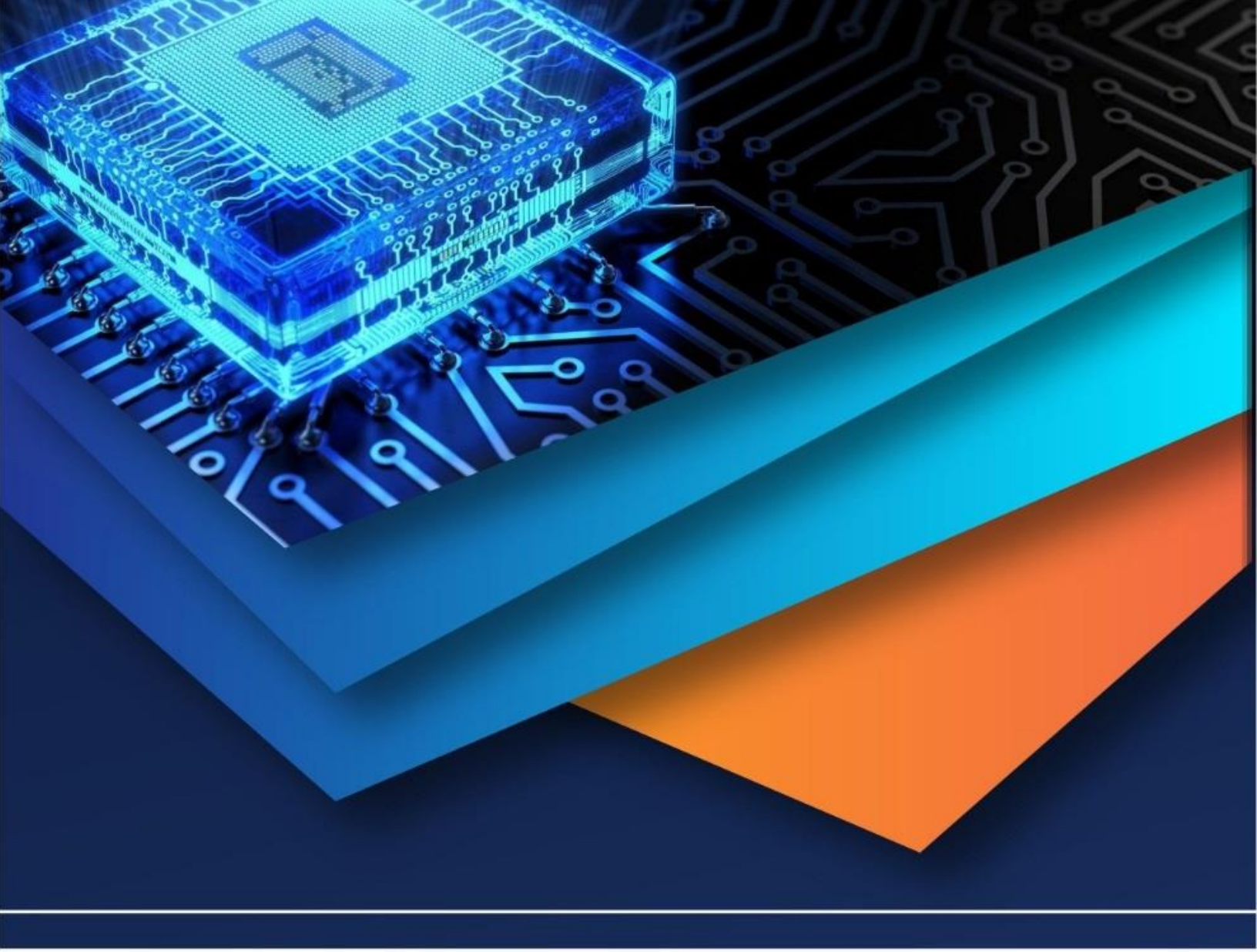
Figure 4. System Architecture

VII. CONCLUSION

In this paper, we survey different approaches to implement CaaS base model and a two-level auditing structure to help users verify whether the cloud service provider is providing valid consistent data or not. With the CaaS model, the users can assess and verify the quality of cloud services, by this help user to choose a right Cloud service provide.

REFERENCES

[1] Static Verification of Data-Consistency Properties, by Nicholas A. Kidd University of wisconsin– Madison, 2009.
[2] Privacy-Preserving Public Auditing for Secure Cloud Storage Cong Wang, Student Member, IEEE, Sherman S.- M. Chow, Qian Wang, Student Member, IEEE, Kui Ren, Member, IEEE, and Wenjing Lou, Member, IEEE.
[3] Learning with Local and Global Consistency Dengyong Zhou, Olivier Bousquet, Thomas Navin Lal,Jason Weston, and Bernhard Sch¨olkopf Max Planck Institute for Biological Cybernetics, Tuebingen.
[4] Service Level Agreement (Sla) Based Scheduling Algorithms For Wireless Networks Mehdi Alasti, Farrokh R. Farrokhi, Masoud Olfat, And K. J. Ray Liu Electrical And Computer Eng. Department,University Of Maryland, College Park.
[5] Benchmarking Eventual Consistency: Lessons Learned From Long-Term Experimental Studies David Bermbach Karlsruhe Institute of Technology, Karlsruhe.
[6] P. Mell and T. Grance, "The NIST definition of cloud computing,"NIST Special Publication, 2011.
[7] W. Lloyd, M. Freedman, M. Kaminsky, and D. Andersen, "Don't settle for eventual: scalable causal consistency for wide-area storage with COPS," in Proc. 2011 ACM SOSP.
[8] Consistency as a Service: Auditing Cloud Consistency Qin Liu, Guojun Wang, IEEE, Member, and Jie Wu,IEEE Fellow IEEE Transactions On Network And Service Management, 2014.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY