ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Two Level Auditing Framework in a large Cloud Environment for achieving consistency as a Service.

Aruna V
MTech Student
Department of CSE
St.Peter's Engineering College,
Hyderabad, TS, INDIA

N Mahipal Reddy
Assistant Professor
Department of CSE
St.Peter's Engineering College,
Hyderabad, TS, INDIA

J Vijay Gopal
Assistant Professor
Department of CSE
St.Peter's Engineering College,
Hyderabad, TS, INDIA

Abstract:
Cloud storage is a model of data storage where the digital data is stored in logical pools, the physical storage spans multiple servers (and often locations), and the physical environment is typically owned and managed by a hosting company. Most of the cloud service provider provides services like infrastructure management, data storage services on 24/7 through any devices at anywhere. To provide this ubiquitous always on service most of the cloud service provider (CSP) maintains each piece of data on geographically distributed servers. The main key problem with this technique is that, it is very expensive and some to fail to provide required consistency of service. To overcome this problem, we propose to use a new approach of service (i.e. Consistency as a Service (CaaS)) this paper, firstly concentrate on a consistency as a service (CaaS) model, which has a large data cloud and multiple small audit clouds. In the CaaS model, a data cloud is formed by a CSP, and a group of users form an audit cloud that can verify whether the data cloud provides the promised level of consistency i.e. quality of service or not, for that make use of two-level auditing strategy which require loosely synchronized clock for ordering operations in an audit cloud. Then perform global auditing by global trace of operations through randomly electing an auditor from an audit cloud. Finally, use a heuristic auditing strategy (HAS) to display as many violations as possible.

Keywords:
Cloud Service Provider (CSP), Consistency as a Service (CaaS), Heuristic Auditing strategy, Service Level Agreement, User Operation Table, Directed Acyclic Graph, Network Time Protocol.

Introduction:

Cloud computing relies on sharing of resources to achieve coherence and economies of scale, similar to a utility (like the electricity grid) over a network. At the foundation of cloud computing is the broader concept of converged infrastructure and shared services.

Cloud computing, or in simpler shorthand just "the cloud", also focuses on maximizing the effectiveness of the shared resources. Cloud resources are usually not only shared by multiple users but are also dynamically reallocated per demand. This can work for allocating resources to users. For example, a cloud computer facility that serves European users during European business hours with a specific application (e.g., email) may reallocate the same resources to serve North American users during North America's business hours with a different application (e.g., a web server). This approach should maximize the use of computing power thus reducing environmental damage as well since less power, air conditioning, rack space, etc. are required for a variety of functions. With cloud computing, multiple users can access a single server to retrieve and update their data without purchasing licenses for different applications.

Cloud computing is the lease of the resources through which the users can use the resources depending upon the requirement and pay based on the usage. Trough cloud computing the user can decrease the cost and can use the resource at any time. There are three types of cloud as shown in fig1

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

i) Public cloud
ii) Private cloud
iii) Hybrid cloud

Public cloud:

Public cloud or external cloud is one in which the resources are leased on self service basis over the internet, via web applications/web services, from an off-site third-party provider who shares resources and bills on a fine-grained utility computing basis.

Private cloud:

Private cloud or internal cloud is used to describe the offerings of private network.

Hybrid cloud:

Hybrid cloud is one which contains multiple internal or external clouds. AMES is based on platform as a service. Platform as a service (PaaS) is a category of cloud computing services that provides a computing platform and a solution stack as a service. Along with software as a service (SaaS) and infrastructure as a service (IaaS), it is a service model of cloud computing. In this model, the consumer creates the software using tools and/or libraries from the provider. The consumer also controls software deployment and configuration settings. PaaS offerings facilitate the deployment of applications without the cost and complexity of buying and managing the underlying hardware and software and provisioning hostingcapabilities.

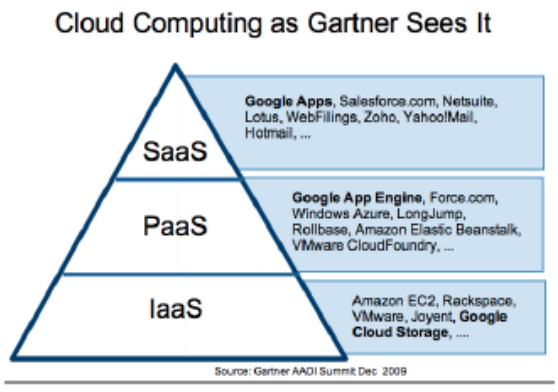

Fig 1 : Types of services

Clouds computing has become more popular choice, because it has succeeded in giving guaranteed basic services like virtualized infrastructure system and providing data storage, etc. e.g. Amazon, SimpleDB are example of such systems. The customers or end users by making use of these services, become authorized users and able to access the data from anywhere and at any time using any device and getting confidence that the capital investment is going to less. The cloud service provider popularly known as CSP promising the users data is going to be available as 24/7, and they can access it efficiently.

The CSP stores the different copies of data in a distributed fashion on different servers, which geographically present in different places. The main issue with distributing multiple copies of data called as replication technique is resultant into a very expensive process to provide strong consistency operation. In the coming days user is assured to see the latest updates about this service or operation.

Many cloud service providers provide week consistency, we call such consistency as eventual consistency, where a user can read the data for particular time. Now-a-days stronger consistency assurance is getting importance. Consider the following figure. In the above figure data is stored in multiple copies on five cloud servers (CS1, CS2…, CS5), users specified in the figure share data through a cloud storage service. Here the cloud should provide casual consistency service, where a user Alice uploads a data on the cloud server CS4. Here the user update should be reflected in all the servers. If cloud service provider provides only eventual consistency then receiver user is going to receive the old version of data. Such a integrated design based on traditional version may not satisfy customer requirements.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

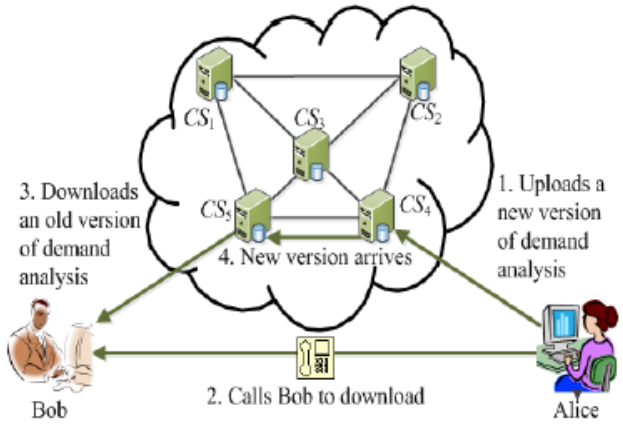

Fig. 2. An application that requires causal consistency

Hence we can conclude that different applications need different level of consistency operation. In this work, we propose CaaS model as a ideal consistency operation, which the applications of today's date are demanding. The standard CaaS model consists of large cloud data created by CSP and complete the operation scheme it contains many audit clouds which are formed by authorized group of users. These group of users working on a project and making a document, which constantly checking whether data cloud offers a guaranteed level of consistency or not. This standard model incorporating two-level auditing structure, which involves a synchronized clock assigning tasks to audit cloud and then performing global auditing with global trace of standard operations periodically by auditor chosen from an audit cloud. Local auditing is going to be performed and focuses on continuous read and read-your write process consistencies, which is going to be done online by a light-weight algorithm, while global auditing operation focuses on casual consistency, where it constructs a directed graph.

Whatever graph has been constructed, if it directed acyclic graph, which is also called as precedence graph, we can say that casual consistency is maintained. We confirm the severity of violations by calculated two metrics for the standard CaaS model: One is called commonality of violations and other is staleness of value of read operation. Finally in this work, we propose a approach called Heuristic Auditing Strategy(HAS), which proves cloud consistency and required cost i.e. actual cost per transaction. The two level auditing structures basically contain 2 auditing types: 1. Local Auditing 2. Global Auditing 1. Local Auditing: structure each user can perform local auditing with local trace operation periodically .this auditing focuses on monotonic read and read your write consistency .which can be perform by light-weight online algorithm the local auditing algorithm is online algorithm. 2. Global Auditing: the auditor can be selected from audit cloud .the main works of the auditor is to perform global auditing with global trace operation .this auditing focuses on causal consistency because causal consistency perform by constructing directed graph .the directed acyclic graph is constructed then causal consistency is obtain .Finally we propose analytical auditing strategy which appropriate reads to reveal many unsuccessful result.

Although the infrastructures under the cloud are much more powerful and reliable than personal computing devices, they are still facing the broad range of both internal and external threats for data integrity.Second, there do exist various motivations for CSP to behave unfaithfully toward the cloud users regarding their outsourced data status.

In particular, simply downloading all the data for its integrity verification is not a practical solution due to the expensiveness in I/O and transmission cost across the network. Besides, it is often insufficient to detect the data corruption only when accessing the data, as it does not give users correctness assurance for those un accessed data and might be too late to recover the data loss or damage.

Encryption does not completely solve the problem of protecting data privacy against third-party auditing but just reduces it to the complex key management domain. Unauthorized data leakage still remains possible due to the potential exposure of decryption keys.

Related Work:

As per our survey there are many previously worked done in the field of cloud data consistency. In paper

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[1] they reduce that confusion by clarifying terms, providing simple figures to quantify comparisons between of cloud and conventional Computing, and identifying the top technical and non-technical obstacles and opportunities of Cloud Computing. We believe the only plausible solution to very high availability is multiple Cloud Service Providers (CSP).We predict Cloud Computing will grow, so developers should take it into account. Regardless whether a cloud provider sells services at a low level of abstraction like EC2 or a higher level like AppEngine, we believe that computing, storage and networking must all focus on horizontal scalability of virtualized resources rather than on single node performance. In paper [2] they provide the better way to store any file on cloud storage. A key contribution of COPS is its scalability, which can enforce causal dependencies between keys stored across an entire cluster, rather than a single server like previous systems. But storing a file on a cluster is producing huge problem for providing consistency. Measuring consistency is a very important task in this system because monitoring and controlling consistency is major goal of proposed system. By the time various benchmarking techniques are offered.

In [3] they are providing whole new perspective to see the need of consistency as a service (CaaS).In cloud computing storage services, every service request has an associated cost. In particular, it is possible to assign a very precise monetary cost to consistency protocols (i.e., the number of service calls needed to ensure the consistency level times the cost per call). Therefore, in cloud storage services, consistency not only influences the performance and availability of the systems but also the overall operational cost. In [4] to know the consistency models we have studied local and global consistency model of Dengyong Zhou, Olivier Bousquet, and Thomas Navin Lal. 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.

S. Esteves [15] has highlighted the work on the critical data information, which is stored in cloud data centers across the globe, and getting increased in great way. And they are using different replication methods or approaches to deliver high-availability of services, demand of high performance, and mainly to maintain the consistency among multiple copies of data i.e. replicas. The proposed technique targets or focuses on data stored in tabular format, provides rationalization of resources, here bandwidth means bandwidth and which also requires improvement in the QOS parameters like latency value, performance in the network and availability of resources.

H. Wada [16] has introduced a new class of data storage systems, called NoSQL (Not Only SQL), has emerged to complement traditional database systems, with rejection of general ACID transactions as one common feature. Here the authors have brought a new area of for study, where a new class has been introduced in the data storage system, which is called as NoSQL (Not Only SQL), which have been introduced to support the traditional or general classic database systems, where it come across the removal of ACID transactions properties from these systems, considered as the one common general feature.

M. Rahman [18] highlighting the study of storage systems, these systems with large-scale key-value standard storage systems, compromise with consistency for the interest of dependability, i.e. availability of resources and partition tolerance systems, as well as performance of network, considering latency parameter. The system under this study provides eventual consistency, which is difficult to implement in real time systems. The authors have attempted to measure such consistency empirically, but these systems suffer from some drawbacks. But

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

their accuracy has been limited due to some state-of-the are systems considering consistency benchmarks.
D. Kossmann [19] discussed about now-a-days the cloud computing systems have many advantages for deploying applications in real time systems, example for such applications are data-intensive applications. Under this system, the user follows pay-as-you-go service model, under which user pay to the services, which he has used. The system provides a promising service with reduced cost. Another promising feature service to add here is, the system provides unlimited throughput by adding servers. They have focused on the transaction processing work, such as read and update workloads, instead of on the other processing like analytics operation or OLAP workloads operation.

E. Brewer [3] proposed the work on common current distributed systems, even the ones sytem that work, tend to be very generally fragile: they are very hard to keep up, very hard to manage, hard to grow, very hard to evolve, and very hard to program. Here in this talk, the author looks at several problems in an attempt to clear the way we think about these general systems. These problems include the general fault model, very high availability, considered graceful degradation, specific data consistency, specifying evolution, composition operation, and autonomy process.

L. Lamport [20] has introduced the concept of formalism for generally specifying and particular reasoning about concurrent systems has been described. It is not going to be based upon specific atomic actions. A general definition of a higher-level system is given and also justified correspondingly. And considering in Part II, the generic formalism is going to be used to specify several specific classes of inter-process communication and algorithms have been proven to be correct for implementing them.

Scope of The Project

The scope of this project is to upload and download a file from cloud. While providing cloud consistency, the following objectives are to be met:

1] Understanding the novel consistency as a service (CaaS) model provided by the cloud service provider.
2] The cloud computing solution should provide basic consistency as service.
3] Maintain synchronized clock at audit clouds that responsible for checking weather cloud provide promised consistency or not.
4] Service Availability.

Existing System

- By using the cloud storage services, the customers can access data stored in a cloud anytime and anywhere using any device, without caring about a
- large amount of capital investment when deploying the underlying hardware infrastructures.
- The cloud service provider (CSP) stores data replicas on multiple geographically distributed servers.
- Where a user can read stale data for a period of time. The domain name system (DNS) is one of the most popular applications that implement eventual consistency. Updates to a name will not be visible immediately, but all clients are ensured to see them eventually.

Disadvantages:

- The replication technique in clouds is that it is very expensive to achieve strong consistency.
- Hard to verify replica in the data cloud is the latest one or not.

Proposed System

- In this paper, we presented a consistency as a service (CaaS) model and a two-level auditing structure to help users verify whether the cloud service provider (CSP) is providing the promised consistency, and to quantify the severity of the violations, if any.
- With the CaaS model, the users can assess the quality of cloud services and choose a right CSP among various candidates, e.g, the least

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

expensive one that still provides adequate consistency for the users' applications.

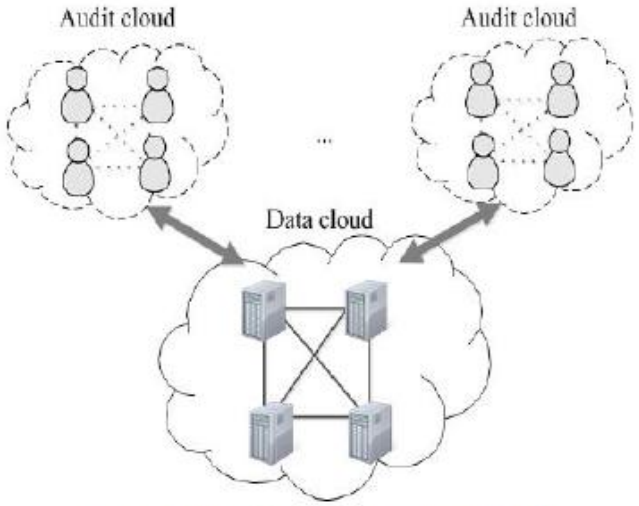

Fig. 3. Consistency as a service model.

Advantages :

- Do not require a global clock among all users for total ordering of operations.
- The users can assess the quality of cloud services.
- choose a right CSP
- Among various candidates, e.g, the least expensive one that still provides adequate consistency for the users' applications.
- As a rising subject, cloud consistency is playing an increasingly important role in the decision support activity of every walk of life.
- Get Efficient Item set result based on the caas.

Implementation

1. System Module:

In the first module, we develop the System Module with User Module, Admin Module, Auditor Module and data Upload Module. In user module, the authorized user should undergo registration operation andregister their detailed information and collect the secret key information for login operation and user can able to upload the file regarding the operation like auditing. Next in user module, system user already uploaded files can be stored in system cloud database in systematic manner. Next Auditor can view the file or locate the file present in the database and it can be very secured. In admin module admin will be able to view all the register user details; and also user uploads information details, and third party TPA activities regarding the auditing strategy. In auditor module, the generally selected auditor can do the auditing operation based on the strategy operation called as heuristic auditing strategy. This operation is related to the basic operation of document verification. Then special unit Auditor can collect and check the auditing file, the he decide whether to reject or accept the file. In the Data module, the user uploaded files can be stored in cloud database. It can be very secure auditor can view the file from the database it can be very secure.

On this he is going to make a report and enter all the details and about the decision like whether it's good or bad. Also in this module an auditor can submit revision report. In this report information like accept or waiting. After the decision if status is present as a accept means then the user can view the file else if the status is waiting condition means then user cannot view the particular file.

2.User Operation Table: In this module each user is going to maintain a UOT for the operation of recording local operations in systematic manner. Then each present record in the UOT unit is shown or explored by three components: first is an operation parameter, next the present logical vector, and finally more important one is physical vector. When the user is working on any operation, he is going to record his complete activity and also the current logical vector and final value as a physical vector, in his own UOT. In this module each user of the system is going to maintain a special logical vector and a basic physical vector to track the complete logical and physical time when an operation is going to take place correspondingly.

3. Local Consistency Auditing: Local consistency auditing technique is an online algorithm. The operation in this module or unit, in which each user is

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

going to record all his complete activities and store in his UOT. During the read operation, the authorized user is going to perform local consistency operation in an independent manner.

4. Global Consistency Auditing: Global consistency auditing technique is going be considered as an offline algorithm. Next to consider is that an auditor periodically will be selected from the audit cloud system to perform the special operation like global consistency auditing technique. Hence in this case the auditor is going to collect all users' UOTs for obtaining a special global trace of all activities. Then later executing global auditing technique, selected auditor is going to send results of auditing operation as well as its vectors values to all other authorized users . Now given the auditor's vector values, then each user will come to know other users' new clocks up to next global auditing.

Conclusion

In this paper, The presented system is a consistency as a service (CaaS) model and a two-level auditing scheme to help users validate whether the cloud service provider (CSP) is providing the promised consistency, and to enumerate the occurrences of the violations. The CaaS model used in the system helps the users can assess the superiority of cloud services and decide a right CSP among various services. For example the less costly one that still provides satisfactory consistency for the users' applications.

References:

[1] Qin Liu ; Changsha, China ; Guojun Wang ; Jie Wu, Consistency as a Service: Auditing Cloud Consistency, IEEE Transactions on Network and Service Management, (Volume:11 , Issue: 1)

[2] P. Mell and T. Grance, "The NIST definition of cloud computing (draft)," NIST Special Publication 800-145 (Draft), 2011.

[3] E. Brewer, "Towards robust distributed systems," in Proc. 2000 ACM PODC.

[4] "Pushing the CAP: strategies for consistency and availability," Computer, vol. 45, no. 2, 2012.

[5] M. Ahamad, G. Neiger, J. Burns, P. Kohli, and P. Hutto, "Causal memory: definitions, implementation, and programming," Distributed Computing, vol. 9, no. 1, 1995.

[6] 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.

[7] 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.

[8] C. Fidge, "Timestamps in message-passing systems that preserve the partial ordering," in Proc. 1988 ACSC.

[9] W. Golab, X. Li, and M. Shah, "Analyzing consistency properties for fun and profit," in Proc. 2011 ACM PODC.

[10] A. Tanenbaum and M. Van Steen, Distributed Systems: Principles and Paradigms. Prentice Hall PTR, 2002. [11] W. Vogels, "Data access patterns in the Amazon.com technology platform," in Proc. 2007 VLDB.

12] "Eventually consistent," Commun. ACM, vol. 52, no. 1, 2009.

[13] T. Kraska, M. Hentschel, G. Alonso, and D. Kossmann, "Consistency rationing in the cloud: pay only when it matters," in Proc. 2009 VLDB.

[14] 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.

15] 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.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[16] D. Bermbach and S. Tai, "Eventual consistency: how soon is eventual?" in Proc. 2011 MW4SOC.

17] M. Rahman, W. Golab, A. AuYoung, K. Keeton, and J. Wylie, "Toward a principled framework for benchmarking consistency," in Proc. 2012Workshop on HotDep.

[18] D. Kossmann, T. Kraska, and S. Loesing, "An evaluation of alternative architectures for transaction processing in the cloud," in Proc. 2010 ACM SIGMOD.