International Journal of Engineering & Technology, 7 (1.1) (2018) 223-225

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Techniques of providing data integrity in cloud computing

K David Raju, K Vijay Kumar *, K Anthony Rahul Showry, B Lohit Krishn

Computer science and engineering, Koneru Lakshmaiah educational foundation
**Corresponding author E-mail: vijaykondetikumar@gmail.com*

Abstract

The Data Integrity is simply termed as no corruption in the data that can be assured with consistency and accuracy over the time Precisely it can be defined as the data should be recorded as the Original and at the time of retrieval it should ensure that it send the Original recorded data. Data Integrity is the fundamental component of Information Security. Every technique of data integrity ensures the no loss in flow of data. We start with briefing about Data Integrity and Cloud Computing and then briefing Data Integrity models. After this, we examine General strategies that guarantee information uprightness, Challenges in Cloud Computing, Techniques in cloud to guarantee information honesty to be specific provable information ownership and verification of retrievability and their disadvantages and their confinements over some particular cases. This paper is Standardized investigation of existing system for the guaranteeing the information trustworthiness in cloud and another strategy is proposed. This paper is Standardized analysis of existing mechanism for the ensuring the data integrity in cloud and a new method is proposed.

Keywords: *Data Integrity; Retrievability; Computing; Trustworthiness.*

1. Introduction

Data Integrity is the basic key component in acquiring the Information Security. The Data Integrity is simply termed as no corruption in the data that can be assured with consistency and accuracy over the time. Precisely it can be defined as the data should be recorded as the Original and at the time of retrieval it should ensure that it send the original recorded data. Every technique of data integrity ensures the no loss in flow of data. Cloud Computing is the latest and present trending envision architecture of IT Enterprise. It increases the capacity and add capabilities to the objective what industries are in needed. Many were following and developing in Cloud. The main problem is the user has to take his own risk to keep sensitive data in the cloud. The Cloud Service Provider can change or Delete the information without recognizing of the customer. There are numerous strategies are accessible yet having numerous confinements and downsides in the current procedures. The mapping of the client to the specialist organization as follows in three parts [1].

2. Challenges/issues found on cloud

Though having many advantages it also having many concerns in the cloud. The issues were stated below [2].
Accessibility: Information should be available for clients all over the time. There shouldn't be any issues that would lead to data storage problem and leads to the crash/loss of user data.
Network Load: The over load capacity may result in fail of data integrity. There will be problem in transfer of Information between systems and servers.
Integrity (No Corruption): Consistency and accuracy of the information is threatened with the loops having in the cloud techniques
Data Location: Some of the storages follow will be like Centralized storage method. If it fails, there will be no chance of retrieval of data.

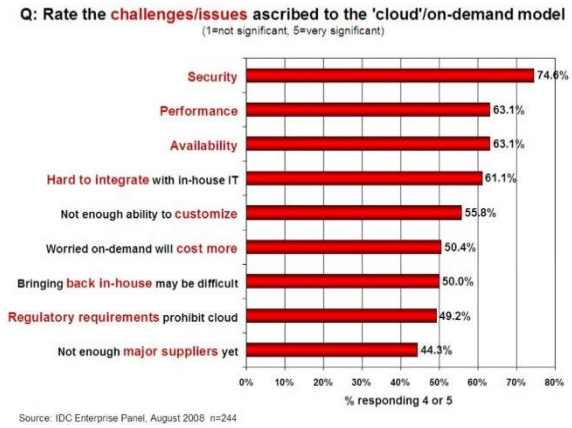

Fig. 1: Challenges and Issues Found on Cloud [3].

3. General techniques used to maintain data integrity

a) Generating hashes

Comparing the hash values can check/verify the uniformity of data. A hash value also known as message digest. The hash value is calculated based on the chosen mathematical function. The input will be the length of the string that as to be transmitted. Some techniques like sha and md5 are used to generate hashes and verify integrity. This was the basic and common methodology to ensure the client's data integrity

b) Using Trusted Third parties (TTP)

Copyright © 2018 K David Raju et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Trusted Third Parties (TTP) like are the supporting vendors that take care of our data transmissions. We can fully rely on them. The existing were more secure but if we go with new TTP it may have some risks. It is secure and more expensive. Some of TTPs are VISA, Bradstreet,Banks etc.

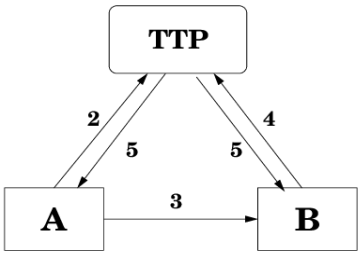

Fig. 2: Using TTP.

4. Techniques in cloud to ensure integrity

There are few techniques that are better and more secure with some drawbacks/limitation so far that could maintain the Stability of information in the online storage. The fundamental procedure for information consistency in cloud are Proof of Retrievability (POR) and Provable Data Ownership Possession (PDP) that are most normally used for ensuring data dependability.

4.1. Provable data possession (PDP)

It assures no occurrence of corruption of data even the data stored in unfaithful storage. It is done with the remote server. It can check the data in the storage without retrieving it.
The principal behind PDP involves in 2 stages [4].
Setup Stage:

- Setup Stage Pair of coordinating keys are produced i.e secret and open keys with utilization of probabilistic key Generating Algorithm Open key nearby the record will be transfer to the server for limit by user and customer removes the report
- Open key alongside the record will be sent to the server for capacity by customer and client erases the document.

Challenge Stage:

- The customer challenges for a proof of ownership for a subset of the pieces in the document.
- The customer verifies the reaction. Fig 3.

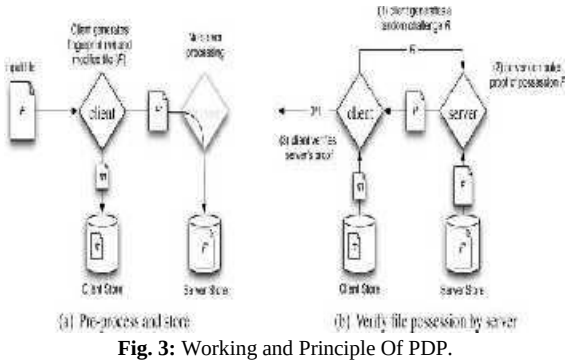

Fig. 3: Working and Principle Of PDP.

Limitations:
Absence of mistake revising codes to address worries of debasement

- Lack of security conservation.
- Boundless number of questions

4.2. PDP based on MAC

Guarantee information trustworthiness of document F put away on distributed storage in extremely straightforward way. The information proprietor registers a Message Authentication Code (MAC) of the entire record with an arrangement of mystery keys and stores them locally before outsourcing it to CSP. It keeps just the registered MAC on this nearby stockpiling, sends the record to the CSP, and erases the neighborhood duplicate of the document F. At whatever point a examiner needs to check the Data respectability of record F, Person sends a demand to recover the document from CSP, uncovers a mystery key to the cloud server and solicits from the entire record, and contrasts the re-figured and the beforehand put away esteem [4].
Limitations:

- The information proprietor needs to recover the whole document of F from the server keeping in mind the end goal to process new MACs, which isn't workable for huge record.
- Public auditability isn't upheld.

4.3. Scalable PDP

Author in [4] proposed Scalable PDP which is an enhanced variant of the first PDP. The primary contrast is Scalable PDP utilizes while unique PDP utilizes open key to lessen calculation overhead. Adaptable PDP can have active operation on distant information. Versatile PDP has every one of the difficulties and answers are preprocessed and predetermined number of updates. It depends on the symmetric-Key which is more effective than open Key encryption. So it doesn't give open obviousness.
Limitations:

- Doesn't functions square inclusions; just affix write additions are conceivable.
- This plan is risky for extensive documents as each refresh requires re-making all the rest of the difficulties

4.4. Proof of retrievability (POR)

POR [4] is strategy without keeping a duplicate of the client's unique records in nearby capacity. In a plan, client reinforcements his information document together with some confirmation information to a conceivably deceptive distributed storage server. Client can verify the information for its corresponding put away with CSP utilizing confirmation key.
Principal of POR:
Author in [4] proposed Scalable PDP which is an enhanced variant of first PDP. The primary contrast is Scalable PDP utilizes the symmetric encryption while unique PDP utilizes open key to lessen calculation. Versatile PDP has every one of the difficulties and answers are pre-processed and predetermined count of updates. Versatile PDP does not require mass encryption. It depends on the symmetric- Key is more effective than open Key encryption. It doesn't offer open obviousness.

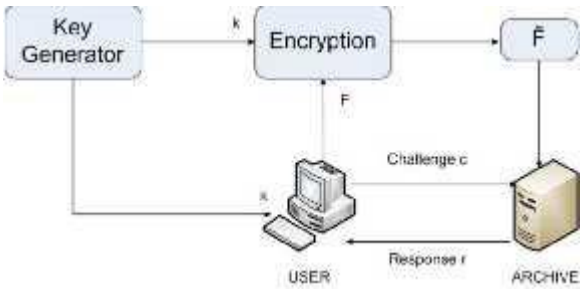

Fig .4: (Schematic View of POR) [5].

Limitations:

- A customer can perform set number of updates and difficulties.
- It doesn't perform square inclusions; just affix write additions are conceivable.
- This plan is risky for extensive documents as each refresh requires re-making all the rest of the difficulties.

4.5. High availability and integrity layer (HAIL)

Proposed HAIL [4] distributed storage, in which HAIL enables the client's Information on various servers so there is a repetition of the information. Basic central of this strategy to guarantee information uprightness of record through information repetition. HAIL utilizes message verification codes (MACs), the pseudorandom capacity, and all-inclusive hash capacity to guarantee trustworthiness process. The evidence is produced by this strategy is autonomous size of information and it is minimized in estimate.
Limitations:

- This system is relevant for the static information as it were.
- It requires more calculation control.
- Not reasonable for thin customer.

5. Drawbacks in existing techniques

- Lack of security conservation.
- Doesn't performs perfect squared inclusions; just affix write additions are conceivable.
- This plan is risky for extensive documents as each refresh requires re-making all the rest of the difficulties.
- The information proprietor needs to recover the whole document of F from the server keeping in mind the end goal to process new MACs, Which isn't workable for huge record.
- Public auditability isn't upheld as the private keys are required for confirmation.

Proposed Model
If a person stores a file in the cloud, once if they want to retrieve the file back from the cloud then they need to verify the file whether the file retrieved is matching with the file what they have sent or did it get corrupted. This should be confirmed by the client. Demonstrate we are proposing is utilizing an outsider rather than the customer's PC or framework. Here, when the client transfers the record the cloud then the document is put away in an outsider (trusted) and after that the trusted outsider produces the hash of the document sent by the customer and the hash will be put away inside the trusted outsider and when client needs to recover the record back to their framework then first the cloud sends the record to the outsider first and what happens is the outsider again creates the hash for the sent record and utilizing a similar hash work. Then, the trusted third party verifies if the currently generated hash matches with the antecedently generated hash. In the event that the hash value the record is in place is matched then respectability is guaranteed and if the hash does not match then the result would be negative. This is how data] integrity is verified and ensured used this model. Creating hashes should be possible with any mechanisms like md5, sha-512 and whatever other systems which are utilized to produce hashes. Making hashes is the primary feasible way to deal with check data respectability and to give the essential administrations to the customer. The client or customer before sending the record to the outsider hosts to validate to the third get together and send solicitations to the outsider for transferring the documents into the cloud. At that point the outsider sends solicitations to the cloud saying that the client has validated and store this document into the cloud.

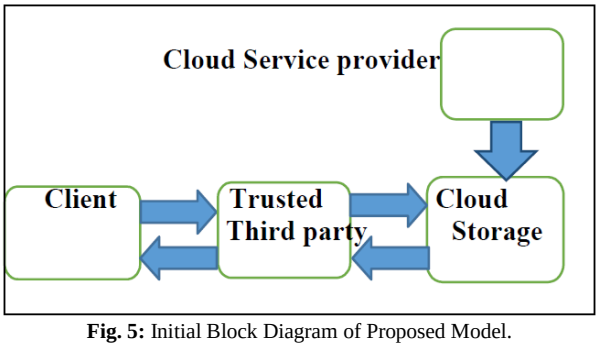

Fig. 5: Initial Block Diagram of Proposed Model.

References

[1] M. Armbrust A Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski,G . Lee, D. Patterson, A. Rabkin, I. Stoica, et al., "Above the clouds: ABerkeley view of cloud computing," University of California, Berkeley, Tech. Rep, 2009
[2] Chandran S. and Angepat M., "Cloud Computing: Analyzing the risks involved in cloud computing environments," in Proceed-ings of Natural Sciences and Engineering, Sweden, 2010.
[3] Minqi Zhou, Rong Zhang, Wei Xie, Weining Qian, Aoying Zhou "Security and Privacy in Cloud Computing: A Survey" in 2010 Sixth International Conference on Semantics, Knowledge and Grids. https://doi.org/10.1109/SKG.2010.19.
[4] Mahesh S.Giri, Bhupesh Gaur, Deepak Tomar "A Survey on Data Integrity Techniques in Cloud Computing" in International Journal of Computer Applications.
[5] Sravan Goud Utkam,David Raju Kuppala, Amudhavel J, Raviteja Parasa," A Secured Symmetric Key Enccryption Technique Using Images as Secret Keys" International Journal of Pure and Applied Mathematics, Volume 116 No. 6 2017, 149-153.