ISSN 2347 - 3983
Volume 8. No. 9, September 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter53892020.pdf
https://doi.org/10.30534/ijeter/2020/53892020

Public Auditing Mechanism toVerify Data Integrity in Cloud Storage

Mattapalli Anil Kumar[1], Dr.Prasadu Peddi[2],Dr.P.M. Yohan[3]
[1]Research scholar, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan.anilkumar11282@gmail.com
[2]Assistant Professor, Dept of CSE, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan.
[3]Professor & Principal of CSI wesley Institute of Technology and Sciences, Secunderabad, Telangana.

ABSTRACT

Cloud remote servers. It gives users a flexible approach to access data that are outsourced remotely via the internet. Cloud users can manage their data without deploying and maintaining storage servers and devices in the local system. Data integrity and confidentiality are the two most vital security concerns over unreliable cloud service providers (CSP). However, the management of data and services by CSP may not fully be trusted by industries/users. While the integrity of data checks whether data stored in cloud server remains intact, confidentiality of data ensures whether the stored data is secure and not disclosed to anyone. In this paper, an approach for ensuring the integrity data in cloud storage server proposed to strengthen the trust of cloud service provider, because of cloud service provider might delete rarely accessed file of users or may try to hide data loss incidents to maintain their reputation and claim that there is no data loss.

Key words: Data integrity, auditing, CSP, TPA.

1. INTRODUCTION

Now, cloud computing support continues to be broadly embraced by industries or associations. This monumental rise of cloud-storage offers industries or businesses to outsource their data and IT services into this cloud service provider. In accordance with recent poll [1] IT outsourcing has increased by79 percent since employers want to decrease cost and pay attention to building their Core Competencies (or) business. This data out sourcing is of good use to their users at feeling they are relieved by the load of storage administration and also maintenance of data. However, users need to rely upon their own providers for continued access to the data,Computing is just a technology which employs both the world wide web and remote central servers to keep software and info. Computing also systems makes it possible for organizations and users to utilize software and get their own files. This technology allows with centralizing memory storage, bandwidth and processing.

Cloud Storage Infrastructure

The cloud storage infrastructure consists of a large number of storage devices and it has four layers to operate their function in a sequential manner. The layer of the cloud storage infrastructure includes

1. Data storage layer
2. Data management layer
3. Application interface layer
4. User access layer

Data storage layer: It's a pool from the cloud-storage infrastructure which is made up of storage companies and apparatus of this cloud supplier. The info and also the services have been organized in a systematic manner to deliver the essential services with no interruption. The unified data direction is offered by incorporating the storage apparatus. The-status of these apparatus is tracked and scalability is achieved inside this coating by utilizing the distributed service- oriented storage approach. The complex operations for example virtualization and audience technology are performed. Additionally, it optimizes the resource usage in computing infrastructure that is cloud.

Data management layer: The data management layer is in charge of security, replication and report administration. Several apparatuses from data storage layer are all incorporated together to present the services with higher end. The distributed file system is controlled and maintained by the data management layer. It gives integrity, reliability and security towards this data that is stored. The procedures of information encryption and replication are performed within this coating. The backup surgeries which are the steps of tragedy retrieval process would be the principal activities of data management coating.

Application layer: The application coating functions as a frontend to those users in offering services throughout the World Wide Web. The end users of this cloud-storage infrastructure interact with the cloud-storage provider via the application layer. The access control mechanisms are offered by the application layer that can also be accountable for user authentication and approval. The flexibility of this cloud storage system is augmented by the combo of Program Interface (API) and the file systems of the storage apparatus.

User Access Layer: The cloud-storage services can gain access from anywhere with the consumer access coating. The consumer access coating supports several standalone devices like personal computers, laptops, palmtops and portable phones to gain access to the cloud-storage platform. The access list inside the application form interface coating is utilized to give use of the licensed users from the person access coating.

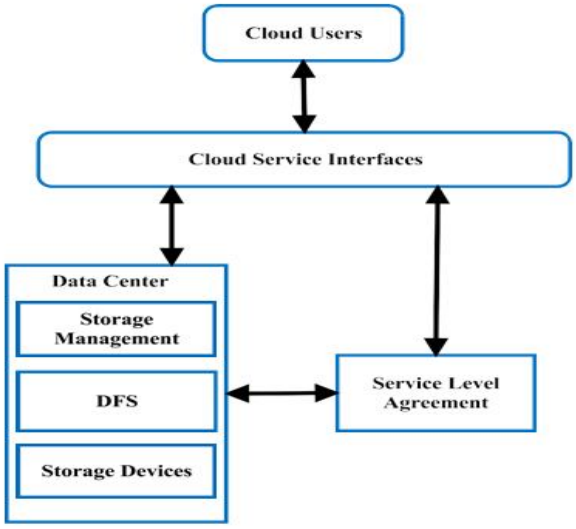

Figure 1: Cloud storage infrastructure

2. ROLE OF AUDITING IN CLOUD

Auditing [3] should aid in organizing and planning, acquisition and execution, delivery, service, observation and investigation of technology selection, regulatory compliance, selection and operation of third-party providers and providers [2]. Information system auditing tests should be utilized to examine confidentiality, data integrity, accessibility, security, authentication and reliability etc... It will take increasing responsibility and ensure value inclusion in key tactical domains such as customer connections, cost reduction, revenue maximization, fraud-detection, control and avoidance and data governance and caliber, keeping in pace with fast changing business environment and also how industry is carried out at a cloud agency environment. Auditing should concentrate on value inclusion by encouraging strategic initiatives, providing premium excellent business insights within a fundamental piece of the process and ought to also actively engage with continuous observation, evaluation and advancement of control environmental and regulatory compliance.

Need of Auditing in Cloud

Cloud doesn't need many regional resources and systems to store and process the info. Whilst the cloud offers an efficient means to store data for many users, most users like to save data from cloud server so as to reduce their own community server difficulties in storage and they also are able to save as-much data as you can with no constraint. However, the cloud transfer that the applying computer into centered in fact direction might trust worthy as you can find numerous security issues like obsolete ip-addresses etc. Auditing may be procedure of assessing and collecting evidence to determine if or not a pc system protects advantage, keeps data integrity and absorbs resource economically. As a way to keep and also verify [4] the ethics of data stored by the users in cloud, then there has to be an auditor to get auditing an individual's data stored at the cloud. You will find various measures which can be used with the intention of producing the data secured. To conserve the data stored at the cloud; so many approaches are proposed under different security and systems models. [7,8] recorded out the advantages and limitations of both auditing in computing. This is likely to soon be handy to researchers to designing a fresh auditing solution centered on the poll.

Types of Audit ability

Considering the role of the auditor, the auditing mechanism falls into two categories:

1. Private auditability
2. Public auditability

In cloud, even the customers themselves are undependable or might not be in a position to pay for the overhead of performing periodic ethics checks. A cloud user always simplifies and upgrades their own data; therefore, those energetic data also will need to be protected and have to be performed by people auditing schemes. In this phase, an effective process to look at their inside suggested and also this procedure uses public audit ability by introducing Trusted third-party Auditor (TTPA). Simultaneously users (U) desire to upgrade the data stored in cloud usually. A process has been introduced to carry out energetic operations like insertion, deletion and alteration with Index Table Construction (ITS). Even though strategies with private audit ability can reach increased strategy efficacy, people audit ability lets anybody, not merely the information proprietor to battle to get whilst confidential details. Then,

customers have the ability to assign the test of their service operation to a different TPA, without loyalty of these computation resources.

Design Goals of auditing

Public Auditability – To allow a third-party auditor (TPA) to check the correctness of cloud data at regular intervals (or) based on demand of the user without downloading the whole data.

Dynamic Operations - To Support dynamic data operation such as insertion, deletion and modification of data.

Timely detection - To be capable of detecting data losses (or) errors in outsourced storage and other abnormal behaviors etc.

Storage Correctness – To ensure that the integrity is verified by server, the one who is providing space for storing and maintaining data.

3. PRIVACY ISSUES IN CLOUD STORAGE

Privacy is considered as one major issues in data storage infrastructure of cloud. As cloud storage is a distributed and shared infrastructure, it is challenging to safeguard data privacy. The users "information's is exposed to unauthorized access that has an impact on sensitive data. The data privacy is violated by following issues such as

- Lack of user control
- Regulatory compliance

Other privacy issues include information disclosure, uncontrolled secondary storage and data proliferation, and dynamic provisioning.

Lack of user control: Once the info will be redirected cloud-storage machine, declines that control within info. This may possibly cause data theft and unauthorized access. The information owner doesn't need any knowledge about the positioning of process and storage from the cloud-storage machine. There are problems in data transmission round the states as a result of lively law enforcements in numerous nations. The battles in data storage and sharing can be worked out by the investigation of regulatory compliances. The alarms concerning the privacy breaches aren't intimated into the information operator as it may spoil the Standing of the service supplier.

Regulatory compliance: The dispersed cloud computing infrastructure stores data in various remote servers which can be found in different geographic locations. The legal limitations differ from place to set and thus it's hard to assign a specific host for use for data transmission systems at the boundaries of a place. The practice of confirming the information storage with the whole legislations from the planet is very dull. Replication of data eliminates the single point of collapse, but at precisely the exact same time that it creates sophistication of information control at distinct locations. Data proliferation and lively provisioning induces the sophistication in maintaining data solitude. Both information outsourcing and off shoring contributes to acute privacy Problems.

The privacy breaches occur in the public cloud environment due to the following reasons:

- Inadequate policies and rules
- Insufficient confidentiality and integrity constraints
- Lack of availability

Deficiency of accessibility: A highly efficient cloud-storage platform needs to provide 99.99% accessibility to the users. The deficiency of safe guard measures inside the cloud platform affects the information accessibility in the cloud-storage infrastructure. The protects for example hazard prevention and control from data disclosure supplies a high accessibility. Inadequate rules and policies: The strikes and also the intrusions aren't accurately discovered due to improper follow-up of polices filed into the cloud provider. There's not any straightforward definition of access and responsibilities functions of their users. The battle between the policy of this person and the provider contributes to deficiency of ethics. The processing of sensitive data besides the coverage's results in the solitude loss. The unauthorized entrance into this remote server ends in the debut of malwares and dangers from the cloud infrastructure.

4. PROBLEM ON HAND

Now-a-days businesses are moving towards keeping their data into cloud servers supplied by the cloud companies (CSPs), according to our present system data ownership in cloudcomputing systems agreements using stored data in a lively method into the host. Multicopy Means, an individual data to be reproduced to multiple servers. The consumer to manually upload the info to a host with mechanically initial data to shoot numerous copies afterward those backups is stored on multiple servers. Restoring the info in multi-server will be in order to stay clear of the data loss from server and Hacking wreck, and also the fee needed in keeping several copies of this data document increases. Currently auditor [12] will divide that data and also carry out double encryption on encoded chunks now stored on various servers.Our suggested system offers highend security and data integrity [9,10] with AES Algorithm. Our bodies additionally fixes modified data with ethics checking mechanism completed by Client and Reputable auditor,Throughout the recovery of this document it's re joined by the merging that the divide file parts and left as one apply for access.

5. RESEARCH ISSUES

1. After outsourcing the data to the cloud, the user no longer has the physical possession of his data, it requires data integrity protection.

2. Ateniese et al. and Juels and Kaliski proposed Provable data possession (PDP) model and POR model [11] for single server model.
3. Existing data integrity mechanisms are intended for private audit, only the data owner is allowed to verify the integrity.
4. The auditing schemes in existing imply the problem that users need to always stay online
5. There is no auditing mechanism for block level in distributed servers.

6. IMPLEMENTATION OF AUDITING

Initially file F to be outsourced is preprocessed and is divided into$n$ blocks.

$F = \{b_1, b_2, \ldots . b_n\}$ where $\{bi\} 1 \le i \le n$ (1)

We can further divide each block into s sectors

$bi = \{bi, 1, bi, 2, \ldots . bi, s\}$ (2)

Where $\sum_{j=1}^{s} m_i, j$in order to minimize the cost required for extra storage andspace.

Our scheme is categorized into two phases based on Index Table Structure (ITS). They are setup phase and proof verification phase. The setup phase is explained as below.

Key generation

The client generates a key pair (Ssk, Spk) for signing. Then it chooses a random value $y \leftarrow Zp$. From this, user computes $a \leftarrow g^y \in G$. The secret key is $SK = \{y, Ssk\}$ and the public keys are chosen as $PK = \{Spk, a, g, u\}$ where $g\ and\ u$ are random elements.

Signature generation

First, we compute file tag for the given file $F = bi\ (i = 1, 2, \ldots . , n)$.

The file tag can be used to identify a particular file during integrity verification. The tag can be calculated as

$б = filename||n||sigSsk\ (filename||n)$ (3)

where n is the number of blocks. $sigSsk\ (filename||n)$ denotes the file name and no of blocks are signed using secret key. SHA-512 algorithm has been used to generate a hash value. In addition to this, the user sends the following information (&, VERi, TSi) to the TTPA during setup phase. Then client computes signature (σ_i) for each block $m_i \epsilon Zp$ with the help of public key u.

$\sigma_i = (H(m_i||VER_i||Ts_i).u^{m_i||VER_i||Ts_i})^y$ (4)

Algorithm 1: Setup phase

1. $begin$
2. $data\ owner\ chooses\ a\ key\ pair\ (S_{sk}, S_{pk})\ for\ signing$
3. $choose\ a\ random\ value\ y \leftarrow Z_p$; compute $sa \leftarrow g^y \in G$.
4. $secrete\ key\ SK = \{y, S_{sk}\};\ public\ key\ PK = \{S_{pk}, a, g, u\}$
5. $compute\ file\ tag\ б = filename||n||sigSsk\ (filename||n)$
6. $send\ (б, VER_i, TS_i)\ to\ TTPA.TTPA\ adds\ it\ to\ ITS$
7. $For\ round\ block\ (i \leftarrow 1\ to\ n)\ do$
8. $begin$
9. Compute signaturetag $\sigma_i = (H(m_i||VER_i||Ts_i).u^{m_i||VER_i||Ts_i})^y$
10. end
11. $denote\ set\ of\ signature\ \Phi = \{\sigma_i\}, 1 \le i \le n$
12. $send\ \{F, \Phi = \{\sigma_i\}_{1 \le i \le n}, б\}\ to\ CSP$
13. end

Let Φ denote the set of signatures for all the blocks as $\Phi = \{\sigma_i\}, 1 \le i \le n$. Then client sends the following information $\{F, \Phi = \{\sigma_i\}_{1 \le i \le n}, б\}$ to cloud service provider (CSP). User deletes the file and corresponding signatures from the local storage.

7. PROOF OF CORRECTNESS VERIFICATION

TTPA verifies the validity of proof as follows, The TTPA verifies the sign using public key. If it is valid, then it computes challenged blocks hash values $H = \prod H\ (VERi, TSi)\ Xi$ and further TTPA verifies the validity ofproof by checking the equation given below,

$$R.e(\sigma^{\gamma}, g) = e((\prod_{i=s_1}^{s_c} H(m_i||VER_i||TS_i)^{x_i})^{\gamma}.u^{\text{ɱ}}, a)$$

It outputs TRUE if the equation holds, otherwise emits FALSE.The correctness of the verification equation can be verified as follows.

$$R.e(\sigma^{\gamma}, g)$$

$$= e(u, a)^r.e((\prod_{i=s_1}^{s_c} H(m_i||VER_i||TS_i).u^{(m_i||||VER_i||TS_i)x_i y})^{\gamma}, g)$$

$$= e(u^r, a).e((\prod_{i=s_1}^{s_c} H(m_i||VER_i||TS_i)^{x_i}.u^{(m_i||||VER_i||TS_i)^{\gamma x_i}}), g)^y$$

$$= e(u^r, a).e((\prod_{i=s_1}^{s_c} H(m_i||VER_i||TS_i)^{x_i})^y.u\ r^{\text{ɱ}}, a)$$

$$= e((\prod_{i=s_1}^{s_c} H(m_i||VER_i||TS_i)^{x_i})^y.u^{\text{ɱ}\gamma + r}, a)$$

$$= e((\prod_{i=s_1}^{s_c} H(m_i||VER_i||TS_i)^{x_i})^y.u^{\text{ɱ}'}, a)$$

Algorithm 2 will provide details steps of verification process.

Verification phase

1. $begin$
2. $TTPA\ generates\ challenge\ request$
3. $begin$
4. $TTPA\ construct\ a\ challenge\ message\ as\ a\ chal = (INDEX = \{INDEX_I\}\ 1 \le i \le c, S = \{S_1 | 1 \in INDEX\})$
5. $Send\ challenge\ message\ \text{“}challenge\text{”}\ to\ CSP$
6. end

CSP computes proof

1. $begin$
2. $\mathfrak{m} = \sum_{i=s_1}^{s_c} X_i m_i \; \Psi = \prod_{i=s_1}^{s_c} \sigma_i{}^{x_i}$
3. $choose\ a\ random\ value\ r \leftarrow Z_p$
4. $calculate\ R = e(u,a)^r.$ Apply blind factor $\mathfrak{m} = r + \gamma\mathfrak{m}$
5. $generate\ proof\ \S = \{\mathfrak{m}, \Psi, R\}$
6. $sign\ the\ proof\ and\ send\ \{\mathfrak{m}, \Psi, R, \S\}\ to\ TTPA$
7. end

TTPA verifies the proof

1. $checks\ the\ signature\ on\ proof\{\mathfrak{m}, \Psi, R, \S\}$
2. $if(Result = PASS)\ then$
3. $begin$
4. $checks\ Hash\ value\ H = \prod H(VER_i, TS_i)^{x_i}$
5. $verifies\ signature\ by\ checking\ Equation\ (4.8)$
6. $if(output == TRUE)\ emits\ PASS$
7. $else\ \{emits\ FALSE\}$
8. $end\ if$
9. $QUIT$
10. end

The workflow of integrity verification scheme is presented inFigure 2. It shows the three involved entities and the flow process betweenthe entities.

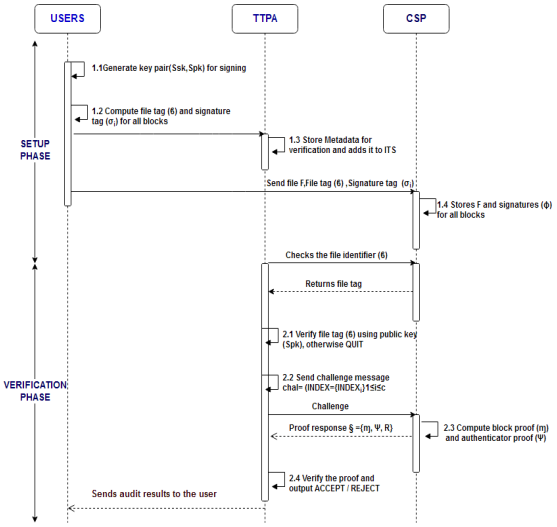

Figure 2: Integrity verification workflow

8. EXPERIMENTAL RESULT

We have conducted our experiment in a Eucalyptus environment. A private cloud has been setup using a eucalyptus open source infrastructure. This open source platform can be used to test IaaS (Infrastructure as a service) related cloud computing problems. We can experience the real cloud set up and benefits by conducting experiments in eucalyptus tool. We have installed eucalyptus fast start version 3.4.1 onCentos6 in an Intel core i5-3520 CPU at 2.2 GHz, 500 GB SATA drive and8 GB RAM. The auditing algorithms have been implemented in this set up and results are discussed here.

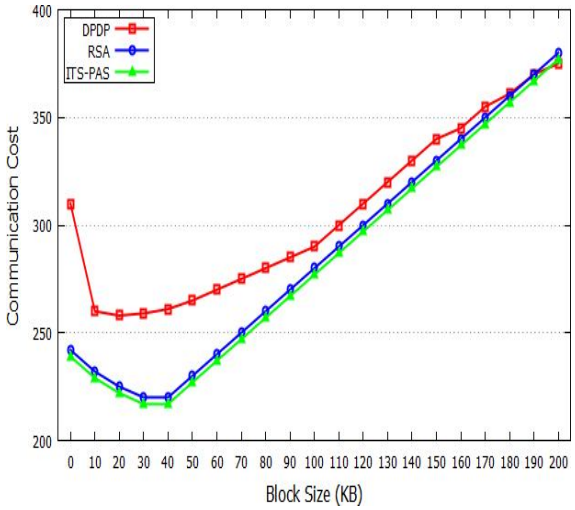

Figure 3: shows the communication cost involved in verification phase.

9. CONCLUSION

The customers or data owner should usually test their data outsourced to some distant server. Thus, this auditing action is assigned to Trusted Third Party Auditor (TTPA) that is capable of assessing the ethics proof data and confidentiality by running affirmation algorithm. TTPA performs auditing process at regular intervals to find the ethics violation of data as well as also the analysis document is delivered to the user. This sort of auditing is known as public auditing. Inside this paper we achieved public audit and lively data strategy encourage. Ergo, our proposed system works predicated on Index Table Construction (ITS) into perform auditing. The next phase addresses the support of energetic statistics operations in the cloud.

REFERENCES

1. Bowers & Opera, A (2009), "Hail: A high-availability and integrity layer for cloud storage", pp. 187-198.
2. Chang& Xu, Jia (2008), "Remote Integrity Check with Dishonest Storage Server", pp. 223-237.
3. Hao, Z, Zhong, S & Yu, N 2011, "A privacy-preserving remote data integrity checking protocol with data dynamics and public verifiability", Vol:23, Issue:9, pp:1432-1437.
4. Jianbing(2014), "On the security of an efficient dynamic auditing protocol in cloud storage", Vol:25, Issue:10, pp:2760 – 2761.

5. Prasadu peddi (2017), "Design of Simulators for Job Group Resource Allocation Scheduling In Grid and Cloud Computing Environments", vol 6, issue 8, pp: 17805-17811.
6. Prasadu peddi (2018), "Data sharing Privacy in Mobile cloud using AES", ISSN:2319-1953, Vol 7, issue 4.
7. Swaminathan & Baker (2008), "Privacy-preserving audit and extraction of digital contents", pp.1 – 21.
8. Wang, Ren, K & Lou (2010), "Privacy-Preserving Public Auditing for data Storage Security in Cloud Computing", pp: 522-533.
9. Wang, B, Li, B & Li, H 2015, "Panda: Public auditing for shared data with efficient user revocation in the cloud", Vol: 1, pp: 92-106.
10. Yuan, Jiawei & Yu, S 2015, "Public integrity auditing for dynamic data sharing with multiuser modification", Vol: 10, Issue: 8, pp: 1717-1726.
11. Yan Zhu et.al (2012), "Secure Collaborative Integrity Verification for Hybrid Cloud Environments", vol. 21, no. 3, pp. 165–.197.
12. Zhong &Nenghai Yu (2011), "A privacy-preserving remote data integrity checking protocol with data dynamics and public verifiability", vol.23, no.9, pp. 1432-1437.
13. Brian William Koloba Malubaya and Gunawan Wang (2020). "Real-time Parking Information System with Cloud Computing Open Architecture Approach", ISSN 2347 – 3983, Vol 8, N0 1, pp: 18-22.
14. K.JoseTriny, G.Anjuka, C.Dhanapal, S.Kavibharani ,C.Kowsaly (2019), "A Bigdata processing with Hadoop Map Reduce in Cloud Systems", ISSN 2347-3983, vol 8, No 3, pp: 752-759.
15. Amita Dhankhar ,Kamna Solanki (2019), "A Comprehensive Review of Tools & Techniques for Big Data Analytics", ISSN 2347-3983, vol 7, No 11, pp: 556-562.