Privacy Preserving Public Auditing and Data Integrity for Secure Cloud Storage Using Third Party Auditor

Ankush R. Nistane[1], Shubhangi Sapkal[2], Dr. R. R. Deshmukh[3]

[1]Department of CSE, Government College of Engineering, Aurangabad, India
[2]Department of MCA, Government College of Engineering, Aurangabad, India
[3]Department of CSE and IT, Dr BAMU, Aurangabad, India

Abstract – *Using cloud services, anyone can remotely store their data and can have the on-demand high quality applications and services from a shared pool of computing resources, without the burden of local data storage and maintenance. Cloud is a commonplace for storing data as well as sharing of that data. However, preserving the privacy and maintaining integrity of data during public auditing remains to be an open challenge. In this paper, we introducing a third party auditor (TPA), which will keep track of all the files along with their integrity. The task of TPA is to verify the data, so that the user will be worry-free. Verification of data is done on the aggregate authenticators sent by the user and Cloud Service Provider (CSP). For this, we propose a secure cloud storage system which supports privacy-preserving public auditing and blockless data verification over the cloud.*

Keywords – Blockless data verification, data integrity, cloud storage, third party auditor (TPA), privacy preserving, public auditing.

NOMENCLATURE

F	-	Data file is divided into blocks m_i ; $i \in \{1,2,\ldots n\}$
F_i	-	Set of files
m_i	-	i^{th} block of data file
h_i	-	Hash on block
Σ		Signature

I. INTRODUCTION

As we know cloud computing is booming nowadays. It is considered as the next generation information technology (IT) architecture for enterprises. Cloud service providers manage the data over the cloud. From users' perspective, storing data remotely to the cloud is beneficial, because it can be accessed on-demand and in a flexible way. It brings relief of the burden high level infrastructure that provides a scalable, secure and reliable environment for users at a much lower cost. Most of the cloud storage likes Google Drive and Dropbox offering space to the users which has become a routine for users to share for storage management, new and challenging security threats toward users' data. As the CSP are distributed, data outsourcing is actually relinquishing user's ultimate control over the fate of their data. As the data is stored in an untrusted cloud, it can be easily be lost or it can also get corrupted due to disasters or failures or human errors [1]. To verify the integrity of the data over the cloud, we introduce a third party auditor (TPA) for public auditing. TPA offers its auditing service with more powerful computation and communication abilities.

Privacy preserving public auditability has following advantages -

A. *Public Auditability* – It allows TPA to check integrity of data without retrieving it. TPA or external auditor should not have any knowledge about data i.e. blockless data verification.
B. *Storage Correctness* – User's data should correctly store on cloud.
C. *Privacy Preserving* – This ensures that TPA cannot derive any data content.
D. *Lightweight* – Auditing should be performed with minimum overhead.

Specifically, the contribution can be summarized as the following three aspects.

a. Our scheme achieves batch auditing where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA.
b. Our scheme provides a privacy-preserving auditing protocol.
c. Our scheme provides the better security and justifies the performance of our proposed schemes through concrete experiments.

II. LITERATURE SURVEY

There are many techniques that are used to provide security to the user's data over cloud. These techniques are also used for data correctness, data integrity and its security over the cloud. But there earlier techniques are not efficient to work on dynamic cloud and there are some disadvantages with these existing systems.

Following are some systems, with their pros and cons-

Ateniese et al. [2] is the one who took the public auditing into consideration, that is, they used "provable data possession" (PDP) model for possession of data over the untrusted cloud storage. They used homomorphic linear authenticator (HLA) scheme for public auditing. But there are some problems in this system related security. This system achieves the public auditability but exposes the data to the external auditors. So the privacy of the data is compromised in this system.

Juels et al. s[3] has described a model called "proof of retrievability" (PoR). In this model, for "possession" and "retrievability" of data error correcting codes and spot checking is used. This model does not support public auditability. This is the disadvantage of this model. Also, this model does not support external auditor.

Later, to support dynamic operations of data over the cloud servers, Wang [4] proposed a dynamic auditing protocol. This protocol has some disadvantages; to send the data blocks to the auditor it requires a server and which may leak the privacy of the data to the auditor.

Waters et al [5] proposed a publicly verifiable homomorphic authenticator based scheme which is nothing but an improved version of "proof of retrievability" scheme. In this scheme they used authenticators to achieve the public auditability. But this scheme is not useful for dynamic data.

Likewise Message Authentication Code (MAC) based scheme is also not useful for dynamic data. It has some disadvantages such as user has to recalculate new MAC.

In HLA based scheme [6], cloud provider reveals user's data to the TPA which is the disadvantage of this scheme. This scheme is same as MAC based scheme but the only difference between MAC and HLA is that HLA can be aggregated.

By observing different existing system, there is need of such a system which provides public auditing services that will fulfill almost all the threats to the data over the cloud.

To do this, we suggested certain requirements for public auditing services-

A. Accountability:

Auditing should be done is proper manner. That is it should identify the problems as well as the particular entity responsible for that problem if any unreliability occurs. Therefore there is need of system's accountability.

B. Performance:

The major aspect of any system is performance. In cloud computing also security of data storage and its integrity is important task.

C. Dynamic Support:

Cloud provides dynamic support for runtime system to access and share the data. The challenge is the legacy users. User has access to data and user can modify the data in the cloud. So, dynamic support in runtime system is the major challenge for public auditing system.

In this paper, we proposed a secure and efficient system for public auditing which covers all the requirements mentioned above. In this system we use external auditor which is used for checking the integrity of the user's data. At the same time external auditor should be unaware of the data so that the privacy will be preserved and the communication overhead will be less. Also we use the blockless data verification scheme, which verifies the correctness of the data without having its knowledge.

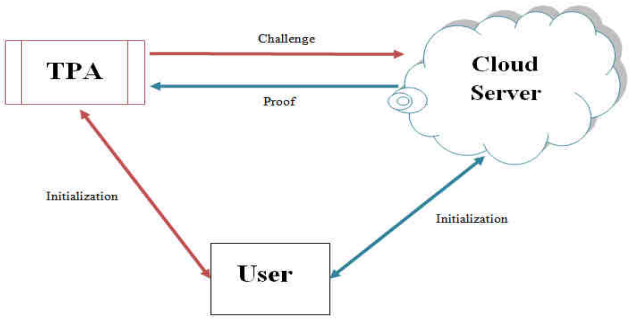

Fig.1: System model of the data auditing

III. PROPOSED SYSTEM

To save the user's computation time, storage resources and online burden, it is very important to introduce public auditing service for cloud data. This ensures the integrity of the data over the cloud and helps to reduce the online burden. Users may use the TPA to audit the data whenever they needed. Normally users don't have such expertise as TPA has. TPA has the capability and expertise to check the integrity of the data stored on the cloud on behalf of that user. This way it makes the integrity verification of the data easier and affordable for the user.

In the proposed scheme there are three algorithms for integrity verification -

a. Key generation: It is a process of generating keys (Secret and public keys).
b. Signing: Signing means generation of proof for verification.
c. Verification: The proof generated by the cloud service provider will be verified by the TPA.

In this paper, we use Boneh–Lynn–Shacham (BLS) signature for the integrity verification purpose. The main purpose of using this scheme is it creates less overhead over the network which automatically decreases the

communication cost. The BLS signature is only of 160 bits. As it is very short the size of authenticators is also reduced that means it requires less storage space over the cloud. Using this scheme aggregate authenticator is calculated on every block, and later all the individual authenticators are aggregated and calculated which is also of 160 bits. BLS signature scheme not only reduces the communication cost and required storage space but also this scheme is secure and unforgeable. [7]

Cloud architecture has three modules as shown in fig.1, User, Cloud Service Provider (CSP) and Third Party Auditor (TPA). User is responsible for storing the data over the cloud. CSP has the large space to store the user's data and has the resources to manage the user's data, whereas TPA which is an external auditor is responsible for auditing.

A. User:

a. User first divides the file into blocks, i.e. $F = (m_1, m_2, m_3 \ldots m_n)$.

b. Once the file is divided into blocks hash value is calculated on each block, i.e. $\text{Hash}(m_i) \longrightarrow h_i$

c. After that digital signature is calculated, i.e. $\text{SignGen}(m_i) \longrightarrow \sigma_i$, here 'i' denotes the i^{th} block.

d. Finally the aggregate authenticator is calculated, i.e. $\text{Aggregate_auth}(\sigma_i) \longrightarrow \sigma$

This aggregate authenticator is sent to the third party auditor (TPA) for checking the correctness of the data.

B. Cloud Service Provider (CSP):

a. Calculate digital signature, i.e. $\text{SignGen}(m_i) \longrightarrow \sigma_i'$

b. Claculate aggregate authenticator, i.e. $\text{Aggregate_auth}(\sigma_i') \longrightarrow \sigma'$

CSP sends the calculated aggregate authenticator to the TPA for verification of data.

C. Third Party Auditor (TPA):

a. Send file to check its integrity (F_i) Where Fi is a eset of files and $i \in \{1, 2, \ldots n\}$

b. Signature verification $\sigma = \sigma'$

Finally TPA is responsible for verifying the integrity of the data.

As shown in fig.2, following steps are performed for integrity verification on the user's data (single auditing).

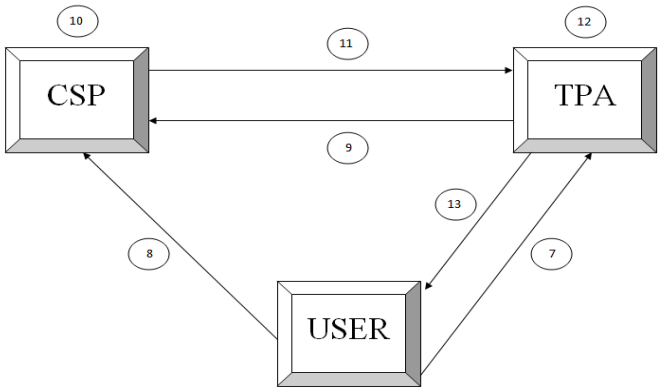

Fig.2: Architecture of cloud for integrity verification

1. First is the key generation process. User is responsible for generating public and private keys
2. User divides the file into individual blocks
3. User encrypts those blocks using 64 bit DES algorithm with the help of private key. This phase is called as signing.
4. Calculate hash for each block using MD5 and public key. MD5 is applied on 512 bit blocks and it produces hash of 128 bit.
5. Calculate digital signature which encrypts hash by using private key
6. User calculates aggregate authenticator
7. User sends calculated aggregate authenticator to the TPA
8. User sends encrypted data blocks to the cloud server and delete its local copy
9. TPA requests for the authenticator to the CSP
10. CSP calculates the aggregate authenticator which is calculated on the encrypted blocks. So this system provides more security as compared to earlier systems.
11. CSP sends aggregate authenticator to the TPA in response to the request.
12. TPA compares both the authenticators, the one which is sent by the user and another which is sent by the CSP. Based on that verification is done.
13. Depending on the result calculated by the TPA, the security message is sent to the user. This message is used to indicate the integrity of the file. If both the authenticators are same then this means the integrity of the file is maintained. If both the authenticators are not same then this means the file is altered by the intruder.

In this scheme, for integrity verification of the data user and CSP do not send the original data to the TPA. So in this case TPA has no knowledge about the data which improves the security of the user's data. Depending on the aggregate authenticators sent by the user and CSP, TPA compares

both the authenticators and gives the result accordingly. Hence, the proposed scheme achieves both i.e. privacy preserving and blockless verification.

IV. MODULES

A. Public Auditing:

In this paper, we proposed a unique privacy preserving public auditing technique which achieves the blockless data verification. At the CSP, the aggregate authenticator is calculated on the already encrypted data blocks. CSP doesn't decrypt the data blocks to calculate aggregate authenticator. This way the security of user's data over the cloud is achieved. Based on the authenticators calculated by both user and CSP on individual blocks are aggregated and compared at the TPA for its correctness. TPA has no knowledge of the data; it has only the authenticators obtained from user and CSP. This way the blockless data verification is achieved.

B. Batch Auditing:

Users may request for auditing service concurrently to the TPA. Auditing each task for individual user can be very inefficient and this can create the burden on the TPA. Using the batch auditing, TPA can simultaneously perform the multiple auditing tasks for different users. In this phase, multiple users send the aggregate authenticators to the TPA. Later TPA batch together all those requests and send it as a single request to the CSP. CSP then calculate the aggregate authenticator and sends it to the TPA. Finally TPA verifies the data. As compared to single auditing, batch auditing is better as multiple auditing requests are handled at a time. This improves the performance of the whole system.

C. Data Dynamics:

Dynamic support for public auditing is very important. User may need to update, delete or add the data. Allowing dynamic support over the cloud improves the efficiency of the public auditor. External auditor has to manage the integrity of the data file where user may wish to do some block-level operations on data like update, delete and modify the file in the running system. The proposed system provides the dynamic support. [8]

V. EXPERIMENTAL RESULTS

Following are the experimental results. Fig 3 shows the user dashboard which shows efficiency report graph of different files uploaded by the user.

Fig. 3: User Home Page (Dashboard)

In Fig 4, user uploads the file which is divided into certain number of blocks. Each block is encrypted with secret key and sent to CSP. At the same time one unique aggregate authenticator is calculated on blocks and is sent to the TPA.

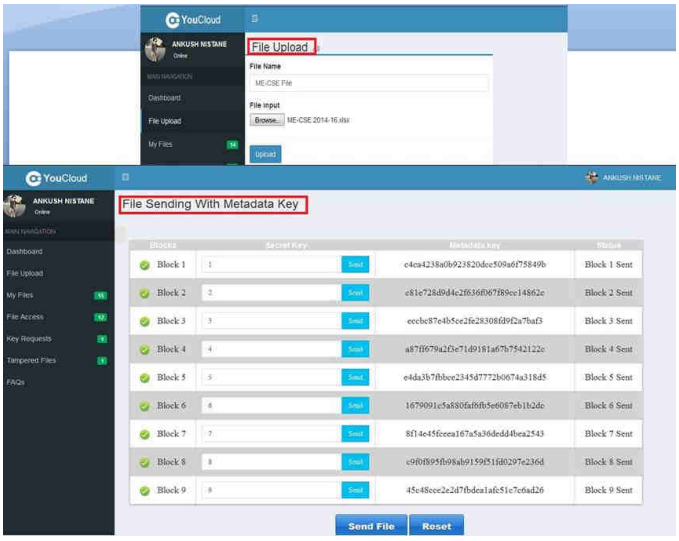

Fig. 4: File Uploading and Sending encrypted file blocks with metadata key for each block

In Fig 5, TPA requests aggregate authenticators to the CSP for that file. Then CSP calculates aggregate authenticator and send it back to the TPA for further processing. TPA then audits the file and security message send to the file owner. If the file is tampered by any other user then user gets the File Tampered message.

In Fig 6, to download the file user needs to enter the secret key. If secret key is valid then user need to enter valid metadata key for each block. If entered key is invalid then file will not be available download. More than 3 wrong attempts will send an intruder alert to the TPA as shown in Fig 7.

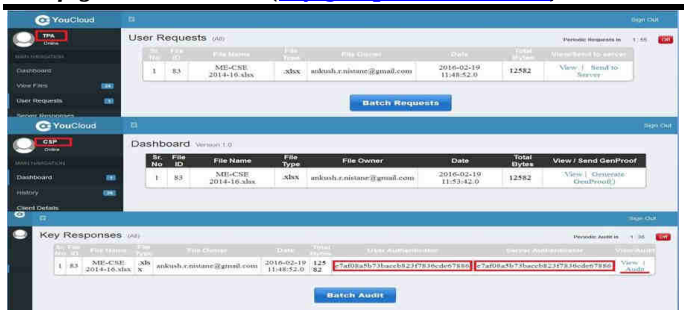

Fig. 5: TPA's and CSP's Communication for Aggregate Authenticator

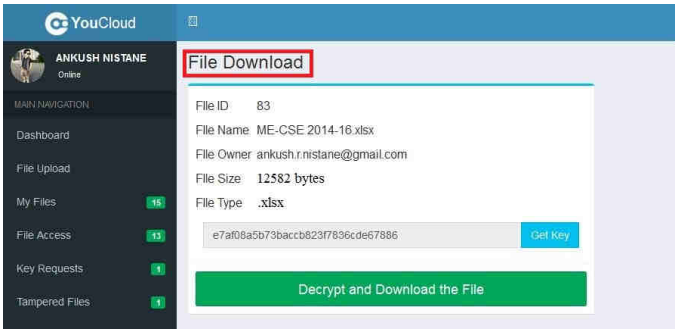

Fig. 6: Decrypt and download the file with Secret key

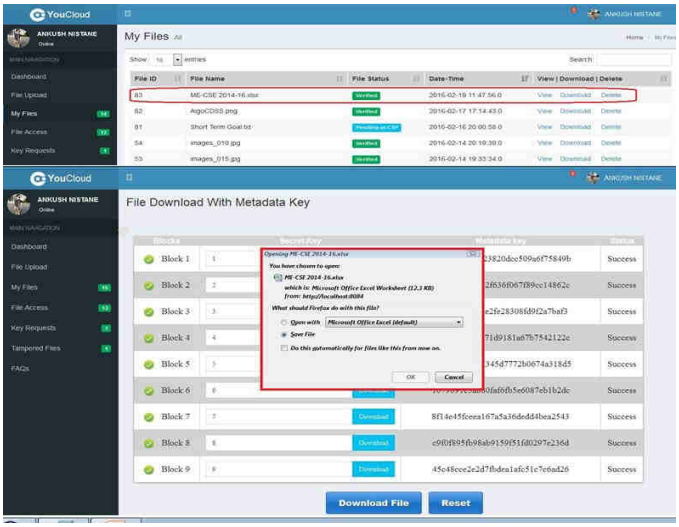

Fig. 7: File status and Downloading file with the help of metadata

VI. CONCLUSION AND FUTURE SCOPE

Cloud storage is increasing day by day. Public auditing over the cloud is of critical importance. As the user doesn't have such capabilities and expertise as third party auditor has, user resorts to the TPA for the integrity verification of the data. This work studies the importance of integrity verification over the cloud with dynamic support. Also proposed system achieves the privacy preserving public auditing and blockless data verification. Batch auditing improves the efficiency of the TPA as multiple requests are handled at the same time, which reduces the burden of TPA. Since this system is effective and efficient for precise public auditing for integrity verification of user's data.

Using different schemes the performance and security of this system can be improved.

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D.Joseph, R. H.Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, Apirl 2010.

[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," *Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07)*, pp. 598-609, 2007.

[3] Juels and J. Burton, S. Kaliski, "PORs: Proofs of Retrievability for Large Files," *Proc. ACM Conf. Computer and Comm. Security (CCS '07)*, pp. 584-597, Oct. 2007.

[4] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," *IEEE Trans. Parallel Distributed Systems*, vol. 22, no. 5, pp. 847-859, May 2011.

[5] H. Shacham and B. Waters, "Compact proofs of retrievability," in *Proc. of Asiacrypt 2008*, vol. 5350, Dec 2008, pp. 90–107

[6] Cong Wang, qian wang, kui ren, wenjing lou , "Privacy – Preserving Public Auditability for Secure Cloud Storage", *IEEE transaction on Cloud Computing Year 2013.*

[7] Boneh, B. Lynn, and H. Shacham, "Short signatures from the Weil pairing," *J. Cryptology*, vol. 17, no. 4, pp. 297–319, 2004.

[8] Imran Ahmad, Prof. Hitesh Gupta, "Privacy-Preserving Public Auditing & Data Intrgrity for Secure Cloud Storage", *International Conference on Cloud, Big Data and Trust 2013, Nov 13-15, RGPV 100.*