IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727
PP 27-34
www.iosrjournals.org

Public Auditing Mechanisms to Protect the Integrity of Data Shared in the Cloud

Gayathri Dili, Assoc. Prof. Anu V.R
Final Year M.TechDept. of Computer Science & Engineering Sree Narayana Gurukulam College of Engineering Kerala, India

***Abstract**: Clients are capable of storing large amount of data on a storage space that can be either a trusted or untrusted server. This storage outsourcing has resulted in a number of security issues. This could happen at the time when the data shared by one user is being used by another user and modify it and again the data. Cloud offers data storage and sharing facilities that provides better scalability. Apart from share the advantages offered by Cloud it find difficulties in maintaining the integrity of the shared data. Public auditing is a mechanism by which the integrity of data could be maintained so that the correctness of data could be verified. In this paper, we discuss about the different Public auditing mechanisms that has contributed towards the maintenance of integrity of data and the use of public auditing for data shared in the cloud.*
***Keywords:** Shared data, Public auditing, User-revocation, Cloud computing.*

I. Introduction

Public auditing started its first phase with the concept of checking whether the server whether trusted or an untrsuted one contain the original data, that is to ensure the correctness of data and could be done without retrieving entire data from the data stored server. The early version of public auditing is capable of utilizing homomorphic authenticators which provide an authenticable signature scheme for checking the integrity of data and there by verifying the correctness of data [2], without retrieving the entire data from the storage space.

The public auditing can be further combined with user revocation process [11].Revocation could be done at the time when a user within a group who share data within that group or that database and its members, is being removed by the admin of that group. The data being shared and used by the revoked user will be further need to be signed by another user called re-signature generation process. So that all the data used by the revoked user in that group is no longer allowed to access that data. To further check the integrity of data, there added a public verifier who verifies whether the data in the database is correct or not. Cloud computing offers a good option for storage of data same as a normal database, and the data stored in the Cloud could be shared with large number of users. A Cloud group can have large number of users, who share data, modify the data etc. Each user will be signing the data for security reasons to shoe the identity of each block of data shared by the users within that group. Same as a normal database storage the users within that group or users within that database can be revoked if necessary by the admin, here Cloud can also revoke the users and further re-sign the blocks with any of the existing users. Just like for the database systems the public auditing could bring integrity for the shared data to some extent.

II. Literature Review

In 2007, Atniese proposed a method called Provable Data Possession (PDP) [2] which allowed a public verifier to check the correctness of data which was being stored by the user or a client on an untrusted server .Even it offered high privacy for private data of the user from the auditor, it was good for only the static data.

An extension to the PDP was introduced in 2009 [3].In this extension model Ateniese implemented PDP using some symmetric keys which could provide support for the dynamic data. But it couldn't do much for verifying the integrity of data as verifier could only provide limited number of verification request.

Later Q.Wang introduced the Merkel Hash Tree for supporting the public auditing mechanism by providing a complete support for fully dynamic operations.

Users or clients who share the data on a storage space was so much worried about how to maintain the integrity of data, as the data became larger and larger the idea of checking the integrity of data by users itself need to get changed and C.Wang in 2010, suggested the idea to bring the Third Party Auditor (TPA) [5], the workload or complexity felt by the users or clients could be overcome to a greater extent. But protecting the private or confidential data of users from TPA came forward as an issue, but Wang solved it in a better way by random masking.

In 2012, B.Wang proposed a model ―Oruta‖ which could help in identifying the each of the signers who have signed on the data blocks being shared in that storage space and keep the signer's identity private form the public verifiers and thus provide integrity of shared data without retrieving the entire file [6].Apart from the other previously discussed mechanisms this could perform multiple auditing tasks.

Later in June 2012,B.Wang proposed another model called ―Knox‖ ,even if there is large number of users, it is not affecting the auditing of large amounts of data shared by a client and time taken to audit those data. This could be considered as a privacy –preserving mechanism too. But the user revocation and public auditing couldn't be implemented so successfully here.

In 2015, B.Wang another concept called ―Panda‖ [12] which could provide better public auditing mechanism and with the user revocation capability ,so that the admin or group manager could be capable of removing the users and further go for checking the integrity of data and perform the re-signing of data with less time. They brought a concept of batch auditing that helps to perform multiple auditing.

III. Architectural Designand Mechanisms

A. Design

The general architecture of the system model that could perform auditing or verification of data comprises of the cloud server-which act as the database server, the public verifier and the users as clients who share the data Fig 1.

The Provable Data Possession (PDP) utilized RSA-based homomorphic authenticators and some sampling strategies for maintaining the correctness of data being stored in untrusted servers [2].The homomorphic authenticators were used as building blocks in the public auditing mechanisms, thus helps to verify the correctness without downloading the entire data.

The architecture describes about Cloud which could act like a Cloud storage that is more efficient and offers good scalability. The public verifier is just like a client who performs the computation, data mining or even the search operation. Users are the clients who share the data and further use them, modify them.

B. Mechanisms

The Cloud storage is very large and provides the users the ability to share the data, modify them. To keep an identity on each block of data, the user keeps a signature on the blocks of data shared by him. Therefore the data shared by different user has different signature and those changes when they go for modifying it. The data shred in a group is easily accessible for the original user who shares it and also the other members of the group. Once the data is being modified by a user or being re-signed at the time of user revocation, checking the integrity of stored data or data re-signed by the user is important.

There are many mechanisms for performing public auditing of this data.

1) Provable Data Possession (PDP)

Provable Data Possession (PDP) [2] allows a verifier to check the correctness of data, so that they can confirm that the database server (untrusted) contains the original data. Here the client or user will be performing this verification. The client or the verifier go for a challenge-response protocol as shown in Fig.2, for verification process and to confirm the verification appropriate proofs are also maintained. Challenge-response protocol helps the verifier to go for verification based on linear combination of blocks of data.

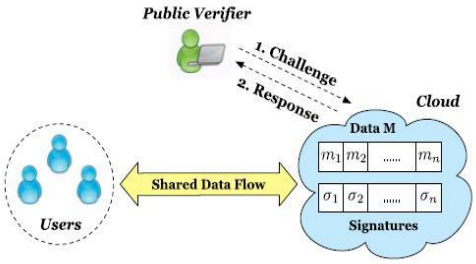

Fig 1.Our system model includes the cloud server, a group of users and a public verifier.

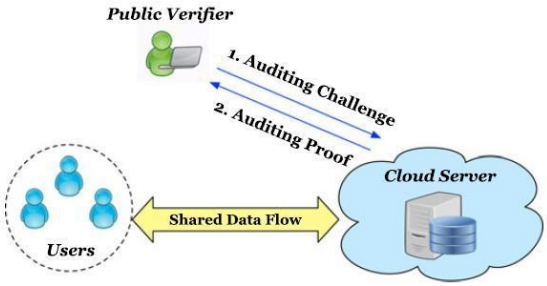

Fig .2.Verifier performs challenge-response with Cloud (database server) to perform the verification of data.

Instead of verifying data by considering linear combinations of data , it is advised to go for public auditing mechanism which cloud avoid the time being consumed for performing this linear way of verification that too by considering the entire data anyways, so by considering homomorphic authenticators that were based on RSA the verifiers (public verifier) to check the correctness of data and there by ensure the data integrity without downloading entire data .The homomorphic authenticators should satisfy two properties that is

- Block less verifiability allows the verification process or checking the correctness of data by considering linear block of data, without downloading the entire data.

- Non malleability shows that users who do not possess private keys, cannot generate valid signatures.

Advantages
1. Minimum network communication overhead.
2. Performs remote data checking.
3. Support larger datasets in widely distributed storage systems.
4. Good for static databases.

Disadvantages
1. Do not support Dynamic operations.
2. Time consuming.
3. No data privacy.
4. No identity privacy.

2) PDP based on symmetric keys

To get a proof of data possession a public key technique [2] was used by the verifier which goes for querying server to get data possession, which is also termed as public verifiability. This way of interactive proof of data possession can be repeated a many times.

PDP also termed as POR-Proof of Data Retrivability [3] used sentinels based approach. Here during verification sentinels are randomly picked and checked. And these special blocks called sentinels must be encrypted too.

Disadvantage of POR:
1. Not suitable for public databases like archives, libraries etc.
2. Use limited confidential data.

Here the PDP consider the security feature too which helps in providing efficient and secure ways of outsourcing of personal digital contents with two requirements,

(1).outsourcing data in clear text.
(2).bandwidth and computation efficiency

In symmetric key based PDA [3], before outsourcing the data the owner of the data go for pre-computing some short possession verification tokes that could cover some set of data blocks. After that actual data is handed over to server. The owner goes for asking the proof of possession subsequently and challenges the server with a set of random-looking block indices. Server performs some integrity checks on the block of data and returns the result to owner. For the proof of hold the returned integrity check must match with the value pre-computed by the owner.

Advantages
1. Pre-computed tokens are kept locally or outsourcing.
2. Owner's storage overhead is constant.
3. Efficient in terms of bandwidth and computation.
4. No bulk encryption of outsourced data.
5. No data expansion due to additional sentinel blocks.
6. Support dynamic operations like modification, deletion, append.

Disadvantages
1. It is not publicly verifiable.
2. Provides a user with limited number of verification requests.

3) Third Party Auditor

When the data is kept locally, users will have the burden of keeping the control of data, its maintenance, performing verification and checking the integrity of data. But if this auditing role is assigned to another person Fig.3, these difficulties could be solved, thus the auditing will be done by a Third Party Auditor [8], [5].

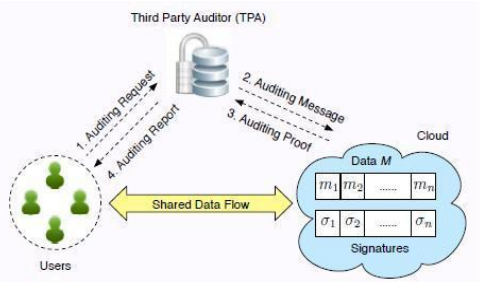

Fig .3.The Third Party Auditor (TPA)

performs public auditing and checks the integrity of data.

There are two requirements
- TPA should be capable of auditing cloud data storage without demanding any local copy of data, without affecting user data privacy.
- TPA should not introduce any additional online burden to the cloud user.

To enable the privacy preserving public auditing for Cloud storage and preserve a user's confidential data from TPA, some random masking techniques are integrated with homomorphic authenticators [5] .Later some bilinear aggregate signature technique is used to extent the result of this to multiuser settings, to operate multiple auditing tasks from different users and further extended to support batch auditing.

Advantage
1. Provide support for preserving the privacy of user's confidential data.
2. Extensions can be given to support multiple auditing and batch auditing.
3. Provide support for data dynamics including block level operations of modification, deletion, insertion.

Disadvantages
1. Cannot reveal the identity of signers on the block of shared data

4) Public Auditing for shared data in the Cloud using "Oruta"

Users are capable of sharing data in the

Cloud, so that it could be used by many users within that group of users. In the previous works data owners and public verifiers need to perform public verifiability without downloading entire data from the Cloud by dividing the data in to small blocks, where each of the blocks are independently signed by the owner and they take random combination of blocks instead of whole data retrieving during integrity checking. With those existing methods of public auditing data integrity could be maintained and further extended for verifying shared data integrity.

In order to protect user's confidential data, it is essential and critical to preserve identity privacy from public verifiers during public auditing. Privacy issues on shared data are solved by privacy-preserving public auditory mechanism called ―Oruta‖ [6]. In Oruta ring signature are used to construct homomorphic authenticators so as to perform public auditing properly. Oruta used a Homomorphic Authenticable Ring Signature scheme which helped in preserving identity privacy and support block less verifiability. Oruta also helps in integrity check and the identity of signers on each block in shared data is kept private from pubic

verifier.

Advantages
1. Support public auditing, identity privacy and data privacy. must be generated and securely shared among rest of the group.

Ring signature technique [6] is used in Oruta with which the verifier is convinced that the signature is computed using one of the group member's private key, but the verifier couldn't determine which one. To make the ring signature scheme to satisfy block less verifiability, ring signatures can be combined with homomorphic authenticators that cloud help in preserving the identity privacy.

Advantages
1. Support integrity check.
2. Identity of signer on each block in shared data is kept private from public verifier during auditing.

Disadvantge
1. Failed to scale well to large number of users who share data in the group.

5) Public Auditing for shared data in the Cloud using "knox"

Knox‖ is a privacy preserving auditing mechanism for data stored in the Cloud and shared among a large number of users in a group [7]. In knox the amount of information used for every verification and the time taken to audit with it are not affected by the number of users in the group. Knox is capable of auditing the data stored in untrusted server or Cloud and those data that is shared among large number of users in the group.

The idea of group signature which aim to provide anonymity of signers, who are from a same group and is used to construct homomorphic authenticators [7] .To verify the integrity of shared data without retrieving the entire data, TPA was introduced. But TPA lack in revealing the identity of signers on the block of shared data. Apart from adding new users in to the group without re-computing any verification information the original user is also capable of tracing group signatures on shared data and thus reveals the identity of signers. To check the integrity of data Homomorphic Authenticable Group Signature (HAGS) scheme is used that satisfy block less verifiability [7]. The HAGS helps in preserving the identity of signers form the TPA is possible.

The original member, who is the owner of the data share the data with other users of the group,
2. On extension support multiple can reveal the signer's identity. Whenever user auditing and batch auditing.

To preserve the identity of the signer on each block during public auditing, one alternative approach was used to ask all the users of the group to share global private key [9],[10]. Every user signs the block with a global private key so that even if one user leaves the group new global private key wanted to protect the privacy of shared data user can encrypt data using some encryption technique as the combination of symmetric encryption and attribute based encryption before outsourcing data to the Cloud server.

Advantage
1. Audit correctness of data stored among large number of users.
2. Preserve identity privacy for large number of users

Disadvantage
1. Do not support public auditing, since in Knox TPA needs to share a secret key pair with all group users, referred to as authorized editing.

6) Public Auditing for shared data in the Cloud using "Panda"

A user in a Cloud group is capable of sharing the data within that group. For identity maintenance each user keep a signature on the data shared by them, which changes if any other user accesses it, modifies it etc. Thus each of blocks of shared data will be having different signatures. For security reasons , when a user leaves a group the data being shared by him is needed to be re-signed by some other user, and further block the access of the revoked user to that Cloud group and access to the data shared by him [12]. The user revocation removes the user from the group rather than affecting the data being shared.

According to the straight forward method as in Fig.4 , at the time of user revocation the signatures are needed to be re-computed by asking an existing user to first download the blocks of data signed by the revoked user and then verify the data and further upload the data with their signature. This process creates complexity at the time when large amount of data is needed to be re-signed by the existing user [12].

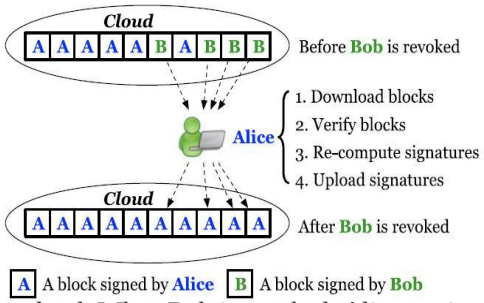

Fig.4. Alice and Bob share data in the cloud. When Bob is revoked, Alice re-signs the blocks that were previously signed by Bob with her private key.

Panda is public auditing mechanism that handles efficient user revocation by using proxies, that is proxy re-signature scheme. Using proxies for re-signing at the time of user revocation helps to protect integrity of data and thereby avoiding the need for downloading the entire data. When a user is revoked from the group, the Cloud acts like a proxy and take the responsibility for re-signing the blocks of data with some re-signing keys. The proxy re-signatures allow a semi-trusted proxy to act as a translator of signatures between two users and verifications could be done with the public keys [1].

Once the data is shared in to the Cloud by the original user who is the owner of the data, the data blocks will be signed by him. After that once a user modifies a data block the signatures are needed to be re-computed and the new user go for re-signing with their private key. Re-signing the shared data of the revoked user by an existing user helps in verifying the integrity of data and helps in checking the correctness of data. This re-signing task can be accomplished by considering Cloud as a semi trusted proxy as shown in Fig.5.This method of re-signing could be done without downloading entire data.

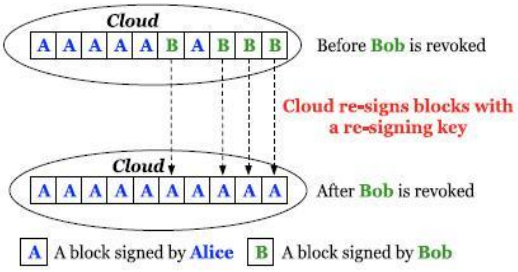

Fig.5. When Bob is revoked, the Cloud re-signs the blocks that were previously signed by Bob with a re-signing key.

To ensure the block less verifiability property and correctness of signatures Homomorphic Authenticable Proxy Re-signature scheme could be used [12]. The HAPS are also responsible for supporting the non malleability. Here Cloud is assumed to have a server to store shared data and another to manage re-signing keys.

Advantage

1. Protect integrity of data.
2. Performs public auditing.
3. offers scalability and reliability.
4. Support dynamic data.

5. Support large number of users to share the data.
6. Handle multiple auditing and batch auditing.

IV. Comparitive Study

So far we have seen many mechanisms for performing the public auditing for data shared in the Cloud. Now we can make a comparative study about those mechanisms as shown in Table.1.

Table.1 Comparative Study of different public auditing mechanisms

S. No.	Mechanisms	Features	Advantage	Disadvantage	PA*	IC*	IP*	DH*
1	Provable Data Possession	Using RSA-based techniques and sampling strategy.	1. Minimum network communication overhead. 2. Performs remote data checking. 3. Support larger datasets. 4. Good for static data.	1. Do not support Dynamic operations. 2. Time consuming. 3. No data privacy and identity privacy. .	Y	Y	N	N
2	Provable Data Possession	Using symmetric keys.	1. Pre-computed tokens are kept locally outsourcing. 2. Efficient in terms of bandwidth and computation. 3. No bulk encryption of outsourced data. 4. Support dynamic operations.	1. It is not publicly verifiable. 2. Provides user with limited number of verification requests.	N	N	N	Y
3	Third Party Auditor	Using random masking or bilinear signature technique or homomorphic authenticators.	1. Provide support for preserving the privacy of user's confidential data. 2. Provide support for data dynamic, multiple auditing and batch auditing	1.Cannot reveal the identity of signers on the block of shared data	Y	N	Y	Y
4	Public auditing for shared data in the Cloud-Oruta	Homomorphic Authenticable Ring Signature.	1. Support public auditing, identity privacy and data privacy. 2. Support multiple and batch auditing. 3. Support integrity check.	1. Failed to scale large number of users who share data in the group.	Y	Y	Y	Y
5	Public auditing for shared data in the Cloud-Knox	Homomorphic Authenticabel Group Signature.	1. Audit correctness of data. 2. Preserve identity privacy.	1. Do not support public auditing.	N	N	Y	Y
6	Public auditing for shared data in the Cloud-Panda	Homomorphic Authenticabel Proxy Re- Signature.	1. Protect integrity of data. 2. Perform public auditing. 3. Support dynamic data. 4. Handle multiple and batch auditing.	1. Identity privacy is not considered.	Y	Y	N	Y

*PA-Public Auditing *IC-Integrity Check *IP-Identity Privacy *DH-Dynamic Data Handling

V. Conclusions

Cloud storage helps the clients to store the data, to share the data and to modify the data. Cloud can consider many groups inside it. Each group consists of many users. Users of these groups share data between members of that group. Apart from providing scalability and reliability, the Cloud should be capable of protecting the integrity of data. Another factor to be considered is achieving privacy, especially for the situations where data is stored in untrusted servers. Public auditing could be done to verify correctness of data without downloading entire data. This paper presents some public auditing mechanism that provides features for protecting identity privacy, supporting user revocation, and providing re-signing facilities and based on that a survey was conducted.

References

[1] M. Blaze, G. Bleumer, and M. Strauss, ―Divertible Protocols and Atomic Proxy Cryptography,‖ Proc. Int'l Conf. the Theory and Application of Cryptographic Techniques (EUROCRYPT'98), pp. 127-144, 1998.
[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. 598-610, 2007.
[3] G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, ―Scalable and Efficient Provable Data Possession, ‖Proc. Fourth Int'l Conf. Security and Privacy in Comm. Networks (ICST SecureComm'08),2008.
[4] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, ―Enabling Public Verifiability and Data Dynamic for Storage Security in Cloud Computing,‖ Proc. 14th European Conf. Research in Compute Security(ESORICS'09, 2009.
[5] Wang, Q. Wang, K. Ren, and W. Lou, ―Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing,‖ Proc IEEE INFOCOM, pp. 525-533, 2010.
[6] B. Wang, B. Li, and H. Li, ―Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud,‖ Proc. IEEE CLOUD, pp. 295-302, 2012.
[7] B. Wang, B. Li, and H. Li, ―Knox: Privacy-Preserving Auditing for Shared Data with Large Groups in the Cloud,‖ Proc. 10th Int'l Conf. Applied Cryptography and Network Security (ACNS'12), pp. 507-525, June 2012.
[8] B. Wang, B. Li, and H. Li, ―Public Auditing for Shared Data with Efficient User Revocation in the Cloud,‖ Proc. IEEE INFOCOM, pp. 2904-2912, 2013.
[9] B. Wang, H. Li, and M. Li, ―Privacy-Preserving Public Auditing for Shared Cloud Data Supporting Group Dynamics,‖ Proc. IEEE Int'l Conf. Comm. (ICC'13), pp. 539-543, 2013.
[10] B. Wang, S.S. Chow, M. Li, and H. Li, ―Storing Shared Data on the Cloud via Security-Mediator,‖ Proc. IEEE 33rd Int'l Conf. Distributed Computing Systems (ICDCS'13), pp. 124-133, 2013.
[11] C.Wang,Q.Wang, K.Ren,W.Lou‖Privacy Preserving Public Auditing for Secure Cloud Storage‖, IEEE Transactions on Computers Vol.62,2013.
[12] B.Wang,BaochinLi, Hui Li,‖Panda:Public Auditing for Shared Data with Efficient User Revocation in the Cloud‖, IEEE Transactions on Services computing, Vol.8,2015.