International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Review Article

Survey Paper on Privacy Preserving Auditing Protocol for Cloud Storage

Nandini P.Wasnik†* and Mahip M.Bartere†

†Dept. of Comp. Science and Engineering, G.H. Raisoni Collage, Amravati, India

Accepted 28 Jan 2015, Available online 01 Feb 2015, **Vol.5, No.1 (Feb 2015)**

Abstract

Cloud storage provides users to easily store their data and enjoy the good quality cloud applications need not install in local hardware and software system. So benefits are clear, such a service is also gives users' physical control of their outsourced data, which provides control over security problems towards the correctness of the storage data in the cloud. The main goal of cloud computing concept is to secure, the data protection and the process which come under the property of users. As the data is stored at the remote place how users will get the confirmation about data which are stored. That is why cloud storage should have some technique which will specify storage correctness and integrity of data stored on cloud. Users can resort to a third-party auditor (TPA) to check the integrity of outsourced data and be worry free. TPA should be able to efficiently audit the cloud data storage without demanding the local copy of data. To introduce an effective and secure TPA, the auditing process should bring in no new vulnerabilities toward user data privacy, and no online additional burden to user. In this paper, we propose a cloud storage system which is secure and supporting privacy-preserving public auditing. The proposed system perform audits for multiple users simultaneously and efficiently.

Keywords:. *Data storage, privacy preserving, public audit ability, cloud computing, delegation, batch verification, zero knowledge.*

1. Introduction

In the history of IT, cloud computing has brought unprecedented benefits to the computing world. It has made it possible to have a different computing model that does not suffer with scarcity of resources. Cloud computing enables to share computing resources without the need for investment in pay as you use fashion. Cloud service providers such as Microsoft, Oracle, Amazon, Google etc. are able to provide huge clouds which are nothing but computing resources that are provided on demand through Internet(P. Mell and T. Grance *et al* 2009).The way on that IT infrastructure has been used; is changing with the emergence of cloud computing. One important part of cloud computing is that data is stored in a centralized server which is linked to cloud data centre. The storage and other services provided by cloud can be utilized by individuals and organizations alike without the need for capital investment. To organizations and individuals cloud provides very useful benefits as they are relieved from storage management, investment, and maintenance .(M. Armbrust, A. Fox, R. Griffith, A. D. Joseph *et al* 2009).

*Corresponding author **Nandini P.Wasnik** is a M.E. final year CSE Student and **Mahip M.Bartere** is working as Assistant Professor

Along with the advantages, it also has challenges in terms of security threats. This is because the users' data is stored in a remote server which is considered untrusted. Users are losing control over their data and the storage facilities are under control of cloud service providers. Thus the correctness or integrity of the data is questioned. The cloud data storage might be subjected to internal and external threats.(A. Juels and B.S. Kaliski *et al* 2007) Security problems surfaced in cloud computing were known to the walorld M. Arrington *et al* 2006. On the other hand CSPs might have intentions to be unfair towards cloud users and their outsourced data besides hiding security flaws in their storage infrastructure (CongWang,Chow,S.S.M, QianWang *et al* 2013)

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 not enough to detect the data corruption only when accessing the data, as it does not give correctness assurance for those unaccessed data and might be too late to recover the data loss or damage. Moreover, the overhead of using cloud storage should be minimized as soon as possible, such that a user need not to perform too many operations to use the data (in additional to retrieving the data). In particular, user does not want to go through the

complexity in verifying the data integrity. Besides, there may be many user accesses the same cloud storage, in a frame setting. For easier management, it is desirable that cloud only entertains verification request from a single designated party.

To fully ensure the data integrity and save the cloud users' computation resources as well as online burden, it is of critical importance to enable public auditing service for cloud data storage, so that users may employ to an independent third-party auditor (TPA) to audit the outsourced data when needed. The TPA, who has expertise capabilities that users do not have, it is periodically check the integrity of all data which are stored in the cloud instead of the users, which provides a much more easier and affordable way for the users to ensure their storage correctness in the cloud. Moreover, in addition it help users to evaluate the risk of their subscribed services of cloud data, the audit result from TPA would also be important for the cloud service providers to improve their cloud-based service platform, and even serve for independent arbitration purposes (Shrinivas *et al* 2011). In a word, enabling public auditing services will play an important role for this nascent cloud economy to become fully established; where users will need ways to assess risk and gain trust in the cloud. Recently, the notion of public audit ability has been proposed in the context of ensuring remotely stored data integrity under different system and security models.

Public auditability allows an external party, in addition to the user himself, to verify the correctness of remotely stored data. However, most of these schemes do not consider the privacy protection of users' data against external auditors. Indeed, they may potentially reveal user's data to auditors, this severe drawback greatly affects the security of these protocols in cloud computing. From the perspective of protecting data privacy, the users, who own the data and rely on TPA just for the storage security of their data, do not want this auditing process introducing new vulnerabilities of unauthorized information leakage toward their data security (Q. Wang, C. Wang, K. Ren *et al* 2011).The first ones to support privacy-preserving public auditing in cloud computing, with a focus on data storage. As the individual auditing of these growing tasks can be tedious, a natural demand is then how to enable the TPA to efficiently perform multiple auditing tasks in a batch manner, i.e., simultaneously.

To address these problems ,utilizes the technique of public key-based homomorphic linear authenticator (or HLA for short) , which enables TPA to perform the auditing without demanding the local copy of data and thus drastically reduces the communication and computation overhead as compared to the straightforward data auditing approaches. By integrating the HLA with random masking, our protocol guarantees that the TPA could not learn any knowledge about the data content stored in the cloud server (CS) during the efficient auditing process. The aggregation and algebraic properties of the authenticator further benefit our design for the batch auditing.

2. Literature survey

1) Wang *et al*.[2011] propose to combine BLS-based HLA with MHT to support both public auditability and full data dynamics, none of them me*et al*l the requirements for privacy preserving public auditing in cloud computing.
2) Ateniese *et al*.[2007] are the first to consider public auditability in their provable data possession (PDP) model for ensuring possession of data files on untrusted storages.
3) Shah *et al*.[2008] Propose introducing a TPA to keep online storage honest by first encrypting the data then sending a number of pre computed symmetric-keyed hashes over the encrypted data to the auditor. The auditor verifies the integrity of the data file and the server's possession of a previously committed decryption key.
4) Juels *et al*[2007] describe a proof of retrievability (PoR) model, where spot-checking and error-correcting codes are used to ensure both possession and retrievability of data files on remote archive service systems.
5) Shacham and Waters *et al*. [2008] design an improved PoR scheme built from BLS signatures with proofs of security in the security model defined in, they use publicly verifiable homomorphic linear authenticators that are built from provably secure BLS signatures.
6) Erway *et al*.[2009] developed a skip lists based scheme to enable provable data control with fully dynamics support. However, all their protocol requires the linear combination of sampled blocks and thus does not support privacy preserving auditing on users outsourced data.
7) Sebe *et al*.[2008] thoroughly study a set of requirements which ought to be satisfied for a remote data possession checking protocol to be of practical use. Their proposed protocol supports unlimited times of file integrity verifications and allows preset tradeoff between the protocol running time and the local storage burden at the user.
8) Schwarz and Miller *et al*.[2006] propose the first study of checking the integrity of the remotely stored data across multiple distributed servers. Their approach is based on erasure-correcting code
9) Bowers *et al*.[2008] utilize a two-layer erasure-correcting code structure on the remotely archived data and extend their POR model to distributed scenario with high-data availability assurance.
10) Curtmola *et al*.[2009] aim to ensure data possession of multiple replicas across the distributed storage system. They extend the PDP scheme into cover multiple replicas without encoding each replica separately.

3. Related work

Different factors such as data integrity, data dynamics and privacy of data affects. Each and every approach

has merits and demerits which make them suitable for different applications. Following are the different approaches which are already carried out for cloud data security.

1) **PDP Method**: The author Ateniese *et al*. [2011] are the first who have considered the public adaptability in their defined provable data possession. (PDP) method which protect possession of data files on mistrustful storages. For auditing out going sourced data these technique utilizes the RSA-based homomorphic authenticators and which suggests to randomly sample a few blocks of the file. However, in these scheme the public auditability demands the linear combination of sampled blocks which exposed to the external auditor. The goal of a PDP scheme that achieves probabilistic proof of data possession is to detect server misbehavior when the server has deleted a fraction of the file.
2) **Proof of retrievability:** Juels *et al*. [2007] describe a proof of retrievability (PoR) model, where spot-checking and error correcting codes are used to ensure both possession and retrievability of data files on remote archive service systems. (PoR) model requires that the prover access only a small portion a file F.The portion of F touched by the prover is essentially independent of the length of F.POR scheme encrypts F and randomly embeds a set of randomly-valued check blocks called *sentinels*. The use of encryption here is that the sentinels indistinguishable from other of blocks the file. The user challenges the archive by specifying the positions of a collection of sentinels and asking the prover to return the associated sentinel values. If the prover has found change or deleted a *substantial* portion of F, then with high probability it will also have suppressed a number of sentinels. It is therefore unlikely respond to the verifier for correcting their data file.

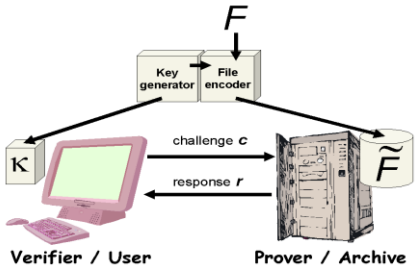

Fig 1: POR System

3) **MAC (Message Authentication Code):** It can be used to protect the data integrity. Data owners will initially locally maintain a small amount of MACs [2008] for the data files which are to be outsourced. The data owner can verify the integrity by recalculating the MAC of the received data file when he/she wants to retrieve data and will compare it to the local pre computed value but if the data file is large, MACs cannot be use. It is used for the data authentication. In this, user uploaded Block of data and MAC to CS provide its secret key SK to TPA. The TPA will randomly retrieve blocks of data & MAC uses secret key to check correctness of stored data on the cloud. Problems with this system are listed below as it introduces additional online burden to users due to limited use and stateful verification.

4. Proposed system

The proposed system contain following three entities, as show in Fig. 1: cloud user (U), which contain the amount of data files which are stored in the cloud;
cloud server (CS), managed by the cloud service provider (CSP) for providing data storage service and has significant storage space as well as computation resources; third party auditor (TPA), who has expertise and capabilities that cloud users does not have and it is trustful for assess the cloud storage service reliability on behalf of the user request. Users rely on the CS for cloud data storage and maintenance, they may also dynamically interact with the CS for accessing and update the stored data for purpose of various application . To save the computation resource as well as the online burden, the cloud users may resort to TPA for ensuring their outsourced data storage integrity, which hoping to keep their data private from TPA.

To ensure the data integrity and save the cloud users' computation resources as well as online burden, it is most importance to enable public auditing service for cloud data storage, so that users resort to an third party auditor (TPA) which is independent to audit the outsourced data whenever needed. The TPA,has expertise and capabilities that cloud users do not have it can periodically check the integrity of all the data stored in the cloud on behalf of the users, which make it a much more easier and efficient way for the users to ensure their storage correctness in the cloud. Moreover, for evaluate the risk of the cloud user the audit result from TPA would also be beneficial for the cloud service providers to improve their cloud based service platform, and even serve for independent negotiation purposes. In a word, enabling public auditing services will play an important role for this nascent cloud economy to become fully established, where users will need ways to assess risk and gain trust in the cloud.

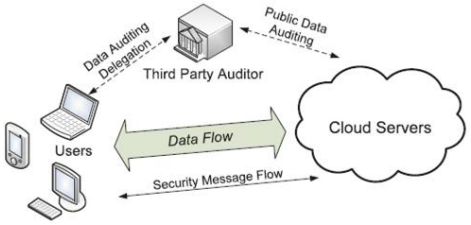

Fig 2: Architecture of cloud data storage service.

5. Advantages

1. Public auditability: TPA to verify the correctness of the cloud data on demand without retrieving a copy of the whole data or introducing additional online burden to the cloud users.
2. Storage correctness: To ensure that there exists no cheating cloud server that can pass the TPA's audit Without indeed storing users' data intact.
3. Privacy preserving: To ensure that the TPA cannot derive users' data content from the information collected during the auditing process.
4. Batch auditing: To enable TPA with secure and efficient auditing capability to cope with multiple auditing delegations from possibly large number of different users simultaneously.
5. Lightweight: To allow TPA to perform auditing with minimum communication and computation Overhead.

Conclusions

Proposed system introduced a privacy-preserving public auditing for data storage security in cloud computing. Proposed system utilize the homomorphic linear authenticator and random masking to guarantee that the TPA would not learn any knowledge about the data content stored on the cloud server during the efficient auditing process, which not only reduces the burden of cloud user from the tedious and possibly expensive auditing task. The process as data user can check the integrity of their data stored in cloud server using TPA which can be done in efficient manner. If any changes find out in data by the TPA, TPA will immediately intimate to the owner of the file and so security and data integrity is secured properly. The system further extend our privacy-preserving public auditing protocol into a many user setting, where the TPA can perform multiple auditing tasks in a batch manner for better efficiency. Third party auditor can be a trusted third party to resolve the conflicts between the cloud service provider and the client.

References

P. Mell and T. Grance[2009] Draft NIST working definition of cloud computing,

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,[Feb 2009] Above the clouds: A Berkeley view of cloud computing, University of California, Berkeley, Tech. Rep. UCBEECS-2009-28

A. Juels and B.S. Kaliski Jr.[2007] Pors: Proofs of Retrievability forLarge Files,Proc. 14th ACM Conf. Computer and Comm. Security(CCS '07),pp. 584-597.

CongWang ; Chow,S.S.M. ; QianWang ; KuiRen ;WenjingL Ou[2013] Privacy_preserving Public Auditing for Secure CloudS torage, IEEE Transactions on ComputersVolume: 62 , ,PP no : 362 - 375,.

R. Curtmola, O. Khan, and R. Burns, Robust Remote Data Checking[2008], Proc. Fourth ACM Int‟l Workshop Storage Security and Survivability (StorageSS ‟08), pp. 63-68,

Shrinivas, Privacy-Preserving Public Auditing in Cloud Storage security[2011], International Journal of computer science nad Information Technologies, vol 2, no. 6, pp. 2691-2693.

C. Wang, Q. Wang, K. Ren, and W. Lou[2012] Towards Secure and Dependable Storage Services in Cloud Computing, IEEE Trans. Service Computing, vol. 5, no. 2, pp-220-232.

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

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

M.A. Shah, R. Swaminathan, and M. Baker[2008] Privacy-Preserving Audit and Extraction of Digital Contents, Cryptology ePrint Archive, Report 2008/186.

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

H. Shacham and B. Waters[2008], Compact Proofs of Retrievability, Proc. Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (Asiacrypt), vol. 5350, pp. 90-107.

C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia[2009] Dynamic Provable Data Possession, Proc. ACM Conf. Computer and Comm. Security (CCS '09), pp. 213-222.

F. Sebe, J. Domingo-Ferrer, A. Martı´nez-Balleste, Y. Deswarte, and J.-J. Quisquater[2008] Efficient Remote Data Possession Checking in Critical Information Infrastructures, IEEE Trans. Knowledge and Data Eng., vol. 20, no. 8, pp. 1034-1038.

T. Schwarz and E.L. Miller[2006] Store, Forget, and Check: Using Algebraic Signatures to Check Remotely Administered Storage, Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS '06).

R. Curtmola, O. Khan, R. Burns, and G. Ateniese, [2008]MR-PDP: Multiple-Replica Provable Data Possession, Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS '08), pp. 411-420.

K.D. Bowers, A. Juels, and A. Oprea,[2009]HAIL: A High-Availability and Integrity Layer for Cloud Storage, Proc. ACM Conf. Computer and Comm. Security (CCS '09), pp. 187-198.