International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 12, December 2015
Available at http://internationaljournalofresearch.org

A Novel Approach for Auditing for Data in the Cloud by Using Privacy-Preserving Techniques

Boppudi Ravikumar [1] & Pradeep Venuthurumilli[2]
[1] **M.Tech** PG Scholar,Dept Of CSE, Nova's Institute of Technology, Eluru,West Godavari,AP.
[2] **(Ph.D.,)**Associate Professor, Dept Of CSE Nova's Institute of Technology, Eluru,WG Dist,AP.

Abstract:
By using Cloud storage, users can access applications, services, software whenever they requires over the internet. Users can put their data remotely to cloud storage and get benefit of on-demand services and application from the resources. The cloud must have to ensure data integrity and security of data of user. The issue about cloud storage is integrity and privacy of data of user can arise. To maintain to overkill this issue here, we are giving public auditing process for cloud storage that users can make use of a third-party auditor (TPA) to check the integrity of data. Not only verification of data integrity, the proposed system also supports data dynamics. The work that has been done in this line lacks data dynamics and true public auditability. The auditing task monitors data modifications, insertions and deletions. The proposed system is capable of supporting public auditability, data dynamics and Multiple TPA are used for the auditing process. We also extend our concept to ring signatures in which HARS scheme is used. Merkle Hash Tree is used to improve block level authentication. Further we extend our result to enable the TPA to perform audits for multiple users simultaneously through Batch auditing.

Index Terms: Cloud Storage; Data Dynamics; Public Auditing; Privacy Preserving; Ring Signatures

I. INTRODUCTION

Cloud computing is an internet based computing which enables sharing of services. Cloud computing allows users to use applications without installation any application and access their personal file and application at any computer with internet or intranet access. Cloud applications are e-mail, e-commerce, web conferencing-shopping, customer relationship management (CRM) etc. Cloud computing is widely developed technology used in IT industries to provide services like resources, rapid resource elasticity, network access control and platform as per user require. In cloud computing the user data is centralized to the cloud. The user can access the cloud services within the help of mobile devices and internet connection. Cloud storage is a online storage in which the data store in format of file or block pattern .Cloud data stored in virtualized pools of storage that are generally given by the Third Party Auditor(TPA). The cloud provides its application, software and data services are in remotely and temporarily, user can access it's by using personal computers, mobile phones or other internet access devices. In IT industries, individuals storing their data into the cloud in flexible manner having some benefits like relief from Hardware, software, online burden of data storage, reduce the cost of capital expenditure on personal maintenance. Cloud containing majority problems are related to data sharing resources between multi-owners and group membership changes. There are many methods are using for overcome from these issues as ring signature and encryption techniques.

Today everyone depending network to developing their knowledge in any area at any time so many users from remote locations use network services continuously so there may arise some issues. The cloud containing main issues are depending

privacy, security, data integrity, dynamic updates. This problem is addressed and solving by using public auditing for secure cloud. The third party auditor (TPA), who has capabilities and expertise that can periodically check the integrity of the data, which is stored in the cloud. The users cannot have the auditing capabilities than TPA. The TPA check the correctness of data stored in cloud on behalf of user and maintain the data integrity. Enabling public auditing service will play an important role for privacy data security & minimizing the data risk from misuses. The TPA is act as an external party .Which can also view the data stored in cloud and does not give the guarantee of data privacy. The auditing data in cloud can access the original data owner at any time. The cloud computing architecture contains a Third Party Auditor (TPA) for auditing the system which is connected with the particular group of the cloud storage. Group member or user is Cloud users where they store their private data into the cloud sever and also share that data with other user of Cloud system as a group member. Cloud is a system operated by the Cloud service provider, which allow to store and share data of cloud user in a system and also access service on a demand basis as pay. The cloud contains two types of storage, private and public type.

In the public anyone can access and anyone can change the cloud containing data and in the private the particular user can only access the data and the user cannot change the data without the owner's permission.

II. Literature Survey

Ateniese et al. [6] are the first to consider public auditability in their defined "provable data possession" (PDP) model for ensuring possession of files on untrusted storages. In their scheme, utilize RSA based homomorphic tags for auditing outsourced data, thus public auditability is achieved. However, Ateniese et al. do not consider the case of dynamic data storage, and the direct extension of their scheme from static data storage to dynamic case may suffer design and security problems. In their subsequent work [7], Ateniese et al. propose a dynamic version of the prior PDP scheme. However, the system imposes a priori bound on the number of queries and does not support fully dynamic data operations, i.e., it only allows very basic block operations with limited functionality, and block insertions cannot be supported. In [17], Wang et al. consider dynamic data storage in a distributed scenario, and the proposed challenge-response protocol can both determine the data correctness and locate possible errors. Similar to [7], they only consider partial support for dynamic data operation. Juels et al. [10] 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 archive service systems. Specifically, some special blocks called "sentinels" are randomly embedded into the data file F for detection purpose, and F is further encrypted to protect the positions of these special blocks. However, like [7], the number of queries a client can perform is also a fixed priori, and the introduction of precomputed "sentinels" prevents the development of realizing dynamic data updates.

Shacham et al. [16] design an improved PoR scheme with full proofs of security in the security model defined in [10].

They use publicly verifiable homomorphic authenticators built from BLS signatures , based on which the proofs can be aggregated into a small authenticator value, and public retrievability is achieved. Still, the authors only consider static data files. Erway et al. [9] was the first to explore constructions for dynamic provable data possession. They extend the PDP model in [6] to support provable updates to stored data files using rank-based authenticated skip lists. The scheme is essentially a fully dynamic version of the PDP solution. To support updates, especially for block insertion, they eliminate the index information in the "tag" computation in Ateniese"s PDP model [6] and employ authenticated skip list data

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 12, December 2015
Available at http://internationaljournalofresearch.org

structure to authenticate the tag information of challenged or updated blocks first before the verification procedure. However, the efficiency of their scheme remains unclear. Shan et al.[13] introduce TPA concept to maintain data integrity and preserve privacy. It reduces online burden and keeps the privacy preserve. Chen et al.[8] gives mechanism for auditing the correctness of data with multiple server. Frenz et al.[11] introduce a new strategy ,an Oblivious out-sourced storage which is based on Oblivious RAM technique. This idea used to conceal user access pattern and preserve the identity

III.Problem Definition

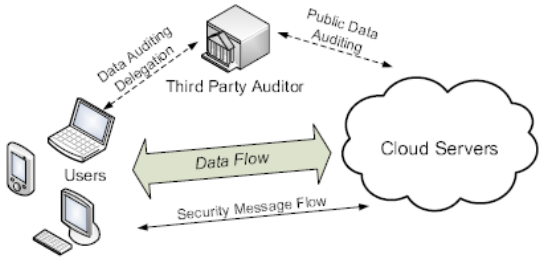

Figure 1: The architecture of Cloud Data Service

Consider a cloud data storage service involving three different entities, as illustrated in Fig. the cloud user (U),who has large amount of data files to be stored in the cloud; the cloud server (CS), which is managed by the cloud service provider (CSP) to provide data storage service and has significant storage space and computation resources, the third party auditor (TPA), who has expertise and capabilities that cloud users do not have and is trusted to assess the cloud storage service reliability on behalf of the user upon request.

Users rely on the CS for cloud data storage and maintenance. They may also dynamically interact with the CS to access

and update their stored data for various application purposes. To save the computation resource as well as the online burden, cloud users may resort to TPA for ensuring the storage integrity of their outsourced data, while hoping to keep their data private from TPA.

Assuming that the data integrity threats towards user data can come from both internal and external attacks at CS. These may include: software bugs, hardware failures, bugs in the network path, economically motivated hackers, malicious or accidental management errors, etc. Besides, CS can be self-interested. For their own benefits, such as to maintain reputation, CS might even decide to hide these data corruption incidents to users. Using third-party auditing service provides a cost-effective method for users to gain trust in cloud. Considering the TPA, who is in the business of auditing, is reliable and independent. However, it may harm the user if the TPA could learn the outsourced data after the audit.

IV. The Proposed Schemes

In this section we are giving public auditing Scheme for ensuring the data integrity. Firstly giving the notation and preliminaries, after that the framework and later overview of public auditing system.

A. Notation and Preliminaries

1: F- The data file to be fetched, denoted as a sequence of n blocks m1,….mi….,mn Є Z p for some large prime p.

2: MAC (.) (.) – Message authentication code (MAC) function, given as : $K\times\{0,1\}^*\rightarrow\{0,1\}l$
Where K denotes the key space.

3: H (.), h (.) - Hash function

B. Framework

The framework for privacy-preserving public auditing system maintains the data integrity. Public auditing schemes consist of four algorithms. **KeyGen, SigGen, GenProof, VerifyProof.** In **KeyGen** the Key is generated called as Key generation algorithm, which is run by the user to

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 12, December 2015
Available at http://internationaljournalofresearch.org

set up scheme. In **SigGen** verification metadata is generated by the user which consists of digital signature. **GenProof** is run by the cloud server to generate a proof of data storage. **VerifyProof** algorithm is run by TPA to audit and verify the proof. During the public auditing process, it consists of two phases as explain below:

Setup: By executing KeyGen algorithm, the user initializes the public and secret parameters of the system .By using SigGen, it generate verification metadata by preprocessing the data file F. The data file F & verification metadata is stored at cloud server by the user and delete its local copy. User may also alter the data file F by expanding it.

Audit: TPA send audit challenge or message to cloud server to become sure that the cloud server has keep data file F during the auditing process. By executing Gen Proof the cloud server gets response message by using file F and verification metadata as an input. And lastly the TPA verifies the response given by the cloud server by performing verify proof algorithm.

C. How process Works?

Here we are given the block diagram for the process flow. Fig 2 show the process flow for multiple users by using multiple TPA.

Algorithm for Data Integrity Verification

Algorithm for Data Integrity Verification
1. Start
2. TPA generates a random set like public key pk, private key sk and signature σ on each block (Verification metadata).
3. CS computes root hash code based on the filename/blocks input.
4. CS computes the originally stored value.
5. TPA decrypts the given content and compares with generated root hash.
6. After verification, the TPA can determine whether the integrity is breached.
7. Stop

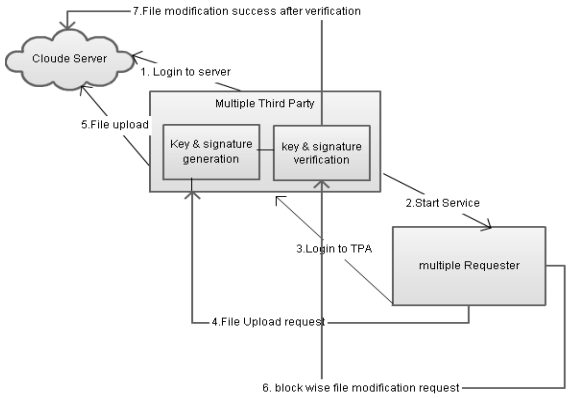

Fig 2. Process flow

D. Supports For Data Dynamics

In cloud computing, users update their data continuously for various application purposes. [14][15][16].So here privacy preserving public auditing supports for data dynamics in which user can do modifications on stored data. This data dynamics supports for update, delete, inert operation. For data dynamics we are using here Merkle Hash Tree (MHT). As data file F is divided into number of blocks m1,m2, . .,mn. Suppose user wants to modify the ith block mi to mi‟. At that time client generates new signature on block σi‟ = (H(mi‟).umi‟)α. When the CS receives the request it runs ExecUpdate(F,ϕ,update). Specifically, the server 1) replaces mi with mi‟ and outputs F‟; 2) replaces σi with σi' ; 3) replaces H(mi) withH(mi‟) in the MHT construction and generates the new root R‟. On this new root hash signature is generated and that signature is stored newly.

For insert operation , data gets inserted block wise manner into the file. And if user wants to delete file or particular block user can delete it by the same procedure.

E.Privacy-preserving Public Auditing Using Ring Signature scheme

As per existing work, public auditing scheme utilizes the technique of public key-based homomorphic linear authenticator (HLA), which

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 12, December 2015
Available at http://internationaljournalofresearch.org

allows TPA to perform the auditing without expecting the local copy of data and thus minimizes the communication. To maintain more data integrity and security Ring Signatures concept we are implementing here. The concept of ring signatures is first proposed by Rivest et al. in 2001[12]. The ring signature is the type of digital signature which can be performed by any group member of users that each have keys. Therefore, a message signed with a ring signature is endorsed by someone in a particular group of people. The best characteristic of a ring signature is that it should be difficult to identify which of the group members' keys was used to produce the signature. In this, the signature is computed using one of the group member‟s private key, but the verifier is not able to determine which one. This property can be used to preserve the identity of the signer from a verifier.

KeyGen RingSign, **RingVerify**. In KeyGen algorithm each user in the group generates their public key and private key. In RingSign algorithm user in the group is related to sign a block with her private key and all group members‟ public keys. In Ring verify algorithm the verifier is used to check whether the given block is signed by the group member. The ring signatures for public auditing consist of following steps for auditing.

1: Each user generates its public and private key.

2: user in the group sign a block with her private key and all group member‟s public key. Pk1 is public key of the user; Sk1 is private key of the user; (Pk1 ……Pkd) is „d‟ number of users of data block m Є ZP

3: User randomly selects data block m Let id is identifier of data block m

4: User ui encrypts with all user‟s public key, so only private key of the group user‟s i Є [1,d] would be able to decrypt it. This ensures privacy of data.

5: To ensure auditing by third-party user (ui), where i Є [1,d] signs the data block using his private key.

6: TPA (Third-party auditor) , using a Pk1 …….Pkd Where d is number of users in the group.

TPA calculates signature of data blocks but unaware of who sign it .Therefore calculates signature using each given public key (Pk1 ……Pkd) from this set.

Gsign = signature set for (Pk1 ……Pkd)

If Gsign = {sign1,sign2,……signd} matches with original sign then data block is intact.

By using this scheme user can also do the data dynamic operation. As there is group of users which share their data to each other , they can do modification on data of CS.

VI. CONCLUSION

The problem of data security in cloud data storage is being investigated, which is essentially a distributed storage system. i.e, Propose privacy preserving public auditing system that support data dynamics for data storage security in cloud. To achieve the assurances of cloud data integrity and availability and enforce the quality of dependable cloud storage service for users, my work propose an effective and flexible distributed scheme with explicit dynamic data support, including block update, delete, and append. By utilizing the homomorphic token with distributed verification of erasure coded data, proposed scheme is expected to achieve the integration of storage correctness insurance and data error localization, i.e., whenever data corruption has been detected during the storage correctness verification across the distributed servers, it almost guarantee the simultaneous identification of the misbehaving server(s).

Homomorphic authenticator guarantees that TPA will not learn the user content which eliminates the user from the burden of auditing task. Considering the time, computation resources, and even the related online burden of users, proposed system is expected to be highly efficient and secure Homomorphic linear authenticator is utilized to check integrity and random masking to guarantee that the TPA would not learn any knowledge about the data content stored. The full-fledged implementation of the mechanism on commercial public cloud is an important future extension, which is expected to robustly cope with very large scale data and thus encourage users to adopt cloud storage services more confidently. Design could be implemented for cloud storage of an organisation where the administrator provides access credentials to various personnel of the organisation who can concurrently access the data privileged to them with increased rate of security and integrity.

VII. REFERENCES

[1] C. Wang, Q. Wang, K. Ren, and W. Lou, "Towards Secure and expandable Storage Services in Cloud Computing," Proc. 17th Int'l Workshop Quality of Service, pp. 1-9, July 2009.

[2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Ensuring Data Storage Security in Cloud Computing," Proc. 17th Int'l Workshop Quality of Service , pp. 1-9, July 2009.

[3] Amazon.com, "Amazon Web Services (AWS)," http://aws. amazon.com, 2009

[4] Sun Microsystems, Inc., "Building Customer Trust in Cloud Computing with Transparent Security,"https://www.sun.com/offers/details/sun_transparency.xml, Nov. 2009.

[5] K. Ren, C. Wang, and Q. Wang, "Security Challenges for the Public Cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69- 73, 2012.

[6] K.D. Bowers, A. Juels, and A. Oprea, "Proofs of Retrievability: Theory and Implementation," Proc. ACM Workshop Cloud Computing Security (CCSW '09), 2009.

[7] C. Wang, S.S.M. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy- Preserving Public Auditing for Secure Cloud Storage," IEEE Trans. Computers, preprint, 2012,

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

[9]Ashish Bhagat, Ravi Kant Sahu.,"Using Third Party Auditor for Cloud Data Security: A Review" International Journal of Advanced Research in Computer Science and Software Engineering.

[10] Akkala.Saibabu, T.Satyanarayana Murthy" Security Provision in Publicly Auditable Secure Cloud Data Storage Services Using SHA-1 Algorithm" International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012,4084-4088.

[11] 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, Oct. 2007.

[12]. Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security

[13] M.A. Shah, R. Swaminathan, and M. Baker, "PrivacyPreserving Audit and Extraction of Digital Contents," Cryptology ePrint Archive, Report 2008/186, http://eprint.iacr.org, 2008.

[14] 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 (SecureComm '08), pp. 1-10, 2008.

[15] L. Carter and M. Wegman, "Universal Hash Functions," J.Computer and System Sciences, vol. 18, no. 2, pp. 143-154, 1979.

[16] J. Hendricks, G. Ganger, and M. Reiter, "Verifying Distributed Erasure-Coded Data," Proc. 26th ACM Symp. Principles of Distributed Computing, pp. 139-146, 2007.