IPHV8I1066X
International Journal Of Advanced Research and Innovation -Vol.8, Issue .I
ISSN Online: 2319 – 9253
Print: 2319 – 9245

ACHIEVING SECURE AND DYNAMIC DATA STORAGE IN CLOUD

[#1]Dr.K. RAMESHWARAIAH, Professor and H.O.D, Dept of CSE,

[#2]A.PRASHANTHI, Assistant Professor, Dept of CSE,

[#3]VISHNU VARDHAN REDDY, M.Tech Student, Dept of CSE,

NALLA NARSIMHA REDDY GROUP OF EDUCATIONAL SOCIETY, HYD, T.S., INDIA.

ABSTRACT---Using Cloud Storage, users can remotely store their data and enjoy the on-demand high quality applications and services from a shared pool of configurable computing resources, without the burden of local data storage and maintenance. However, the fact that users no longer have physical possession of the outsourced data makes the data integrity protection in Cloud Computing a formidable task, especially for users with constrained computing resources. Moreover, users should be able to just use the cloud storage as if it is local, without worrying about the need to verify its integrity. Thus, enabling public auditability for cloud storage is of critical importance so that users can resort to a third party auditor (TPA) to check the integrity of outsourced data and be worry-free. To securely introduce an effective TPA, the auditing process should bring in no new vulnerabilities towards user data privacy, and introduce no additional online burden to user. In this paper, we propose a secure cloud storage system supporting privacy-preserving public auditing. We further extend our result to enable the TPA to perform audits for multiple users simultaneously and efficiently. Extensive security and performance analysis show the proposed schemes are provably secure and highly efficient. Our preliminary experiment conducted on Amazon EC2 instance further demonstrates the fast performance of the design.

Keywords- *privacy preserving, Data Integrity, Cloud Storage.*

I.INTRODUCTION

CLOUD computing has been envisioned as the next generation information technology (IT) architecture for enterprises, due to its long list of unprecedented advantages in the IT history: on-demand self-service, ubiquitous network access, location independent resource pooling, rapid resource elasticity, usage-based pricing and transference of risk [2]. As a disruptive technology with profound implications, cloud computing is transforming the very nature of how businesses use information technology. One fundamental aspect of this paradigm shifting is that data are being centralized or outsourced to the cloud. From users' perspective, including both individuals and IT enterprises, storing data remotely to the cloud in a flexible on-demand manner brings appealing benefits: relief of the burden for storage management, While cloud computing makes these advantages more appealing than ever, it also brings new and challenging security threats toward users' outsourced data. Since cloud service providers (CSP) are separate administrative entities, data outsourcing is actually relinquishing user's ultimate control over the fate of their data. As a result, the correctness of the data in the cloud is being put at risk due to the following reasons. First of all, although the infrastructures under the cloud are much more powerful and reliable than personal computing devices, they are still facing the broad range of both internal and external threats for data integrity [4]. Examples of outages and security breaches of noteworthy cloud services appear from time to time [5], [6], [7]. Second, there do exist various motivations for CSP to behave unfaithfully toward the cloud users regarding their outsourced data status. For examples, CSP might reclaim storage for monetary reasons by discarding data that have not been or are rarely accessed, or even hide data loss incidents to maintain a reputation [8], [9], [10]. In short, although outsourcing data to the cloud is economically attractive for long-term large-scale storage, it does not immediately offer any guarantee on data integrity and availability. To address these problems, our work utilizes the technique of public key-based homomorphism linear authenticator (or HLA for short) [9], [13], [8], 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. Specifically, our contribution can be summarized as the following three aspects:

1) We motivate the public auditing system of data storage security in cloud computing and provide a privacy-preserving auditing protocol. Our scheme enables an external auditor to audit user's cloud data without learning the data content.

2) To the best of our knowledge, our scheme is the first to support scalable and efficient privacy-preserving public storage auditing in cloud. Specifically, our scheme achieves batch auditing where multiple delegated auditing tasks from

different users can be performed simultaneously by the TPA in a privacy preserving manner.

3) We prove the security and justify the performance of our proposed schemes through concrete experiments and comparisons with the state of the art.

II. LITERATURE SURVEY

Atomies et al. [9] are the first to consider public audit ability in their "provable data possession" PDP) model for ensuring possession of data files on entrusted storages. They utilize the RSA-based homomorphism linear authenticators for auditing outsourced data and suggest randomly sampling a few blocks of the file. However, among their two proposed schemes, the one with public audit ability exposes the linear combination of sampled blocks to external auditor. When used directly, their protocol is not provably privacy preserving, and thus may leak user data information to the external auditor. Jules et al. [11] describe a "proof of irretrievability" (PoR) model, where spot-checking and error-correcting codes are used to ensure both "possession" and "irretrievability" of data files on remote archive service systems. However, the number of audit challenges a user can perform is fixed a priori, and public audit ability is not supported in their main scheme. Although they describe a straightforward Merkle-tree construction for public PoRs, this approach only works with encrypted data. Later, Bowers et al. [18] propose an improved framework for POR protocols that generalizes Jules' work. Dodis et al. [29] also give a study on different variants of PoR with private audit ability. Sachems and Waters [13] design an improved PoR scheme built from BLS signatures [19] with proofs of security in the security model defined in [11]. Similar to the construction in [9], they use publicly verifiable homomorphism linear authenticators that are built from provably secure BLS signatures. Based on the elegant BLS construction, a compact and public verifiable scheme is obtained. Again, their approach is not privacy preserving due to the same reason as [9]. This problem, if not properly addressed, may impede the success of cloud architecture. As users no longer physically possess the storage of their data, traditional cryptographic primitives for the purpose of data security protection cannot be directly adopted [11]. 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 often insufficient to detect the data corruption only when accessing the data, as it does not give users correctness assurance for those uncased data and might be too late to recover the data loss or damage. Considering the large size of the outsourced data and the user's constrained resource capability, the tasks of auditing the data correctness in a cloud environment can be formidable and expensive for the cloud users [12], [8]. Moreover, the overhead of using cloud storage should be minimized as much as possible, such that a user does not need to perform too many operations to use the data (in additional to retrieving the data).

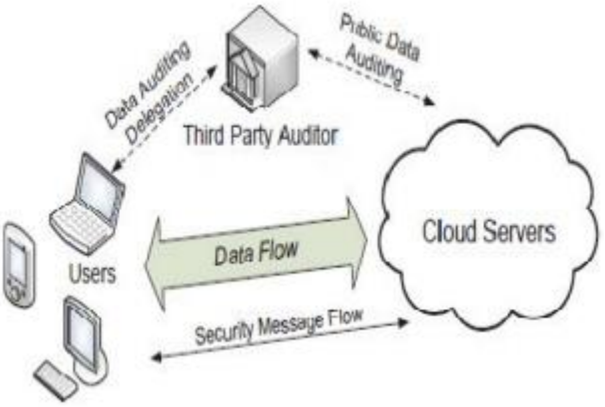

Fig.1 the architecture of cloud data storage service

III. RELATED WORK

The public audit ability in their defined ―provable data possession‖ model for ensuring possession of data files on untrusted storages. Their scheme utilizes the RSA based homomorphism non-linear authenticators for auditing outsourced data and suggests randomly sampling a few blocks of the file. However, the public audit ability in their scheme demands the linear combination of sampled blocks exposed to external auditor. When used directly, their protocol is not provably privacy preserving, and thus may leak user data information to the auditor. Jules etal describe a ―proof of irretrievability‖ model, where spot-checking and error-correcting codes are used to ensure both ―possession‖ and ―irretrievability‖ of data files on remote archive service systems. However, the number of audit challenges a user can perform is fixed a priori, and public audit ability is not supported in their main scheme. Although they describe a straightforward Merkle-tree construction for public PoRs, this approach only works with encrypted data. Dodis et al.give a study on different variants of PoR with private audit ability. Shacham et al.design an improved PoR scheme built with full proofs of security in the security model defined. Similar to the construction, they use publicly verifiable homomorphism non-linear authenticators that are built from provably secure BLS signatures. Construction, a compact and public verifiable scheme is obtained. Again, their approach does not support privacy preserving auditing for the same reason. The propose allowing 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 both the integrity of the data file and the server's possession of a previously committed decryption key. This scheme only works for encrypted files, and it suffers from the auditor state fullness and bounded usage, which may potentially bring in online burden to users when the keyed hashes are used up. The dynamic version of the prior PDP scheme, using only symmetric key cryptography but with a bounded number of audits. Consider a similar support for partial dynamic data storage in a distributed scenario with additional feature of data error localization. In a subsequent work, Wang et al. propose public audit ability and full data dynamics. Almost simultaneously developed a skip lists based scheme to enable provable data possession with full dynamics support. However, the verification in these two protocols requires the linear combination of sampled blocks and thus does not support privacy preserving auditing. While all the above schemes provide methods for efficient auditing and provable assurance on the correctness of remotely stored data, none of them meet all the requirements for privacy preserving public auditing in cloud computing. More importantly, none of these schemes consider batch auditing, which can greatly reduce the computation cost on the TPA when coping with a large number of audit delegations.

IV. EXISTING SYSTEM

The cloud data storage service contains 3 different entities as cloud user, Third party auditor & cloud server / cloud service provider. Cloud user is a person who stores large amount of data or files on a cloud server. Cloud server is a place where we are storing cloud data and that data will be managed by the cloud service provider. Third party auditors will do the auditing on users request for storage correctness and integrity of data. The proposed system specifies that user can access the data on a cloud as if the local one without worrying about the integrity of the data. Hence, TPA is used to check the integrity of data. It supports privacy preserving public auditing. It checks the integrity of the data, storage correctness. It also supports data dynamics & batch auditing. The major benefits of storing data on a cloud is the relief of burden for storage management, universal data access with location independent & avoidance of capital expenditure on hardware, software & personal maintenance.

In cloud, data is stored in a centralized form and managing this data and providing security is a difficult task. TPA can read the contents of data owner hence can modify. The reliability is increased as data is handled by TPA but data integrity is not achieved. It uses encryption technique to encrypt the contents of the file. TPA checks the integrity of the data stored on a cloud but if the TPA itself leaks the user’s data. Hence the new concept comes as auditing with zero knowledge privacy where TPA will audit the users’ data without seeing the contents. It uses public key based homomorphism linear authentication (HLA) [1], [2] which allows TPA to perform auditing without requesting for user data. It reduces communication & computation overhead. In this, HLA with random masking protocol is used which does not allow TPA to learn data content.

A. Goals.

It allows TPA to audit users’ data without knowing data content. It supports batch auditing where multiple user requests for data auditing will be handled simultaneously. It provides security and increases performance through this system.

B. Design Goals

1) Public audit ability: Allows third party auditor to check data correctness without accessing local data.

2) Storage Correctness: The data stored on a cloud is as it. No data modification is done.

3) Privacy preserving: TPA can’t read the users’ data during the auditing phase.

4) Batch Auditing: Multiple users auditing request is handled simultaneously.

5) Light Weight: Less communication and computation overhead during the auditing phase.

C. Batch Auditing

It also supports batch auditing through which efficiency is improved. It allows TPA to perform multiple auditing task simultaneously and it reduces communication and computation cost. Through this scheme, we can identify invalid response. It uses bilinear signature (BLS proposed by Boneh, Lynn and Sachems) to achieve batch auditing. System performance will be faster.

D. Data Dynamics

It also supports data dynamics where user can frequently update the data stored on a cloud. It supports block level operation of insertion, deletion and modification. Author of proposed scheme which support simultaneous public audibility and data dynamics. It uses Merkle Hash Tree (MHT) which works only on encrypted data. It uses MHT for block tag authentication.

V.PROPOSED SYSTEM

In this paper, the TPA will be fully automated and will be able to properly monitor confidentiality and integrity of the data and uniquely integrate it with random mask technique to achieve a privacy-preserving public auditing system for cloud data storage security while keeping all above requirements in mind. Extensive security and performance analysis shows the proposed schemes are provably secure and highly efficient. We are encrypt the data using RSA algorithm cloud computing can be applied to the data transmission security. Transmission of data will be encrypted, even if the data is stolen, there is no corresponding key that cannot be restored. Only the user knows the key, the clouds do not know the key. Also, because the properties of encryption, the cloud can operate on cipher text, thus avoiding the encrypted data to the traditional efficiency of operation. User's privacy is protected because user's files are encrypted in cloud storage. The main issue with the cloud is data integrity, in this paper we are going to use MD5 algorithm for maintain the integrity of data. This MD5 algorithm has more expensive and more secure than other algorithms. MD5 Message Digest is a widely used hash technique, such that it will produce 128-bit hash value we need to convert the input data into bytes in order to convert it to hash value. This is useful in many security applications and it ensures data integrity. Sender creates input message (M) and computes its message digest. Then he uses his private key and encrypts message digest. Encrypted message digest is attached to the input message and the whole message is sent to receiver. Receiver gets the message and extracts the encrypted message digest. Then he computes his own message digest of the received message. He also decodes received message digest with sender's public key and gets decoded message digest. Then he compares both message digests. When both message digests are equal, the message was not modified during the data transmission.

VI. ADVANCED ISSUES IN CLOUD COMPUTING SECURITY

In the previous section, we have discussed generic set of security concerns observed in public and hybrid clouds. We now turn our focus to some atypical cloud specific security issues. In particular, cloud does bring out a set of unique challenges like:

Abstraction: Cloud provides an abstract set of service end-points. For a user, it is impossible to pin-point in which physical machine, storage partition (LUN), network port MAC address, switches etc. are actually involved. Thus, in event of security breach, it becomes difficult for a user to isolate a particular physical resource that has a threat or has been compromised.

***Lack of execution controls*: **The external cloud user does not have fine-gained control over remote execution environment. Hence the critical issues like memory management, I/O calls, access to external shared utilities and data are outside the purview of the user. Client would want to inspect the execution traces to ensure that illegal operations are not performed.

***Third-party control of data*: **In cloud, the storage infrastructure, and therefore, the data possession is also with the provider. So even if the cloud provider vouches for data integrity and confidentiality, the client may require verifiable proofs for the same.

***Multi-party processing*: **In multi-cloud scenario, one party may use part of the data which other party provides. In absence of strong encryption (as data is being processed), it becomes necessary for participating cloud computing parties to preserve privacy of respective data.

Data breach is a big concern in cloud computing. A compromised server could significantly harm the users as well as cloud providers. A variety of information could be stolen. These include credit card and social security numbers, addresses, and personal messages. The U.S. now requires cloud providers to notify customers of breaches. Once notified, customers now have to worry about identify theft and fraud. While providers, have to deal with federal investigations, lawsuits, and bad reputation. Customer lawsuits and settlements have resulted in over $1 billion in losses to cloud providers. Cloud collaboration brings together new advances in cloud computing and collaboration that are becoming more and more necessary in firms operating in an increasingly globalised world. Cloud computing is a marketing term for technologies that provide software, data access, and storage services that do not require end-user knowledge of the physical location and configuration of the system that delivers the services. A parallel to this concept can be drawn with the electricity grid, where end-users consume power without needing to understand the component devices or infrastructure required to provide the service. Collaboration, in this case, refers to the ability of workers in a company to work together simultaneously on a particular task. In the past, most document collaboration would have to be completed face to face. However, collaboration has become more complex, with the need to work with people all over the world in real time on a variety of different types of documents, using different devices. While growth in the collaboration sector is still growing rapidly, it has been noted that the uptake of cloud collaboration services has reached a point where it is less to do with the ability of current technology, and more to do with the reluctance of workers to collaborate in this way. A report by Erica Rugullies mapped out five reasons why workers are reluctant to collaborate more. These are: People resist sharing their knowledge. Users are most comfortable using e-mail as their primary electronic collaboration tool. People do not have incentive to change their behavior. Teams that want to or are selected to use the software do not have strong team leaders who push for more collaboration. Senior management is not actively involved in or does not support the team collaboration initiative.

As a result, many providers of cloud collaboration tools have created solutions to these problems. These include the integration of email alerts into collaboration software and the ability to see who is viewing the document at any time. All the tools a team could need are put into one piece of software so workers no longer have to rely on email based solutions. Recently, cloud collaboration has seen rapid evolution. In the past, cloud collaboration tools have been quite basic with very limited features. Newer packages are now much more document-centric in their approach to collaboration. More sophisticated tools allow users to "tag" specific areas of a document for comments which are delivered real time to those viewing the document. In some cases, the collaboration software can even be integrated into Microsoft Office, or allow users to set up video conferences. Furthermore, the trend now is for firms to employ a single software tool to solve all their collaboration needs, rather than having to rely on multiple different techniques. Single cloud collaboration providers are now replacing a complicated tangle of instant messengers, email and FTP. Cloud collaboration today is promoted as a tool for collaboration internally between different departments within a firm, but also externally as a means for sharing documents with end-clients as receiving feedback. This makes cloud computing a very versatile tool for firms with many different applications in a business environment. The best cloud collaboration tools Use real-time commenting and messaging features to enhance speed of project delivery Leverage presence indicators to identify when others are active on documents owned by another person. Allow users to set permissions and manage other users' activity profiles. Allow users to set personal activity feeds and email alert profiles to keep abreast of latest activities per file or user. Allow users to collaborate and share files with users outside the company firewall.

VII. CONCLUSIONS

In this paper, we proposed watermarking technique for Privacy Preserving Public Auditing for cloud data storage security. Cloud computing security is a major issue that needs to be considered. Using TPA, We can verify the correctness and integrity of data stored on a cloud. It uses public key based homomorphic linear authentication (HLA) protocol with random masking to achieve privacy preserving data security. We achieved zero knowledge privacy through random masking technique. It supports batch auditing where TPA will handle multiple users request at the same time which reduces communication and computation overhead. It uses bilinear signature to achieve batch auditing. It also supports data dynamics. It uses Merkle Hash Tree (MHT) for it. We are introducing Privacy Preserving Public Auditing with watermark process for secure cloud Storage.

REFERENCES

[1] C wang, Sherman S. M. Chow, Q. Wang, K Ren and W. Lou, “Privacy-Preserving Public Auditing for Secure Cloud Storage”,IEEE Trasaction on Computers I, vol. 62, no. 2, pp.362-375 , February 2013.

[2] C. Wang, Q. Wang, K. Ren, and W. Lou, “PrivacyPreserving Public auditing for storage security in cloud computing,” in Proc.of IEEE INFOCOM’10, March 2010.

[3] Wang Shao-hu, Chen Dan-we, Wang Zhi-weiP, Chang Su-qin, “Public auditing for ensuring cloud data storage security with zero knowledge Privacy” College of Computer, Nanjing University of Posts and Telecommunications, China, 2009

[4] KunalSuthar, Parmalik Kumar, Hitesh Gupta, “SMDS: secure Model for Cloud Data Storage”, International Journal of Computer applications, vol56, No.3, October 2012

[5] AbhishekMohta, Lalit Kumar Awasti, “Cloud Data Security while using Third Party Auditor”, International Journal of Scientific & Engineering Research, Volume 3, Issue 6, ISSN 2229-8 June 2012.

[6] Q. Wang, C. Wang,K.Ren, W. Lou and Jin Li “Enabling Public Audatability and Data Dynamics for Storage Security in Cloud Computing”, IEEE Transaction on Parallel and Distributed System, vol. 22, no. 5, pp. 847 – 859,2011.

[7] D. Shrinivas, “Privacy-Preserving Public Auditing in Cloud Storage security”, and International Journal of computer science nad Information Technologies, vol 2, no. 6, pp. 2691-2693, ISSN: 0975-9646, 2011

[8] K Govinda, V. Gurunathprasad and H. sathishkumar, “ Third Party Auditing for Secure Data Storage in Cloud Through Digital Signature Using RSA”, International Journal of Advanced science and Technical Research, vol 4,no. 2, ISSN: 2249-9954,4 August 2012

[9] S. Marium, Q. Nazir, A. Ahmed, S. Ahthasham and Aamir M. Mirza, “Implementation of EAP with RSA for Enhancing The Security of Cloud Computig”, International Journal of Basic and Applied Science, vol 1, no. 3, pp. 177-183, 2012

[10] XU Chun-xiang, HE Xiao-hu, Daniel Abraha, “Cryptanalysis of Auditing protocol proposed by Wang et al. for data storage security in cloud computing”, http://eprint.iacr.org/2012/115.pdf, and cryptologyeprintarchieve: Listing for 2012.

[11] B. Dhiyanesh “A Novel Third Party Auditability and Dynamic Based Security in Cloud Computing” , International Journal of Advanced Research in Technology, vol. 1,no. 1, pp. 29-33, ISSN: 6602 3127, 2011

[12] C. Wang, Q. Wang and K. Ren, “Ensuring Data Storage security in Cloud Computing”, IEEE Conference Publication, 17th International Workshop on Quality of Service (IWQoS), 2009

[13] Balkrishnan. S, Saranya. G, Shobana. S and Karthikeyan. S, “Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud”, International Journal of computer science and Technology, vol. 2, no. 2, ISSN 2229-4333 (Print) | ISSN: 0976-8491(Online), June 2012.