INTERNATIONAL JOURNAL OF RESEARCH
In Advanced Engineering Technologies

Volume 2, Issue 2 SEP 2014

PUBLIC AUDITING SCHEME IN SECURE CLOUD STORAGE

[1]B. Harika PG Scholar in CSE

[1]harikahvs@gmail.com

[2] C.V. Chiranjeevi Kumar M.Tech CSE,

[2]cvchiru@gmail.com

ABSTRACT— Using Cloud Storage, users can distantly store their data and benefit from the on-demand high class 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, facilitating public auditability for cloud storage is of significant importance so that users can resort to a third party auditor (TPA) to ensure the integrity of outsourced data and be worry-free. To securely start 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.

Index Terms— Cloud Computing, Cloud Storage, Privacy, Privacy-Preserving

INTRODUCTION

Security is a necessary component for strong privacy safeguards in all online computing scenarios, but security alone is not enough. Consumers and businesses are ready to use online computing only if they have the reliance that their data will stay private and secure. So to create a trusted environment for customers, we need to develop software, services, and processes with privacy in mind. Cloud computing is the biggest buzz in the computer world these days. Cloud computing is everywhere. The locality of physical resources and devices being accessed are in general not known to the end user. It also provides services for users to build up, deploy and manage their applications „on the cloud", which involves virtualization of resources that maintains and manages by itself [1]. NIST definition of cloud computing:

"Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction".

One of the first cloud offerings was cloud storage and it remains a popular answer. Cloud storage is a model of networked online storage in which the data is stored in virtualized pools of storage that are generally being hosted by the third parties. Cloud storage allows data stored remotely to be temporarily cached on mobile phones, desktop computers, or other Internet-linked devices. Security and cost are the top issues in this field and vary greatly, depending on the vendor one choose. Despite the first success and recognition of the cloud computing model and the extensive availability of providers and tools, a number of challenges and risks are innate to this new model of computing.

Privacy in cloud storage

Newly, many services in the cloud, e.g., healthcare, online marketing, banking & payment, and social media depend on the use of personal information. Those privacy-sensitive data are residing in the other side of the globe. This movement highlights concerns on privacy in the cloud like how privacy of users is perceived and protected. For these

growing privacy concerns, many technologies have been proposed, and governments in the world are preparing lawful frameworks to protect privacy. Nevertheless, there are still gaps between practices and proposed solutions, conflicts of interests, and disagreement on requirements and concepts.

What is privacy?

Privacy means that the person to be free from all interference. Privacy control allows the person to maintain a degree of intimacy. Privacy is the protection for the truthful use of personal information of cloud user. Privacy breaches may create a lot of troubles to cloud users.

The American Institute of Certified Public Accountants (AICPA) and Canadian Institute of Charted Accountants (CICA) define that, *"Privacy is the right and obligation of individuals and organizations with respect to the collection, use, retention, and disclosure of personal information"*.

Privacy issues in cloud storage

When considering privacy risks in the cloud context, it is very important as privacy threats differ according to the type of cloud scenario. The papers [2][3] presents some of those issues in privacy which is as follows: lack of user control, lack of training and expertise, unauthorized secondary usage, complexity of regulatory compliance, addressing transborder data flow restrictions, litigation, legal uncertainty, compelled disclosure to the government, data security and disclosure of breaches, data accessibility, location of data, transfer and retention.

Related Work

Ateniese et al. [9] are the first to consider public auditability in their "provable data possession" PDP) model for ensuring possession of data files on untrusted storages. They utilize the RSA-based homomorphic 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 auditability 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. Juels et al. [11] 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. However, the number of udit challenges a user can perform is fixed a priori, and public auditability 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 Juels' work. Dodis et al. [29]
also give a study on different variants of PoR with private auditability. Shacham 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 homomorphic 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 unaccessed 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).

System Design

Cloud Computing is presently one of the hottest topics in information technology (IT). Since the outsourcing of all the essential data is available with a third party, there is always having a concern of cloud service provider‟s trust-worthiness. Due to data privacy, it is essential for users to encrypt their sensitive data before storing them into the

cloud. Yet, there exist some shortcomings in the situation of traditional encryption. When a secret key owner wants to look for some data that are stored in the cloud storage, he may be needed to download all encrypted data from the cloud server, and then decrypts and searches them. If the encrypted data are huge or the client is a mobile user, then it will be very inefficient and is not convenient. Otherwise he must send his key to the cloud server which performs the decryption and search procedures. It causes a serious trouble that the cloud server obtains the secret key So many models were existed to ensure the integrity of data file.

In "Provable Data Possession" (PDP) model [4] ensures the possession of data files on untrusted storages. It uses a RSA based homomorphic linear authenticator for auditing outsourced data, but this model leaks the data to external auditors and hence was not provably privacy preserving. Juels et.al [5] describes a "Proof of Retrievability" (PoR) model, where spot-checking and error correcting codes are used in order to ensure the possession and retrievability. But this approach works only with encrypted data. Improved versions of PoR protocols had been proposed which guarantees private auditability and one which make use of BLS signatures. But these approaches were not privacy-preserving. Then comes the TPA based approach to keep online storage honest. This scheme only works for encrypted files which requires the auditor to keep state, and suffers from bounded usage, which potentially brings in online burden to users when the keyed hashes are used up.

Thus to provide secure cloud storage supporting privacy-preserving many methodologies, frameworks and protocols have been proposed. This paper examines those existing methodologies that guarantee the privacy in cloud storage by categorizing it into four types by which the privacy in cloud storage is achieved and performs the analysis on the existing methodologies which best suits to deal with the privacy issue.

Encryption Methods

There are approaches that make use of encryption techniques to achieve privacy in cloud and the papers [6]-[10] state about it. RuWei et.al [6] proposed the design of privacy-preserving cloud storage framework to solve privacy security problem, this comprises the design of data organization structure, the generation and management of keys, the interaction between participants and the handling of change of user‟s access right and also supports the dynamic operations of data. It uses an interactive protocol and an extirpation based key derivation algorithm. It ensures data confidentiality, solve ineffectiveness of key derivation, reduces the burden of encryption and decryption, can be able to manage numerous keys, saves owners storage space, reduce run-time overheads of the system, gives excellent privacy security and can apply to multiple users, data owners and service providers. But it needs to have techniques to reduce owner‟s encryption burden and to work on ciphertext.

A method for improving user privacy with secret key recovery in cloud storage that allows users‟ to encrypt their files in the cloud storage has been proposed in [7]. A Secret sharing Algorithm to Key Recovery Mechanism is used. AES-128 to encrypt user‟s file, the key length is set to 128 bits is used. Key Recovery scheme partially trusted because no one has the full information about the encryption key except the user himself. The compression algorithm used here is ZIP. The user‟s privacy is protected and it decreases the risk of encryption key lose. But it puts a big computation burden for users. It has concerns about transforming speed. Renewing user‟s key is a challenge here, users can‟t search words and there is dispersal of information. RuWei et.al in [8] provides a privacy-preserving cloud storage framework supporting ciphertext retrieval, it is to solve the problems while operating on an encrypted data and to reduce the data owner‟s workload on management of data and support data sharing. Interaction protocol, Key derivation Algorithm, combination of symmetric and asymmetric encryption and Bloom Filter is used here. It can operate on encrypted data; reduce data owner‟s workload on managing the data and storage space, reduce communication, computation and storage overhead. It can manage numerous keys and is efficient, safe and economic. But it supports only owner-write-user-read and lacks in technique that support cipher text-based computing.

The paper [9] is about controllable privacy preserving search functionalities which include revocable delegated search and un-decryptable delegated search that are based on symmetric predicate encryption in the cloud storage. Thus the Owner of cloud can easily control lifetime and search privileges of data which is suitable for delegation-based business applications. But it cannot support complex access control and search privileges. A method using discretion algorithm [10] for preserving privacy through data control in a cloud computing architecture, which provides security solution

that requires more than user authentication and digital certificate are discussed. Here the SP can directly use data without any key and is more flexible and safe to protect individuals‟ privacy. But the use of Encryption limits data usage and needs communication and compatibility with heterogeneous host.

The main problem in using encryption based technique is that it limits the data usage and puts into an additional burden. The access control mechanisms are available which will overcome the burden of the above overheads.

Access Control Mechanisms

The access control mechanisms that provide privacy has been discussed in papers [11] and [9]. A privacy preserving access authenticated access control scheme for securing data in clouds that verifies the authenticity of the user without knowing the user‟s identity before storing information has introduced in [11]. Here only valid users are able to decrypt the stored information. It prevents reply attack, achieves authenticity and privacy. It is decentralized and robust which allow multiple read and write, distributed access control and the identity of user is protected. But in [9] the access policy for each record stored in the cloud should be known and should be based on assumption that cloud administrator is honest but it does not support complex access control.

Query Integrity/Keyword Searches

There are approaches that make use of queries and keyword search scheme to check the privacy in cloud and papers [12]-[13] discusses those schemes. Qin Liuy et.al [12] proposed an efficient privacy preserving keyword search scheme in cloud computing that allows a service provider to participate in partial decipherment and enables them to search the keywords on encrypted files. It makes use of an efficient privacy preserving keyword search scheme (EPPKS). It provides protection of user data privacy, queries privacy and support key word search on encrypted data. It is found efficient, practical and provably and semantically secure. But the computation on encrypted data was a challenge. A privacy preserving approach for data outsourcing in cloud environment which make use of Fragmentation and heuristic algorithm is used by Sayi et.al [13]. It proves to be efficient and effective but confidentiality is not achieved.

Auditability Schemes

Auditing reduces risk for customers as well as give incentives to providers to improve their services [14].

higher scheme efficiency, public auditability permits Auditability falls under two categories as follows when we

3.4.2 Public Verifiability for Storage Security

A work has been done in [17] for studying the problem in ensuring the integrity of data storage in Cloud. To ensure the correctness of data they allow a third party auditor to work on behalf of the cloud consumer, to check the integrity of the stored data in the cloud. This scheme ensures that the storage at the client side is minimal which will be helpful for thin clients.

3.4.3 Remote Data Checking Using Provable Data Possession

This paper [19] Ateniese et.al, introduces a model for provable data possession which can be used for remote data checking. By having a sampling random set of blocks from the server, this model produces probabilistic proofs of possession which will significantly reduce I/O costs. In order to minimize network communication the challenge/response protocol transmits a small and constant amount of data. The model incorporates some mechanisms for mitigating arbitrary amounts of data corruption and it is robust. It offers two efficient secure *PDP* schemes and the overhead at the server is low. To add robustness to any remote data checking scheme based on spot checking it proposes a generic transformation.

3.5.4 Privacy Preserving Data Integrity Checking

A privacy preserving remote data integrity checking protocol with data dynamics and public verifiability [20] make use of a Remote Data Integrity Checking Protocol. The protocol provides public verifiability without the help of a third party auditor. It doesn‟t leak any privacy information to third party, which provides good performance without the support of the trusted third party and provides a method for independent arbitration of data retention contracts. But it gives unnecessary computation and communication cost.

3.5.5 Privacy Preserving Public Auditability for Storage Security

The studies about the problem that ensures integrity of the data storage in Cloud Computing has been analyzed in paper [18]. It allows a third party auditor to confirm the integrity of dynamic data stored in the cloud. This scheme achieves both public auditability and dynamic data operations. The authors in [21] propose privacy-preserving public auditing for secure cloud storage. The protocol design to achieve the security and performance guarantees like: Public Auditability, Storage consider the available schemes in auditability: private auditability and public auditability. Even though schemes with private auditability can attain

anyone, not just the client (data owner), to deal with the cloud server for correctness of data storage while keeping no private information. Then, clients are able to pass on the evaluation of the service performance to an independent third party auditor (TPA), without giving their computation resources. So we can denote the types of auditing protocols as Data Owner Auditing and Third Party Auditing.

According to [15] the methods of data storage auditing methods can be categorized into three: Message Authentication Code (MAC) - based methods, RSA- based Homomorphic methods and Boneh-Lynn-Shacham signature (BLS) – based Homomorphic methods. The challenging issues of data storage auditing include Dynamic Auditing, Collaborative Auditing and Batch Auditing. We need to meet the three performance criteria when comes to designing of auditing protocols as: low storage overhead, low communication cost and low computational complexity.

This paper compares the schemes available in public auditability. The papers [16]-[22] are some works related to public auditability in cloud. But the papers [16]-[19] are provably not privacy preserving but they lead to the development of efficient privacy-preserving methodologies in papers [20]-[22].

3.4.1 Remote Data Possession at Untrusted Host

The paper [16] has been proposed with the goal of remote data possession checking schemes. This paper proposes an efficient RDPC scheme which is efficient in terms of computation and communication; it allows verification without the need for the challenger to compare against the original data; it uses only small challenges and responses, and users need to store only two secret keys and several random numbers. Finally, a challenge updating method is proposed based on Euler‟s theorem.

Correctness, Privacy-Preserving, Batch auditing, and to be Lightweight. The method is found to be scalable and efficient which provides complete outsourcing solution, integrity checking and thus saves amount of auditing time. It relies on third party auditors and has the use of expensive modular exponentiation operations which leads to storage overhead on server and extra communication cost. Zhu et.al [22] proposes an efficient audit service outsourcing for data integrity in clouds. It is based upon the creation of an interactive PDP protocol to inhibit the dishonesty of prover (soundness property) and the leakage of verified data (zero-knowledge property). It describes the periodic verification for improving the performance of audit services. Here the approach adopts a way of sampling verification. The scheme not only prevents the deception and forgery of cloud storage providers, but also prevents the leakage of outsourced data in the process of verification. It supports an adaptive parameter selection.

The system shows only lower computation cost as well as a shorter extra storage and the scheme is less complex due to fragment structure. It achieves Audit-without-downloading, Verification-correctness, Privacy-preserving and High-performance.

The TPA Contains the information about the users sessions as shown below :

CONCLUSION

Cloud data security is an important aspect for the client while using cloud services. Third Party Auditor can be used to ensure the security and integrity of data.

Third party auditor can be a trusted third party to resolve the conflicts between the cloud service provider and the client. Various schemes are proposed by authors over the years to provide a trusted environment for cloud services. Encryption and Decryption algorithms are used to provide the security to user while using third party auditor.

This paper provides an abstract view of different schemes proposed in recent past for cloud data security using third party auditor. Most of the authors have proposed schemes which rely on encrypting the data using some encryption algorithm and make third party auditor store a message digest or encrypted copy of the same data that is stored with the service provider.

The third party is used to resolve any kind of conflicts between service provider and client.

REFERENCES

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

[2] 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

[3] 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.

[4] K. Kiran Kumar, K. Padmaja, P. Radha Krishna, "Automatic Protocol Blocker for Privacy-Preserving Public Auditing in Cloud Computing", International Journal of Computer science and Technology, vol. 3 pp, ISSN. 0976-8491(Online), pp. 936-940, ISSN: 2229-4333(Print), March 2012

[5] Lingaraj Dhabale, Priti Pavale, "Providing Secured Data Storage by Privacy and Third Party Auditing In Cloud", International Conference on Computing and Control Engineering, ISBN 978-1-2248-9, 12 & 13 April, 2012

[6] Jachak K. B., Korde S. K., Ghorpade P. P. and Gagare G. J. ,"Homomorphic Authentication with Random Masking Technique Ensuring Privacy & Security in Cloud Computing", Bioinfo Security Informatics, vol. 2, no. 2,pp. 49-52, ISSN. 2249-9423, 12 April 2012

[7] Dr. P. K. Deshmukh, Mrs. V. R. Desale, Prof. R. A. Deshmukh, "Investigation of TPA (Third Party Auditor Role) foe Cloud Data Security", International Journal of Scientific and Engineering Research, vo. 4,no. 2,ISSn 2229-5518, Feb 2013.

[8] Gayatri. R, "Privacy Preserving Third Party Auditing for Dynamic Data", International Journal of Communication and engineering, vol. 1, no. 1, issue: 03, March 2012

[9] Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009, http://www. cloudsecurityalliance.org. [20] Juels, B. Kaliski. "Pors: proofs of retrievability for large files[C]", Proceedings of CCS 2007. Alexandria, VA, USA, 2007. 584-597.

[10] Honywei Li, Yuanshun Dai, Bo Yang. "Identity-Based Cryptography for Cloud Security".

[11] C. Hota, S. Sanka, M. Rajarajan, S. Nair, "Capabilitybased Cryptographic Data Access Control in Cloud Computing", in International Journal of Advanced Networking and Applications, Volume 01, Issue 01, 2011.

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

[13] C. Wang, Q. Wang, K.Ren, and W.Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage", in Proc. Of IEEE INFOCOM‟10,March 2010.

[14] Y. Zhu,Z. Hu,Gail-J Ahn, H. Hu,Stephen S. Yau, Fellow, IEEE, Ho G. An, and Shimin Chen,"Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds",in Proc.of IEEE SAC‟11 March 2011.

[15] Q. Wang, C. Wang, Kui Ren, W.Lou and Jin Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing", in IEEE transaction on parallel and distributed system May 2011.

[16] Nandeesh.B.B, Ganesh Kumar R, Jitendranath Mungara"Secure and Dependable Cloud Services for TPA in Cloud Computing" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-1, Issue-3, August 2012.

[17] M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage service honest", in Proc. Of HotOS‟07, CA, USA: USENIX Association, 2007, pp.1-6.

[18] Muralikrishnan Ramane and Bharath Elangovan, "A Metadata Verification Scheme for Data Auditing in Cloud Environment", International Journal on Cloud Computing: Services and Architecture(IJCCSA), Vol.2, no.4, August 2012.

[19] Dalia Attas and Omar Batrafi "Efficient integrity checking technique for securing client data in cloud computing", October 2011.

[20] S. Balakrishnan, G. Saranya, S. Shobana, S. karthikeyan, "Introducing Effective Third Party Auditing(TPA) for Data Storage in Cloud" IJCST Vol. 2, Issue 2, June 2011.

[21] Cryptography and Network Security Chapter 12 – *Hash* Algorithms.http://vlsi.byblos.lau.edu.lb/classes/csc736/Notes/Lecture12.pdf