RESEARCH ARTICLE OPEN ACCESS

Towards Secrecy on Public Auditability for Secure Storage in Cloud Computing

Jency. D. J*, II M. Tech, D. Joseph Pushparaj[#],
Assistant Professor Department of Information Technology, Francis Xavier Engineering College.

Abstract
Cloud storage helps the users to save their large amount of data remotely without any online burden. But there is no guarantee for the physical ownership of data to the users. This makes the data integrity protection a terrifying chore. Moreover, users have the rights to just use the cloud storage as if it is local, without the need to check its integrity. So, permitting public auditability for cloud storage is essential. For that, third party auditor (TPA) is introduced. In order to host an effective TPA, the auditing process doesn't affect the user data privacy and should prevent the online burden to user. In this paper, a secure cloud storage system supporting privacy preserving public auditing is proposed. We further extend our result to facilitate the TPA to perform audits for multiple users simultaneously. Also we propose a protocol supporting for fully dynamic data operations, especially to support block insertion, and other modification of data which is missing in most existing schemes. Extensive security and performance analysis show the proposed schemes are provably secure and highly efficient.

I. INTRODUCTION

CLOUD computing has been considered as the next generation information technology (IT) architecture for ventures, due to its long list of unparalleled benefits in the IT history. Cloud computing is an Internet based development and use of computer technology. Cloud computing is a general term for anything that involves delivering the services over the Internet. Cloud computing explains the basic applications that are limitless to be accessible through the Internet. These cloud applications use large data centres and powerful servers that host the Web applications and Web services. Cloud computing and storage, takes advantage from years of development and testing of large scale infrastructure. Cloud helps to store the data of the user in the cloud. "The cloud" is so overused by the startups desperate for VC money, and by the big companies desperate to look like hip startups, that IT professionals are increasingly cautious. Cloud computing is used for computation, and accessible in all platforms.

The infra structure affects from internal and external threats. Examples of outages and the security breaches of noteworthy cloud services appear from time to time. Second, there do live various motivations for CSP to behave unfaithfully toward the cloud users regarding their outsourced data status. For examples, CSP may delete the user files that are rarely accessed and even hide the data loss to maintain a reputation. Simply downloading the files from the cloud are not convenient because of the IO cost. The data corruption can be detected when the data should be accessed, so the storage correctness may not be guaranteed. The tasks of auditing the large data files correctness in a cloud environment can be formidable and expensive for the cloud users.

To tackle all the above said problems, a third-party auditor (TPA) is introduced to audit the outsourced data when needed. The TPA, who has expertise and have capabilities that users do not, It can periodically check the integrity of all the data stored in the cloud on behalf of the users. Usually, the CSP produce a copy of data for the users. This becomes an information leakage from the cloud.

There are several legal regulations like the US Health Insurance Portability and Accountability Act (HIPAA) further helps the outsourced data not to be leaked to external parties. The encryption of outsourced data before storing it in the cloud is the only solution but it creates the online burden to the user.

To overcome these problems, Homomorphic Linear Authenticator is used here, which does not need any copy of outsourced data. By integrating the HLA with random masking, our protocol guarantees that the TPA does not contain any knowledge about the contents stored in the CSP. The aggregation and algebraic properties of the authenticator benefit our design for the batch auditing. Our contribution can be summarized as the following aspects: 1. we provide the public auditing system of data storage security in cloud computing and also provide a privacy-preserving auditing protocol. Our system provides an external auditor which does not have any knowledge about the data of the user. 2. This scheme is used to support scalable and efficient privacy-preserving public storage auditing in cloud. Specifically, it achieves batch auditing where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA.

II. PROBLEM STATEMENT

The System and Threat Model:

We consider a cloud data storage service involving three different entities, as illustrated in the following Fig. The cloud user contains a large amount of data files to be stored in the cloud. The cloud server is managed by the cloud service provider to offer the data storage service and has significant storage space and computation resources. The third-party auditor has expertise and has capabilities that cloud users do not have and it is trusted to assess the cloud storage service reliability on behalf of the user upon request. Users depend on the CS for cloud data storage and maintenance. They may also energetically interact with the CS to access and update their stored data for various application purposes.

As the user does not have their data locally, it is very important for the users to ensure that their data are being correctly stored and maintained. To save the computation resource and the online burden potentially brought by the periodic storage correctness verification, cloud users may route to TPA for certifying the storage integrity of their outsourced data, while eager to remain their data private from TPA.

We presume the data integrity threats toward the user's data can arrive from both internal and external attacks at CS. CS can be self-interested. For their advantages, such as to maintain reputation, CS may hide these data corruption incidents to users. Using third-party auditing service, it provides a cost-effective method for users to make trust in cloud.

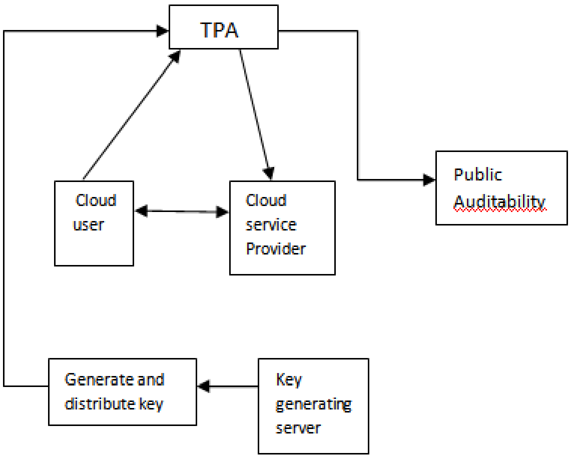

Fig: The architecture of Cloud data storage services

We assume the TPA is in the business of auditing and it is reliable and independent. However, it may damage the user if the TPA could learn the outsourced data after the audit. On the one hand, there are several regulations, e.g., HIPAA, requesting CS to maintain the user's data privacy. On the other hand, as user's data belong to their business asset, they also survive financial incentives for CS to protect it from any external parties. Therefore, here we assume that neither CS nor TPA has motivates to collude with each other during the auditing process. To approve the CS to respond to the audit delegated to TPA's, the user may issue a certificate on TPA's public key, and all audits from the TPA are validated against such a certificate.

III. DESIGN GOALS

To enable confidentiality on public auditing for cloud computing, the following steps are taken.

1. Public auditability: allows the TPA to check the data of the user periodically in the cloud.

2. Storage correctness: to ensure that there survives no cheating cloud server that can pass the TPA's audit without indeed storing user's data are unbroken.

3. Privacy preserving: to ensure that the TPA cannot extract the user's data from the information collected together during the auditing process.

4. Batch auditing: to enable the TPA with secure and efficient auditing capability to manage with multiple auditing delegations from possibly large number of the different users simultaneously.

5. Lightweight: to allow TPA to perform auditing with low communication and computation overhead.

IV. THE PROPOSED SCHEMES

A public auditing scheme consists of four algorithms. They are KeyGen, SigGen, GenProof, VerifyProof.

KeyGen is a key generation algorithm that is run by the user to setup the plan. SigGen is used to make verification of metadata by the user, which contains of digital signatures. GenProof is run by the cloud server to generate a proof of data storage correctness, and VerifyProof is run by the TPA for auditing the proof. Running a public auditing system consists of two phases. They are Setup and Audit:

Setup: The user initializes the public and secret parameters of the system by processing KeyGen, and preprocesses the data file F with SigGen to produce the verification metadata. The user next stores the data file F and the verification metadata at the cloud server is done, and deletes its local copy. As a part of preprocessing, the user may change or modify the data file F by escalating it or including supplementary metadata to be stored at server.

Audit: The TPA issues an audit message or challenge to the cloud server to make sure that the cloud server has reserved the data file F properly at the time of audit. The cloud server will conclude a response message by processing GenProof using F and the verification metadata as an inputs. Then the TPA verifies a response via VerifyProof. This systematic framework assumes that the TPA is stateless, i.e., TPA may not need to maintain and update the state between audits. It is easy to extend the framework to capture a stateful auditing system, essentially by dividing the verification of metadata into two parts which are stored by the TPA and the cloud server. This systematic design does not take up any additional property on the data file. If the user needs to have more error resilience, then he can first redundantly encodes the data file and next uses our system with the data that contains error correcting codes integrated.

V. THE BASIC SCHEMES

We have to concentrate upon two classes of schemes. The first one is a MAC-based solution which suffers from undesirable systematic demerits: bounded usage and stateful verification, which may pose additional online burden to users, in a public auditing setting. This also shows that the auditing problem is still not easy to solve even if we have introduced a TPA. The second one is a system based on homomorphic linear authenticators, which covers much recent proof of storage systems. Our main plan to be presented is based on a specific HLA scheme. There are two possible methods to make use of MAC to validate the data. A small way is to upload the data blocks with their MACs to the server, and sends its corresponding secret key to the TPA. Later, the TPA can randomly get back the blocks with their MACs and check the correctness through the secret key. Apart from the high communication and computation complexities, the TPA needs the knowledge of the data blocks for verification. The TPA can disclose a secret key to the cloud server and ask for a fresh keyed MAC for comparison in each audit. This is confidential as long as it is impossible to recover F. However, it suffers from the following severe drawbacks:

1) The number of times a particular data file can be audited is limited by the number of secret keys that must be fixed a priori. Once all possible secret keys are exhausted, the user then has to retrieve data in full to recomputed and republish new MACs to TPA;

2) The TPA also has to maintain and update state between audits, i.e., keep track on the revealed MAC keys. Considering the potentially large number of audit delegations from multiple users, maintaining such states for TPA can be difficult and error prone;

3) It can only support static data, and cannot efficiently deal with dynamic data at all. However, supporting data dynamics is also of critical importance for cloud storage systems. For the reason of brevity and clarity, our main protocol will be presented based on static data.

HLA-BASED SOLUTION

To effectively support public auditability without having to retrieve the data blocks themselves, the HLA technique can be used. HLAs, like MACs, are also some unforgivable verification metadata that authenticate the integrity of a data block. The difference is that HLAs can be aggregated. It is possible to compute an aggregated HLA which authenticates a linear combination of the individual data blocks.

Privacy-Preserving Public Auditing Scheme:

To achieve privacy-preserving public auditing, we propose to uniquely integrate the homomorphic linear authenticator with random masking technique. In our protocol, the linear combination of sampled blocks in the server's response is masked with randomness generated by the server. With random masking, the TPA no longer has all the necessary information to build up a correct group of linear equations and therefore cannot derive the user's data content, no matter how many linear combinations of the same set of file blocks can be collected. On the other hand, the correctness validation of the block-authenticator pairs can still be carried out in a new way which will be shown shortly, even with the presence of the randomness. Our design makes use of a public key-based HLA, to equip the auditing protocol with public auditability. Specifically, we use the HLA proposed in which is based on the short signature scheme. The TPA verifies the storage correctness and then checks for verification.

Properties of our protocol:

It is easy to see that our protocol achieves public auditability. There is no secret keying material or states **for** the TPA to keep or maintain between audits, and the auditing protocol does not pose any potential online burden on users. This approach ensures the privacy of user data content during the auditing process by employing a random masking to hide, a linear combination of the data blocks. Besides, the HLA helps to achieve the constant communication overhead for server's response during the audit. For certain cloud storage providers, it is possible that certain information dispersal algorithms (IDA) may be used to breakup and distribute the user's outsourced data for increased availability. Here, the cloud side operations would not affect the behaviour of our planned mechanism, as long as the

IDA is systematic. This is from user's point of view, as long as there is a complete unchanged copy of outsourced data in cloud, the pre-evaluated verification of metadata; As a result, the metadata can be used in our auditing mechanism to verify the correctness of user's outsourced cloud data.

Support for Batch auditing:

With the establishment of confidentiality on public auditing, the TPA may simultaneously control multiple auditing upon different users delegation. The individual auditing of the tasks for the TPA may be tedious and inefficient. Given K auditing delegations upon the K distinct data files from the K different users, it is more benefit for the TPA to group these multiple tasks together and audit at one time. Keeping the basic requirement in mind, we change the protocol in a single user case, and achieve the aggregation of K verification in to single one.

Support for Data Dynamics:

In cloud computing, the data are not only accessed by the user but also frequently uploaded. So, the data dynamics support are very important. The support of data dynamics is achieved by the classic data structure—Merkle hash tree (MHT) for the underlying block sequence enforcement. Here the data dynamics helps to find the data modification. Our system prevents the unauthorized modification.

VI. CONCLUSION

In this system, A privacy-preserving public auditing system for data storage security in cloud computing is proposed. Here, the homomorphic linear authenticators with random masking technology are used. So, the TPA would not have any knowledge about the data content of the user stored on the cloud server during the auditing process, which not only decrease the burden of cloud user from the auditing task but also lighten the user's fear of their outsourced data leakage. Considering TPA may simultaneously handle multiple audit sessions from different users for their outsourced data files, the system further extends the confidentiality on public auditing protocol into a multiuser setting for audit, where the TPA can execute multiple auditing tasks in a group manner for better efficiency. The System also supports for the data dynamics. In the future, we plan to increase the speed of the auditing task for increasing the performance of multiple auditing tasks. Further, service level agreements will be provided to improve the security measures.

REFERENCES

[1] "Privacy-Preserving Public Auditing for Secure Cloud Storage" Cong Wang, Sherman S.M. Chow, Qian Wang, Kui Ren, , and Wenjing Lou,IEEE TRANSACTIONS ON COMPUTERS,Feb 2013

[2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy Preserving Public Auditing for Storage Security in Cloud Computing," Proc. IEEE INFOCOM '10, Mar. 2010.

[3] P. Mell and T. Grance, "Draft NIST Working Definition of Cloud Computing," http://csrc.nist.gov/groups/SNS/cloudcomput ng/ index.html, June 2009.

[4] 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, "Above the Clouds: A Berkeley View of Cloud Computing," Technical Report UCB-EECS 2009-28, Univ. of California, Berkeley, Feb. 2009

[5] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "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, May 2011.

[6] 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, 2007.

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