Privacy-Preserving Public Examining for Protected Cloud Storage

M. JAYAPAL REDDY
Dept of Computer Science & Engineering
Siddartha Educational Academy Institutions, Tirupathi
Tirupathi (India)

D. LAVANYA
Professor,
Dept. of Computer Science & Engineering
Siddartha Educational Academy Institutions, Tirupathi
Tirupathi (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 audit ability 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 toward 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: *Cloud Storage, Extensive security, Privacy, TPA*

1. Introduction

Cloud computing is the next stage in the Internet's evolution, providing the means through which everything from computing power to computing infrastructure, applications, business processes to personal collaboration can be delivered as a service wherever and whenever need. 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.

1.1 Cloud Services

Cloud computing is anything that involves services over the internet. These services are broadly classified into three categories: software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS). Cloud software as a service (SaaS) is the on-demand service developed for end users; provider will license the software for their own use. As the software is managed over the central location over the web, the user need not required to handle the upgrades. E.g.- Gmail. And the next service is cloud platform as a service (PaaS) is designed for the

application developers, which provide all the facilities for developing the web applications easily with more features without the complexity of buying and maintaining the software and the infrastructure. E.g.-Google App Engine. Finally the cloud infrastructure as a service (IaaS) is way of delivering the cloud computing infrastructure which provisions the storage, service and network. As it is fully outsources service it is not necessary to purchase the server, software and other equipment's for the business and the service providers benefit from cost saving.

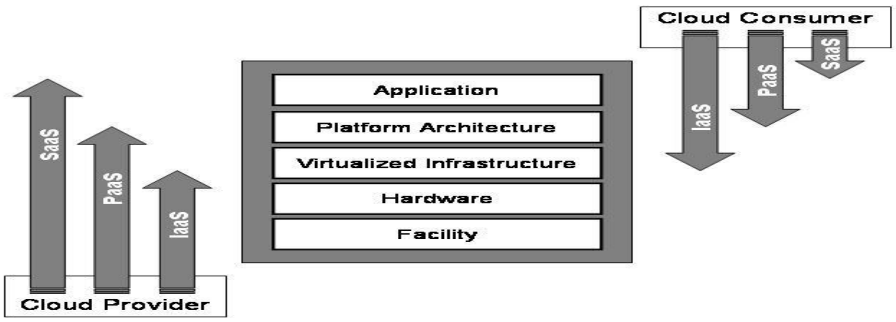

Fig. 1 Differences in Scope and Control among Cloud Service Models

1.2 Cloud Storage

Cloud storage is an important service of cloud computing, which allows data owners (owners) to move data from their local computing systems to the cloud. More and more owners start to store the data in the cloud. However, this new paradigm of data hosting service also introduces new security challenges. Owners would worry that the data could be lost in the cloud. This is because data loss could happen in any infrastructure, no matter what high degree of reliable measures cloud service providers would take. Sometimes, cloud service providers might be dishonest. They could discard the data which has not been accessed or rarely accessed to save the storage space and claim that the data are still correctly stored in the cloud. Therefore, owners need to be convinced that the data are correctly stored in the cloud.

In existing system, the clients store the data in server that server is trustworthy and after the third party auditor can audit the client files. So, the third party auditor can stolen the files. The main Disadvantage of the Existing system can support both features with the help of a third party auditor. Consider a cloud storage system in which there is a client and a un trusted server. The user stores their data in the server without keeping a local copy. Hence, it is of critical importance that the client should be able to verify the integrity of the data stored in the remote un trusted server. If the servers modify any part of the client's data, the user should be able to detect it; furthermore, any third party auditor should also be able to detect it. In case a third party auditor verifies the integrity of the client's data, the data should be kept private against the third party auditor. Advantages of the proposed scheme have the following main contributions: Remote data integrity checking protocol for cloud storage. The proposed system inherits the support of data dynamics, and supports public verifiability and privacy against third-party verifiers, while at the same time it doesn't need to use a third-party auditor. Data correctness and security analysis of the proposed system which shows that data is secure against the un trusted cloud service provider and private against Third Party Auditor.

2. Statement of the Problem

2.1 The Cloud and Threat Model

The Cloud security responsibilities can be taken on by the customer, if he is managing the cloud, but in the case of a public cloud, such responsibilities are more on the cloud provider and the customer can just try to assess if the cloud provider is able to provide security. cloud data storage service involving three different entities. 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 *cloud service provider*

(CSP) to provide data storage service and has significant storage space and computation resources (we will not differentiate CS and CSP hereafter.); 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 security on behalf of the user upon request. Cloud users dynamically interact with the CS to access and update their stored data for various application purposes. The traditional cryptographic technologies for data integrity and availability, cannot work on the outsourced data without a local copy of data. It is not a practical solution for data validation by downloading them due to the expensive communications, especially for large size files. The ability to audit the correctness of the data in a cloud environment can be formidable and expensive for the cloud users. Therefore, it is crucial to realize public audit ability for CSS, so that data owners may resort to a third party auditor (TPA), who has expertise and capabilities that a common user does not have, for periodically auditing the outsourced data. This audit service is significantly important for digital forensics and credibility in clouds. The users may resort to TPA for ensuring the storage security of their outsourced data, while hoping to keep their data private from TPA. We consider the existence of a semi-trusted CS as [11] does. Namely, in most of time it behaves properly and does not deviate from the prescribed protocol execution. However, during providing the cloud data storage based services, for their own benefits the CS might neglect to keep or deliberately delete rarely accessed data files which belong to ordinary cloud users. Moreover, the CS may decide to hide the data corruptions caused by server hacks or Byzantine failures to maintain reputation. We assume the TPA, who is in the business of auditing, is reliable and independent, and thus has no incentive to collude with either the CS or the users during the auditing process. TPA should be able to efficiently audit the cloud data storage without local copy of data and without bringing in additional on-line burden to cloud users. However, any possible leakage of user's outsourced data towards TPA through the auditing protocol should be prohibited. The audit delegation and authorize CS to respond to TPA's audits, the user can sign a certificate granting audit rights to the TPA's public key, and all audits from the TPA are authenticated against such a certificate.

2.2 Design Goals

The privacy-preserving public auditing for cloud data storage under the aforementioned model, our protocol design should follow the security and performance.

Public Audit: It allows TPA to verify the correctness of the cloud data on demand without retrieving a copy of the whole data

Storage Consistency: the data in cloud server that can pass the audit from TPA without indeed storing users' data intact.

Privacy-Preserving: to ensure that there exists no way for TPA to derive users' data content from the information collected during the auditing process.

Batch Auditing: It enable TPA with secure and efficient auditing capability to cope with multiple auditing delegations from possibly large number of different users simultaneously.

Light Weight: It allow TPA to perform auditing with minimum communication and computation overhead.

3. The System and Hazard Model

We consider a cloud data storage service connecting three different network entities, the cloud user (U), who has bulky 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 knowledge and capabilities that cloud users do not have and is trusted to assess the cloud storage service dependability on behalf of the user upon call. Users rely on the CS for cloud data storage and Protection. 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. Namely, in most of time it behaves correctly and does not move away from the prescribed protocol execution. However, for their own benefits the CS might ignore to keep or purposely delete rarely accessed data files which belong to normal cloud users. Moreover, the CS may decide to hide the data corruptions caused by server hacks or Byzantine failures to continue reputation. We assume the TPA, who is in the production of auditing, is reliable and independent, and thus has no incentive to collude with either the CS or the users during the auditing process. However, it harms the user if the TPA could learn the outsourced data after the audit. To authorize the CS to respond to the audit delegated to TPA's, the user can sign a certificate granting audit rights to the TPA's public key, and all audits from the TPA are authenticated against such a certificate. These authentication handshakes are omitted in the following presentation.

3.1 System Architecture

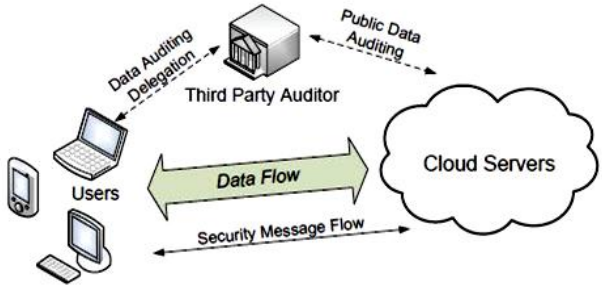

Fig. 2 System Architecture

4. Privacy-Preserving Public Auditing

The privacy-preserving public auditing, we propose to uniquely integrate the holomorphic non-linear authenticator with random masking technique. In our protocol, the non-linear blocks in the server's response is masked with randomness generated the server. With random masking, the TPA no longer has all the necessary information to build up a correct group of non-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 audit ability. Specifically, we use the HLA proposed in [13], which is based on the short signature scheme.

4.1 Periodic Sample Audit

In the Cloud Server environment random "sampling" checking greatly reduces the workload of audit services, while still achieve an effective detection of misbehaviour. Thus, the probabilistic audit on sampling checking is preferable to realize the abnormality detection in a timely manner, as well as rationally allocate resources. This algorithm relies on holomorphic properties to aggregate data and tags into a constant size response, which minimizes network communication. Since the single sampling checking may overlook a very small number of data abnormalities, we propose a periodic sampling approach to audit outsourcing data, which is called as Periodic Sampling Audit. In this way, the audit activities are efficiently scheduled in an audit period, and a TPA needs merely access small portions of file to perform audit in each activity. Therefore, this method can detect the exceptions in time, and reduce the sampling numbers in each audit.

4.2 Security Consistency for Batch Auditing

The way to describe the result to a multi-user setting will not affect the aforementioned security insurance, as shown in the Theorem.

Theorem: The batch auditing protocol achieves the same storage correctness and privacy preserving guarantee as in the single-user case.

Solution*:* The privacy-preserving guarantee in the multi- user setting. The storage correctness guarantee, we are going to reduce it to the single-user case. We use the forking technique for the verification equation for the batch audits involves K challenges from the random block. This time we need to ensure that all the other K − 1 challenges are determined before the forking of the concerned random oracle response. This can be done using the idea in [4]. As soon as the adversary issues the very first random oracle query for i = h(R||vi||L) for any i ε [1,K], the simulator immediately determines the values j = h(R||vj ||L) for all j ε [1,K]. This is possible since they are all using the same R and L. Now, all but one of the k's are equal, so a valid response can be extracted similar to the single-user case.

5. Conclusion

In this paper, we propose a privacy-preserving public auditing system for data storage security in cloud computing. We utilize the holomorphic linear authenticator and random masking to guarantee that the TPA would not learn any knowledge about the data content stored on the cloud server during the efficient auditing process, which not only eliminates the burden of cloud user from the tedious and possibly expensive auditing task, but also alleviates the users' fear of their outsourced data leakage. Considering TPA may concurrently handle multiple audit sessions from different users for their outsourced data files, we further extend our privacy-preserving public auditing protocol into a multiuser setting, where the TPA can perform multiple auditing tasks in a batch manner for better efficiency. Extensive analysis shows that our schemes are provably secure and highly efficient. Our preliminary experiment conducted on Amazon EC2 instance further demonstrates the fast performance of our design on both the cloud and the auditor side. We leave the full-fledged implementation of the mechanism on commercial public cloud as 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.

References

1. Amazon.com (2008). "Amazon s3 availability event: July 20,"Online at http://status.aws.amazon.com/s3-20080720.html.
2. Arrington, M.(2006). "Gmail disaster: Reports of mass email deletions," Online at http://www.techcrunch.com/2006/12/28/gmail- disaster reports-of-mass-email-deletions/, December.
3. Ateniese, G. R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, (2007). "Provable data possession at untrusted stores," in Proc. of CCS'07, Alexandria, VA, October, pp. 598–609.
4. Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009, http://www. cloudsecurityalliance.org.
5. Juels, A. and J. Burton S. Kaliski, (2007). "Pors: Proofs of retrievability for large files," in Proc. of CCS'07, Alexandria, VA, October, pp.584–597.
6. Kincaid, J.(2008). "Media Max/The Linkup Closes Its Doors," Online at http://www.techcrunch.com/2008/07/10/ mediamax the linkup-closes- its-doors/, July
7. Krebs, B. (2009). "Payment Processor Breach May Be Largest Ever," Online at http://voices.washingtonpost.com/securityfix/2009/01/payment processor breach may b.html, Jan.
8. Shah, M. A. R. Swaminathan, and M. Baker, (2008). "Privacy preserving audit and extraction of digital contents," Cryptology e-Print Archive, Report, 2008/186.
9. Wang, Q. C. Wang, J. Li, K. Ren, and W. Lou, (2009). "Enabling public verifiability and data dynamics for storage security in cloud computing," in Proc. of ESORICS'09, volume 5789 of LNCS. Springer-Verlag, Sep. pp. 355–370.
10. Wilson, S. (2008). "Appengine outage," Online at http://www.cio-weblog.com/50226711/appengine outage.php, June.