Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud

[1]Balakrishnan.S, [2]Saranya.G, [3]Shobana.S, [4]Karthikeyan.S
[1,2,3,4]Dept. of CSE, SNS Colloge of Technology, Comibatore, Tamilnadu, India

Abstract
Cloud computing is a internet based computing which enables sharing of services. Many users place their data in the cloud, so correctness of data and security is a prime concern. This work studies the problem of ensuring the integrity and security of data storage in Cloud Computing. Security in cloud is achieved by signing the data block before sending to the cloud. Signing is performed using BLS algorithm which is more secure compared to other algorithms. To ensure the correctness of data, we consider the task of allowing a third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the data stored in the cloud. By utilizing public key based homomorphic authenticator with random masking privacy preserving public auditing can be achieved. The technique of bilinear aggregate signature is used to achieve batch auditing. Batch auditing reduces the computation overhead. As the data in the cloud is used by many industries, modification of data cannot be avoided. Unlike most prior works, the new scheme further supports secure and efficient dynamic operations on data blocks, including: data update, delete and append. We explore the efficient technique for error correction called reed Solomon technique which ensures the correctness of data.

Key words
Third Party Auditor, Boneh–Lynn–Shacham, Bilinear aggregate signature, Reed Solomon technique.

I. Introduction

Cloud computing is a general term for anything that involves delivering hosted services over the internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). A cloud service has three distinct characteristics that differentiate it from traditional hosting. It is sold on demand, typically by the minute or the hour; it is elastic - a user can have as much or as little of a service as they want at any given time and the service is fully managed by the cloud service provider (the consumer needs nothing but a personal computer and Internet access).The advantage of cloud is cost savings. The prime disadvantage is security. Cloud computing is used by many software industries nowadays. Since the security is not provided in cloud, many companies adopt their unique security structure. For eg) Amazon has its own security structure. Introducing a new and uniform security structure for all types of cloud is the problem we are going to tackle in this paper. Since the data placed in the cloud is accessible to everyone, security is not guaranteed. To ensure security, cryptographic techniques cannot be directly adopted. Sometimes the cloud service provider may hide the data corruptions to maintain the reputation. To avoid this problem, we introduce an effective third party auditor to audit the user's outsourced data when needed. The security is achieved by signing the data blocks. Signing is performed using BLS algorithm. We utilized public key based homomorphic authenticator with random masking to achieve privacy preserving auditing protocol.TPA performs the auditing task for each user i.e. single auditing. This increases the auditing time and computation overhead. The technique of Bilinear Aggregate Signature is used to achieve batch auditing i.e.) multiple auditing tasks simultaneously. Earlier works performs auditing only for static data. We enhance the system with dynamic operations on data blocks i.e.) data update, append and delete. Reed Solomon technique is a best error correction technique which ensures the correctness of the data stored in cloud. Our contribution in this paper is summarized as follows:

- We provide a privacy preserving auditing protocol i.e. our scheme supports an external auditor to audit the users outsourced data without learning knowledge on the data content.
- Our scheme supports dynamic operations on data blocks i.e. data update, append and delete.
- We perform a technique called Reed Solomon technique which is a best error correction method.

A. Characteristics

Cloud computing is cost-effective. Here, cost is greatly reduced as initial expense and recurring expenses are much lower than traditional computing. Maintenance cost is reduced as a third party maintains everything from running the cloud to storing data. Cloud is characterized by features such as platform, location and device independency, which make it easily adoptable for all sizes of businesses, in particular small and mid-sized. However, owing to redundancy of computer system networks and storage system cloud may not be reliable for data, but it scores well as far as security is concerned. In cloud computing, security is tremendously improved because of a superior technology security system, which is now easily available and affordable. Yet another important characteristic of cloud is scalability, which is achieved through server virtualization. Some of the most important five key characteristics are,

1. On-demand Self Service

A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.

2. Broad Network Access

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms.

3. Resource Pooling

The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

4. Measured Service

Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service. Resource usage can be

monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service.

5. Selection of Provider

A good service provider is the key to good service. So, it is imperative to select the right service provider. One must make sure that the provider is reliable, well-reputed for their customer service and should have a proven track record in IT- related ventures.
As cloud computing has taken hold, there are six major benefits that have become clear,
1) Anywhere/anytime access - It promises "universal" access to high-powered computing and storage resources for anyone with a network access device.
2) Collaboration among users -cloud represents an environment in which users can develop software based services and from which they can deliver them.
3) Storage as a universal service - the cloud represents a remote but scalable storage resource for users anywhere and everywhere.
4) Cost benefits - the cloud promises to deliver computing power and services at a lower cost.

II. System Design

Cloud computing components are classified as 1) Cloud User (CU),2) Cloud Service provider (CSP) and 3) Cloud server (CS)

1. Third Party Auditor

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. 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. 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 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.

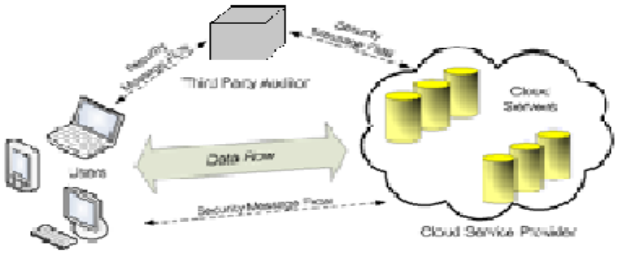


Fig. 1 : The architecture of cloud as a storage service

The Cloud Computing model of computing is a distributed application structure that partitions tasks or workloads between the providers of a resource or service, called Cloud servers, and service requesters, called clients. Often clients and servers communicate over a computer network on separate hardware, but both client and server may reside in the same system. A server machine is a host that is running one or more server programs which share their resources with clients. A client does not share any of its resources, but requests a server's content or service function. Clients therefore initiate communication sessions with servers which await incoming requests.

2. Existing System

To introduce an effective third party auditor (TPA) for privacy and security, the following fundamental requirements have to be met: TPA should be able to efficiently audit the cloud data storage without demanding the local copy of data, and introduce no additional on-line burden to the cloud user. The third party auditing process should bring in no new vulnerabilities towards user data privacy. They utilized and uniquely combined the public key based homomorphic authenticator with random masking to achieve the privacy-preserving public cloud data auditing system, which meets all above requirements.
This scheme is the first to support scalable and efficient public auditing in the Cloud Computing. In particular, this scheme achieves batch auditing where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA. The security and performance is justified through concrete experiments and comparisons with the state-of-the-art.
In cloud service providers, for monetary reasons, reclaiming storage by discarding data that has not been or is rarely accessed or even hiding data loss incidents so as to maintain a reputation. In short, although outsourcing data into the cloud is economically attractive for the cost and complexity of long-term large-scale data storage, it does not offer any guarantee on data integrity and availability. This problem, if not properly addressed, may impede the successful deployment of the cloud architecture. Another problem is that data stored in the cloud does not remain static.

3. Proposed System

We enhance the scheme with explicit and efficient dynamic data operations for data storage security in Cloud Computing. Therefore, it is crucial to consider the dynamic case, where a user may wish to perform various block-level operations of update, delete and append to modify the data file while maintaining the storage correctness assurance. The straightforward and trivial way to support these operations is for user to download all the data from the cloud servers and re-compute the whole parity blocks as well as verification tokens.

A. Update Operation

In cloud data storage, sometimes the user may need to modify some data block(s) stored in the cloud, from its current value fij to a new one, fij + Δ fij. We refer this operation as data update.

B. Delete Operation

Sometimes, after being stored in the cloud, certain data blocks may need to be deleted. The delete operation we are considering is a general one, in which user replaces the data block with zero or some special reserved data symbol. From this point of view, the delete operation is actually a special case of the data update operation, where the original data blocks can be replaced with zeros or some predetermined special blocks.

C. Append Operation

The user may want to increase the size of his stored data by adding blocks at the end of the data file, which we refer as data

append. We anticipate that the most frequent append operation in cloud data storage is bulk append, in which the user needs to upload a large number of blocks (not a single block) at one time. Dynamic operations are performed by constructing the matrix, where 0's indicate the blocks we need to change and 1's indicate the unchanged blocks [8].

We create a cloud environment where user, TPA and cloud server are connected each other. In public auditing system, the correctness of the data is checked by keygen, sagging, gen proof and verify proof algorithms. Homomorphism authenticator with random masking is used to achieve privacy preserving auditing scheme. The technique of bi-linear aggregate signature is used to achieve batch auditing. In cloud, the data does not remain static. We enhance the system with explicit dynamic operations in data blocks.

III.Construct Public Auditing System

The public auditing system can be constructed in two phases, Setup and Audit.

Setup: The user initializes the public and secret parameters of the system by executing KeyGen, and preprocesses the data file F by using SigGen to generate the verification metadata. The user then store the data file F at the cloud server, deletes its local copy, and publishes the verification metadata to TPA for later audit. As part of pre-processing, the user may alter the data file F by expanding it or including additional metadata to be stored at server.

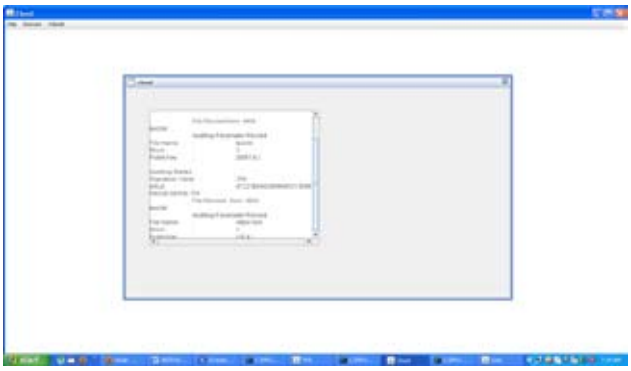


Fig. 2: The data are received by sign on it

Audit: The TPA issues an audit message or challenge to the cloud server to make sure that the cloud server has retained the data file F properly at the time of the audit. The cloud server will derive a response message from a function of the stored data file F by executing GenProof. Using the verification metadata, the TPA verifies the response via Verify Proof.

Keygen Process

We propose a batch signature scheme based on the BLS signature. The BLS signature scheme uses a cryptographic primitive called pairing, which can be defined as a map over two cyclic groups G1 and G2.

The BLS signature scheme consists of three phases,

1. In the key generation phase, a sender chooses a random integer $x \in \mathbb{Z}_p$ and computes $y = {g_1}^x \in G_1$. The private key is x and the public key is y.

2. Given a message $m \in \{0.1\}$ in the signing phase, the sender first computes $h = h(m) \in G_1$, where h () is a hash function, and then computes $\sigma = h^x \in G_1$. The signature of m is σ.

3. In the verification phase, the receiver first computes $h = h(m) \in G_1$ and then check whether $e(h,y) = e(\sigma,g_1)$.

If the verification succeeds, then the message m is authentic.

Merits:

- It can generate a very short signature.
- It can solve communication overhead.

IV. Privacy-Preserving Public Auditing Scheme

We propose to uniquely integrate the homomorphic 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 a pseudo random function (PRF). 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. Meanwhile, due to the algebraic property of the homomorphic authenticator, the correctness validation of the block-authenticator pairs will not be affected by the randomness generated from a PRF.



Fig. 3: User performs the dynamic operations

V. Multiple Batch Auditing

TPA may concurrently handle multiple auditing delegations upon different users' requests. The individual auditing of these tasks for TPA can be tedious and very inefficient. Given K auditing delegations on K distinct data files from K different users, it is more advantageous for TPA to batch these multiple tasks together and audit at one time [4]. Keeping this natural demand in mind, we propose the technique of bilinear aggregate signature, which supports the aggregation of multiple signatures by distinct signers on distinct messages into a single signature and thus provides efficient verification for the authenticity of all messages. Using this signature aggregation technique and bilinear property, we can now aggregate K verification equations into a single one, so that the simultaneous auditing of multiple tasks can be achieved.

VI. Multiple Batch auditing using

Block-level operations

We enhance the scheme with explicitly and efficiently handle dynamic data operations public auditing system of data storage security in Cloud Computing. Therefore, it is crucial to consider the dynamic case, where a user may wish to perform various block-level operations of update, delete and append to modify the data file while maintaining the storage correctness assurance. The straightforward and trivial way to support these operations is for user to download all the data from the cloud servers and re-compute the whole parity blocks as well as verification tokens.

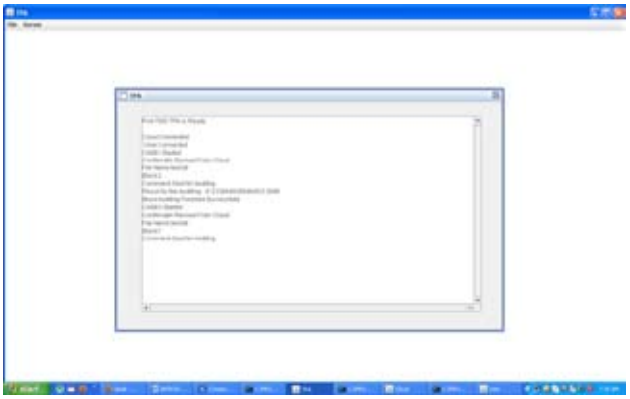


Fig. 4: TPA auditing sessions

A. Update Operation

In cloud data storage, sometimes the user may need to modify some data block(s) stored in the cloud, from its current value fij to a new one, fij + Δ fij. We refer this operation as data update.

B. Delete Operation

Sometimes, after being stored in the cloud, certain data blocks may need to be deleted. The delete operation we are considering is a general one, in which user replaces the data block with zero or some special reserved data symbol. From this point of view, the delete operation is actually a special case of the data update operation, where the original data blocks can be replaced with zeros or some predetermined special blocks.

C. Append Operation

In some cases, the user may want to increase the size of his stored data by adding blocks at the end of the data file, which we refer as data append. We anticipate that the most frequent append operation in cloud data storage is bulk append, in which the user needs to upload a large number of blocks at one time.

VII.Conclusion

The public auditability for cloud data storage security is of critical importance so that users can resort to an external audit party to check the integrity of outsourced data when needed. This work studies the problem of ensuring the integrity of data storage in Cloud Computing. In particular, we consider the task of allowing a third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the dynamic data stored in the cloud. We utilize and uniquely combine the public key based homomorphic authenticator with random masking to achieve the privacy-preserving public cloud data auditing system. This scheme is the first to support scalable and efficient public auditing in the Cloud Computing. The technique of Bilinear Aggregate signature is used to achieve batch auditing, where TPA can perform multiple auditing tasks simultaneously. The data in the cloud does not remain static. Unlike most prior works, the new scheme further supports secure and efficient dynamic operations on data blocks stored in the cloud, including: data update, delete and append.

VIII. Future Enhancements

We envision several possible directions for future research on this area. The most promising one we believe is a model in which public verifiability is enforced. Public verifiability, allows TPA to audit the cloud data storage without demanding users' time, feasibility or resources. An interesting question in this model is if we can construct a scheme to achieve both public verifiability and storage correctness assurance of dynamic data. Besides, along with our research on dynamic cloud data storage, we also plan to use Trapdoor Commitment Scheme.

References

[1] A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, M. Zaharia, "Above the clouds: A berkeley view of cloud computing," University of California, Berkeley, Tech. Rep, 2009.

[2] Cong Wang,Qian Wang,Kui Ren,Wenjing Lou (2009),"Ensuring Data Storage Security in Cloud Computing".

[3] Cong Wang,Qian Wang,Kui Ren,Wenjing Lou (2010),"Privacy Preserving Public Auditing for Data Storage Security in Cloud Computing".

[4] A. L. Ferrara, M. Greeny, S. Hohenberger, M. Pedersen (2009), "Practical short signature batch verification", in Proceedings of CT-RSA, volume 5473 of LNCS. Springer-Verlag, pp. 309–324.

[5] H. Shacham, B. Waters (Dec 2008), "Compact proofs of retrievability", in Proc. of ASIACRYPT 2008, vol. 5350, pp. 90–107

[6] M.A.Shah, R.Swaminathan, M. Baker (2008), "Privacy-preserving audit and extraction of digital contents", Cryptology ePrint Archive.

[7] Q. Wang, C. Wang, J. Li, K. Ren, W. Lou (2009), "Enabling public verifiability and data dynamics for storage security in cloud computing", in Proc. of ESORICS'09, Saint Malo.

[8] Wang, K. Ren, W. Lou (2010), "Achieving secure, scalable, and fine-grained access control in cloud computing", in Proc. of IEEE INFOCOM'10, San Diego, CA, USA.

[9] [Online] Available : Amazon.com, "Amazon s3 availability event: (2008)," Online at http://status.aws.amazon.com/ s3-20080720.html.

[10] Cloud Security Alliance, (2009) "Security guidance for critical areas of focus in cloud computing," 9, [Online] Available : http://www.cloudsecurityalliance.org.

[11] P. Mell, T. Grance, (2009) "Draft nist working definition of cloud computing," Referenced on June. 3rd, 2009 Online at http://csrc.nist.gov/ groups/SNS/cloud-computing/index.html.

[11] Craig Gentry, Dan Boneh, (2004) "Aggregate and verifiably encrypted signatures from bilinear maps".



KARTHIKEYAN.S is an Assistant Professor of Computer Science and Engineering at SNS College of Technology, Anna University. He received B.Tech degree in Information Technology from Anna University, India in 2006 and M.E degree in Computer Science and Engineering from Anna University, India in 2009.

BALAKRISHNAN.S, SARANAYA.G, SHOBANA.S are currently pursuing their final year computer science and engineering at SNS College of Technology, Coimbatore-35.