Privacy Preserving Integrity Verification for Securing Cloud Storage

Sonali D. Thosar
PG Student, Department of Computer Engineering,
Sinhgad Technical Education Society's, Smt. Kashibai Navale College of Engineering
Pune, Maharashtra, India

Nalini A. Mhetre
Asst. Prof. Department of Computer Engineering,
Sinhgad Technical Education Society's, Smt. Kashibai Navale College of Engineering
Pune, Maharashtra, India

Abstract

User stores data on cloud and ask for data whenever required. User uses resources provided by cloud server for storage and computation purpose. So, user is free from burden of storage and maintenance. User no longer possesses data locally. So, cloud storage brings great challenge in the field of security and privacy preservation. It is important that clouds should be secure .User should use cloud data without worrying its integrity. To check integrity of data without having its local copy is an important issue. Data is not only stored on cloud, but also shared across different users. Public auditing for this data by achieving privacy is an open challenge. In public auditing, user resort third party auditor to check integrity of data. This paper presents different vulnerabilities in cloud storage and schemes used to overcome these vulnerabilities.

1. Introduction

Cloud computing involves delivery of services over internet.

These services are divided in 3 terms – Platform as a Service(PaaS), Infrastructure as a Service(IaaS), Software as a Service(SaaS).

PaaS - It is used for developing websites on the system without installing any software. It can be executed without any administrative expertise.

IaaS - If the revenue for cloud services primarily comes from charging for infrastructure. This System can be referred to as Infrastructure as a Service (IaaS). IaaS is the hardware and software that powers it all servers, storage, networks, operating systems.

SaaS - It is run by cloud service provider and mostly used by organizations. It is available to users through internet.[1].

Cloud computing have following advantages to store data on cloud –

1] On demand self-service – Cloud is pool of resources. User should buy them as per demand.

2] Cost effective – There is centralized management of resources which reduces burden of client for processing.

3] Rapid elasticity – The scale of cloud is dynamic, which is in accordance with demand of the users.

4] High commonality – Cloud is not aimed at specific application. Different applications can utilize same cloud resources.

There are many benefits to store data on cloud. For example, users do not have to care about hardware maintenance. As users store their data on the cloud, it means that they will lose the control of them and worries will come out about the data security. Data security is an important issue in cloud computing and it is an important factor for quality of service. [2]

2. Issues in Cloud Computing

A. Security issues

Security issues are considered as most important issues in cloud computing. Following are some major security issues –

- Access Control – Unauthorized access may exist if security mechanism is not adequate. As user's data reside on cloud for long time so risk of illegal access is also more.
- Authentication and Identification – Cloud serves many clients that is cloud allows one

single instance of software to serve many clients. So, there is problem of authentication and identification.

- Availability – User stores data on cloud and deletes its local copy. If service or data on cloud is not available due to some problem then it is hard to retrieve data.
- Policy Control – The Cloud is heterogeneous, which means that different Cloud servers may have different mechanisms to ensure the client's data security. Thus policy integration is one of the concerns.
- Audit – Cloud service provider has control of data. There is possibility that cloud service provider can alter data without client's permission. So, there is need of audit. [3]

B. Privacy issues

Privacy issues include protection of identity information, transaction histories and sensitive data. Idea of cloud computing is store user data on shared infrastructure. So, there is risk of unauthorized access. Following are some privacy issues –

- Unauthorized secondary usage – Users data can be utilized by cloud service provider. Unauthorized usage of data may cause serious security problems, which becomes one significant concern.
- Lack of user control – As user no longer possesses its data. Data is not transparent to users which means user have lost control over data. So there is need of protection mechanism.
- Unclear responsibility – There is one problem related to the privacy that it is sometimes unclear about which CSP is responsible for privacy protection, detecting who use and modify the user data or ensuring user data privacy requirements. [3]

Solution for ensuring data safety is utilizing a trustful third party auditor to assess and expose risk of cloud storage services on behalf of the users upon request.

3. System Model

Cloud data storage consists of 3 entities –
1] User – who has large amount of data which is to be stored on cloud.
2] Cloud Server – which has computing resources to manage cloud data.
3] Third party auditor – to challenge cloud server to check correctness of data on behalf of user.

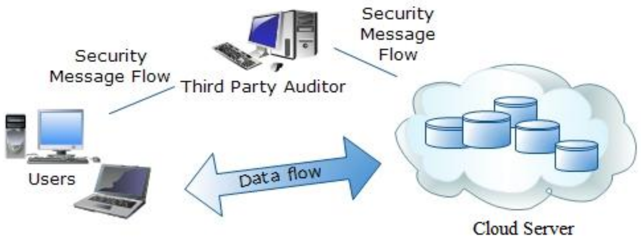

Fig – 1 : Architecture of cloud storage[8]

Public auditability is used to ensure data integrity. Public auditability allows an external party, in addition to the user to verify the correctness of remotely stored data. As third party auditor is used for verifying integrity of data, time is saved.[4]

A. Privacy Preserving Scheme

When third party auditor is used for verification purpose privacy should be preserved. In privacy preserving scheme, third party auditor does not have any knowledge about data.
Privacy preserving public auditing scheme has following advantages –
1] Public Auditability – Public auditability allows TPA to check integrity of data without retrieving it.
2] Storage Correctness – User's data should be stored correctly on cloud.
3] Privacy Preserving – This ensures that TPA cannot derive any data content.
4] Lightweight – To allow TPA to perform auditing with minimum overhead.
5] Batch auditing – To allow TPA to challenge server for checking integrity of data for multiple clients at the same time. [5]

4. Related Work

In past, different techniques were used to provide security to cloud data but there are some disadvantages of these system. Common methods for protecting user data include encryption prior to storage, user authentication procedures prior to storage or retrieval, and building secure channels for data transmission. These protection methods normally require cryptography algorithms and digital signature techniques.

Ateniese et al is the first to consider public auditability in provable data possession model for ensuring possession of files on untrusted storages. In this model RSA based homomorphic tags are used. In this public auditability is achieved. But this model does not support dynamic data operation and also suffer security problems.[6]

Wang considered dynamic data storage in a distributed scenario, and the proposed challenge response protocol can both determine the data correctness and locate possible errors but this model only considered partial support for dynamic data operation.[7]

Juels and Kaliski describe a proof of retrievability model. Disadvantage of this model is it does not support public auditability. Shacham and Waters design an improved PoR scheme with full proofs of security in the security model. They use publicly verifiable homomorphic authenticators built from BLS signatures based on which the proofs can be aggregated intoa small authenticator value, and public retrievability is achieved. Still, the authors only consider static data files.[8]

MAC based scheme – MAC based scheme has following disadvantages

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.[8]

2] The TPA also has to maintain and update state between audits i.e. keep track on the revealed MAC keys.

3] It can only support static data and cannot efficiently deal with dynamic data at all.

HLA based scheme – There is need of system which can verify integrity of data without retrieving data block. So HLA scheme was used for this purpose. HLA is similar to MAC only difference is that HLA can be aggregated. Disadvantage of this system is that data can be retrieved if linear combinations of same block is used.[8]

5. Proposed System

Disadvantage of HLA is linear combination of block can reveal user data information. To achieve privacy preserving public auditing HLA with random masking is used. TPA cannot retrieve information if random masking is used. Public key based HLA is used to achieve public auditability.

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.[8]

6. Conclusion

This scheme uses third party auditor which achieves public auditabilty, stateless verification and also supports data dynamics. Third party auditor verifies integrity of user data as well as privacy is preserved.

As block of data is not retrieved for verification purpose, there is less communication overhead.

7. References

1] V. Nirmala, R.K.Sivanandhan, Shanmuga Lakshmi,"Data Confidentiality and interity verification using user authenticator scheme in cloud", Proc of 2013 International Conference on Green High Performance Computing

2] Ayad F. Barsoum **,** M. Anwar Hasan **"**Integrity Verification of Multiple Data Copies over Untrusted Cloud Servers", 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing

3] Ziyuan Wang, "Security and privacy issues within the Cloud computing", 2011 International Conference on Computational and Information Sciences

4] Qian Wang, Cong Wang, Kui Ren, wenjing Lou,Jin Li **"**Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing**",** IEEE transactions on parallel and distributed systems, 2011

5] Chang Liu, Jinjun Chen, Laurence T. Yang, Xuyun Zhang, Chi Yang, Rajiv Ranjan,Kotagiri Ramamohanarao, **"** Authorized Public Auditing of Dynamic Big Data Storage on Cloud with Efficient Verifiable Fine-grained Updates**"**

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] A. Juels and B.S. Kaliski Jr., "Pors: Proofs of retrievability for Large Files," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 584-597, 2007.

8] "Privacy preserving public auditing for Secure Cloud Storage", Cong Wang, Sherman S.-M. Chow, Qian Wang, Kui Ren, and Wenjing Lou.