

A Survey on Panda: Public Auditing for Shared Data with Efficient User Revocation in the Cloud

Gauri Thorat

PG Student ME (Computer)

JSPM's IMPERIAL COLLEGE OF ENGINEERING
AND RESEARCH WAGHOLI, PUNE – 412 207

Prof.S.R.Todmal

(Assistant Professor)

Department of Information Technology
JSPM's IMPERIAL COLLEGE OF ENGINEERING
AND RESEARCH WAGHOLI, PUNE – 412 207

Abstract— In today's Computing world Cloud computing is one of the biggest innovation which uses advanced computational power and it improves data sharing and data storing capabilities. Main difficulty in cloud computing was issues of data integrity, data privacy and data access by unauthorised users. TTA (Trusted Third Party) is used to store and share data in cloud computing. Modification and sharing of data is quite simple as a group. To verify integrity of the shared data, members in the group needs to compute signatures on all shared data blocks. Different blocks in shared data are generally signed by different users due to data modifications performed by different users. User revocation is one of the biggest security threats in data sharing in groups. During user revocation shared data block signed by revoked user needs to download and re-sign by existing user. This task is very inefficacious due to the large size of shared data blocks on cloud. PANDA Plus is the new public auditing mechanism for the maintaining integrity of shared data with efficient user revocation in the cloud. This mechanism is based on proxy re-signatures concept which allows the cloud to re-sign blocks on behalf of existing users during user revocation, so that downloading of shared data blocks is not required. PANDA Plus is the public auditor which audits the integrity of shared data without retrieving the entire data from the cloud. It also monitor batch to verify multiple auditing tasks simultaneously.

Keywords— Cloud computing, Data integrity, Public auditing, User revocation

I. INTRODUCTION

[A] Cloud Computing

Cloud computing is nothing but internet based computing which made revolution in today's world. It is the biggest innovation which uses advanced computational power and improves data sharing and data storing capabilities. Cloud is a large group of interconnected computers, which is a major change in how we store information and run application. Cloud computing is a shared pool of configurable computing resources, on-demand network access and provisioned by the service provider [1].The advantage of cloud is cost savings. The prime disadvantage is security. The cloud computing security contains to a set of policies, technology & controls deployed to protect data, application & the associated infrastructure of cloud computing.

Some security and privacy issues that need to be considered. The only thing was the cloud computing lacks

regarding the issues of data integrity, data privacy, and data accessed by unauthorised members.

[B] Data integrity

Integrity is nothing but consistency. It is a major factor that affects on the performance of the cloud. Data integrity contains protocols for writing of the data in a reliable manner to the persistent data storages which can be retrieved in the same format without any changes later. Maintaining integrity of shared data is quite difficult task. Numbers of mechanisms have been proposed [2]-[15] to protect integrity of data. Concept of attaching Signature to each block of data is used in these mechanisms. Data Integrity is most important of all the security issues in cloud data storages as it ensures completeness of data as well as that the data is correct, accessible, consistent and of high quality.

Data model consist of three types of integrity constraints:

- Entity integrity
- Referential integrity
- Domain integrity

[C] Public Data Auditing in Cloud

On cloud we can able to store data as a group and share it or modify it within a group. In cloud data storage contains two entities as cloud user (group members) and cloud service provider/ cloud server. Cloud user is a person who stores large amount of data on cloud server which is managed by the cloud service provider. User can upload their data on cloud and share it within a group. A cloud service provider will provide services to cloud user. The major issue in cloud data storage is to obtain correctness and integrity of data stored on the cloud. Cloud Service Provider (CSP) has to provide some form of mechanism through which user will get the confirmation that cloud data is secure or is stored as it is. No data loss or modification is done by unauthenticated member. To achieve security data auditing concept is come into picture. This can be achieved in 2 ways as

- without trusted third party
- With trusted third party based on who does the verification.

In cloud computing architecture data is stored centrally and managing this centralised data and providing

security to it is very difficult task. TPA is used in this situation. The reliability is increased as data is handled by TPA but data integrity is not achieved. TPA uses encryption to encrypt the contents of the file. It checks data integrity but there is threat of TPA itself leaks user's data.

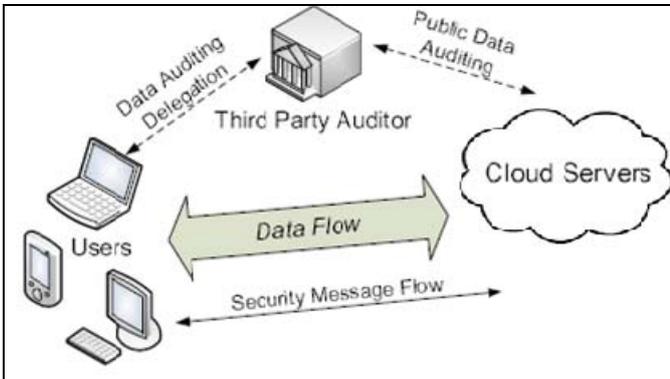


Fig.1 Architecture of Cloud Data Storage Service

Researchers of [3] specify way to achieve storage correctness without Trusted Third Party (TTP). They achieve this by using secure key management, Flexible access right managements and light weight integrity verification process for checking the unauthorised change in the original data without requesting a local copy of the data.

II. LITERATURE REVIEW

[A] Techniques used in Public Auditing on Cloud

There are some different techniques which used in different auditing mechanisms. This section introduce some the techniques like MAC, HLA etc. which are used for different purposes like data authentication, data integrity in auditing schemes on cloud.

1. MAC Based Solution

This technique used for data authentication. In this mechanism user upload data blocks with MAC and Cloud provider provides Secret key SK to TPA. Here TPA's task is to retrieve data blocks randomly and MAC uses SK to check correctness of data. Limitations of this technique are:

- Online burden to users due to limited use (i.e. Bounded usage) and stateful verification.
- Complexity in communication and computation
- Maintaining and updating TPA states is difficult.
- User need to download all the data to recompute MAC and republish it on CS
- This technique only supports for static data.

2. HLA Based Solution

This technique performs auditing without retrieving data block. HLA is nothing but unforgettable verification meta data that authenticate. It checks integrity of data block by authenticating it in linear combination of the individual blocks. This technique allows efficient data auditing and consuming only constant bandwidth, but its time consuming as it uses linear combination for authentication.

3. Using Virtual Machine

Abhishek Mohta proposed Virtual machines concept which use in case of Software as a Service (SaaS) model of the cloud computing. In this mechanism as shown in Fig when user request CSP for service CSP authenticate the client and provide a virtual machine by means of Software as a service. Virtual Machine (VM) uses RSA algorithm for cryptography, where client encrypt and de-encrypt the file. A SHA-512 algorithm is also used for making the message digest and check the integrity of data. This also helps in avoiding unauthorised access and providing privacy and consistency. Limitation to this technique is it is useful only for SaaS model.

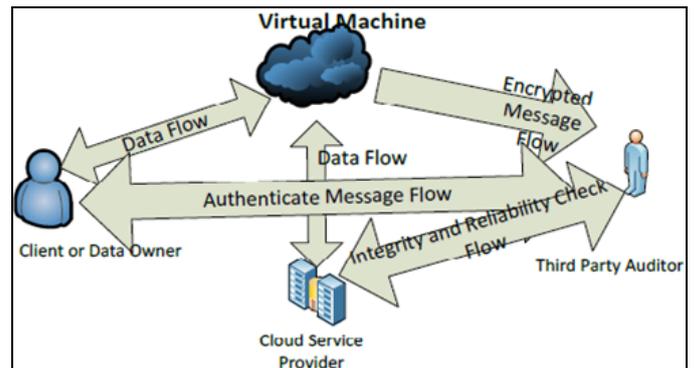


Fig. 2 Architecture of Cloud Data Storage Service using Virtual Machine

4. Using EAP

As mentioned by S. Marium Extensible authentication protocol (EAP) can also use through three ways hand shake with RSA. Using EAP they proposed identity based signature for hierarchical architecture. They provide an authentication protocol for cloud computing (APCC) [4]. As compare to SSL authentication protocol APCC is more lightweight and efficient. It also used Challenge – handshake authentication protocol (CHAP) for authentication.

The steps are as follows

- 1) When Client request for any service to cloud service provider, SPA send a CHAP request / challenge to the client.
- 2) The Client sends CHAP response/ challenges which is calculated by using a hash function to SPA
- 3) SPA checks the challenge value with its own calculated value. If they are matched then SPA sends CHAP success message to the client.

5. Using Automatic Protocol Blocker

Balkrishna proposed efficient Automatic Protocol Blocker technique for error correction which checks data storage correctness [4].Kiran Kumar proposed automatic protocol blocker to avoid unauthorized access [5]. When an unauthorized user access user data, a small application runs which monitors user inputs, It matches the user input, if it is matched then it allow user to access the data otherwise it will block protocol automatically. It contains five algorithms as keygen, SinGen, GenProof, VerifyProof,

Protocol Verifier. Protocol Verifier is used by CS. It contains three phases as Setup, Audit and Pblock.

6. Random Masking Technique

Jachak K. B. proposed privacy preserving Third party auditing without data encryption. It uses a linear combination of sampled block in the server's response is masked with randomly generated by a pseudo random function (PRF) [7].

[B] Different Public auditing mechanisms on Cloud

This section consist different mechanisms, different system proposed by authors which are used for auditing in cloud computing.

1. Compact Proofs of Retrievability

Hovav Shacham and Brent Watersy[9] proposed proof-of-retrievability system. In this system, data storage center must prove to a verier that he is actually storing all of a client's data. They have proposed two homomorphic authenticators the first, based on PRFs, gives a proof-of-retrievability scheme secure in the standard model. The second, based on BLS signatures [8], gives a proof-of-retrievability scheme with public variability secure in the random oracle model. Frameworks explained by them allow to argue about the systems unforgeability, extractability, and retrievability with these three parts based respectively on cryptographic, combinatorial, and coding-theoretical techniques.

2 Provable Data Possession at Untrusted Stores

Giuseppe Ateniese et all introduce a model which based on provable data possession (PDP)[10]. This is used for verifying that server is processing the original data without retrieving it. In this model probablistic proof of possession is generated by sampling random sets of blocks from the server.This helps to reduces I/O cost.

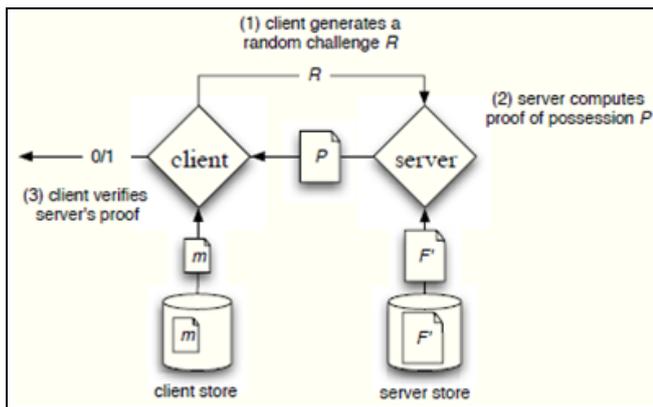


Fig.3 Provable Data Possession at Untrusted Stores

As shown in Fig.3 client maintains a constant amount of metadata to verify the proof. The challenge/response protocol transmits a small, constant amount of data, which minimizes network communication. DP model for remote data checking supports large data sets in widely-distributed storage systems. A key component of this mechanism is the homomorphic verifiable tags.

3. Privacy Preserving Public Auditing

Cong Wang Proposed Privacy Preserving Public Auditing technique [11].In this technique public auditing allows TPA along with user to check the integrity of the outsourced data stored on a cloud & Privacy Preserving allows TPA to do auditing without requesting data. Here TPA can audit the data by maintaining cloud data privacy. They have used the homomorphic 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 prevent the users from fear of the outsourced data leakage.

This mechanism is based on 4 algorithms:

- Keygen: It is a key generation algorithm for setup the scheme.
- Singen: It is used by the user to generate verification metadata which may consist of digital signature.
- GenProof: It is used by CS to generate a proof of data storage correctness.
- Verifyproof: Used by TPA to audit the proofs

4. LT Codes-based Secure and Reliable Cloud Storage Service

Ning Cao et all explore the problem of secure and reliable cloud storage with the efficiency consideration of both data repair and data retrieval, and design a LT codes-based cloud storage service (LTCS)[12]. LTCS provides efficient data retrieval for data users by utilizing the fast Belief Propagation decoding algorithm, and releases the data owner from the burden of being online by enabling public data integrity check and employing exact repair. LTCS is much faster data retrieval than the erasure codes-based solutions. It introduces less storage cost, much faster data retrieval, and comparable communication cost comparing to network coding-based storage services.

5. Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud

Boyang Wang et all proposed Oruta, the first privacy preserving public auditing mechanism for shared data in the cloud in [13].They have used ring signatures to construct homomorphic authenticators, so the TPA is able to audit the integrity of shared data, without retrieving the entire data. They have used HARS and its properties for constructing Oruta.

III. CONCLUSIONS`

Cloud computing is world's biggest innovation which uses advanced computational power and improves data sharing and data storing capabilities. It increases the ease of usage by giving access through any kind of internet connection. As every coin has two sides it also has some drawbacks.Privacy security is a main issue for cloud storage. To ensure that the risks of privacy have been mitigated a variety of techniques that may be used in order to achieve privacy. This paper showcase some privacy

techniques and different methods for overcoming the issues in privacy on untrusted data stores in cloud computing. There are still some approaches which are not covered in this paper. This paper categorizes the methodologies in the literature as encryption based methods, access control based mechanisms, query integrity/keyword search schemes, and auditability schemes. Even though there are many techniques in the literature for considering the concerns in privacy, no approach is highly developed to give a privacy-preserving storage that overcomes all the other privacy concerns. Thus to handle all these privacy concerns, we need to develop privacy-preserving framework which handle all the worries in privacy security and strengthen cloud storage services.

REFERENCES

- [1] P. Mell and T. Grance, "Draft NIST working definition of cloud computing".
- [2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Towards Secure and Dependable Storage Services in Cloud Computing," *IEEE Transactions on Services Computing*, vol. 5, no. 2, pp. 220–232, 2011.
- [3] Y. Zhu, G.-J. Ahn, H. Hu, S. S. Yau, H. G. An, and S. Chen, "Dynamic Audit Services for Outsourced Storage in Clouds," *IEEE Transactions on Services Computing*, accepted.
- [4] S. Mariam, Q. Nazir, A. Ahmed, S. Ahasham and Aamir M. Mirza, "Implementation of EAP with RSA for Enhancing The Security of Cloud Computing", *International Journal of Basic and Applied Science*, vol 1, no. 3, pp. 177-183, 2012
- [5] Balkrishnan. S, Saranya. G, Shobana. S and Karthikeyan.S, "Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud", *International Journal of computer science and Technology*, vol. 2, no. 2, ISSN 2229-4333 (Print) | ISSN: 0976-8491(Online), June 2012
- [6] 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
- [7] 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
- [8] J. Yuan and S. Yu, "Proofs of Retrievability with Public Verifiability and Constant Communication Cost in Cloud," in *Proceedings of ACM ASIACCS-SCC'13*, 2013
- [9] H. Shacham and B. Waters, "Compact Proofs of Retrievability," in the *Proceedings of ASIACRYPT 2008*. Springer-Verlag, 2008, pp.90–107.
- [10] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," in the *Proceedings of ACM CCS 2007*, 2007, pp. 598–610.
- [11] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," in the *Proceedings of IEEE INFOCOM 2010*, 2010, pp. 525–533.
- [12] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. T. Hou, "LT Codes-based Secure and Reliable Cloud Storage Service," in the *Proceedings of IEEE INFOCOM 2012*, 2012, pp. 693–701.
- [13] H. Wang, "Proxy Provable Data Possession in Public Clouds," *IEEE Transactions on Services Computing*, accepted.
- [14] B. Wang, B. Li, and H. Li, "Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud," in the *Proceedings of IEEE Cloud 2012*, 2012, pp. 295–302.
- [15] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamic for Storage Security in Cloud Computing," in the *Proceedings of ESORICS 2009*. Springer-Verlag, 2009, pp.355–370.
- [16] B. Wang, B. Li, and H. Li, "PANDA: Public Auditing for Shared Data with Efficient User Revocation in the Cloud," in the *Proceedings of IEEE INFOCOM 2014*, 2014, pp.sss