Public Auditing for Secure Data Storage in Cloud Computing

Shrutika Nikhar, Author[2], Author[3], Author[4]

Abstract

Cloud storage is growing in use and popularity with its appealing advantages and economic benefits. It offers users on-demand network access to a large shared pool of computing resources in a pay-as-you-go manner. This brings a relief of the burden when it comes to storage management, location-independent data access, and asset expense on maintenance. On the other hand, the use of cloud storage means ultimate faith on the CSP (Cloud Service Provider), since the user has less or no physical control over their cloud data. Data integrity is an important factor when it comes to data security and privacy, thus periodic integrity check of the cloud data stored in cloud for users is extremely important and necessary which depends on public auditing services. This paper describes a systematic review of problems and solutions presented for secure data storage in cloud computing or in simple words various types of privacy preserving methods developed through the years that involve Third Party Auditors (TPA) as a solution for incompetent user control due to deficient computing resources and expertise.

Keywords- cloud computing, data integrity, cloud storage, data privacy preserving, cloud data auditing, security.

1. Introduction

Cloud computing enables convenient, on-demand network access to the shared pool resources that can be configured according to the requirements and can be supplied with minimal management efforts or service provider interaction. The resources include network, servers, storage, application and services [1]. In addition to on-demand self-service and ubiquitous network access, some features provided by cloud computing include location independent resource pooling, rapid resource elasticity, usage-based pricing and transference of risk [2]. The widespread adoption of cloud computing brings several new challenging security threats towards the users' outsourced data. One of the major concerns is integrity of data stored in cloud. The outsourcing of data leave users with no physical control over their data irrespective of the sensitivity of the data outsourced. The third-party auditors are suggested as a solution and are introduced for the process of data integrity checking have expertise and capabilities that users do not possess, and thus can check the integrity of data stored in cloud on the behalf of the users to ensure their storage correctness in the cloud. In a word, TPS is introduced to audit the integrity of cloud data which is termed as public cloud storage auditing [2]. Users usually have deficient computing resources and expertise in order to complete the task of verifying the integrity of the stored data. Thus, by bringing in the third-party auditor (TPA) to check the integrity of outsourced

data, users will be able to resort to its expertise and capabilities that make data integrity checking process more competent flexible for both client and CSPs [16], [17]. The third-party auditing process if enabled, should not bring any profound vulnerabilities and additional burden to the user. The additional problems that have been a major concern are tried to be addressed including both internal and external threats for data integrity, occasional unfaithful behaviour of CSP, unreliable or semi-independent TPA, Sybil identities creation by malicious CPA [16]. In order to remedy the difficulty of data integrity checking, various schemes are proposed for different systems considering different security models [3], [6], [12], [13], [15], [18], [19], [20], [21], [22], [23], [24], [25], [26].

2. Previous Work in the Domain

Wang et el. [3] the dynamic auditing scheme is extended to support batch auditing for multiple owners. In their scheme, public key based homomorphic linear authenticator (HLA) and random masking technique is used to ensure data privacy against TPA. The scheme also improves efficiency by allowing TPA to perform multiple auditing tasks in batch manner.

Ateniese et al. [6] are the first to consider public auditability in their “provable data possession” (PDP) model for ensuring possession of data files on untrusted storages. They utilize the RSA-based homomorphic linear authenticators for auditing outsourced data and suggest randomly sampling a few blocks of the file.

Shah et al. [12], [13] propose introducing a TPA to keep online storage honest by first encrypting the data then sending a number of precomputed symmetric-keyed hashes over the encrypted data to the auditor. The auditor verifies the integrity of the data file and the server’s possession of a previously committed decryption key. This scheme only works for encrypted files, requires the auditor to maintain state, and suffers from bounded usage, which potentially brings in online burden to users when the keyed hashes are used up.

Rui et al. [18] propose a lightweight certificateless RDIC scheme to ensure integrity of data by authorizing TPA. This scheme solves the shortcoming of certificate management of PKI cryptography and avoids the key escrow drawback. Previous RDIC schemes relied on PKI which were inefficient due to communication and storage overhead caused by certificate management. In their model, the data privacy is taken into account from the side of TPA as well, which means the TPA cannot analyze the data information from the proof received from CSP.

Han et al. [19] address the problems of block corruption checking and inefficient authenticated data structure for achieving accurate auditing in the case of frequent data updating. In their model, they proposed an authorized dynamic public auditing scheme by designing a dynamic index table (DIT) which is updated without the need of moving, insertion or deletion of elements. TPA use

DIT to view block properties and then achieve data auditing. The scheme can verify and check the blocks that are lost or corrupt during data integrity fails.

Shivarajkumar and Sanjeev [20] propose an auditing scheme which uses AES algorithm for encryption and Secure Hash Algorithm (SHA-2) to generate verification metadata or message digest for data integrity check. In their model, TPA performs the task of auditing by verifying the message digest and takes constant time for file auditing for files of different sizes. The process of auditing involves encryption of user data before uploading in cloud then storing it in encrypted form. Thus, TPA performs auditing without retrieving the data copy of cloud user.

Swapnali et al. [21] propose a privacy preserving public auditing scheme, in which the data owner splits the file into blocks and encrypts using AES algorithm followed by generation of SHA-2 hash value for each. RSA signature generated by TPA and the one stored in TPA is matched. The matching of signature indicates that data is intact, tempered otherwise.

Yong et al. [22] propose an attribute-based cryptography scheme to simplify the complex key management issue in data integrity auditing protocols. In their model, for private keys, users can choose arbitrary attributes to upload files in cloud and owners can also specify the set of auditors assigned to check data integrity of outsourced data.

Wenting et al. [23] propose a framework of public cloud storage auditing using Authenticators Generation Center (AGC) which generates lightweight authenticators for users that consumes little computation. The AGC is unaware of users' data content due to cloud data blinding by the owners. The correctness of authenticators generated can be verified by the cloud hence reducing the users' computational burden.

Suganthi et al. [24] propose a privacy-preserving and public auditing mechanism for shared data in multi-cloud. In their model, aggregate signature scheme is used to construct homomorphic authenticators to perform auditing of shared data by TPA within a group, in which TPA is unaware of the block signer hence the privacy of the owner is fulfilled. The system proposed here can identify malicious users.

Wang et al. [15] proposed a dynamic auditing scheme based on Merkle Hash Tree to support dynamic operations of data stored in the cloud. They combine BLS-based HLA with MHT to support fully data dynamics. This scheme needs high computation and communication overhead during verification process.

S Ezhil et al. [25] propose a technique to verify integrity of data shred between two parties by using keyed Hash Message Authentication Code (HMAC) with homomorphic tokens to improve security of TPA. The parties agree on a shared secret key to transmit data. If this shared key is compromised the system will fail.

C C Erway et al. [26] extended the PDP model suggested by [6] in his DPDP scheme and designed the first public cloud storage auditing scheme by introducing a skip list structure

3. Data Auditing in Cloud

3.1. Components

Cloud Users/ Clients

Cloud user is an authorized person or party that have access rights to the resources from the cloud, in other word access to the cloud service. Cloud users/ Clients can be termed as end users/customers of cloud services. A third-party company called CSP (Cloud Service Provider) offers users a cloud-based platform, infrastructure, application and storage services.
Cloud Client can be hardware device or software that is used to access a cloud service.
Some examples of cloud clients are computer systems, tablets, home automation devices, operating systems, navigation devices and so on.

Cloud Storage Servers

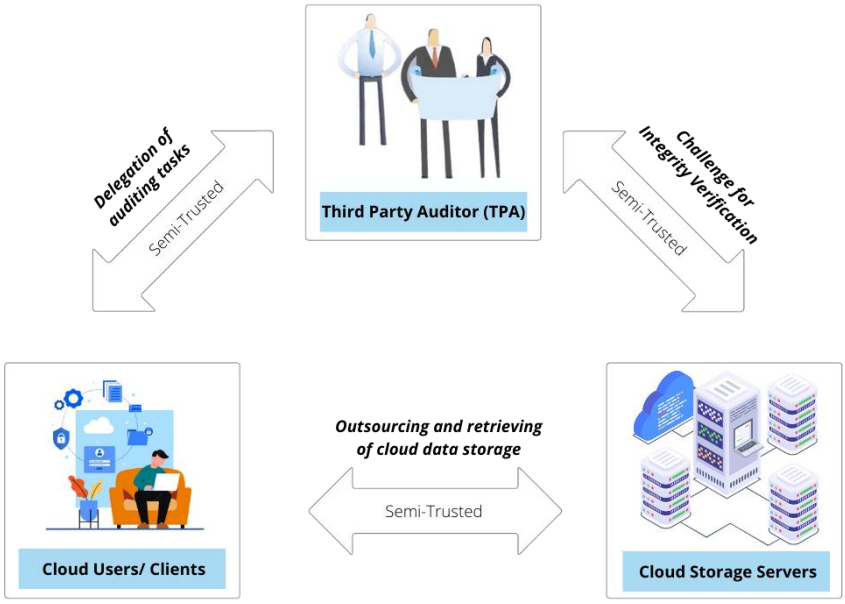
Cloud Storage Server, CSS are virtual servers that emulate physical servers which can host multiple virtual servers to provide cloud-based storage solutions to multiple users/customers. It is a pooled, centralized server resource that is hosted and delivered over a network (internet) and accessed on demand by multiple users.

Cloud Storage Providers

Cloud Storage Providers also known as Managed Service Provider (MSP) are responsible for availability and accessibility of stored data to the users, as well as making the physical environment protected and running. The users can lease cloud storage capacity per month or on demand.

Third Party Auditor (TPA)

These are external independent Auditors that perform the task of data integrity check. TPA is involved in hash values generation for encrypted blocks received from cloud server, then concatenating them and signature generation. TPA then compares both the signatures to check whether the data is intact or is being tampered. Users with constrained computing resources can resort to TPA to check the integrity of outsourced data and be worry free.



3.2 Comparative Analysis

Scheme	Method used	Authorized Auditing	Maintains Data confidentiality	Maintains data Integrity	Supports Dynamic Auditing
[3] Privacy-preserving public auditing for data storage security in cloud computing	HLA with BLS Signature	Yes	No	Yes	Yes
[6] Provable Data Possession at Untrusted Stores	PDP	Yes	No	Yes	Yes
[12], [13] Auditing to Keep Online Storage Services Honest, Privacy-Preserving Audit and Extraction of Digital Contents	HMAC	Yes	Yes	Yes	No

[15] Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing	Merkle Hash Tree (MHT)	No	No	Yes	Yes
[18] Certificateless Public Auditing Scheme with Data Privacy Preserving for Cloud Storage	Certificateless RDIC	Yes	Yes	Yes	Yes
[19] An Authorized Public Auditing Scheme for Dynamic Big Data Storage in Cloud Computing	Dynamic Index Tables (DIT)	Yes	No	Yes	Yes
[20] A Novel Data Auditing Approach to Achieve Data Privacy and Data Integrity in Cloud Computing	AES and SHA-2	Yes	Yes	Yes	No
[21] Third Party Public Auditing Scheme for Cloud Storage	AES and SHA-2	No	Yes	Yes	No
[22] Attribute-Based Cloud Data Integrity Auditing for Secure Outsourced Storage	Attribute Based Signature	Yes	Yes	Yes	No
[23] A Public Cloud Storage Auditing Scheme with Lightweight Authenticator Generation	Authenticators Generation Centre (AGC)	Yes	Yes	Yes	No
[24] Privacy preservation and public auditing for cloud data using ass in multi-cloud	Aggregate Signature Scheme	Yes	Yes	Yes	Yes

[25] Privacy-Preserving Public Auditing in cloud using HMAC Algorithm	Hash Message Authentication code (HMAC)	Yes	Yes	Yes	No
[26] Dynamic provable data possession	DPDP	Yes	Yes	Yes	Yes

4. Contribution

This paper is an in-depth survey of al the current fashions and old developments related to Public Auditing in Cloud Computing with Secure Data Storage in focus. This paper is organised in chronological manner from very first scheme on public auditing in cloud and is written keeping in mind all the referenced papers experiments, results, outcomes and problems.
This includes proposed schemes and their working of referenced research and survey papers. The models suggested should be secure and efficient, most importantly must provide data confidentiality to the customers if required. The third-party auditing process if enabled, should not bring any profound vulnerabilities and additional burden to the user. The additional problems that have been a major concern are tried to be addressed including both internal and external threats for data integrity, occasional unfaithful behaviour of CSP, unreliable or semi-independent TPA, Sybil identities creation by malicious CPA.

5. Conclusion

Cloud computing being a fast-developing topic, data integrity verification is an important area which is attracting more and more research interest and there still lots of ongoing research in this field. As we can see from the above that existing research has achieved great goals, there is still a room for evolution in integrity verification mechanisms which will continue to evolve alongside development of cloud.

5.1 Future Work / Potential Research

For future developments we can look into the following components:

Security: Even though currently suggested schemes seem potent and rigorous, with new exploits coming into the picture every day, finding security holes and fixing them can be a long-lasting solution.
Efficiency: While designing new techniques, the processes of integrity verification/data auditing like storage, computation and communication cost can be made more efficient to bring cost efficiency in this pay-as-you-go model.
Scalability/elasticity: Scalability of cloud is another factor that is considered in programming models for parallel and distributed systems. On the other hand, elasticity is one of the biggest reasons why big companies are moving their business to the cloud. Scalability and elasticity can turn out to be highly resourceful for big data applications in cloud environment.

6. References

[1] P. Mell and T. Grance, “Draft NIST Working Definition of Cloud Computing,” http://csrc.nist.gov/groups/SNS/cloudcomputing/index.html, June 2009.

[2] Cong Wang, Member, IEEE, Sherman S.M. Chow, Qian Wang, Member, IEEE, Kui Ren, Senior Member, IEEE, and Wenjing Lou, Senior Member, IEEE "Privacy-Preserving Public Auditing for Secure Cloud Storage", february 2013.

[3] Wang, C., Wang, Q., Ren, K., Lou, W.: ‘Privacy-preserving public auditing for data storage security in cloud computing’. InfoCom2010, IEEE, March 2010,

[4] Shacham, H., Waters, B.: ‘Compact proofs of retrievability’. Proc. 14th Int. Conf. on the Theory and Application of Cryptology and Information Security (ASIACRYPT’08), 2008

[5] Xu, J.: ‘Auditing the auditor: secure delegation of auditing operation over cloud storage’. Proc. IACR Cryptology ePrint Archive, 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) 2007.

[7] A. Juels and J. Burton, S. Kaliski, “PORs: Proofs of Retrievability for Large Files,” Proc. ACM Conf. Computer and Comm. Security (CCS ’07), Oct. 2007

[8] K.D. Bowers, A. Juels, and A. Oprea, “Proofs of Retrievability: Theory and Implementation,” Proc. ACM Workshop Cloud Computing Security (CCSW ’09) 2009.

[9] Y. Dodis, S.P. Vadhan, and D. Wichs, “Proofs of Retrievability via Hardness Amplification,” Proc. Theory of Cryptography Conf. Theory of Cryptography (TCC), 2009.

[10] H. Shacham and B. Waters, “Compact Proofs of Retrievability,” Proc. Int’l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (Asiacrypt), vol. 5350, Dec. 2008.

[11] D. Boneh, B. Lynn, and H. Shacham, “Short Signatures from the Weil Pairing,” J. Cryptology, vol. 17, no. 4, pp. 297-319, 2004.

[12] M.A. Shah, M. Baker, J.C. Mogul, and R. Swaminathan, “Auditing to Keep Online Storage Services Honest,” Proc. 11th USENIX Workshop Hot Topics in Operating Systems (HotOS ’07), pp. 1-6, 2007.

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

[14] C. Wang, Q. Wang, K. Ren, and W. Lou, “Towards Secure and Dependable Storage Services in Cloud Computing,” IEEE Trans. Service Computing, vol. 5, no. 2, Apr.-June 2012.

[15] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, “Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing”, May 2011.

[16] Kun Huang, Ming Xian, Shaojing Fu, Jian Liu “Securing the cloud storage audit service: defending against frame and collude attacks of third-party auditor”, Oct 2013.

[17] Mingxiao Ma, Jos Weber and Jan van den Berg “Secure Public-Auditing Cloud Storage Enabling Data Dynamics in the Standard Model”, 2016.

[18] Rui Zhou, Mingxing He, Zhimin Chen, "Certificateless Public Auditing Scheme with Data Privacy Preserving for Cloud Storage", April 2021.

[19] Han Yu, Xiuqing Lu, Zhenkuan Pan, "An Authorized Public Auditing Scheme for Dynamic Big Data Storage in Cloud Computing", July 2020.

[20] Shivarajkumar Hiremath,Sanjeev Kunte "A Novel Data Auditing Approach to Achieve Data Privacy and Data Integrity in Cloud Computing", Dec 2017.

[21] Swapnali Morea, Sangita Chaudhari , “Third Party Public Auditing Scheme for Cloud Storage”, 2016.

[22] Yong Yu,Yannan Li, Bo Yang, Willy Susilo, Guoming Yang, Jian Bai, "Attribute-Based Cloud Data Integrity Auditing for Secure Outsourced Storage", Aug 2015

[23] Wenting Shen, Jia Yu, Rong Hao, Xu an Wang, "A Public Cloud Storage Auditing Scheme with Lightweight Authenticator Generation", Nov 2015.

[24] Dr. J. Suganthi, Ananthi , S. Archana, “Privacy preservation and public auditing for cloud data using ass in multi-cloud”, 2015.

[25] S Ezhil Arasu, B Gowri, and S Ananthi. “Privacy-Preserving Public Auditing in cloud using HMAC Algorithm”, March 2013.

[26] Erway C C, Kupcu A, Papamanthou C, R. Tamassia “Dynamic provable data possession”, 2015