RESEARCHERID
THOMSON REUTERS

[Vincent* *et al.*, 6(6): June, 2017]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

ROBUST AND EFFICIENT PRIVACY PRESERVING PUBLIC AUDITING FOR REGENERATING-CODE-BASED CLOUD STORAGE

Tessy Vincent*, Mrs.Krishnaveni.V.V

* M-Tech Student, Dept of Computer Science College of Engineering, Kidangoor, Kottayam, Kerala, India

Assistant Professor, Dept of Information Technology College of Engineering, Kidangoor, Kottayam, Kerala, India

DOI: 10.5281/zenodo.802838

ABSTRACT

Cloud computing is gaining more popularity because of its guaranteed services like online data storage and backup solutions, Web-based e-mail services, virtualized infrastructure etc. User is allowed to access the data stored in a cloud anytime, anywhere using internet connected device with low cost. To provide security to outsourced data in cloud storage against various corruptions, adding fault tolerance to cloud storage together with data integrity checking and failure reparation becomes critical. Existing methods for remote regenerating-coded data checking only provide private auditing, which requires data owners to always stay online and do auditing, as well as repairing, which is sometimes difficult and impractical. Here proposes a public auditing scheme for the regenerating-code-based cloud storage in which data is splitted and encrypted before outsourcing. To solve the regeneration problem of corrupted files in the absence of data owners, a proxy is introduced, which have the right to regenerate the files. This scheme can release data owners completely from online burden. The cloud server is used only to save the encrypted blocks of data. In addition, the encode coefficients are randomized with a random function to preserve data privacy.

KEYWORDS: cloud storage, public audit, privacy preserving, regeneration, proxy

INTRODUCTION

Cloud computing provides service to the user over the internet. Cloud is interconnected with group of computers, which is used to store information and run their applications in cloud platform. Cloud computing is very promising for the Information Technology (IT) applications. It provides infrastructure, platform and software as services to cloud user. Through cloud computing, we can access any file, document of user from anywhere in the world. Mainly, cloud can be used for cost savings, high scalability and large storage space; however, there are still some issues to be solved for personal users and enterprises to store data and deploy applications in the Cloud computing environment. A major issue in cloud computing is security include data privacy, data protection, data availability, data location, and secure transmission. Threats, data loss, service disruption, outside malicious attacks, and multi tenancy issues are the security challenges included in the cloud. The users concerns for security should be rectified first to make cloud environment trustworthy, so that it helps the users and enterprise to adopt it on large scale. Data integrity in the cloud system means preserving the integrity of stored information. The data should not be lost or modified by unauthorized users. Cloud computing providers are trusted to maintain data integrity and accuracy of data. Data confidentiality is also important aspect from user's point of view because they store their private or confidential data in the cloud. Authentication and access control strategies are used to ensure data confidentiality. The data confidentiality could be addressed by increasing the cloud reliability and trustworthiness in Cloud computing. Therefore security, integrity, privacy and confidentiality of the stored data on the cloud should be considered and are important requirements from user's point of view [9]. To achieve all of these requirements, new methods or techniques should be developed and implemented.

Data auditing is introduced in Cloud computing technology to deal with security in data storage. Auditing is a process of user data verification which can be carried out either by the data owner or by a TPA. It helps in maintaining the integrity of data stored on the cloud. The verifier's role are categorized into two: first one is

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

private auditability, in which only user or data owner is allowed to check the integrity of the stored data. No other person has the authority to question the server regarding the data. But it tends to increases verification overhead of the user. Second is public auditability, which allows anyone, not just the client, to challenge the server and performs data verification check with the help of TPA. The TPA is an entity which is used so that it can act on behalf of the client. It has all the necessary expertise, capabilities, knowledge and professional skills which are required to handle the work of integrity verification and it also reduces the overhead of the client. It is necessary that TPA should efficiently audit the cloud data storage without requesting for the local copy of data. It should have zero knowledge about the data stored in the cloud server. It should not introduce any additional on-line burden to the cloud user [7].

The rest of the paper is organized as follows. Section II gives a brief overview of related work. Section III introduces preliminary and system model. Section IV gives the implementation overview and section V presents the security analysis and performance evaluation. Finally, we draw our conclusions in the last section.

RELATED WORK

A model for provable data possession (PDP) [1] that provides probabilistic proof that a third party stores a file. Also presents a Sampling Provable Data Possession (SPDP) scheme, which combines the RSA cryptography with Homomorphic Verifiable Tags (HVT). This model gives the provably secure scheme for remote data checking. A PDP protocol checks that an outsourced storage site retains a file, which consists of a collection of blocks. The client (data owner) pre-processes the file, generating a piece of metadata that is stored locally, transmits the file to the server and may delete its local copy. Before deleting its local copy of the file, the client may execute a data possession challenge to make sure the server has successfully stored the file. Clients may encrypt a file prior to outsourcing the storage. For this purpose, encryption is an orthogonal issue; the file may consist of encrypted data and metadata does not include encryption keys. At a later time, the client issues a challenge to the server to establish that the server has retained the file. The client requests that the server compute a function of the stored file, which it sends back to the client. Using its local metadata, the client verifies the response.The goal of a PDP scheme is to detect server misbehavior when the server has deleted a fraction of the file. The performance of PDP is bounded by disk I/O and not by cryptographic computation. Although the SPDP scheme can keep the data privacy, it cannot support the dynamic auditing and the batch auditing for multiple owners.POR, [2] a protocol in which the verifier stores only a single cryptographic key irrespective of the size and number of the files. This scheme requires that the prover access only a small portion of a (large) file F in the course of a POR. This POR protocol encrypts F and randomly embeds a set of randomly-valued check blocks called *sentinels*. The use of encryption here renders the sentinels indistinguishable from other file blocks. The verifier challenges the prover by specifying the positions of a collection of sentinels and asking the prover to return the associated sentinel values. If the prover has modified or deleted a substantial portion of F, then with high probability it will also have suppressed a number of sentinels. It is therefore unlikely to respond correctly to the verifier. To protect against corruption by the prover of a small portion of F, error-correcting codes are also employed. A drawback of POR scheme is the preprocessing / encoding of F required prior to storage with the prover. This step imposes some computational overhead beyond that of simple encryption or hashing as well as larger storage requirements on the prover. The sentinels may constitute a small fraction of the encoded F' (typically, say, 2%); the error-coding imposes the bulk of the storage overhead.*R. Curtmola et al.* [3] represents a storage system replication that increases durability and availability of the data. This provides no stronger evidence that multiple copies of the data are actually stored. It presents MR-PDP techniques that present t replicas to verify the challenges. This unique replica can produce at t unique time and that system uses t time to store the t unique replica. It shows that t replica is much more efficient than using single replica. *Kevin D.Bowers et al*. [4] introduce the High Availability and Integrity Layer (HAIL) for distributed cryptographic system that allows a set of server that prove to client that file stored is retrievable and intact. HAIL verifies and reactively reallocates file shares cryptographically. In cloud computing data owner host their data on cloud server and user can access their data from cloud server. Due to this new paradigm of outsourcing, the new security challenges which require checking the data integrity in cloud storage increases. Some techniques are available for static data storage but for dynamic data storage an efficient and secure dynamic auditing protocol is desired to convince data owners that the data are correctly stored in the cloud. For that [5] designed an auditing framework for cloud storage system and presents an efficient privacy preserving auditing protocol. This auditing protocol to support the dynamic data operations, which is efficient and provably secure in the random oracle model. Also this auditing protocol support for both multiple owners and multiple clouds, without using any trusted organizer. This paper implements secure dynamic auditing protocol.[6] presents a cooperative PDP (CPDP) scheme based on homomorphic verifiable response and hash index hierarchy method. Also verify the security of the scheme based on multiprover zero-knowledge proof system, which can

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

satisfy completeness, knowledge soundness and zero knowledge properties. In addition, performance optimization mechanisms are intellectual for the scheme, and in particular present an efficient method for selecting optimal parameter values to minimize the computation costs of clients and storage service providers. *C. Wang et al.* [7] proposed a flexible distributed storage integrity auditing mechanism, utilizing the homomorphic token and distributed erasure coded data. This scheme allows users to examine the cloud storage with very lightweight communication and computation cost. The auditing result not only ensures correctness guarantee of cloud storage, but also at the same time achieves fast data error localization, i.e., the identification of misbehaving server. Considering the dynamic nature of cloud data, this system design also supports secure and efficient dynamic operations on outsourced data, including block deletion, modification and append. It is based on the observation of linear property of the parity vector blinding process. Recall that the reason of blinding process is for protection of the secret matrix against cloud servers. However, this can be achieved either by blinding the parity vector or by blinding the data vector (assume k < m). Thus, if data vector is made blind before file distribution encoding, then the storage verification task can be successfully delegated to third party auditing in a privacy preserving manner. The privacy-preserving third party auditing is achieved, but the overall computation overhead and communication overhead remains roughly the same. A privacy preserving public auditing scheme [8] that supports public auditing and identity privacy on shared data stored in the cloud storage service for enhancing its security and effectiveness. Using HARS and its properties, designs a privacy preserving public auditing technique for outsourced data in the cloud. Here, the TPA can verify the completeness of shared data for a group of users and the identity of the signer on each block in shared data is kept private from the TPA during the auditing.Here AES algorithm is used for file encryption and uploaded to the cloud and TPA.

PRELIMINARIES AND SYSTEM MODEL

Preliminaries

Encryption: *The translation of data into a secret code. Encryption is the most effective way to achieve data security. To read an encrypted file, you must have access to a secret key or password that enables you to* decrypt *it. Unencrypted data is called* plain text ; *encrypted data is referred to as* cipher text. *There are two main types of encryption: asymmetric encryption (also called public-key encryption) and symmetric encryption. A secret key algorithm (sometimes called asymmetric algorithm) is a cryptographic algorithm that uses the same key to encrypt and decrypt data. AES (acronym of Advanced Encryption Standard) is a symmetric encryption algorithm. AES was designed to be efficient in both hardware and software, and supports a block length of 128 bits and key lengths of 128, 192, and 256 bits.*

Hashing: *hashing is a form of cryptographic security which differs from encryption. Where as encryption is a two step process used to first encryptand then decrypt a message, hashing condenses a message into an irreversible fixed-length value, or hash.Hash of a file include the message-digest hash functions MD2, MD4, and MD5, used for hashing digital signatures into a shorter value called a message-digest, and the Secure Hash Algorithm (SHA), a standard algorithm, that makes a larger (60-bit) message digest and is similar to MD4. There are several hashing methods are available, for example, MD5 hashing (128-bit hash value), SHA-1 (160-bit hash value), SHA-256 (32 byte hash value), SHA-512(64 byte hash value). As MD5 and SHA-1 hashing are expected to be collided, SHA-256 and SHA-1 can be used in the proposed system.*

System Model

The system consider the auditing system model for Regenerating-Code-based cloud storage , which involves four entities as in Fig.1: *the data owner*, who owns the large amounts of data files that to be uploaded into the cloud; *the cloud*, which are managed by the cloud service provider, that provide storage service and computational resources; *the third party auditor* (TPA), who has the abilities to perform public audits on the encoded data in the cloud, the TPA is trusted and its audit result is equal for both data owners and cloud servers; and *a proxy agent*, who is semi-trusted and acts on behalf of the data owner to regenerate data blocks on the failed servers during the repair procedure. Notice that the data owner has restricted computational and storage resources compared to other entities and may become off-line even after the data upload procedure is completed. The proxy, who would always be online, is supposed to have much more power than the data owner but have less computation and memory capacity than that of the cloud servers. To save resources as well as the online burden implicitly brought by the periodic auditing and accidental repairing, the data owners recourse to the TPA for integrity verification and delegate the reparation to the proxy. Compared with the traditional public auditing system model, system model involves an additional proxy agent. For public data auditing, the company relies on a trusted third party organization to check the data integrity; Similarly, to release the staffs from heavy online burden for data

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

regeneration, the company supply a powerful workstation (or cluster) as the proxy and provide proxy reparation service for the staffs' data.

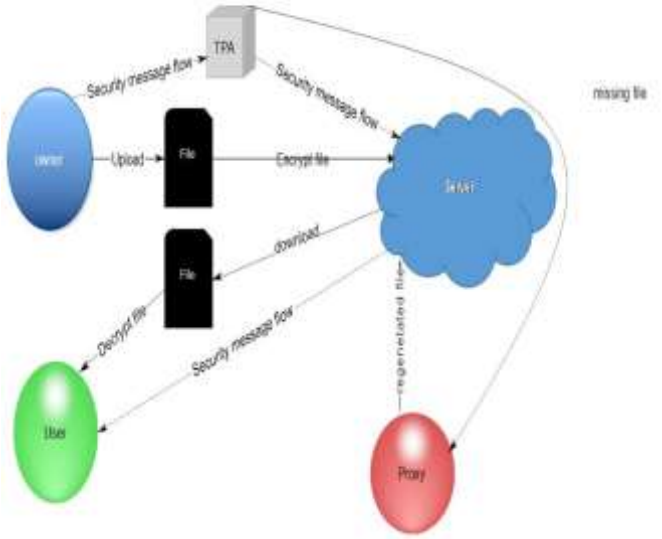

Fig.1 the cloud data storage architecture.

First stage helps the owner to register the details and also include login details. Then the owner to upload Split up file into the cloud account. Split up file hash value and the hash value of file stored in the cloud compared by the TPA to detect the modification and missing files. SHA1 method and file byte conversion are used for verification. After verification the client can view the file by downloading the permitted files with the help of secret key send by data owner to the client email-id. Random key generation method is used. If there is a problem in requested file then a message send to proxy. If the file is corrupted, the proxy interacts with the cloud servers during the repair procedure to repair the corrupted file detected by the auditing process.

IMPLEMENTATION

In existing system there is no method to detect whether the file is present in the cloud or not. Corruption of file is also not solved in the existing system due to which user's always face difficulty while downloading it. Existing system mainly focuses on file security with the help of encryption decryption system. Existing remote checking methods for regenerating-coded data only provide private auditing, requiring data owners to always stay online and handle auditing, as well as repairing, which is sometimes impractical. So the proposed system provide a TPA for avoiding data owner always stay online. To solve the regeneration problem of failed authenticators in the absence of data owners, introduce a proxy, which is privileged to regenerate the files, into the traditional public auditing system model. To fully ensure the data integrity and save the users 'computation resources as well as online burden, proposed public auditing scheme for the regenerating-code-based cloud storage, in which the integrity checking and regeneration(of failed data blocks) are implemented by a third-party auditor and a semi-trusted proxy separately on behalf of the data owner. Instead of directly adapting the existing public auditing scheme to the multi-server setting, design a novel authenticator, which is more appropriate for regenerating codes.

The proposed auditing scheme consists of three procedures: Setup, Audit and Repair.

Setup*:* The data owner maintains setup procedure to initialize the auditing scheme. In setup phase owners who are responsible for uploading professional works. While uploading the work Data is encrypted and split into 4 equal parts and stored to four different cloud accounts. This helps the file to be prevented from hackers. While uploading the four split to cloud account system hash those file using SHA1 algorithm and store the information to database table for further verification. Owner is responsible person who assign work or data to client.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

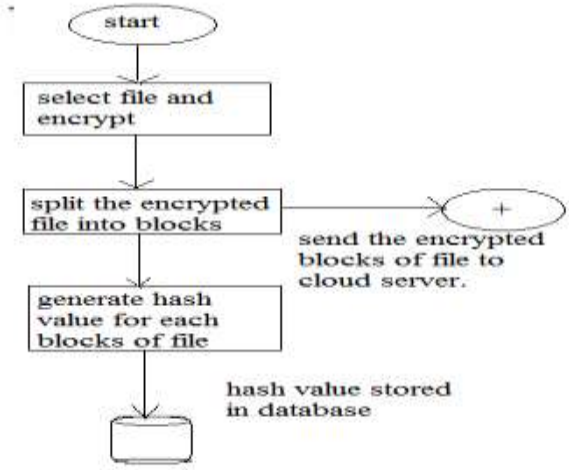

Fig.2 Working Of Data Owner

Audit: The cloud servers and TPA interact with one another to take a random sample on the blocks and check the data intactness in this procedure.TPA means Third Party Auditor. This is the responsible authority who always verify Data Owner file, which is stored in the cloud. Verification helps to know whether the file is modified or deleted from the cloud. If there is any alteration found then the information is passed to proxy. While at verification phase system uses SHA1 algorithm. With respect to the TPA, it to be honest but curious. It performs honestly during the whole auditing procedure but is curious about the data stored in the cloud. After verification (file status set into 1) the client can facility to view the file the permitted file. Client get secret key that is provided by Owner. While downloading a file system send security key for email verification. This assures the right person is accessing the file.

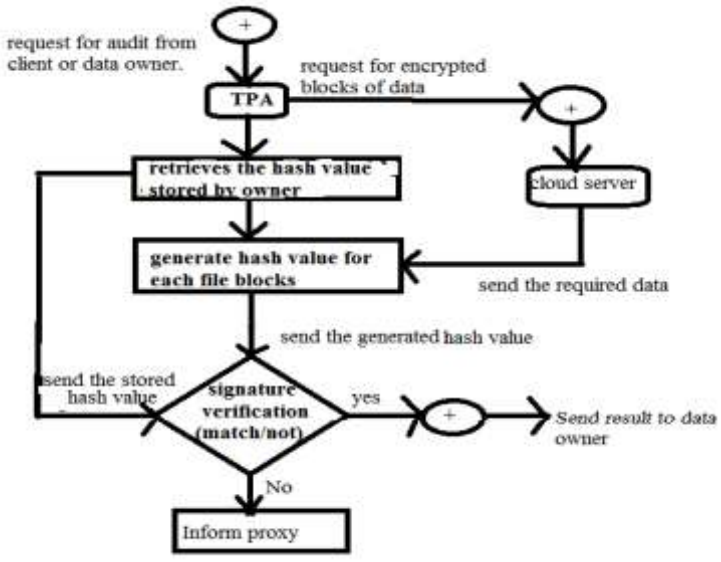

Fig.3 Working of TPA

Repair: As previously introduced, the data owner empowers a proxy to take charge of faulty server reparation and moves off-line once it completes its upload procedure. In the absence of the data owner, the proxy interacts with the cloud servers during this procedure to repair the wrong server detected by the auditing process. When TPA detects a server corruption, an alert will be sent to proxy and then a repair procedure will be triggered.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

SECURITY ANALYSIS AND EVALUATION

In this auditing system, there are two verification process, one during the audit phase by TPA to verify the possession of those data and other during repair phase by proxy to check the integrity of data. Here the data owner is completely released from online burden.Since the hash values are used by the TPA for verification, no original data is made available for audit. Also SHA-256 hashing used in this system is deterministic, fast to compute, resistant to pre-image and second-preimage attacks, and is collision resistant. In the repair phase, Proxy generates the file without reading the original file content. Only encrypted data is available for regeneration. Here for encryption AES algorithm is used which is more secure and efficient which supports larger key sizes of 128,192,256.

Incase of corruption or missing of files from the cloud, the system is more secure and efficient to regenerate the file.

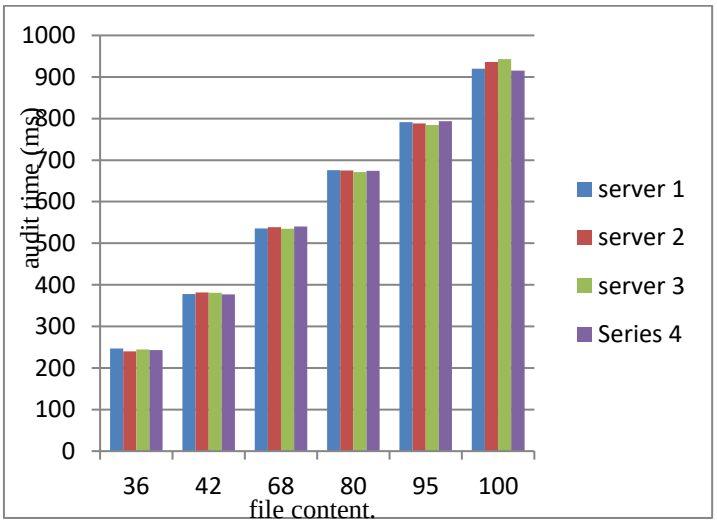

Fig.4 Time for audit with different content size

Fig.4 shows the auditing time for different files with different amount of contents. It can be concluded from the graph that there is no excess time is needed for the verification of a file stored in multiple servers. That is, all the splits that are stored at different servers are verified in a equal amount of time. There is no much overhead in audit process. Only audit result is calculated here. The time for regeneration of file is to be calculated and overall performance of the system is about to be completed.

CONCLUSION

In this paper, we propose a privacy preserving public auditing system where the data owners can allow a TPA for validating and integrity checking their data. Here data owner protect the data by encrypting and spitting files before outsourcing and only permitted users are allowed to access the the files using the key provided for them. Also the system releases the data owner from online burden by introducing a proxy to handle the repair and regeneration of files.

REFERENCES

[1] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, ser. CCS '07. New York, NY, USA: ACM, 2007, pp. 598-609.

[2] A. Juels and B. S. Kaliski Jr, "Pors: Proofs of retrievability for large files," in *Proceedings of the 14th ACM conference on Computer and communications security*. ACM, 2007, pp. 584–597.

[3] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "Mr-pdp: Multiple replica provable data possession," in *Distributed Computing Systems, 2008. ICDCS'08. The 28th International Conference on*. IEEE, 2008, pp. 411–420.

[4] K. D. Bowers, A. Juels, and A. Oprea, "Hail: a high-availability and integrity layer for cloud storage," in *Proceedings of the 16th ACM conference on Computer and communications security*. ACM, 2009, pp. 187–198.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

[5] K. Yang and X. Jia, "An efficient and secure dynamic auditing protocol for data storage in cloud computing," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 24, no. 9, pp. 1717–1726, 2013.
[6] Y. Zhu, H. Hu, G.-J. Ahn, and M. Yu, "Cooperative provable data possession for integrity verification in multicloud storage," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 23, no. 12, pp. 2231–2244, 2012.
[7] C. Wang, Q. Wang, K. Ren, and W. Lou, "Towards secure and dependable storage services in cloud computing," *Service Computing, IEEE Transactions on*, vol. 5, no. 2, pp. 220–232, May 2012.
[8] S. G. Worku, C. Xu, J. Zhao, and X. He, "Secure and efficient privacy preserving public auditing scheme for cloud storage," *Computers & Electrical Engineering*, 2013.Jian Liu, Kun Huang, Hong Rong, Huimei Wang And Ming Xian," Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage", I*nformation And Security, IEEE Transactions on*, vol 1 no 2015.
[9] Zissis, Dimitrios, and Dimitrios Lekkas. "Addressing cloud computing security issues. ",*Future Generation computer systems* 28.3(2012): 583-592.
[10] Swapnali Morea, Sangita Chaudharib," Third Party Public Auditing scheme for Cloud Storage", *bDepartment of Computer Engineering, A.C.Patil College of Engineering, Kharghar, Navi Mumbai 410210, India*

CITE AN ARTICLE

Vincent, T., & V, K. V., Mrs. (2017). ROBUST AND EFFICIENT PRIVACY PRESERVING PUBLIC AUDITING FOR REGENERATING-CODE-BASED CLOUD STORAGE. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 6*(6), 96-102. doi:10.5281/zenodo.802838