Indonesian Journal of Electrical Engineering and Computer Science
Vol. 9, No. 1, January 2018, pp. 107~110
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v9.i1.pp107-110

Cryptographic Hashing Method using for Secure and Similarity Detection in Distributed Cloud Data

A. Mohamed Divan Masood[1], Dr. S.K. Muthusundar[2]
[1]Research scholar, Information Technology, AMET University, Chennai
[2]Department of computer science, Sri Muthukumaran Institute of Technology, Chennai

Article Info

Article history:

Received Jun 29, 2017
Revised Nov 23, 2017
Accepted Dec 17, 2017

Keywords:

Deduplication
Distributed Storage
Reliability
Secure Sharing

ABSTRACT

The explosive increase of data brings new challenges to the data storage and supervision in cloud settings. These data typically have to be processed in an appropriate fashion in the cloud. Thus, any improved latency may origin animmense loss to the enterprises. Duplication detection plays a very main role in data management. Data deduplication calculates an exclusive fingerprint for each data chunk by using hash algorithms such as MD5 and SHA-1. The designed fingerprint is then comparing against other accessible chunks in a database that dedicates for storing the chunks. As an outcome, Deduplication system improves storage consumption while reducing reliability. Besides, the face of privacy for responsive data also arises while they are outsourced by users to cloud. Aiming to deal with the above security challenges, this paper makes the first effort to honor the notion of distributed dependable Deduplication system. We offer new distributed Deduplication systems with privileged reliability in which the data chunks are distributed across a variety of cloud servers. The protection needs a different of using convergent encryption as in foregoing Deduplication systems.

Corresponding Author:

A. Mohamed Divan Masood,
Research scholar, Information Technology,
AMET University, Chennai

1. INTRODUCTION

Aquantity of deduplication methods have future based on varied deduplication strategy such as client and server side deduplications file and Contentlevel deduplications. Mostly, with the start of cloud storage, data deduplication Mechanism occur to new gorgeous and important for the management of budding volumes of data in cloud storage services which cause project and association to ranch out data to third party cloud providers, as evidenced by numerous real life study [1].

There are two types of deduplication in setting of the size: (i) File Namelevel, which discovers Difficulties between various files and remove these Difficulties to reduce faculty strain, (ii) Blocklevel, which finds and remove Difficulties among data blocks. The file can be divided into smaller fixed size or not levelsize blocks. Using presetsize blocks simplify the computations of block restrictions; though using roughsize blocks provides enhanced deduplication use [2].

In accumulation, Secure Cloud also enables secure deduplication. Perceive that the "security" measured in Secure Cloud is the avoidance of leakage of surface direct information. In organize to check the leakage of such side direct information, wepursue the tradition of and mean a proof of privileges procedure among clients and cloud servers, which authorize clients to verify to cloud servers to they closely own the objectdata [3-5].

In addition, the test for data privacy also arises as added and more sensitive data are being outsourced by users to cloud. Encryption mechanisms contain usually been utilized to shield the confidentiality prior to outsourcing data into cloud. As aneffect, identical data copies of diverse users will

lead to different ciphertexts with Survey on Encryption Techniques used to Secure Cloud Storage System [6-8]. To protect both confidentiality and reliability even as achieving deduplication in a cloud storage system is at rest a challenge with generating a digital signature based on new cryptographic scheme for user authentication and security.

2. EXISTING WORKS

Howeverthis technique can remain the storage space for the cloud storage service providers; it decreases the reliability of the scheme. Data reliability is really a very serious issue in a deduplication storage system since there is only one copy for every file stored in the server shared by all the owners. If such a shared file/chunk was missing, a suspiciously large amount of data becomes inaccessible because of the unavailability of every file that shares this file/chunk. If the value of a chunk be measured in expressions of the amount of file data that would be lost in case of losing a particular chunk, followed by the amount of user data lost as a chunk in the storage system is corrupted grows with the number of the unity of the chunk. Thus, how to assurance high data reliability in deduplication system is a critical problem.

In addition, the challenge for data privacy also arises as more sensitive data are being outsourced by users to cloud. Encryption mechanisms have typically been utilized to guard the confidentiality before outsourcing data keen on cloud. Most profitable storage service provider is disinclined to apply encryption over the data for the reason that it makes deduplication impossible. As aneffect, identical data copies of diverse users will direct to different ciphertexts.

3. PROBLEM FORMULATION

3.1. System Model

In this division, we explain our proposed Secure Cloud system. Particularly, we begin with generous the system model of Secure Cloud as well as introducing the design goal for Secure Cloud. In what follows, we show the proposed Secure Cloud in detail as shown in figure 1.

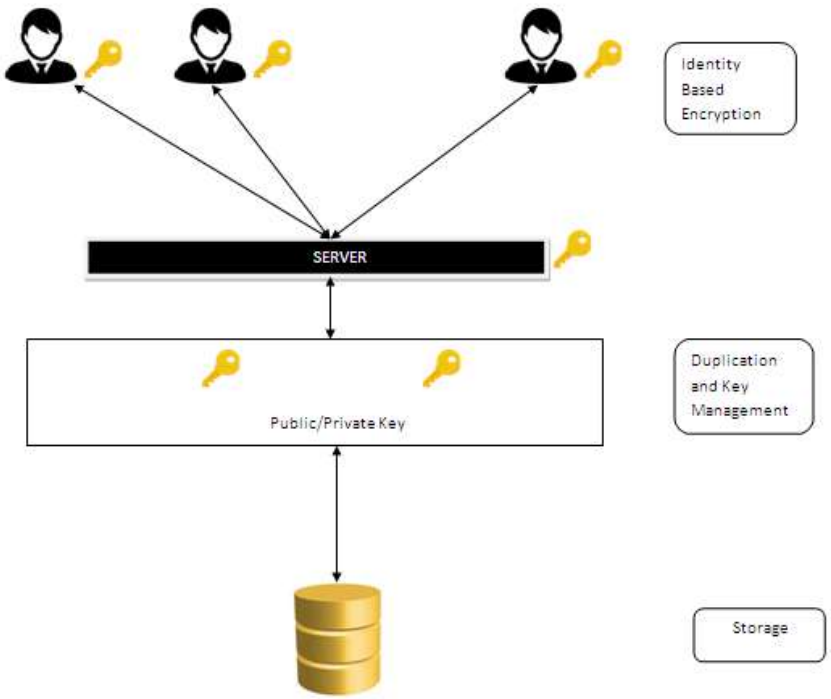

Figure 1. Secure Cloud System

Aim at allow for auditable and deduplicated storage, we present the protected Cloud system. In the system, we embrace three entities:

Clients contain huge data files to be stored on the cloud for data safety and computation. They can be also entity consumers or profitable organizations.

Cloud Servers virtualizes the goods according to the supplies of clients and account them as storage. Usually, the cloud clients may buy or fee storage capacity from cloud servers, and store their individual data in these buy or borrowed spaces for wishuse.

Auditor which assist client upload and audit their data maintain a cloud and acts related to a certificate authority. In addition, all the data has been encrypted prior to they are outsourced. The data is encrypted with the usual symmetric encryption scheme and the key is generated by the key server. The convergent key is encrypted by a new master key and stored in the cloud server.

4. OUR CONTRIBUTION

In this article, we show to design secure deduplication technique through reliability in cloud storage. We launch the distributed cloud servers involved in deduplication methods to offer better blunders. To more keep data confidentiality, the secure sharing methods are utilized, which is also well defined with the distributed storage.

4.1. Distributed Storage

These methods areuse to realistic storage process and can as well be realistic to network data transfers to decrease the bytes that must be sent. In the deduplication method, exclusive chunks of data, or byte patterns, are recognized and stored all through a process of study. Since the study continues, other chunks are compared to the stored duplicate and every time a match occurs, the unnecessary chunk is replaced among a small position that points to the stored chunk. Known that the similar byte pattern possibly will occur dozens, hundreds, or even thousands of era (the competition frequency is needy on the chunk size), the quantity of data that must be stored or transferred can be very much a bridged.

4.2. Deduplication

Data deduplication is a selective data compression method for removing photocopy copies of repeating data. Connected and rather identical terms are intelligent (data) compression and single occurrence (data) storage. In this division we emerge how to derive the well grained block level distributed deduplication. The user divider this files into blocks, if no duplication is found and performs block level deduplication system. The system set up is alike to file level deduplication and also block size restriction will be distinct.

4.3. Integrity Auditing

This attempt is to create available the ability of verifying accuracy of the some what stored data. The integrity proof additional needs two features: i) public proof, which allow each, not instantly the clients at first stored the file, to perform proof; ii) stateless proof, which is able to remove the need for state information defense at the verifier exterior between the trial of auditing and data storage.

5. RESULT AND DISCUSSION

Managing encrypted data with deduplication is significant in practice for running a secure, dependable, and green cloud storage service, especially for big data processes. Future work includes efficient data ownership verification, scheme optimization with hardware acceleration at IoT devices for practical deployment, and development of a flexible solution to support deduplication and data access controlled by either the data owner or its representative agent.

6. CONCLUSION

This paper projected the secure deduplication techniques to development of data while achieve the privacy of the users data an encryption mechanism. Four methods are proposed to embrace file or block level data deduplication. The security of tag and trut his achieved. An auditing thing with protection of a cloud, which offers clients make data tags forward of uploading well audit the honesty of data have been stored in cloud. Implement our deduplication methods using the secure sharing technique and using that it small encoding/decoding transparency compared to the network transmission.

REFERENCES

[1] Li M, Qin C, Lee P P, Li J. Convergent Dispersal: Toward Storage-Efficient Security in a Cloud-of-Clouds. In *HotCloud*. 2014.
[2] Armbrust M, Fox A, Griffith R, Joseph A D, Katz R, Konwinski A, Zaharia M. A view of cloud computing. *Communications of the ACM*, 2010; 53(4), 50-58.
[3] Stanek J, Sorniotti A, Androulaki E, Kencl L. *A secure data deduplication scheme for cloud storage*. In International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg. 2014: 99-118.
[4] Yuan J, Yu S. *Secure and constant cost public cloud storage auditing with deduplication.* IEEE Conference in Communications and Network Security (CNS). 2013: 145-153.
[5] Ateniese G, Burns R, Curtmola R, Herring J, Khan O, Kissner L, Song D. Remote data checking using provable data possession. *ACM Transactions on Information and System Security (TISSEC)*, 2011; 14(1): 12.
[6] Li J, Chen X, Li M, Li J, Lee P P, Lou W. Secure deduplication with efficient and reliable convergent key management. *IEEE transactions on parallel and distributed systems*, 2014; 25(6): 1615-1625.
[7] Kirubakaramoorthi R., Arivazhagan D, Helen D. Survey on Encryption Techniques used to Secure Cloud Storage System. *Indian journal of Science and Technology*. 2015; 8(36).
[8] Ganeshkumar K, & Arivazhagan D. Generating a digital signature based on new cryptographic scheme for user authentication and security. *Indian Journal of Science and Technology*. 2014; 7(S6): 1-5.
[9] Z. Yan, W. Ding, and H. Zhu, "Manage Encrypted Data Storage with Deduplication in Cloud," Proc. Int'l Conf. Algorithms and Architectures for Parallel Processing (ICA3PP), 2015, pp. 547–561.