Cloud Data Auditing and Deduplication by Implementing Bit Exchange Method to the SecCloud+

Sabareesan M[1], Dhivya P[2], Hemalatha S[3], Keerthana M[4]

[1]Assistant Professors, [2, 3 & 4] UG Scholars Department of Computer Science & Engineering,

V.R.S College of Engineering & Technology, Arasur, Villupuram

[1]sabareesan81@gmail.com,[2]rajuranjith1@gmail.com,

[3]hemalatha2710@gmail.com, [4]manikamkeerthana@gmail.com

---***---

Abstract - ***Cloud storage has started getting boom in the market present trend. But considering the problem of storage being handled in a remote server far away from the vision of the clients arises a suspicion of integrity of the uploaded contents. The privacy of the user's is the atmost criterion for the day today cloud scenario. Apart from which another bigger threat is with the duplication of the data which leads to server overflow and thus ultimately expanding servers for duplicate files which could have been otherwise handled in a better manner. At present, the system for today with regards to the primitive cloud architecture has a more efficient system. With SecCloud and then SecCloud+ which has put up with comparative advantages to the primitive system. But considering it's real time application it has proved to be delaying due to the encryption and tagging process. Hence we adopt a faster system by implementing Bit-Exchange Algorithm for encryption and the similar Hash-Based message authentication code for tagging the encrypted messages.***

Keywords: Cloud storage, SecCloud, SecCloud+, Bit-Exchange, Hash Based message authentication, Deduplication.

1. Introduction

Cloud storage is a model of networked enterprise storage where data is stored in virtualized pools of storage which are generally hosted by third parties. Cloud storage provides customers with benefits, ranging from cost saving and simplified convenience, to mobility opportunities and scalable service. These great features attract more and more customers to utilize and storage their personal data to the cloud storage: according to the analysis report, the volume of data in cloud is expected to achieve 40 trillion gigabytes in 2020. Even though cloud storage system [5] has been widely adopted, it fails to accommodate some important emerging needs such as the abilities of auditing integrity of cloud files by cloud clients and detecting duplicated files by cloud servers.

We illustrate both problems below. The first problem is integrity auditing. The cloud server is able to relieve clients from the heavy burden of storage management and maintenance. The most difference of cloud storage from traditional in-house storage is that the data is transferred via Internet and stored in an uncertain domain,[4] not under control of the clients at all, which inevitably raises clients great concerns on the integrity of their data. These concerns originate from the fact that the cloud storage is susceptible to security threats from both outside and inside of the cloud, and the uncontrolled cloud servers may passively hide some data loss incidents from the clients to maintain their reputation. What is more serious is that for saving money and space, the cloud servers might even actively and deliberately discard rarely accessed data [2]files belonging to an ordinary client.

Considering the large size of the outsourced data files and the clients' constrained resource capabilities, the first problem is generalized as how can the client efficiently [14] perform periodical integrity verifications even without the local copy of data files. The second problem is secure deduplication. The rapid adoption of cloud services is accompanied by increasing volumes of data stored at remote cloud servers. Among these remote stored files, most of them are duplicated: according to a recent survey by EMC, 75% of recent digital data is duplicated copies. This fact raises a technology namely deduplication, in which the cloud servers would like to deduplicate by keeping only a single copy[1] for each file (or block) and make a link to the file (or block) for every client who owns or asks to store the same file (or block). Unfortunately, this action of deduplication would lead to a number of threats potentially affecting the storage system for example, a server telling a client that it (i.e., the client) does not need to send the file reveals that some other client has the exact same file, which could be sensitive sometimes. These attacks originate from the reason that the proof that the client owns a given file[10] (or block of data) is solely based on a static, short value (in most cases the hash of the file).

Thus, the second problem is generalized as how can the cloud servers efficiently confirm that the client (with a certain degree assurance) owns the uploaded file (or block) before creating a link to this file[7] (or block) for him/her. In this paper, aiming at achieving data integrity and deduplication in cloud, we propose two secure systems namely SecCloud and SecCloud+. SecCloud introduces an auditing entity with maintenance of a MapReduce cloud, which helps clients generate data tags before uploading as well as audit the integrity of data having been stored in cloud. This design fixes the issue of previous work that the computational load at user or auditor is too huge for tag generation. For completeness of fine-grained, the functionality of auditing designed in SecCoud is supported on both block level and sector level. In addition, SecCoud also enables secure deduplication.

Notice that the “security” considered in SecCoud is the prevention of leakage of side channel information. Motivated by the fact that customers always want to encrypt their data before uploading, [9]for reasons ranging from personal privacy to corporate policy, we introduce a key server into SecCloud and propose the SecCloud+ schema. Besides supporting integrity auditing and secure deduplication, SecCloud+ enables the guarantee of file confidentiality. Specifically, thanks to the property of deterministic encryption in convergent encryption,[11] we propose a method of directly auditing integrity on encrypted data.

The challenge of deduplication on encrypted is the prevention of dictionary attack. As with we make a modification on convergent encryption[12] such that the convergent key of file is generated and controlled by a secret “seed”, such that any adversary could not directly derive the convergent key from the content of file and the dictionary attack is prevented.

2. Proposed system

To overcome this problem we are implementing secCloud+ with bit exchange method and Hash tag algorithm for identifying the duplication file that has been already stored in the cloud[16].Besides supporting integrity auditing and secure deduplication,[15] SecCloud+ enables the guarantee of file confidentiality .This method done integrity auditing directly on encrypted data.

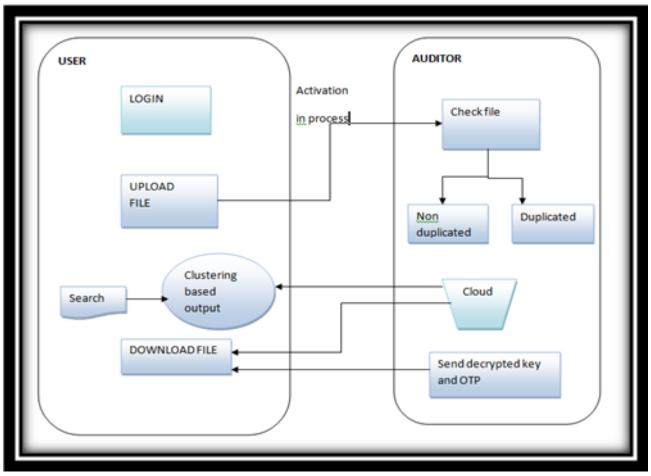

Fig 1: Architecture Diagram

3. System Implementation

A. User module
B. File upload/download protocol
C. secure auditing protocol
D. Block-level deduplication system
E. Bit exchanging method

A .User Module:

In this module a user has to upload their files in a cloud server, [3]they should register first. Then only they can be able to upload their file. For that they have to fill their details in the registration form. These details are maintained in a database. In this module, the user have to login, they should give their name and password to get authenticate from the cloud

B.File uploading/downloading protocol:

Upload:

In this module the user uploads their file into the cloud server and then the uploaded file was converted into cipher text, [6] the encryption process is done by BEM (Bit Exchanging Method).The Auditor audits the user's files for deduplication and then only user's files are uploaded to the cloud server.

Download:

In this module user downloads the files in decrypted format. The downloaded file is in encrypted format. User receives OTP (One Time Password) on their mail. The user enters the correct key then only it is decrypted. Decryption process also we are using Bit Exchanging Method algorithm only.

C. Secure Auditing Protocol:

In this module, auditor have to login, they should login by giving their username and password. A secure auditing protocol will audit the uploaded files as per their status. [8]Auditor approves only non duplicate files, to be stored in cloud.

The System audits user files. If it is duplicate then the auditor will not provide the uploading permission to that file. If uploaded file is non-duplicate then auditor gives the activation for that file.Then only that file is stored in the cloud server

D.Block-level deduplication System:

We consider block level deduplication in that file is divided into blocks and checked for deduplication for blocks. For encryption we are using Bit Exchanging Method. Data deduplication can generally operate at the file,[9] block or byte level thus defining minimal data fragment that is checked by the system for redundancy. Block-level data deduplication methods deliver the benefit of optimizing storage capacity. Hash algorithm generates a unique identifier – hash number - for each analyzed chunk of data. It is then stored in an index and used for figuring out [13]duplicates – the duplicated fragments have the same hash numbers.

E.Bit Exchanging Method:

Encryption taken on the secret message files using simple bit shifting and XOR operation.

The bit exchange method is introduced for encrypting any file.

Steps:

1. Read one by one byte from the secret data and convert each byte to 8 bits. Then apply one bit right shift operation.
2. Divide the 8 bits into two blocks and then perform XOR operation with 4 bits on the left and 4 bits on the right side.
3. The same thing repeated for all bytes in the file.

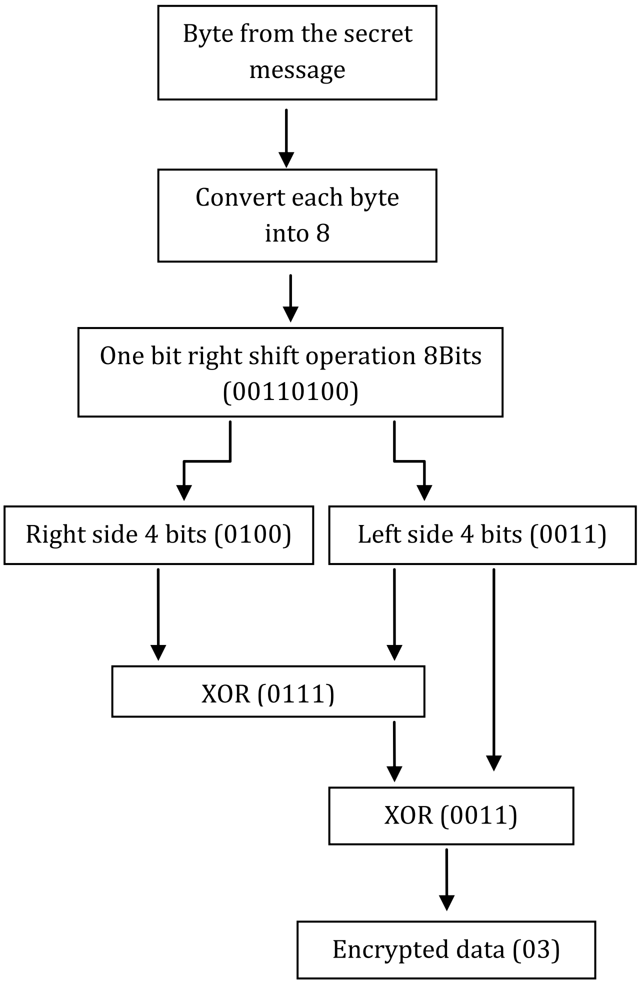

Fig 2: Bit exchange method

4. Conclusion

Aiming at achieving both data integrity and deduplication in cloud, we propose SecCloud and

SecCloud+. SecCloud introduces an auditing entity with maintenance of a MapReduce cloud, which helps clients generate data tags before uploading as well as audit the integrity of data having been stored in cloud. In addition, SecCoud enables secure deduplication through introducing a Proof of Ownership protocol and preventing the leakage of side channel information in data deduplication. Compared with previous work, the computation by user in SecCloud is greatly reduced during the file uploading and auditing phases. SecCloud+ is an advanced construction motivated by the fact that customers always want to encrypt their data before uploading, and allows for integrity auditing and secure deduplication directly on encrypted data. In this project we implement secCloud+ with BIT EXCHANGE METHOD (BEM) for identifying deduplication in the uploaded file.

REFERENCES

[1] Yan Zhua,b, Hongxin Huc, Gail-Joon Ahnc, Stephen S. Yauc ,"Efficient Audit Service Outsourcing For Data Integrity In Clouds",In The Journal of Systems and Software 85 (2012) .

[2] Wang.Q, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling Public Audit Ability And Data Dynamics For Storage Security In Cloud Computing", In IEEE Trans. Parallel Distributed Systems, vol. 22, no. 5, pp. 847-859, May 2011.

[3] Cong Wang, Student Member, IEEE, Sherman S.M. Chow, Qian Wang, Student Member, IEEE, Kui Ren, Member, IEEE, and Wenjing Lou, Member, IEEE, "Privacy-Preserving Public Auditing For Secure Cloud Storage".

[4] Abhishek Mohta* ,Ravi Kant Sahu,Lalit Kumar Awasthi ,Dept. of CSE, NIT Hamirpur (H.P.) India,"Robust Data Security For Cloud While Using Third Partyauditor"

[5] Juels.A and J. Burton, S. Kaliski, "Pors: Proofs Of Retrievability For Large Files",In Proc. ACM Conf. Computer and Comm. Security (CCS‟07), pp. 584-597, Oct. 2007.

[6] Ateniese.G, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z.Peterson, and D. Song, "Provable Data Possession At Untrusted Stores" ,In Proc. 14th ACM Conf. Computer and Comm. Security (CCS‟07), pp. 598-609, 2007.

[7]Govinda.K, V.Gurunathaprasad, H.Sathishkumar, "Third Party Auditing For Secure Data Storage In Cloud Through Digital Signature Using RSA", In International Journal Of Advanced Scientific And Technical Research(Issue 2, Volume 4- August 2012) Issn 2249-9954.

[8] Ezhil Arasu.S, B.Gowri, S.Ananthi ,"Privacy-Preserving Public Auditing In Cloud Using HMAC Algorithm ",In International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-1, March 2013 .

[9] Shingare Vidya Marshal ,"Secure Audit Service by Using TPA for Data Integrity in Cloud System",In International Journal of Innovative Technology and Exploring Engineering (IJITEE)ISSN: 2278-3075, Volume-3, Issue-4, September 2013.

[10] Jiawei Yuan,Shucheng Yu "Secure and Constant Cost Public Cloud Storage Auditing with Deduplication"

[11] Jiawei Yuan, Shucheng Yu,"Proofs of Retrievability with Public Verifiability and Constant Communication Cost in Cloud",University of Arkansas at Little Rock ,USA.

[12] Jeyadevan.S, Dr.S.Basavaraj Patil, S.Saravanan, Naina Kumari, "Introducing Various Algorithms To Make The Data-Storage In Clouds Secure",In International Conference on Computing and Control Engineering (ICCCE 2012), 12 & 13 April, 2012.

[13]Vijeyta Devi & Vadlamani Nagalakshmi,"A Prospective Approach On Security With RSA Algorithm And Cloud SQL In Cloud Computing", In International Journal Of Computer Science And Engineering (Ijcse) Issn 2278-9960 Vol. 2, Issue 2, May.

[14] Vidhisha.S, C.Surekha, S.Sanjeeva Rayudu, U.Seshadri," Preserving privacy for secure and outsourcing for Linear Programming in cloud computing", Computer Science Engineering Jawaharlal Nehru Technological University Ananatapur.

[15] V.Venkatesh, P.Parthasarathi," Enhanced audit services for the correctness of outsourced data in cloud storage ",In International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2.

[16]" Deleting Secret Data with Public Verifiability" Feng Hao, Member, IEEE, Dylan Clarke, Avelino Francisco Zorzo.

BIOGRAPHIES

Mr.M.SABAREESAN is working as a Assistant Professor at V.R.S. College of Engineering & Technology, Villupuram. His Qualification is M.E,.Computer Science & Engineering. His research area is Image Processing, Network Security & Wireless Sensor Networks. He has published six international journals.

Ms.P.Dhivya is a final year student at V.R.S college of Engineering and Technology,Villupuram.

Ms.S.Hemalatha is a final year student at V.R.S. college of Engineering & Technology, Villupuram.

Ms.M.Keerthana is a final year student at V.R.S. college of Engineering & Technology,Villupuram..