International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 23 Issue 8 –NOVEMBER 2016.

Avoiding Reduplications and Providing Secure Auditing In Cloud

SK.NAZIYA BANU[#1] and I. TABHITA[*2]

[#] *M.Tech (CSE) Student, Nimra College of Engineering & Technology, A.P., India.*

[*] *Assistant Professor, Dept. of Computer Science & Engineering, Nimra College of Engineering & Technology, A.P., India.*

***Abstract*— According to some survey, the volume of data in cloud is expected to achieve many trillions of gigabytes. Even though cloud storage system has been extensively used, it ignores some important emerging needs such as the abilities of auditing integrity of cloud file and detecting duplicated files by cloud servers. since the outsourced cloud storage service is not fully trustworthy, it raises security concerns on how to realize data de-duplication in cloud while achieving integrity auditing. In this work, we study the problem of secure de-duplication on cloud data, also ensuring integrity. Specifically, aiming at achieving both data integrity and de-duplication in cloud, we propose a system, namely D-cloud. D-cloud introduces an auditing entity with maintenance of the cloud, which helps generate hash value before uploading as well as audit the integrity of data having been stored in cloud. Compared with previous work, the computation by user in D-Cloud is greatly reduced during the file uploading and auditing phases. D-cloud is designed realizing the fact that users always want to encrypt their data being uploaded, and enables integrity auditing and secure de-duplication on encrypted data. This paper present, the plan of approved data deduplication planned to guard data security by counting discrepancy privileges of users in the duplicate check. Deduplication systems, users with differential privileges are added measured in duplicate check besides the data itself. To maintain stronger security the files are encrypted with differential privilege keys. Users are only permitted to carry out the copy check for files marked with the matching privileges to access. The user can confirm their occurrence of file after deduplication in cloud with the help of a third party auditor by auditing the data. Additional auditor audits and confirms the uploaded file on time. As a result, this paper generates advantages to both the storage provider and user by deduplication system and auditing method correspondingly.**

Index Terms— Data deduplicating, Secure auditing, D-cloud, Integrity, Encrypted Data.

I. INTRODUCTION

Cloud data storage services involves four entities.(i)Administrator controls the user details, file insertion, file access, file deletion and the time of user presents in the network to access the cloud data's. (ii)Third party auditor checks the correctness of cloud data. Some techniques are used to establish the auditing concepts. (iii) Users access the cloud data as per demand services. Users retrieve more useful information from multiple repositories and no limitation to access the particular storage part in the shared pool. Those days the computer was just used forperforming arithmetic and logical operations. But as the world evolved with inventions and innovations, more and more data got generated eventually. Then there was the use of Hard driveto store useful data, which was very costly. Birth of Internet provided various technologies including Cloud Storage. As we all know cloud storage is basically storing of data (Image, Videos, File, etc.) on a virtual server or we can say on a virtual database. Technically a cloud computing/storage is further explaining as, a system for enabling convenient on demand networkaccess to share data between computers. It is an internet based service which helps to store data by managing the storage. Cloud Storage provides the users ranging from cost saving from and simplified convenience, to mobility opportunities and scalable services. According to some survey, the volume of data in cloud is expected to achieve many trillions of gigabytes. Even though cloud storage system has been extensively used, it ignores some important emerging needs such as the abilities of auditing integrity of cloud file and detecting duplicated files by cloud servers .The second problem is solved using deduplication. The rapid adoption of cloud services is accompanied by increasing volumes of data stored at remote cloud servers which also causes similar (duplicate) files being stored at multiple locations, wasting the memory resource. This problem is countered by a technology namely deduplication, in which the cloud servers would like to deduplicateby keeping only a single copy for each file (or block) and attach a link to this file (or block) for every client who owns it . This generates the problem of transparency, which need not be compromised i.e. no two owners of the same duplicate file should be aware of ownership by other client as well and also of the deduplication being performed on his/her data. In this paper, aiming at achieving deduplication with standard security and data integrity, we propose a secure system named Dcloud. Deduplication is a method where the server saves only a single copy of each file, regardless of which clients asked to store that file, such that the disk space of cloud servers as well as network bandwidth are saved. However, trivial client side deduplication leads to the leakage of side channel information.

II. RELATED WORK

Cloud Clients: Cloud Clients have large data files to be stored and rely on the cloud for data maintenance and computation. They can be either individual consumers or commercial organizations. **Public Storage:** Public Storage is an storage disk which permit to store the users data which contains authorization and not permit to upload the duplicate data. Thus save storage space and bandwidth of transmission. This uploaded data is in encrypted form, only a user with individual key can decrypt it. **Private Cloud**: A private cloud acts as a proxy to allow both data owner and user to strongly perform duplicate check along with disparity permissions. **Auditor:** Auditor is a TPA work as proficiency and capabilities where cloud users do not have to faith to assess the cloud storage service reliability on behalf of the user upon request. The set of permissions and the symmetric key for each privilege is allocates and stored in private cloud. The user registers into the system, permissions are assigned to user according to identity given by the user at registration time; means on basis of situation which access by the user. The data owner with permission can upload and share a file to users, further the data owner performs identification and sends the file tag to the private server. Private cloud server checks the data owner and computes the file token and will send back the token to the data owner. The data owner throws this file token and a request to upload a file to the storage provider. If duplicate file is found then user needs to run the PoW protocol with the storage provider to prove that user has an ownership of respective file. In the PoW result; if proof of ownership of file is approved then user will be provided a pointer for that file. And on the next case; for no duplicate is found for the file, the storage provider will be come again a signature for the result of that proof for the particular file. To upload file user sends the privilege set as well as the proof to the private cloud server in the form of a request. The private cloud server verifies the signature first on receiving the request for the user to upload file.

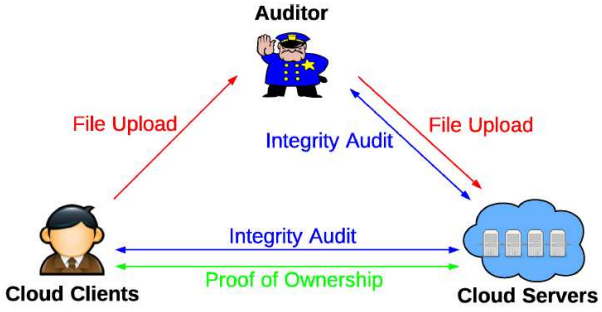

Fig. 1 Overview of System Architecture

III. PROPOSED METHODOLOGY

This paper is an attempt to enhance the Security Model of SecCloud+ with multi-layered cryptosystem based Secure and Authorized Auditing deduplication model. In this paper, we present a scheme that permits a more fine-grained trade-off. The intuition is that outsourced data may require different levels of protection, depending on how popular it is: content shared by many users, such as a popular song or video, arguably requires less protection than a personal document, the copy of a payslip or the draft of an unsubmitted scientific paper. As more corporate and private users outsource their data to cloud storage providers, recent data breach incidents make end-to-end encryption an increasingly prominent requirement. Unfortunately, semantically secure encryption schemes render various cost-effective storage optimization techniques, such as data deduplication, ineffective. We present a novel idea that differentiates data according to their popularity. Based on this idea, we design an encryption scheme that guarantees semantic security for unpopular data and provides weaker security and better storage and bandwidth benefits for popular data. This way, data deduplication can be effective for popular data, whilst semantically secure encryption protects unpopular content. We show that our scheme is secure under the Symmetric External Decisional Diffe-Hellman Assumption in the random oracle model.

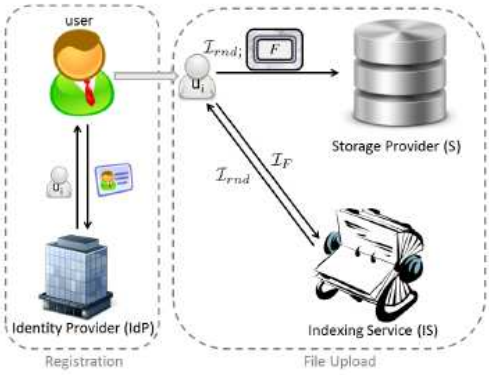

Figure 2: System Model

IV. LITERATURE REVIEW

1.Author: Shai Halevi Cloud storage systems are becoming more and more popular. A promising technology that keeps their cost down is deduplication, which stores only a single copy of duplicatin g data. Client-side deduplication attempts to identify deduplication opportunities already at the client side and save the bandwidth of uploading copies of existing files to the server. In this work we identify attacks that exploit client-side deduplication, granting an attacker to gain access to arbitrary-size files of other users based on a very small hash signature of these files. More specifically, an attacker who knows the hash signature of a file can assure the storage service that it owns that file, hence the server lets the attacker download the entire file.

2.Author: Qian Wang Cloud Computing system has been predicted as the next-generation architecture of IT Enterprise. It moves the application software and databases to the centralized with large data centers, where the management of the data and services may not be fully trustworthy. This unique ensample brings about many new security challenges,

which have not been well understood. Our research work examine the problem of assuring the integrity of data storage in Cloud Computing. In particular, we consider the task of allowing a third party auditor (TPA), on concern of the cloud client, to verify the integrity of the dynamic data stored in the cloud. The introduction of TPA dismiss the involvement of client through the auditing of whether user's data stored in the cloud is truly intact, which can be important in achieving economies of scale for Cloud Computing. The support for data dynamics through the most general forms of data operation, such as block modification, insertion and deletion, is also more powerful step to - ward practicality, since services in Cloud Computing are not limited to archive or backup data only. While presiding work on ensure remote data integrity often lack the supports of either public verifiability or dynamic data operation. Proofs of Ownership in Remote Storage Systems.

3.Authors: Giuseppe Ateniese We suggest a model for provable data possession (PDP) that can be used for remote data checking: A client that has stored data at an untrusted server can verify that the server possesses the original data without retrieving it. The model generates probabilistic proofs of possession by sampling random sets of blocks from the server, which drastically reduces I/O costs. The 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. Thus, the PDP model for remote data checking is lightweight and supports large data sets in distributed storage systems. The model is also robust in that it incorporates mechanisms for mitigating arbitrary amounts of data corruption.

4.Author: Hovav Shacham and Brent Watersy introduced proof-of-retrievability system. In this integrity check system, data storage centre provide proof to a verier that it is actually storing all of a client's data. Here they have explained two homomorphic authenticators the first authenticator is based on PRFs, gives a proof-of-retrievability scheme secure in the standard model. The second, based on BLS signatures [4], which give a proof-of-retrievability scheme with public variability secure in the random oracle model. Frameworks explained can allow arguing about the systems unforgeability, extractability, and retrievability with these parts based on cryptographic, combinatorial, and coding theoretical techniques respectively

V. CONCLUSION

The major goal of this web application is to help the users to store their data on the cloud with confidentiality and security. De-duplication of data is the main focus in the entire web application. Providing storage of data on a large scale with multiple file sharing. Auditing helps the user to check the integrity of the data. Data privacy and data security are the main issues 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 which introduced to maintain integrity of data and different methods for overcoming the issues data deduplication on untrusted data stores in cloud computing. There are still some approaches which are not covered in this paper. This paper categories the different methodologies in the literature as encryption based methods, access control based techniques, query integrity, keyword search schemes, and auditability schemes. Even though there are many techniques in the literature for considering the concerns in data integrity and data deduplication. De-duplication of data is the main focus in the entire web application. Providing storage of data on a large scale with multiple file sharing. Auditing helps the user to check the integrity of the data.

REFERENCES

[1] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, ―Proofs of ownership in remote storage systems,‖ in Proceedings of the 18th ACM Conference on Computer and Communications Security . ACM, 2011, pp. 491– 500.

[2] Chandinee Saraswathy K. , Keerthi D. , Padma G. "HLA Based Third Party Auditing For Secure Cloud Storage" International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1526-1532

[3] H. Shacham and B. Waters, “Compact Proofs of Retrievability,”in the Proceedings of ASIACRYPT 2008. Springer-Verlag, 2008, pp.90–107.

[4] 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.

[5] 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.

[6] G. Ateniese, R. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. Peterson, and D. Song, “Remote data checking using provable data possession,” ACM Trans. Inf. Syst. Secur., vol. 14, no. 1, pp. 12:1–12:34, 2011.

[7] G. Ateniese, R. Di Pietro, L. V. Mancini, and G. Tsudik, “Scalable and efficient provable data possession,” in Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks, ser. SecureComm ’08. New York, NY, USA: ACM, 2008, pp. 9:1–9:10.

[8] https://view.officeapps.live.com/op/view.aspx?src=http%3A%2F%2Fwww.cs.sjsu.edu%2F~stamp%2FCS265%2Fprojects%2Fpapers Spr03%2FMD5.ppt

[9] AnthonyVelte& Robert C.Elsenpeter “Cloud Computing a Practical Approach”, McGraw-Hill, Inc. New York, NY, USA ©2010 [10] R Sravan Kumar &A.Saxena “Data Integrity and Proofs in Cloud Storage”

SK.NAZIYA BANU is a student of NIMRA COLLEGE OF INSTITUTE AND TECHNOLOGY, Ibraheempatnam, VIJAYAWADA. She is presently pursuing his M.Tech degree from JNTU, Kakinada. She has obtained B.Tech, degree from JNTU, Kakinada.

I. TABHITA is presently working as Assistant professor in CSE department in Nimra college of Engineering and Technology, Jupudi, Nimra Nagar,VIJAYAWADA. She has published several research papers in various national and international Journals.