International Journal of Engineering Research and General Science Volume 4, Issue 1, January-February, 2016
ISSN 2091-2730

A Survey Paper on Secure Auditing and Maintaining Block Level Integrity with Reliability of Data in Cloud

Aishwarya R. Kumthekar[1], Prof. Jyoti Raghatwan [2]

[1] M.E. II Computer, aish275@gmail.com, 8806967523

Abstract-- As the cloud computing innovation creates amid the most recent decade, outsourcing information to cloud administration for capacity turns into an alluring pattern, which benefits in saving endeavours on substantial information upkeep and administration, In any case, following the outsourced cloud stockpiling is not completely reliable, it raises security worries on the most proficient method to acknowledge information deduplication in cloud while accomplishing uprightness examining. In this work, we ponder the issue of honesty examining and secure deduplication on cloud information. Specifically, going for accomplishing both information uprightness and deduplication in cloud, we propose two protected frameworks, to be specific SecCloud and SecCloud+. SecCloud presents an examining substance with an upkeep of a MapReduce cloud, which assists customers with producing information labels before transferring and in addition review the honesty of information having been put away in cloud. Contrasted and past work, the calculation by client in SecCloud is enormously lessened amid the file transferring and reviewing stages. SecCloud+ is composed propelled by the way that clients constantly need to scramble their information before transferring, and empowers honesty evaluating and secure deduplication on encoded information.

Keywords— Secure auditing, Deduplication, Reliability, Cloud computing, Third Party Auditor, Diffie-Hellman Key Exchange,

INTRODUCTION

Despite the fact that cloud stockpiling framework has been generally embraced, it neglects to oblige some critical emerging needs, for example, the capacities of auditing integrity of cloud files by cloud customers and detecting copied files by cloud servers. We show both issues underneath. The first issue is integrity auditing. The cloud server has the capacity alleviate customers from the substantial weight of capacity administration and maintenance. The most distinction of cloud stockpiling from customary in-house stockpiling is that the data is exchanged by means of Internet and put away in an uncertain domain, not under control of the customers by any stretch of the imagination, which inevitably raises customers extraordinary worries on the integrity of their data. These worries originate from the way that the cloud stockpiling is defenseless to security dangers from both outside and inside of the cloud, and the uncontrolled cloud servers might inactively conceal some data misfortune incidents from the customers to maintain their notoriety.

In addition genuine is that for saving cash and space, the cloud servers may even effectively and purposely dispose of once in a while got to data files belonging to an ordinary customer. Considering the substantial size of the outsourced data files and the customers' constrained asset abilities, the first issue is summed up as in what manner can the customer efficiently perform periodical integrity verifications even without the neighborhood duplicate of data files.

The cloud storage is powerless to security dangers from both outside and inside of the cloud [1], and the uncontrolled cloud servers might inactively conceal some information misfortune episodes from the customers to keep up their notoriety. Deduplication would prompt various dangers conceivably influencing the stockpiling framework [3][2], for instance, a server telling a customer that it (i.e., the customer) does not require to send the record uncovers that some other customer has the precise same record, which could be touchy sometimes.customers dependably need to encrypt their information before transferring, for reasons extending from individual protection to corporate strategy, we bring a key server into SecCloud as with [4] and propose the SecCloud+ pattern.

LITERATURE SURVEY

1) Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing:

Cloud Computing has been envisioned as the next-generation architecture of IT Enterprise. It moves the database and application software to the centralized large data centers, where the management of the data and services may not be fully trustworthy. This unique factor brings about many new security challenges, which have not been well understood yet. This work studies the problem of ensuring the data integrity of the storage in Cloud Computing. In particular, we can consider the task of allowing a third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the data stored in the cloud storage but dynamic data. The introduction of TPA eliminates the involvement of client through the auditing of whether his data stored in the cloud is indeed intact, which can be important in achieving economies of scale for Cloud Computing as well. The support for data dynamics via the most general forms of data operation, such as block modification, deletion , insertion is also a significant step towards practicality, since services in the Cloud Computing are not limited to archive or backup data. While prior work on ensure remote data integrity often lack the supports of either public verifiability or dynamic data operation.

2) Proofs of Ownership in Remote Storage Systems:

Cloud storage systems are becoming increasingly popular and popular. A promising technology that keeps their cost down is removing duplication of file , which stores only a single copy of the data repeating. Client side deduplication attempts to identify the deduplication opportunities already present at the client and save the bandwidth of uploading copies of existing files to the server which is harmful. In this work we identify attacks that exploit client-side deduplication, allowing 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 if he knows the hash signature of a file can convince the storage service that it owns that file, hence that server lets the attacker download the entire file.

3) DupLESS: Server-Aided Encryption for Deduplicated Storage :

Cloud storage service providers such as Dropbox and others perform the deduplication to save space by only storing one copy of each file uploaded on it. Should clients conventionally encrypt their files, however their savings are lost. Message locked encryption (the prominent manifestation of which is convergent encryption) resolves this all the tension. However it is inherently subject to the brute force attack that can recover files falling into some known set. We propose an architecture which provides secure de-duplicated storage resisting brute-force attacks too, and realize it in a system called DupLESS . In DupLESS, client encrypt under message based keys obtained from key server via an oblivious PRF protocol. It enables the client for storing encrypted data with an existing service,that have the service perform deduplication on their behalf, and yet achieves strong confidentiality guarantees. We show that encryption for the de-duplicated storage can achieve the performance and the space savings close to that of using all the storage service with plaintext data.

4) Provable Data Possession at Untrusted Stores :

Introduce a model for provable data possession that is PDP which allows a client that has stored data at an untrusted server for the verification of the server possesses the original data without retrieving. The model generates the probabilistic proofs of the possession by sampling some random sets of blocks from the server, which drastically reduces I/O costs. The client maintains some constant amount of the metadata to verify the proof. The challenge or the response protocol transmits small and constant amount of the data, which minimizes the network communication. Thus, the PDP model for the remote data checking supports large data sets in widely-distributed storage systems.

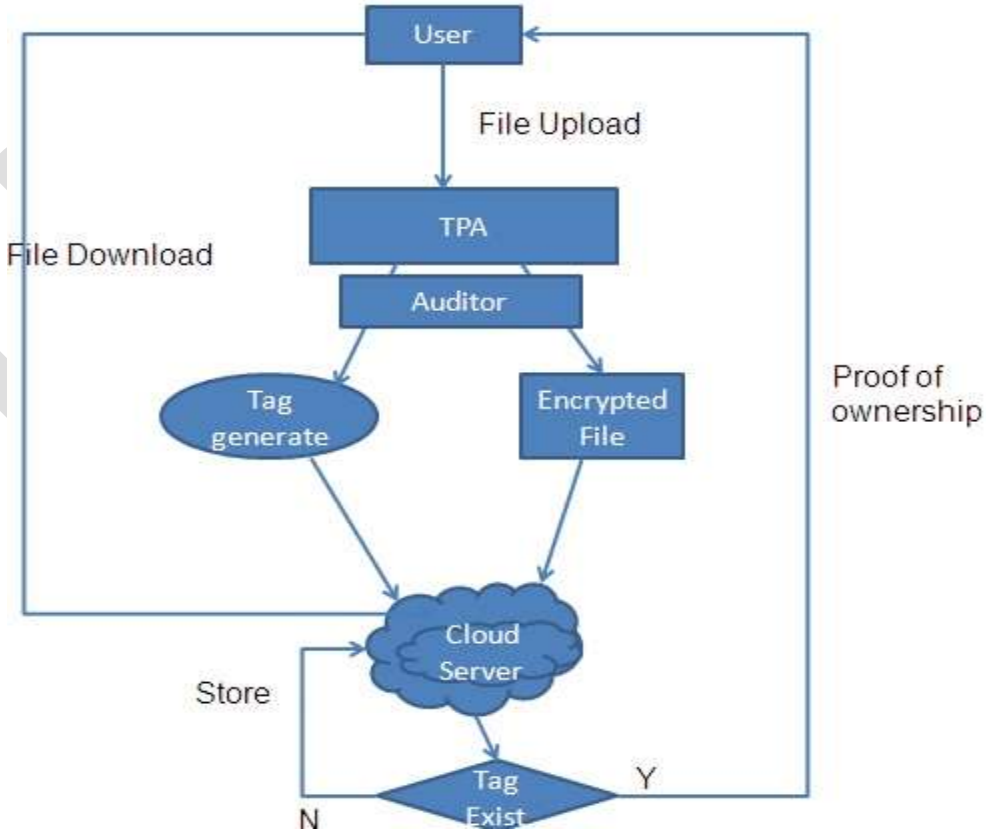
5) Remote Data Checking Using Provable Data Possession :

Introduce a model for provable data possession PDP and that can be used for remote data checking: A client that has stored the data at an untrusted server which can verify that the server possesses the original data without retrieving it. The model generates some probabilistic proofs of the possession by sampling the random sets of the blocks from the server, which drastically reduces the cost of I/O . The client maintains a constant amount of all metadata to verify proof. The challenge/response protocol transmits a small and constant amount of data and which minimizes the network communication. And thus, the PDP model for remote data checking is lightweight and supports large datasets in the distributed storage systems. The model is also robust in that it incorporates mechanisms for the mitigating arbitrary amounts of data corruption.

EXISTING SYSTEM:

We determine that SecCloud framework has accomplished both integrity auditing and file deduplication. Be that as it may, it can't keep the cloud servers from knowing the substance of files having been put away. In other words, the functionalities of integrity auditing and secure deduplication are just forced on plain files. In this area, we propose SecCloud+, which takes into account integrity auditing and deduplication on scrambled files. Framework Model Compared with SecCloud, our proposed SecCloud+ involves an extra trusted element, to be specific key server, which is in charge of assigning customers with mystery key (according to the file content) for encrypting files. This construction modeling is in line with the late work. However, our work is distinguished with the past work by allowing for integrity auditing on encoded data. SecCloud+ takes after the same three protocols (i.e., the file uploading protocol, the integrity auditing protocol and the proof of proprietorship protocol) as with SecCloud. The main distinction is the file uploading protocol in SecCloud+ involves an extra stage for correspondence between cloud customer and key server. That is, the customer needs to speak with the key server to get the merged key for encrypting the uploading file before the phase in SeeCloud.

SYSTEM ARCHITECTURE:



Achievement of integrity and avoiding duplication is the main task we are focusing, we are going enhance it by adding block level deduplication and providing reliability of data also. That is,

1. Improve Reliability:

In this we consider deduplication system improves storage utilization while reducing reliability hence we formalize the notion of distributed reliable deduplication system. A new distributed deduplication systems which are with the higher reliability in which the data chunks are distributed across multiple cloud servers such as block.

2. Block-level deduplication:

We consider block level deduplication in that file is divided into block and check deduplication for block.

For encryption we are going to use Asymmetric Encryption Key Algorithm RSA.

CONCLUSION

Aiming to achieving both data integrity and deduplication in cloud, we propose SecCloud and SecCloud+. SecCloud introduces an auditing substance with maintenance of a MapReduce cloud, which assists customers with generating data labels before uploading and additionally reviews the integrity of data having been put away in cloud. Furthermore, SecCoud empowers secure deduplication through introducing a Proof of Ownership protocol and preventing the leakage of the side channel information in the data deduplication. Contrasted and past work, the calculation by client in SecCloud is incredibly diminished during the file uploading and auditing stages. SecCloud+ is a propelled development persuaded by the way that clients constantly need to encode their data before uploading, and takes into account integrity auditing and secure deduplication straightforwardly on scrambled data.And we are providing block level deduplication and reliability as well.

REFERENCES:

1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," Communication of the ACM, vol. 53, no. 4, pp. 50–58, 2010.

[2] J. Yuan and S. Yu, "Secure and constant cost public cloud storage auditing with deduplication," in IEEE Conference on Communications and Network Security (CNS), 2013, pp. 145–153.

[3] 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.

[4] S. Keelveedhi, M. Bellare, and T. Ristenpart, "Dupless: Server- aided encryption for deduplicated storage," in Proceedings of the 22Nd USENIX Conference on Security, ser. SEC'13. Washington, D.C.: USENIX Association, 2013,pp.179194.[Online].Available:https://www.usenix.org/conference/usenixsecurity13/technicalsessions/presentation/bellare

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

[6] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-locked encryption and secure deduplication," in *Advances in Cryptology – EUROCRYPT 2013*, ser. Lecture Notes in Computer Science, T. Johansson and P. Nguyen, Eds. Springer Berlin Heidelberg, 2013, vol. 7881, pp.296–312.

[7] J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou, "Secure deduplication with efficient and reliable convergent key management," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 6, pp.1615–1625, June 2014.

[8] Jingwei Li, Jin Li, Dongqing Xie and Zhang Cai , "Secure Auditing and Deduplicating Data in Cloud" IEEE TRANSACTIONS

ON COMPUTERS VOL: PP NO: 99 YEAR 2015.

[9] Y. Zhu, H. Hu, G.-J. Ahn, and M. Yu, "Cooperative provable data possession for integrity verification in multicloud storage," IEEE
Transactions on Parallel and Distributed Systems, vol. 23, no. 12, pp. 2231–2244, 2012.

[10] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in *Computer Security – ESORICS 2009*, M. Backes and P. Ning, Eds., vol. 5789. Springer Berlin Heidelberg, 2009, pp. 355–370.

[11] J. Douceur, A. Adya, W. Bolosky, P. Simon, and M. Theimer, "Reclaiming space from duplicate files in a serverless distributed file system," in *22nd International Conference on Distributed Computing Systems*, 2002, pp. 617–624.

[12] Huaqun Wang "Proxy Provable Data Possession in Public Clouds" IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 6, NO. 4, OCTOBER-DECEMBER 2013, 551-559