IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727
PP 65-67
www.iosrjournals.org

Survey On Techniques for data integrity verification in cloud storage

Miss.Prachiti M.Karandikar[1], Dr. Pradeep K. Deshmukh[2]
[1] (Computer Department, RSCOE, Pune/ savitribai phule pune university, India)
[2] (Computer Department, RSCOE, Pune/ savitribai phule pune university, India)

Abstract: *Cloud computing is becoming very popular. Users are choosing cloud as repository for their data. The data in the cloud should be accessible, correct, consistent and high quality. While considering cloud as storage data security and integrity of stored data is burning issue. When users store their data on cloud there is a risk of modification or updation of data. Many researchers had worked and proposed algorithms to solve this issue. This survey paper focuses on two core techniques of proof of storage (POS) that are Proof of data Possession (PDP) and Proof of Retrievability (PoR). Both the techniques are used to ensure the cloud client about integrity of data storage on cloud.*

Keywords: *cloud computing, data security, data integrity, proof of data possession, proof of retrievability,*

I. Introduction

Cloud computing is an emerging technology which provides many services over internet. Many organizations are migrated towards cloud. As cloud is providing multiple characteristics such as on-demand service, location independency, resource pooling and so on. Instead of investing money in new hardware and software and also for the maintaince of resources, users can use servers, storage, applications that are available in cloud. Cloud computing provides you luxury of using all the computer hardware and software from anywhere and at anytime. These softwares and hardwares are not actually installed on your local machine. Few companies provide you services which allow accessing such hardwares/softwares over internet. Users are unaware about where these resources are actually located and how get managed. The information is transparent to end user.

Cloud provides ease of access to all resources needed by client. Storage is most important aspect of this era. Cloud storage represents a model in which service provider give a space in their large scale infrastructure to organizations as well as individuals on rent. It is extensional approach of traditional data centers. Cloud provides unlimited storage with reduced deployment cost. It has many advantages over local storage. Cloud server provides facility to store user's data on a cloud. So users can upload their data on cloud and can access it without any additional burden of time, location, and cost. Many cloud storage service providers such as Google, Microsoft, amazon EC2, and dropbox have attracted users to use cloud storage. Users are enjoying use of these services due to ease of access to their data which is hosted on another infrastructure.

Along with these advantages security and integrity of data stored on cloud storage is burning issue. Users need to verify that their data remain as they stored on cloud .Because data stored on cloud can easily be lost or corrupted due to human errors and hardware and software failures. Traditionally entire data was retrieved from cloud and cryptographic techniques, hash values are used for integrity verification. But this is the wastage of cost, computation of user and communication resources.

As cloud computing provides many things in terms of "*something* –as- a- service" e.g. Platform as a service (PAAS), Software as a service (SAAS), user have to registered himself to the cloud server or to the third party which provide the cloud service. So the privacy of the data and security need to be considered. In order to overcome the problem of integrity verification and security of cloud data many schemes are developed under different systems and models. Privacy of the user data and personal information can be provided by the cryptographic function and technology. Two basic approaches are developed by researchers to check integrity of outsourced data called provable data possession (PDP) and Proof of retrievability (PoR).

Provable data possession (PDP): PDP techniques are used by clients to check their data that is stored on cloud server. It ensures client that their data is untouched. Client maintains some constant amount of metadata to verify proof. It supports large data sets in widely distributed networks.
Proof of retrievability (PoR): In PoR system data storage center must have to give a proof to a data owner (client) that client's data is intact on storage. Also it allows client to recover his outsourced data. In PoR prover and verifier both doesn't needed to have knowledge of file F.

II. Related Work

In [1] authors defined a PDP model. It gives probabilistic proof that third party stored a file. User can access small blocks of file for generating the proof. . This technique uses challenge and response method. Client contains some constant amount of metadata of its data which is stored on cloud storage server. Metadata is locally stored which is further used to verify proof given by server .Client gives challenge of proving possession to server and wait for response. Response must contain a function of stored file computed by server to prove his innocence. To check correctness of response, metadata is used.RSA based Homomorphic variable tags are used as key component.PDP scheme aims to detect servers misbehavior when server has done something wrong with clients data. To achieve this PDP samples servers storage, accessing random sets of blocks. But PDP gives only probabilistic proof not a deterministic proof. It cannot support dynamic data possession.

In[2] a new cryptographic building block is proposed known as proof of retrievability(POR).It helps verifier(user) in determining that whether Prover(server) possesses his file or not. If POR is successfully executed, it gives assurity to verifier that he is retrieving file stored on cloud entirely as it is. Scheme uses disguised blocks (called sentinels) hidden among regular file blocks in order to detect data modification by the server. By specifying locations of collection of sentinels and asking to return associated value verifier challenges prover. Then values are compared to check integrity of data. In this approach verifier computes and stores single cryptographic key. Keyed hash algorithm is used to compute key. Client can check multiple times by storing different key values. Encoding of file before storage is done. Error correcting codes improve error resiliency of their system. But it increases computational overhead and larger storage requirement on prover.

In [3] authors proposed new technique to obtain PoR. Two schemes are implemented here. First scheme implements PoR with pubic verifiability. Shortest query response of any proof of retrievability is obtained which is secure in the random oracle model. Second scheme has shortest response of any PoR scheme with private retrivability and secure in the standard model. Two homomorphic authenticators are uses that are based on PRF’s and second based on BLS signature. Both schemes allow only one authentication value. In this technique, user breaks an erasure encoded file into n blocks. Each file block is accompanied by authenticators of equal length. Use of BLS signature instead of RSA reduces proof size. It also tolerate higher error rate. , This scheme still works on static data only, without support of dynamic data update.

[4] Is an extension of PDP model and supports provable updates on stored data. New version of authenticated dictionaries is used which is based on rank information. Rank information is used to organize dictionary entries. Authentication skip list is used to check the integrity of file blocks. It allows insertion and deletion of blocks within the data structure. File F and its skip list are stored on untrusted server. To prevent replay attacks root metadata is stored at client side. File f is divided into blocks. When client wants to verify integrity of block I he issues query atRank(i) to the server. Server then computes T(i) as its proof and send to client. Clients check integrity by comparing proof of server with stored metadata. Also to update the data client issue atRank(i)(for insertion)and atRank(i-1)(for deletion).Tags used here are more efficient than PDP for static data. It’s computational and communication complexity can be up to O (logt). A limitation of DPDP is that it does not allow for public variability of the stored data. In addition it does not consider data freshness or fairness.

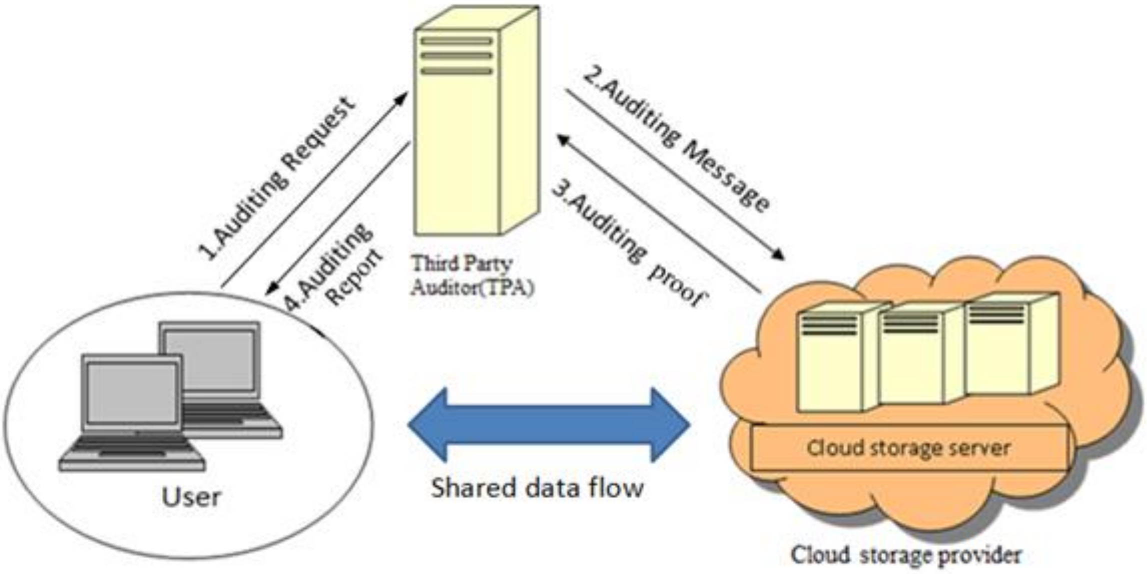
Scheme proposed in Paper [6]provides provable security and desirable efficiency simultaneously. Two servers are used. Particularly one for auditing and other for storage of data. Data integrity is ensured by using efficient verification scheme. Server for auditing is called as Third party Auditor(TPA).TPA monitors stored data in cloud storage as well as transactions between client and cloud storage server(CSS).Public verifiability is proposed here. Public verifiability means allowing cloud customers or TPA to verify integrity by challenging cloud server. In this paper this is done by TPA behalf of cloud customers (data owner).Computation is done by server instead of client while achieving blockless. This reduces computational overhead at client side. This is main advantage of this scheme over previously proposed scheme mentioned above as they have only one server. Security of this scheme is analyzed under variant of[2] which supports public verifiability. The game between challenger C(client) and adversary a(storage server) is played to get proof of retrievability from A. If proof is valid for fraction of challenges, ,F is extracted.

III. Table 1

Overview of the proof of storage(POS) schemes

POS scheme	Confidentiality	Integrity	Availability	Public verifiability	Type
PDP [1]	yes	Yes	Yes	yes	static
POR for large files [2]	yes	Yes	Yes	yes	static
Compact POR [3]	yes	Yes	Yes	yes	static
DPDP [4]	yes	Yes	Yes	no	dynamic
POR with public auditing [5]	yes	Yes	Yes	yes	dynamic

IV. Proposed Architecture



V. Conclusion

This survey gives overview about various integrity verification techniques which gives Proof of storage(POS) in cloud storage. PDP and POR schemes are surveyed. To overcome the limitations of previous work a new POR scheme with public verifiability is proposed. Here audit server(TPA) is trustworthy. TPA preprocesses and uploads data on behalf of clients. Also integrity verification and updation of data stored on cloud is done by TPA. Hence computational burden for tag generation at client side is reduced. Scheme supports public verifiability as well as dynamic data operation.

References

[1]. G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, “Provable data possession at untrusted stores,” in *CCS ’07: Proceedings of the 14th ACM conference on Computer and communications security*. New York, NY, USA: ACM, 2007, pp. 598–609.

[2]. A. Juels and B. S. K. Jr., “Pors: proofs of retrievability for large files,” in *CCS ’07: Proceedings of the 14th ACM conference on Computer and communications security*. New York, NY, USA:ACM, 2007, pp. 584–597.

[3]. H. Shacham and B. Waters, “Compact proofs of retrievability,”in *ASIACRYPT ’08: Proceedings of the 14th International Conference on the Theory and Application of Cryptology and Information Security*. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 90–107.

[4]. K. D. Bowers, A. Juels, and A. Oprea, “Proofs of retrievability: theory and implementation,” in *Proceedings of CCSW 2009*. ACM, 2009, pp. 43–54.

[5]. J.Li,X.Tan.X.Chen,and D.S.Wong,”An efficient proof of retrievability with public auditing in cloud computing,”in/NCoS,2013,pp,93-98