International Journal of Research
Available at https://edupediapublications.org/journal
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 12
August 2016

Data Possession Techniques in Cloud Storage for Deduplicating Data

NIKHAT SULTANA[1] & Dr.K.Ramakrishna[2]

[1]M-Tech Dept. of IT Sridevi Women's Engineering College Hyderabad

[2]Professor & HOD, Dept. of IT Sridevi Women's Engineering College Hyderabad

Abstract-Data integrity and storage efficiency are two important requirements for cloud storage. Proof of Retrievability (POR) and Proof of Data Possession (PDP) techniques assure data integrity for cloud storage. Proof of Ownership (POW) improves storage efficiency by securely removing unnecessarily duplicated data on the storage server. The cloud storage service (CSS) relieves the burden for storage management and maintenance. Fragment Structure, random sampling and index table is used to construct the Audit service. These techniques are supported provable updates to cloud outsourced data. The third party auditing allow to save time and computation resources with reduced online burden of the user. In this work, a method based on Probabilistic query and periodic verification for improving the performance of audit services and also audit system verifies the integrity.

Keywords: Cloud Storage Service, Cloud Service Provider, Third Party Audit, Public Verification Parameter, Provable Data Possession.

1. Introduction

Cloud computing consists a collection of computers and servers that are publicly accessible via the Internet. User access the data’s and will pay as per user basis. Cloud computing has four essential characteristics: elasticity and the ability to scale up and down, self-service provisioning and automatic deprovisioning, application programming interfaces (APIs), billing and metering of service usage in a pay-as-you-go model. While Cloud Computing makes these advantages more appealing than ever, it also brings new and challenging security threats towards users‟ outsourced data. Since cloud service providers (CSP) are separate administrative entities, data outsourcing is actually relinquishing user’s ultimate control over the fate of their data. The correctness of the data in the cloud is being put at risk due to the following reasons. First of all, although the infrastructures under the cloud are much more powerful and reliable than personal computing devices, they are still facing the broad range of both internal and external threats for data integrity and storage management. The cloud storage service (CSS) relieves the burden of storage management and maintenance. However, if such an important service is vulnerable to attacks or failures, it would bring irretrievable losses to users since their data or archives are stored into an uncertain storage pool outside the enterprises. The capability provided to the consumer is to use the provider’s applications running on a cloud infrastructure. The applications are accessible from different client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage data, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

2. Related Work

1. In Yan Zhu [1] used a quantified new audit approach based on probabilistic queries and periodic verification, as well as an optimization method of parameters of cloud audit services. This method greatly reduced the workload on the storage servers. Cloud computing is securely constructed with TPA based on Provable data possession technique. Cloud storage reduced the client's burden for storage management and maintenance. It provided a low cost, scalable, location-independent platform. Audit services avoided the security risks and to achieve digital forensics and credibility on cloud computing. This technique was a cryptographic based for verifying the integrity of data without retrieving it at a UN trusted server, can be used to realize audit services. PDP protocol is used to prevent the fraudulence of hackers and the leakage of verified data (zero-knowledge property).

2. Q. Wang [2] solved the problem of ensuring the integrity of data storage in Cloud Computing. The task is assigned to allow a third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the dynamic data stored in the cloud. The TPA eliminated the involvement of the client through the auditing of whether his/her data stored in the cloud is indeed intact, which can be important in achieving economies of scale for Cloud Computing While prior works on ensuring remote data integrity often lacked the support of either public audit ability or dynamic data operations, this project achieves both. First identify the difficulties and potential security problems of direct extensions with fully dynamic data updated from prior works and then displayed how to construct an elegant verification scheme for the seamless integration of these two salient features in our protocol design. In particular, to achieve efficient data dynamics, improve the existing proof of storage models by manipulating the classic Merle Hash Tree construction for block tag authentication. To support efficient handling of multiple auditing tasks, the technique of bilinear aggregate signature to extend main result into a multi-user setting, where TPA can be performed multiple auditing tasks simultaneously. This paper delivered highly protection for cloud storage.

3. In Cloud Storage, users can be remotely stored their data and enjoy the on-demand high quality applications and services from a shared pool of configurable computing resources, without the burden of local data storage and maintenance. However, the fact that users no longer had physical possession of the outsourced data makes the data integrity protection in Cloud Computing a

formidable task, especially for users with constrained computing resources. Moreover, users should be able to just use the cloud storage as if it is local, without worrying about the need to verify its integrity. Thus, enabling public audit ability for cloud storage was of critical importance so that users can resort to a third party auditor (TPA) to check the integrity of outsourced data and be worry-free. To securely introduce an effective TPA, the auditing process should bring in no new vulnerabilities towards user data privacy, and introduce no additional online burden to user. In this paper, they proposed a secure cloud storage system supporting privacy-preserving public auditing and extend their result to enable the TPA to perform audits for multiple users simultaneously and efficiently. Extensive security and performance analysis showed the proposed schemes were provably secure and highly efficient. Their experiment conducted on Amazon EC2 instance further demonstrated the fast performance of the design. Authors [3] completed their dynamic auditing system to be privacy preserving and it supported the batch auditing for multiple owners.

2.1 Existing System:

In this work, we study the problem of integrity auditing and secure reduplication on cloud data. 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 (or block) for him/her. This design fixes the issue of previous work that the computational load at user or auditor is too huge for tag generation. We specify that our proposed Becloud system has achieved both integrity auditing and file reduplication. However, it cannot prevent the cloud servers from knowing the content of files having been stored. In other words, the functionalities of integrity auditing and secure reduplication are only imposed on plain files. In this section, we propose Becloud+, which allows for integrity auditing and reduplication on encrypted files.

2.2 Proposed System:

In Proposed System we propose two secure systems, namely Becloud and SecCloud+.SecCloud introduces an auditing entity with maintenance of aMapReduce cloud, which helps clients generate data tags before uploading as well as audit the integrity of data having been stored in cloud. Compared with previous work, the computation by user in Becloud is greatly reduced during the file uploading and auditing phases. Becloud+ is designed motivated by the fact that customers always want to encrypt their data before uploading, and enables integrity auditing and secure reduplication unencrypted data.

3. Implementation

After analyzing these papers, in cloud storage system is constructed based on Provable Data Possession technique [6]. It is more efficient compare than other methods. To securely fulfill the two important requirements of cloud storage: data integrity and storage efficiency, a number of schemes have been proposed based on the concepts of POR, PDP. In this work, a constant cost scheme that achieves secure public data integrity auditing and storage reduplication at the same time. It enables the reduplication of both files and their corresponding authentication tags. In addition, to support batch integrity auditing, and thus substantially save computational cost and communication cost for multiple requests scenarios. Cloud service providers to offer an efficient audit service to check the integrity and availability of the stored data [7]

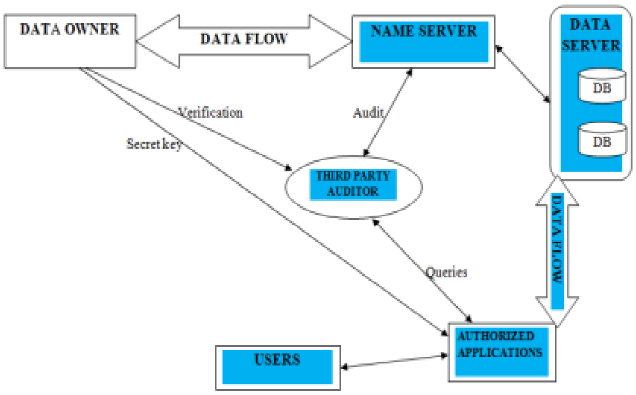


Fig 1.1System architecture of the proposed system

Cloud secure storage is constructed with these modules,

3.1 KEY GENERATION

The owner generates a public/secret key pair (pike, ski) by himself or the system manager, and then sends his public key pike to TPA. Note that TPA cannot obtain the client's secret key ski; secondly, the owner chooses the random secret.

3.2 TAG GENERATION

The client (data owner) uses the secret key ski to pre-process a file, which consists of a collection of n blocks, generates a set of public verification parameters and index-hash table that are stored in TPA, and transmits the file and some verification tags to CSP.

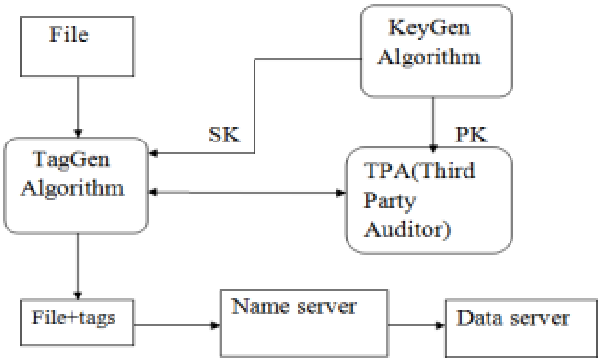


Fig 2 Tag generation.

3.3: PERIODIC SAMPLING AUDIT

TPA (or other applications) issues a "Random Sampling" challenge to audit the integrity and availability of outsourced data in terms of the verification information stored in TPA [9]. The AAs should be cloud application services inside clouds for various application purposes, but they must be specifically authorized by DOs for manipulating outsourced data. Since the acceptable operations require that the AAs must present authentication information for TPA, any unauthorized modifications for data will be detected in audit processes or verification processes.

Based on this kind of strong authorization-verification mechanism, we assume neither CSP is trusted to guarantee the security of stored data, nor a DO has the capability to collect the evidence of CSP‟s faults after errors have been found [14].

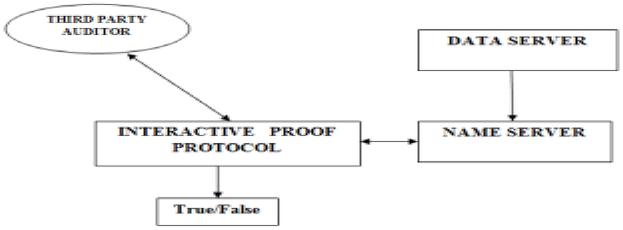


Fig 3 Periodic Sampling Audit Flow.

3.4 AUDITS FOR DYNAMIC OPERATIONS

An authorized application, which holds data owner’s secret key ski, can manipulate the outsourced data and update the associated index hash table stored in TPA[10][11]. The privacy of sk and the checking algorithm ensure that the storage server cannot cheat the authorized applications and forge the valid audit records.

Dynamic data operations:

Update (sk, Xi, mi’) is an algorithm run by AA to update the block of a file m0 i at the index i by using sk, and it returns a new verification metadata.

Delete (sk,Xi,mi) is an algorithm run by AA to delete the block mi of a file mi at the index i by using sk, and it returns a new verification metadata.

Insert(sk,Xi,mi) is an algorithm run by AA to insert the block of a file mi at the index i by using sk, and it returns a new verification metadata.

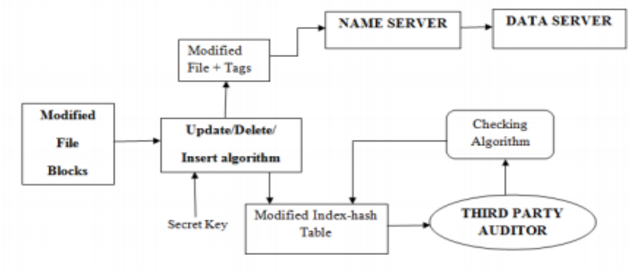
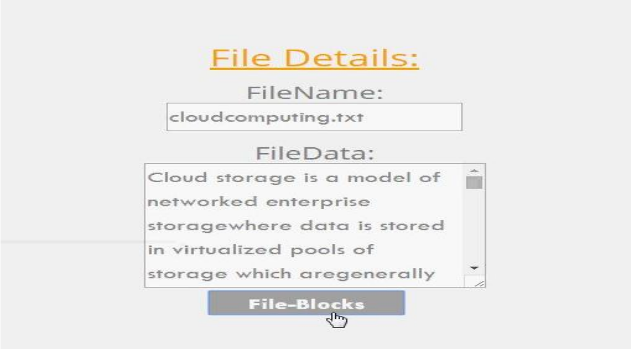


Fig 4 Flow of dynamic data operation.

To ensure the security, dynamic data operations are available only to DOs or AAs, who hold the secret key sk. Here, all operations are based on data blocks. It is necessary for TPA and CSP to check the validity of updated data. First, an AA obtains the public verification information from TPA[12][13]. Second, the application invokes the Update, Delete, and Insert algorithms, and then sends to TPA and CSP, respectively. Next, the CSP makes use of an algorithm Check to verify the validity of updated data. Note that the Check algorithm is important to ensure the effectiveness of the audit. Finally, TPA modifies audit records after the confirmation message from CSP is received and the completeness of records is checked.

4. Experimental Work

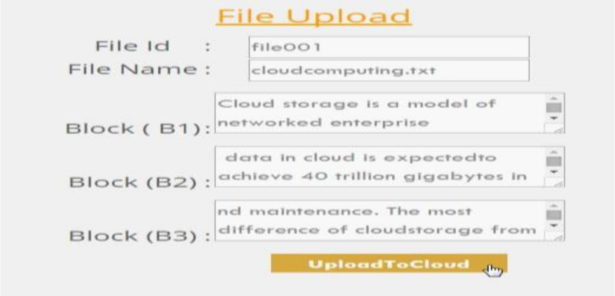


File Details:

FileName: cloudcomputing.txt

FileData: Cloud storage is a model of networked enterprise storagewhere data is stored in virtualized pools of storage which aregenerally

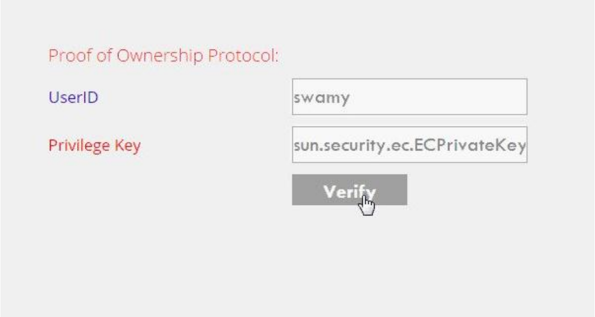
Fig 5: File Upload to Cloud.



File Upload

File Id :	file001
File Name :	cloudcomputing.txt
Block (B1):	Cloud storage is a model of networked enterprise
Block (B2) :	data in cloud is expectedto achieve 40 trillion gigabytes in
Block (B3) :	nd maintenance. The most difference of cloudstorage from

Fig 6: File Uploading with Blocks.



Proof of Ownership Protocol:

UserID	swamy
Privilege Key	sun.security.ec.ECPrivateKey

Fig 7: Proof of Ownership Protocol.

5. Conclusion

To securely fulfill the two important requirements of cloud storage: data integrity and storage efficiency, a number of schemes have been proposed based on the concepts of POR, PDP. In this work, a constant cost scheme that achieves secure public data integrity auditing and storage deduplication at the same time. It enables the reduplication of both files and their corresponding authentication tags. In addition, to support batch integrity auditing, and thus substantially save computational cost and communication cost for multiple requests scenarios. Cloud service providers to offer an efficient audit service to check the integrity and availability of the stored data.

6. 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", University of Arkansas at Little Rock, USA.

[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.