ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

Implementation of Authorized Duplication Checker in Hybrid Cloud

Priyanka K. Shinde, Prof. Avinash P. Wadhe

M.E. Final Year CSE, Dept. of CSE, GHRCEM, Amravati, India

Head of Department, Dept. of CSE, GHRCEM, Amravati, India

ABSTRACT: A hybrid cloud is a combination of public and private clouds together and a technology that enables data and application portability. Proposed system aiming to solving the problem of deduplication efficiently with differential privileges in cloud computing. A hybrid cloud architecture consisting of a public cloud and a private cloud and the data owners only use their data storage by utilizing public cloud while the data operation like storing and retrieving is managed in private cloud. To make data management accessible in cloud computing, deduplication has been a well technique recently is use. Deduplication reduces your bandwidth requirements, speeds up the data transfers, and it keeps your cloud storage needs to a minimum. Proposed system present some new deduplication constructions which supports authorized duplicate check in hybrid cloud architecture. To maintain the confidentiality of data the convergent encryption technique has been used to encrypt the data before outsourcing. Authorized deduplication system support differential authorization duplicate check. As a proof of concept, a prototype is implemented in authorized duplicate check scheme and performs test bed experiments using prototype, authorized duplicate check scheme obtains minimal overhead compared to normal operations.

KEYWORDS: Deduplication; Authorized duplicate check; hybrid cloud.

I. INTRODUCTION

Cloud computing enables new business models and cost effective resource usage. Instead of maintaining their own data center, companies can concentrate on their core business and purchase resources when it will needed. Especially when combining publicly accessible clouds with a privately maintained virtual infrastructure in a hybrid cloud, the hybrid cloud technology can open up new opportunities for businesses. As cloud computing becomes widespread, an increasing amount of data is being stored in the cloud and the data shared with specified privileges by different users, this process define the access rights of the stored data. One serious challenge of cloud storage services is the management of the ever-increasing volume of data on cloud. To make the data management scalable in cloud computing, deduplication [2] has been a well-known technique recently use. The technique is used to improve storage usage and can also be applied to network data transfers which will reduce the number of bytes. In place of keeping multiple data copies with same content, deduplication remove the unnecessary data by keeping only one physical copy and referring other redundant data to that copy. Data deduplication brings a lot of benefits, though security and privacy concerns arise as users sensitive data are exposed to both insider and outsider attacks. Convergent encryption [3] has been proposed to impose data confidentiality while making deduplication possible. It encrypts and decrypts a data copy by using a convergent key and which is gained by computing the cryptographic hash value of the content of the data copy. After the key generation and data encryption, users kept the keys and send ciphertext to the cloud. Since this encryption operation is deterministic and it is derived from the data content, the same convergent key is generated by identical data copies and hence has the same ciphertext. To prevent this from unauthorized access, a secure proof of ownership protocol [4] is also needed to provides the proof that the user undeniably owns the same file when the file duplicate is found and After the proof, consequent users with the same file provide a pointer from the server and need not to upload the same file. With the pointer from the server a user can able to download the encrypted file, which can be decrypted by the only corresponding data owners with their convergent keys. Thus, convergent encryption will enable the cloud to perform deduplication on the ciphertexts and the proof of ownership prevents the unauthorized user to access the file.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

However, the previous deduplication systems cannot support to differential authorization and duplicate check, which is very important in many applications. In an authorized deduplication system each user provide a set of privileges during system initialization. Each file which is uploaded to the cloud is bounded by a set of privileges which clears that which kind of users is permit to perform the duplicate check and access right of the files. The user have to take his file and his own privileges as inputs before submitting his duplicate check request for some file. The user is able to find a duplicate for this file if there is a copy of this file and a matched privilege stored in cloud.

II. PROPOSED SYSTEM

There are three entities defined in system, that is, users, private cloud and storage cloud service provider in public cloud as shown in Fig. 1. The Storage cloud carry out deduplication by checking if the contents of two files are the same and stores only one of them and From set of privileges the access right to a file is defined. Each privilege is represented in the form of a short message called token. Each file is bound with some file tokens, which denotes the tag with specified privileges. A user have to computes and sends tokens for duplicate-check to the public cloud for authorized duplicate check. While when Users access to the private cloud server, for the requesting users a semi trusted third party perform duplicable encryption by generating file tokens.

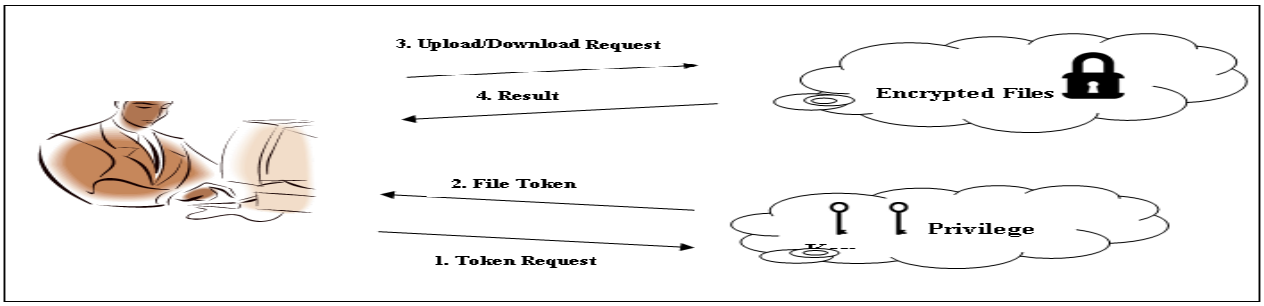

Fig. 1 Architecture for Authorized Deduplication

A. *Storage Cloud*

Storage Cloud provides a data storage service in public cloud. The storage cloud service provider provides the data outsourcing service and it will also stores data on behalf of the users. To reduce the storage cost, the storage cloud remove the storage of simillar data via deduplication and keeps only unique data.

B. *Data User*

Data User outsource data storage to the storage cloud and access the data later when it is needed. In a storage system supporting deduplication, We can save upload bandwidth by uploading unique data but does not upload any duplicate data, which cannot be owned by the same user or the different users. In authorized deduplication system, each user is gets a set of privileges and By convergent encryption key each file is protected and to realize the authorized deduplication with differential privileges use privilege keys.

C. Private Cloud

Private cloud managed the private keys for the privileges, Private Cloud answers the file token requests from the users and this interface allows user to submit files and queries to be securely stored and computed respectively.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

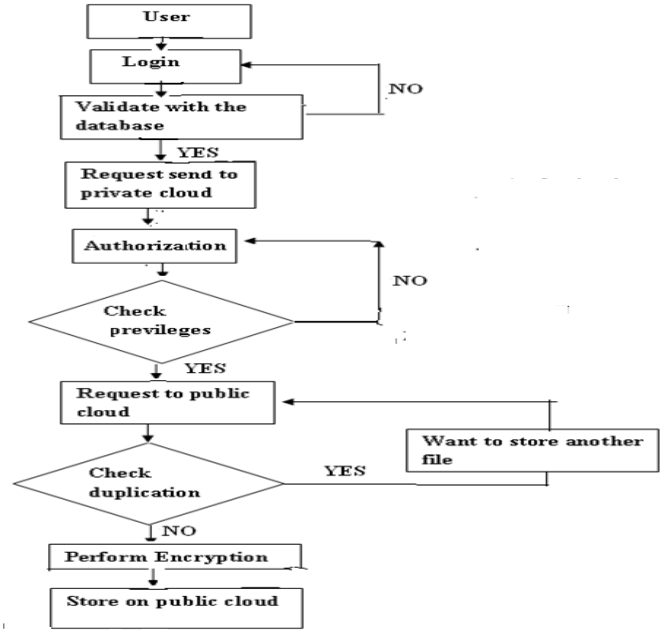

Fig. 2 Flow Diagram of Proposed Method

In deduplication system, a hybrid cloud architecture is introduced to solve the problem of unauthorized deduplication of file. The private keys will be kept and managed by the private cloud server for which will not be issued to users directly to get a file token the user needs to send a request to the private cloud server. User can not able to perform duplicate check for some file the user needs to get the file token from the private cloud server for this. The private cloud server also check the user is authorized or not before issuing the corresponding files token to the user. The user also perform the authorized duplicate check for this file with the public cloud before uploading this file. Based on the results of duplicate check the user either uploads this file or prove that file is owned by the user. If a file duplicate is found, the user needs to run the Proof of ownership protocol with the cloud storage service provider to prove the file ownership. Otherwise, if no duplicate is found then the data owner performs an identification with the private key to prove its identity. If it is passed, user can upload his files and the private cloud server will find the corresponding privileges of the user from its stored table list and send to the user. The same way user can download his file from storage cloud.

III. EXPRIMENTAL RESULT ANALYASIS

We implement the proposed authorized deduplication system, in which we model three entities as separate programs. Microsoft Azure cloud is used to store data. A Client program is used to model the data users to carry out the file upload process. A Private Server program is used to model the private cloud which manages the keys and handles the file token computation.

Evaluation of the proposed system focuses on comparing the overhead induced by authorization steps, including file token generation and the encryption against file upload steps. The upload process breakdown into three steps Token Generation, Duplicate Check and Encryption. For each step, we record the start and end time of it and therefore obtain the breakdown of the total time spent. We present the average time taken in each data set in the figures. To evaluate the effect of file size to the time spent on different steps, we upload unique files. The time spent on encryption linearly increase with the file size, In contrast, other steps such as token generation and duplicate check only use the file metadata for computation and therefore the time spent remains constant. With the file size increasing from 50 KB to 200KB, the overhead of the proposed authorization steps decreases from 1.75% to 1.43%.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

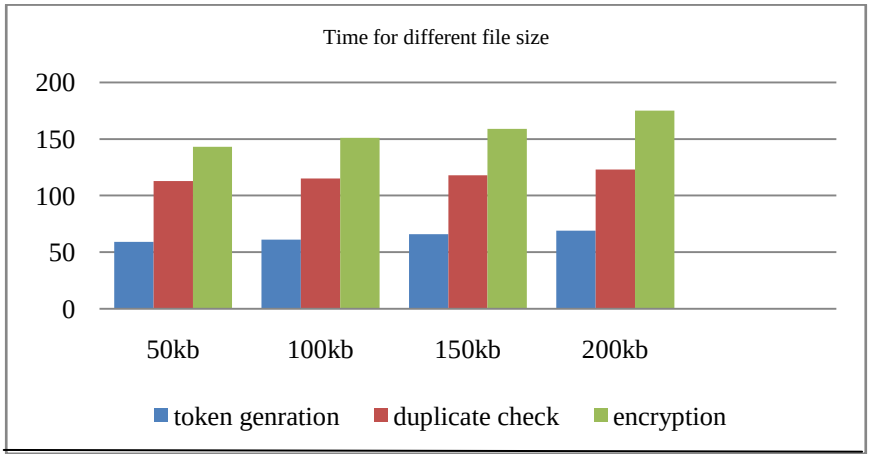

Fig. 3 Time for different file size

IV. CONCLUSION

Hybrid clouds offer a much flexibility to Cloud user while offering choice in terms of keeping control and security. Hybrid clouds are usually utilize by the organizations willing to carry of their workloads to public clouds either for its bursting purposes or for faster implementation projects Base upon company needs and structure of implementation hybrid clouds vary. In proposed system to protect the data security proposed authorized data deduplication which includes differential privileges of user in the duplicate check system. System presented several new deduplication constructions which supports authorized duplicate check in hybrid cloud architecture, the private cloud server generated the duplicate-check tokens of files with private keys. Proposed systems is more secure in terms of insider and outsider attacks which specified by the proposed security model. The proposed authorized duplicate check scheme occurs minimum overhead compared to traditional convergent encryption and network transfer.

REFERENCES

1. Jin Li, Yan Kit Li, Xiaofeng Chen, Patrick P. C. Lee, Wenjing Lou, 'A Hybrid cloud approach for secure authorised deduplication', (*IEEE Transactions on Parallel and Distributed System)s*, 2013.
2. Rahumed, H. C. H. Chen, Y. Tang, P. P. C. Lee, and J. C. S. Lui., 'A secure cloud backup system with assured deletion and version control',(*In 3rd International Workshop on Security in Cloud Computing), 2011.*
3. J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer., 'Reclaiming space from duplicate files in a serverless distributed file system' (*In ICDCS), 2002*.
4. S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg. , 'Proofs of ownership in remote storage systems' (*ACM), 2011*.
5. Elhadj Benkhelifa , Dayan Fernando., 'A Novel cloud hybrid access mechanism for highly sensitive data exchange', *(The Fourth International Conference on Cloud Computing), 2013.*
6. J. Stanek, A. Sorniotti, E. Androulaki, and L. Kencl., 'A secure data deduplication scheme for cloud storage', (*In Technical Report*) 2013.
7. Bellare, M., Keelveedhi, S., Ristenpart, T, 'Message-locked encryption and secure deduplication', (*In: Advances in Cryptology), 2013.*
8. Xu, J., Chang, E.C., Zhou, J., Weak leakage-resilient client-side deduplication of encrypted data in cloud storage, (*In: 8th ACM SIGSAC symposiu).*
9. Bellare, M., Keelveedhi, S., Ristenpart, T, 'DupLESS: server-aided encryption for deduplicated storage, *(In: 22nd USENIX conference on Security), 2013.*
10. Hongwei Li, Yuanshun Dai1,, Ling Tian, and HaomiaoYang,' Identity-Based Authentication for Cloud Computing*' (In Cloud Com*), 2009.
11. S. Bugiel, S. Nurnberger, A. Sadeghi, and T. Schneider, 'Twin clouds: An architecture for secure cloud computing, (*In Workshop on Cryptography and Security in Clouds), 2011*.
12. K. Zhang, X. Zhou, Y. Chen, X.Wang, and Y. Ruan, 'Sedic: privacyaware data intensive computing on hybrid clouds, (*In Proceedings of the 18th ACM conference on Computer and communications security, USA), 2011*.
13. M. Bellare, C. Namprempre, and G. Neven,' Security proofs for identity-based identification and signature schemes' , 2009.
14. R. D. Pietro and A. Sorniotti , 'Boosting efficiency and security in proof of ownership for deduplication', (*ACM),* 2012.
15. Nesrine Kaaniche, Maryline Laurent, 'A Secure Client Side Deduplication Scheme in Cloud Storage Environments',(*6th international conference on new technologies,mobility and security), 2014.*