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

Authorized Deduplication Check for Confidentiality in a Hybrid Cloud

H.Sheela

PG Scholar, Department of CSE, Tagore Institute of Engineering and Technology, Salem, Tamil Nadu, India

ABSTRACT: Cloud computing provides seemingly unlimited recourses to the users services across the whole internet. as cloud computing become prevalent ,an increasing amount of data is being stored in cloud and shared by users with specified privileges ,which access rights of the stored data. One critical challenge of cloud storage services is the management of every-increasing volume of data. To make data management scalable in cloud computing. Deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data in storage.

Data deduplication is the technique which eliminates the duplicate copies of the repeated data or files which is being stored in cloud, so as to reduce the memory and bandwidth which is being used. Another issue is that confidentiality of sensitive data when multiple users get benefited through single cloud, to provide security an encryption and decryption technique (AES) is being employed. In addition to that the users are s et with some privileges where they download their file safely.

Cloud provides token key to the users to view or download the file which is being stored in cloud. The files are stored in encrypted and decrypted format to provide security using advanced encryption technique. Alert should be shown if the file is already present in memory. The download of file should be done by the user only when the key token matches which is an advancement of traditional technique. Authorized duplication check scheme incurs minimal overhead compared to normal operations.

KEYWORDS: Deduplication; different privileges to users;confidentiality;reduced storage space.

I.INTRODUCTION

Cloud computing provides seemingly unlimited “virtualized” resources to users as services across the whole Internet, while hiding platform and implementation details. Today’s cloud service providers offer both highly available storage and massively parallel computing resources at relatively low costs. As cloud computing becomes prevalent, an increasing amount of data is being stored in the cloud and shared by users with specified privileges, which define the access rights of the stored data. One critical challenge of cloud storage services is the management of the ever-increasing volume of data. To make data management scalable in cloud computing, Deduplication has been a well-known technique and has attracted more and more attention recently. Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data in storage.

The technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. Instead of keeping multiple data copies with the same content, deduplication eliminates redundant data by keeping only one physical copy and referring other redundant data to that copy. Deduplication can take place at either the file level or the block level. For file level deduplication, it eliminates duplicate copies of the same file. Deduplication can also take place at the block level, which eliminates duplicate blocks of data that occur in non-identical files.

Although data deduplication brings a lot of benefits, security and privacy concerns arise as users’ sensitive data are susceptible to both insider and outsider attacks. Traditional encryption, while providing data confidentiality, is incompatible with data deduplication. Specifically, traditional encryption requires different users to encrypt their data with their own keys. Thus, identical data copies of different users will lead to different cipher texts, making deduplication impossible. Convergent encryption has been proposed to enforce data confidentiality while making deduplication feasible. Traditional deduplication systems based on convergent encryption, although providing confidentiality to some extent, do not support the duplicate check with differential privileges. In other words, no differential privileges have been considered in the deduplication based on convergent encryption technique. It seems to be contradicted if we want to realize both deduplication and differential authorization duplicate check at the same time.

DOI: 10.15680/IJIRCCE.2015. 0311207

II.RELATED WORK

The users are having authentication and security to access the detail which is presented in the system. Before accessing or searching the details user should have the account in that otherwise they should register first. To support authorized deduplication, the tag of a file F will be determined by the file F and the privilege. To show the difference with traditional notation oftag, we call it file token instead. To support authorized access, a secret key kp will be bounded with a privilege p to generate a file token. Let $\phi' F;p$ = TagGen(F, kp) denote the token of F that is only allowed to access by user with privilege p. In another word, the token $\phi' F;p$ could only be computed by the users with privilege p. As a result, if a file has been uploaded by a user with a duplicate token $\phi' F;p$, . Once the key request was received, the sender can send the key or he can decline it. With this key and request id which was generated at the time of sending key request the receiver can decrypt the message. Files are sending another client throws server. Upload files at that time generates an ciphertext key. In master key gives a file sender. This master key sending for server side they encrypted format. File Sending menu contains a key and upload files.

III.PROPOSED ALGORITHM

To enhance our system in security. Specifically, we present an advanced scheme to support stronger security by encrypting the file with differential privilege keys. In this way, the users without corresponding privileges cannot perform the duplicate check. Furthermore, such unauthorized users cannot decrypt the cipher text even collude with the S-CSP. Security analysis demonstrates that our system is secure in terms of the definitions specified in the proposed security model.

ADVANTAGES

- The user is only allowed to perform the duplicate check for files marked with the corresponding privileges.
- We present an advanced scheme to support stronger security by encrypting the file with differential privilege keys.
- Reduce the storage size of the tags for integrity check. To enhance the security of deduplication and protect the data confidentialitysystem such as portability, performance, security, etc.

IV.PSEUDO CODE

```
Step 1: Create the Main Menu.
Step 2:  Create separate user and admin login.
Step 3:  Set privileges for both user and admin.
Step 4:  Create the database to store the file and set access.
Step 5: Set encryption technique to store and decryption to retrieve the same.
Step 6:  Download can be done once the key is obtained
          else
          Only view option is available
Step 7: End.
```

V.SIMULATION RESULTS

The basic Structure for our proposed system as follows, The main menu which leads to User account and the administrator account.The file which is being uploaded is encrypt and get decrypted accordingly.Once the secret key is obtained one can view the file as well as to download.

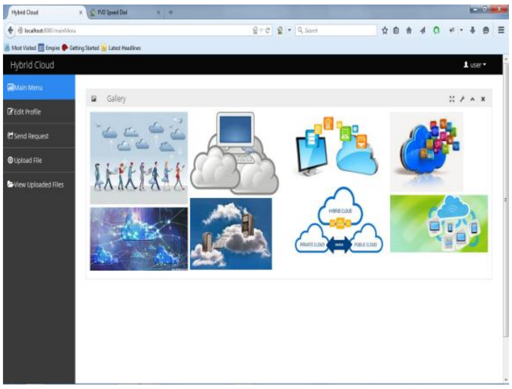
DOI: 10.15680/IJIRCCE.2015. 0311207

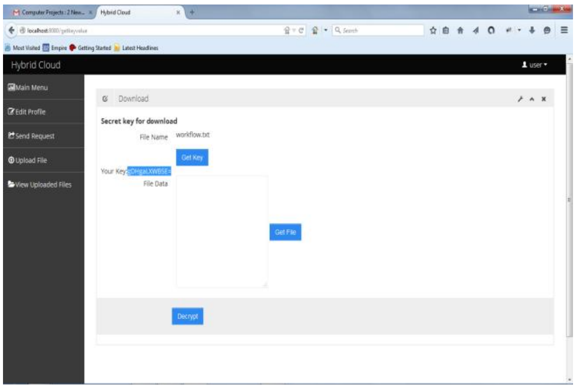
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015





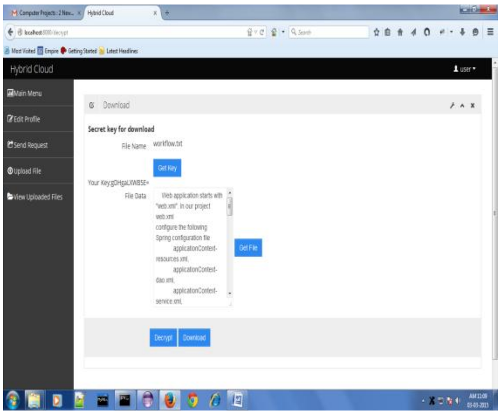
Secret key for download

File Name workflow.txt

Your Key:gDHgaLXWB5E=

File Data

1.1 User Main Menu **1.2 get secret key**

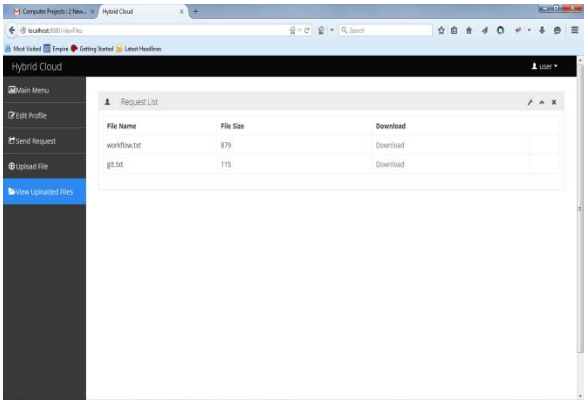


Secret key for download

File Name workflow.txt

Your Key:gDHgaLXWB5E=

File Data Web application starts with "web.xml". In our project web.xml configure the following Spring configuration file applicationContext-resources.xml, applicationContext-dao.xml, applicationContext-service.xml,



File Name	File Size	Download
workflow.txt	879	Download
git.txt	115	Download

1.3 **Decrypt file** **1.4 view uploaded file**

VI. Conclusion and Future Work

The notion of authorized data deduplication was proposed to protect the data security by including differential privileges of users in the duplicate check. We also presented several new deduplication constructions supporting authorized duplicate check in hybrid cloud architecture, in which the duplicate-check tokens of files are generated by the private cloud server with private keys. Security analysis demonstrates that our schemes are secure in terms of insider and outsider attacks specified in the proposed security model. As a proof of concept, we implemented a prototype of our proposed authorized duplicate check scheme and conduct testbed experiments on our prototype. We showed that our authorized duplicate check scheme incurs minimal overhead compared to convergent encryption and network transfer.

Though the above solution supports the differential privilege duplicate, it is inherently subject to bruteforce attacks launched by the public cloud server, which can recover files falling into a known set. More specifically, knowing that the target file space underlying a given ciphertext C is drawn from a message space $S = fF1, \_ \_ \_ , Fng$ of size n, the public cloud server can recover F after at most n off-line encryptions. That is, for each $i = 1, \_ \_ \_ , n$, it simply encrypts Fi to get a ciphertext denoted by Ci. If $C = Ci$, it means that the underlying file is Fi. Security is thus only possible when such a message is unpredictable. This traditional convergent encryption will be insecure for predictable file. We design and implement a new system which could protect the security for predicatable message. The

DOI: 10.15680/IJIRCCE.2015. 0311207

main idea of our technique is that the novel encryption key generation algorithm. For simplicity, we will use the hash functions to define the tag generation functions and convergent keys in this section. In traditional convergent encryption, to support duplicate check, the key is derived from the file F by using some cryptographic hash function $kF = H(F)$. To avoid the deterministic key generation, the encryption key kF for file F in our system will be generated with the aid of the private key cloud server with privilege key kp. The encryption key can be viewed as the form of $kF;p = H0(H(F), kp) \bigoplus H2(F)$, where $H0,H$ and $H2$ are all cryptographic hash functions. The file F is encrypted with another key k, while k will be encrypted with $kF;p$. In this way, both the private cloud server and S-CSP cannot decrypt the ciphertext. Furthermore, it is semantically secure to the S-CSP based on the security of symmetric encryption. For S-CSP, if the file is unpredicatable, then it is semantically secure too.

REFERENCES

[1] OpenSSL Project. http://www.openssl.org/.
[2] P. Anderson and L. Zhang. Fast and secure laptop backups with encrypted de-duplication. In *Proc. of USENIX LISA*, 2010.
[3] M. Bellare, S. Keelveedhi, and T. Ristenpart. Dupless: Serveraided encryption for deduplicated storage. In *USENIX Security Symposium*,
[4] M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication. In *EUROCRYPT*, pages 296–312, 2013.
[5] M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. *J. Cryptology*, 22(1):1–61, 2009.
[6] M. Bellare and A. Palacio. Gq and schnorr identification schemes: Proofs of security against impersonation under active and concurrent attacks. In *CRYPTO*, pages 162–177, 2002.
[7] 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 (WCSC 2011)*, 2011.
[8] 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*, pages 617–624, 2002.
[9] D. Ferraiolo and R. Kuhn. Role-based access controls. In *15th NIST-NCSC National Computer Security Conf.*, 1992.
[10] GNU Libmicrohttpd. http://www.gnu.org/software/libmicrohttpd/.
[11] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg. Proofs of ownership in remote storage systems. In Y. Chen, G. Danezis, and V. Shmatikov, editors, *ACM Conference on Computer and Communications Security*, pages 491–500. ACM, 2011.
[12] J. Li, X. Chen, M. Li, J. Li, P. Lee, andW. Lou. Secure deduplication with efficient and reliable convergent key management. In *IEEE Transactions on Parallel and Distributed Systems*, 2013.
[13] libcurl. http://curl.haxx.se/libcurl/.
[14] C. Ng and P. Lee. Revdedup: A reverse deduplication storage system optimized for reads to latest backups. In *Proc. of APSYS*,

BIOGRAPHY

H.Sheela, PG Scholar, Department of CSE, Tagore Institute of Engineering and Technology ,Salem, Tamil Nadu,India.Did her UG at Sona College of Technology, Salem, TamilNadu, India. Her area of interest are cloud computing, Android technology, Database.

DOI: 10.15680/IJIRCCE.2015. 0311207