Proficient Tape Asset Administration Utilizing Deduplication in Cloud Reinforcement and Recorded Administrations

K. Nagajyothi [1], K. Malathi [2]
[UG Student, Department Of Computer Science and Engineering[1]],
[Assistant Professor, Department Of Computer Science and Engineering[2]],
Saveetha School of Engineering, Saveetha University, Chennai.

Abstract

Data deduplication is one of important data compression techniques for eliminating duplicate copies of repeating data, and has been widely used in cloud storage to reduce the amount of storage space and save bandwidth. To protect the confidentiality of sensitive data while supporting deduplication, the convergent encryption technique has been proposed to encrypt the data before outsourcing. To better protect data security, this paper makes the first attempt to formally address the problem of authorized data deduplication. Different from traditional deduplication systems, the differential privileges of users are further considered in duplicate check besides the data itself. Security analysis demonstrates that our scheme is secure in terms of the definitions specified in the proposed security model. Our main trick is to use the interactive protocol based on static or dynamic decision trees. The advantage gained from it is, by interacting with clients, the server will reduce the time complexity of deduplication equality test from linear time to efficient logarithmic time over the whole data items in the database. A proof of concept, we implement a prototype of our proposed authorized duplicate check scheme and conduct tested experiments using our prototype. We show that our proposed authorized duplicate check scheme incurs minimal overhead compared to normal operations.

Keywords*: Convergent encryption, Deduplication, Confidentiality, Interactive protocol.*

I. INTRODUCTION

Remote distributed storage has turned into a key part for different applications in these days arrange which store a lot of information and give the halfway information required. With the quick developing of distributed storage administrations, for example, distributed storage, encryption turns into a critical procedure for ensuring the classification of information. Despite the fact that information encryption gives an essential certification to the security and protection of customers' information, it restricts the behavior of the openness and accessibility of the scrambled information. The Data deduplication empowers information stockpiling frameworks to discover and expel duplication inside information without trading off its accessibility. The objective of information deduplication is to store more information in less space by putting away and looking after documents (hinders in fine-grained deduplication way) into a solitary duplicate, where the excess duplicates of information are supplanted by a reference

To this duplicate. It implies that information deduplication stockpiling framework could decrease the capacity size of u customers, who share similar information duplicate m, from O (u _ jmj) to O (u + jmj) if some usage subordinate constants are covered up. Additionally, customers don't have to transfer their information to the distributed storage server, when there has been one duplicate put away, which won't just extraordinarily lessen the correspondence cost of customers and cloud server, yet in addition spare the system data transfer capacity. Information deduplication frameworks, the private cloud is included as an intermediary to enable information proprietor clients to safely perform copy check with differential benefits. Such engineering is down to earth and has pulled in much consideration from scientists. The information proprietors just outsource their information stockpiling by using open cloud while the information operation is overseen in private cloud. Customary encryption, while giving information secrecy is incongruent with information deduplication. Indistinguishable information duplicates of various clients will prompt distinctive figure writings, making deduplication unthinkable. In this paper, we upgrade our framework in security. In particular, we show a propelled plan to help more grounded security by encoding the record with differential benefit keys. Thusly, the clients without relating benefits can't play out the copy check. All the more, such unapproved clients can't decode the figure message even conspire with the S%&S'. Security investigation exhibits that our framework is secure as far as the definitions determined in the proposed security show. The client is just permitted to play out the copy check for records set apart with the relating benefits. We introduce a propelled plan to help more grounded security by scrambling the record with differential benefit keys. Diminish the capacity size of the labels for respectability check. To improve the

security of deduplication and ensure the information privacy.

II. EXISTING SYSTEM

Data deduplication systems, the private cloud is involved as a proxy to allow data owner users to securely perform duplicate check with differential privileges. Such architecture is practical and has attracted much attention from researchers. The data owners only outsource their data storage by utilizing public cloud while the data operation is managed in private cloud. Cross-client data deduplication has been widely used to eliminate redundant storage overhead in cloud storage system. Recently, Abadi et al. introduced the primitive of MLE2 with nice security properties for secure and efficient data deduplication. However, besides the computationally expensive non-interactive zero-knowledge proofs (NIZKs), their fully randomized scheme (R-MLE2) requires the inefficient equality-testing algorithm to identify all duplicate cipher texts. The main trick is to use the interactive protocol based on static or dynamic decision trees. The advantage gained from it is, by interacting with clients, the server will reduce the time complexity of deduplication equality test from linear time to efficient logarithmic time over the whole data items in the database. The security analysis and performance evaluation show that our schemes are Path-PRV-CDA2 secure and achieve several orders of magnitude higher performance for data equality test than R-MLE2 scheme when the number of data items are relatively large.

III. PROPOSED SYSTEM

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

IV. SYSTEM ARCHITECTURE

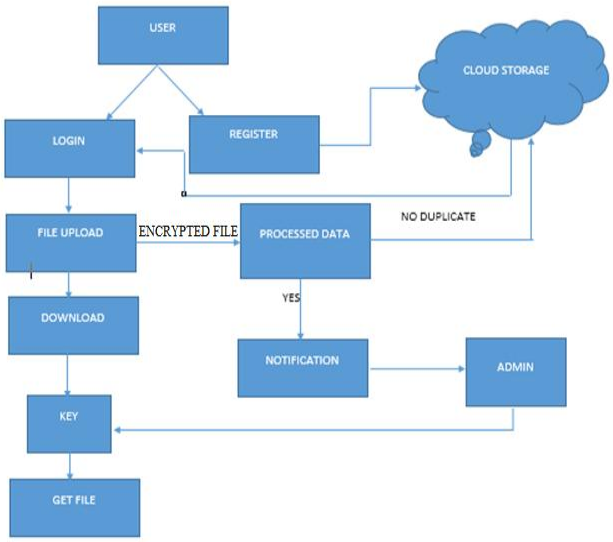

Fig.:1 Cloud Deduplication

V. IMPLEMENTATION

A. User Registration And Login

In this Module the new user needs to enter the required data to create a new account in the form by providing the user details like name, password, E-mail id etc. and the data will be stored in server for future authentication purpose. After registration user will get the username and password for further process. Using Username and Password, user makes a login.

B. File Uploading Process

In file upload process, user chooses the file from the system and makes a specific keyword for the each file uploading into the cloud. In storage area the data in the file are encrypted once it is been uploaded into the cloud, the data user encrypts its data using DES algorithm in order to ensure the security and privacy of data, and stores the encrypted data at database. We implement DES algorithm which converts a file in to a binary format and it gets encrypted and it is stored on to the cloud. The data that is stored on to the cloud will be in encrypted format, which can be accessed by the authorized users.

C. C. Identification Of Deduplicated Files

The file is uploaded into the cloud using a keyword and those files are generated at the backend and stored in the server using its specific key. This avoids the duplication of files into the cloud. If the file is already exist in the cloud the user cannot upload the file, due to duplication of the file in to the cloud, which is automatically intimated to the admin through the notification. Keywords identifies the duplication of files into the cloud, thus keyword

storage for every file would avoid the duplicated files.

D. User File Downloading Process

A user who has registered earlier in the cloud, can download the file by getting the specific key from the admin once if the valid key appears the user can access the requested file from the cloud are decrypted using DES algorithm and turns to original format. Then the file will be downloaded in the user's location. Invalid users cannot download the file without particular authentication to process the file. Only the original users can process the files needed using keyword.

VI. RESULTS

Fig.: 2 User Registration Login

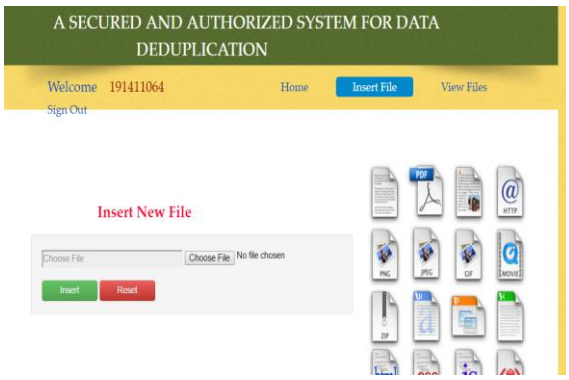

Fig.:3 File Uploading Process

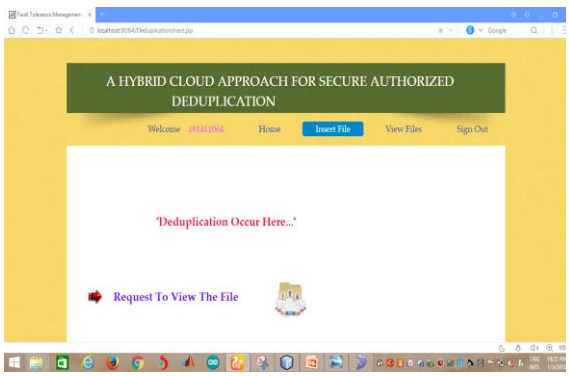

Fig.: 4 Identification of Deduplicated Files

VII. CONCLUSION

In this paper we have proposed convergent encryption technique for user data deduplication. Each data owner can verify the correctness of the convergent key. Also, the session key and encrypted data blocks can be dynamically refreshed when other data owner joins or data block changes. The advantage gained from it is, by interacting with clients, the server will reduce the time complexity of deduplication equality test from linear time to efficient logarithmic time over the whole data items in the database. A proof of concept, we implement a prototype of our proposed authorized duplicate check scheme and conduct tested experiments using our prototype. We show that our proposed authorized duplicate check scheme incurs minimal overhead compared to normal operations.

REFERENCES

[1] T. Jiang, X. Chen, Q. Wu, J. Ma, W. Susilo, and W. Lou. "Towards efficient fully randomized message-locked encryption," in Information Security and Privacy - 21st Australasian Conference, ACISP 2016

[2] C. Batten, K. Barr, A. Saraf, and S. Trepetin"Pstore: A secure peer-to-peer backup system," MIT Laboratory for Computer Science, progress report, 2001.

[3] M. Storer, K. Greenan, D. Long, and E. Miller"Secure data deduplication," in Proc. of the 4th ACM International Workshop on Storage Security and Survivability, VA, USA, Oct. 2008, pp. 1–10.

[4] L. Marques and C. Costa"Secure deduplication on mobile devices," in Proc. of the 2011 Workshop on Open Source and Design of Communication, Lisboa, Portugal, Jul. 2011, pp. 19–26.

[5] D. X. Song, D.Wagner, and A. Perrig "Practical techniques for searches on encrypted data," in Proc. of IEEE Symposium on Security and Privacy, CA, USA, May 2000, pp. 44–55.

[6] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. of the ACM Conference on Computer and Communications Security, VA, USA, Oct. 2006, pp. 79–88.

[7] D. Cash, S. Jarecki, C. Jutla, H. Krawczyk, M. Rosu, and M. Steiner"Highly-scalable searchable symmetric encryption with support for boolean queries," in CRYPTO 2013, ser. Computer Science, R. Canetti and J. A. Garay, Eds. Springer, 2013, vol. 8042 of LNCS, pp. 353–373.

[8] S. Kamara, C. Papamanthou, and T. Roeder "Dynamic searchable symmetric encryption," in Proc. of the ACM Conference on Computer and Communications Security, NC, USA, Oct. 2012, pp. 965–976.

[9] S. Kamara and C. Papamanthou "Parallel and dynamic searchable symmetric encryption," in Proc. of Financial Cryptography, Okinawa, Japan, Apr. 2013, pp. 258–274.

[10] H. Hacigumus, B. Iyer, C. Li, and S. Mehrotra "Executing sql over encrypted data in the database-service-provider model," in Proc. of ACM SIGMOD, Madison, Wisconsin, Jun. 2002, pp. 216–227 .

[11] S. Kamara, C. Papamanthou, and T. Roeder, "Dynamic searchable symmetric encryption," in Proc. of the ACM Conference on Computer and Communications Security, NC, USA, Oct. 2012, pp. 965–976.

[12] S. Kamara and C. Papamanthou, "Parallel and dynamic searchable symmetric encryption," in Proc. of Financial Cryptography, Okinawa, Japan, Apr. 2013, pp. 258–274.

[13] M. Naveed, M. Prabhakaran, and C. Gunter, "Dynamic searchable encryption via blind storage," in Proc. of IEEE Symposium on Security and Privacy, CA, USA, May 2014, pp. 639–654.

[14] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data," in Proc. of ACM SIGMOD, Paris, France, Jun. 2004, pp. 563–574.
[15] H. Hacigumus, B. Iyer, C. Li, and S. Mehrotra, "Executing sql over encrypted data in the database-service-provider model," in Proc. of ACM SIGMOD, Madison, Wisconsin, Jun. 2002, pp. 216–227.
[16] H. Kadhem, T. Amagasa, and H. Kitagawa, "A secure and efficient order preserving encryption scheme for relational databases," in Proc. of the International Conference on Knowledge Management and Information Sharing, Valencia, Spain, Oct. 2010, pp. 25–35.
[17] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "Cryptdb: Protecting confidentiality with encrypted query processing," in Proc. of ACM Symposium on Operating Systems Principles, Cascais, Portugal, Oct. 2011, pp. 85–100.
[18] R. A. Popa, F. Li, and N. Zeldovich, "An ideal-security protocol for order-preserving encoding," in Proc. of IEEE Symposium on Security and Privacy, CA, USA, May 2013, pp. 463–477.
[19] X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, "New algorithms for secure outsourcing of modular exponentiations," IEEE Transactions on Parallel and Distributed Systems, vol. 25(9), pp. 2386–2396, Jul. 2014.
[20] X. Chen, J. Li, J. Weng, J. Ma, and W. Lou, "Verifiable computation over large database with incremental updates," in ESORICS 2014, ser. Computer Science. Springer-Verlag, 2014, vol. 8712 of LNCS, pp.148–162.