International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No.03, April 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

Deduplication on Encrypted Big Data in Cloud

Prof.Prerna Rawat [1], Nilambari Wagh[2], Reshma Gaikwad[3] ,Sadanand Dani[4], Ashok Garje[5]

Department of Computer, Genba Sopanrao Moze College of Engineering, Balewadi ,Pune,(India)

ABSTRACT

Cloud Computing Describes a type of outsourcing of computer services with the continues and exponential increase of the number of users and the size of their data, data Deduplication becomes more and more of necessity for cloud storage providers. The advantages of Deduplication unfortunately come with high cost in terms of new security and privacy challenges.

By storing a unique copy of duplicated data cloud providers greately reduce their storage and data transfer cost. In order to preserve the privacy of data holders, data are store in encrypted form however, encrypted data introduce new challenges for cloud data Deduplication , which becomes crucial for big data storage and processing in cloud.

Traditional Deduplication schemes can not work on encrypted data. Existing solutions of encrypted data Deduplication suffer from security weakeness. They can not flexibly support data access control and revocation.

***Keywords* : Access Control, Big Data, Cloud Computing, Cryptanalysis, Deduplication.**

I. INTRODUCTION

Cloud Computing offers new way of information technology services by rearranging various resources(ex. Storage, computing and providing). Deduplication specialized data compression technique for eliminating duplicate copies of repeating data. The most important and popular cloud service is data storage . Cloud users uploads personal data to the center of CSP and if to maintain this data. But some intrusions and attacks towords this sensitive data at CSP are not available. It leads to high security risks especially data privacy leakages.

At the rapid increase of data mining and other analysis , the privacy issue become serious to outsource encrypted data to the cloud in order to ensure data security and user privacy. But the same or different may upload duplicated data in encrypted form to CSP, especially for scenarios where data shared among many users. Although cloud storage space is huge, data duplication greately wastage network resources , concludes a lot of energy complicates data management. Here we are proposing of scheme to manage encrypted data storage with Deduplication. It is an effective approach to verify data ownership and check duplicate storage with secure challenge and big data support . The result shows us efficiency , effectiveness.

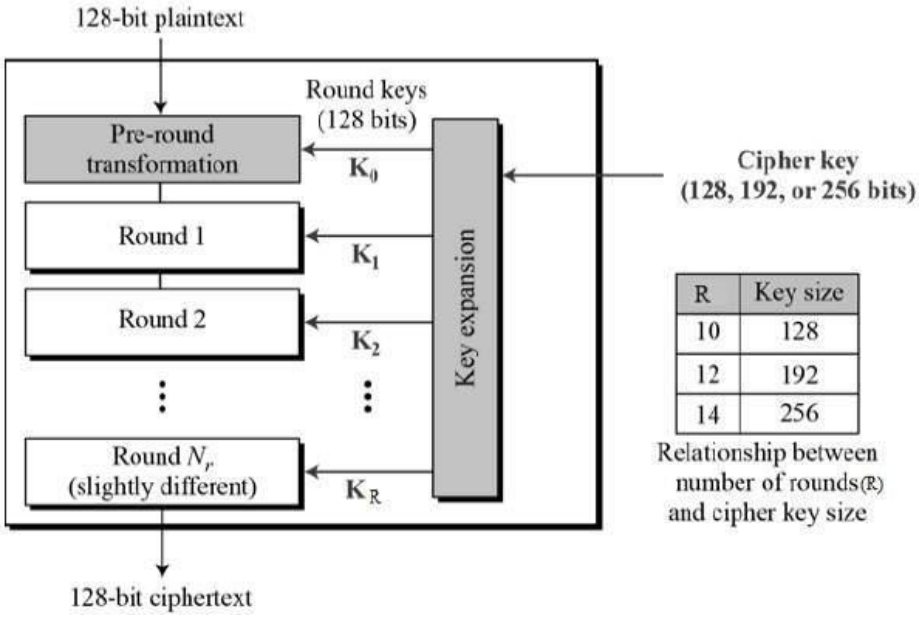
Cryptography: In this encryption, the process of encoding message of information in such a way that only authorized parties can access it.

II. RELATED WORK

Sr No	Paper Title	Work Description	Output
1	A Secure Client Side Deduplication Scheme in Cloud Storage Environments	Merkle-based Tree over encrypted data, in order to derive a unique identifier of outsourced data	Sensitive data leakage is a critical challenge.
2	Deduplication Techniques in Storage System	Supporting the deduplication, to encrypt the databefore outsourcing convergent encryption technique has been proposed	There is a need of data management as back up windows are shrinking due to growth of information.
3	Secure Large File Deduplication Technique Over Distributed Cloud Environment To Store Anonymous User Data	Provides data security using data encryption in cloud environment. For effective usage of storage space we provide de-duplication check at file level as well as block level	Deduplication systems are implemented based on the policies such as, file level, block level deduplications and client-server side de-duplication technique
4	Block-level Deduplication	Cloud Deduplication strengthens convergent	Do not ensure

	with Encrypted Data	encryption by employing a component that implements an additional encryption operation and an access control mechanism.	protection of predictable files against dictionary attacks
5	ClouDedup: Secure Deduplication with Encrypted Data for Cloud Storage	Prevents curious cloud storage providers from inferring the originalcontent of stored data by observing access patterns or accessing metadata	Existing public key encryption methods allow a party to encrypt data to a particular user, but are unable to efficiently handle more expressive types of encrypted access control

.



R	Key size
10	128
12	192
14	256

Relationship between number of rounds(R) and cipher key size

Fig:- AES ALGORITHM

III. SYMMETRIC ENCRYPTION

Encrypt DEK, M - The Encrypt algorithm takes as input data M, the symmetric key DEK. It encrypts M with DEK and outputs the ciphertext CT . This process is conducted at user u to protect its data stored at CSP with DEK.

Decrypt DEK,CT - The Decrypt algorithm takes as input the encrypted data CT , the symmetric key DEK. The algorithm decrypts CT with DEK and outputs the plain data M. A user (data holder) conducts this process to gain the plaintext of stored data at CSP.

IV. SYSTEM ARCHITECTURE

V. CONCLUSION

For a successful cloud storage service, managing encrypted data with Deduplication is important as well as significant. Our scheme can flexibly support data update and storing with Deduplication based on ownership challenge. We propose to manage practical scheme for managing encrypted data in cloud with Deduplication based on ownership and challenge and PRE.

It also support flexibly support data update and sharing with Deduplication even when the data holders are offline. Encrypted data can be securely accessed because only authorized users can obtain the symmetric keys used for data encryption.

VI. ACKNOWLEDGEMENTS

This work is sponsored by the National Key Foundational Research and Development on Network and Space Security, China (grant 2016YFB0800704), the NSFC (grant U1536202), the 111 project (grant B08038), the PhD grant of the Chinese Educational Ministry (grant JY0300130104), the Project Sup-ported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2016ZDJC-06), and Aalto University.

REFERENCES

[1] Zheng Yan; Wenxiu Ding; XiXun Yu; Haiqi Zhu; Robert H.Deng, "Deduplication on Encrypted Big Data in Cloud", in proc. IEEE Transaction on Big Data, 2016.

[2] M. Bellare, S. Keelveedhi, and T. Ristenpart, "DupLESS: Server aided encryption for deduplicated storage," in Proc. 22nd USENIX Conf. Secur., 2013, pp. 179–194.

[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 Proc. IEEE Int. Conf. Distrib. Comput. Syst., 2002, pp. 617–624, doi:10.1109/ICDCS.2002.1022312.

[4] G. Wallace, et al., "Characteristics of backup workloads in produc-tion systems," in Proc. USENIX Conf. File Storage Technol., 2012, pp.1–16.

[5] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distrib-uted storage," ACM Trans. Inform. Syst. Secur., vol. 9, no. 1, pp. 1– 30, 2006, doi:10.1145/1127345.1127346.

[6] D. T. Meyer and W. J Bolosky, "A study of practical deduplication," ACM Trans. Storage, vol. 7, no. 4, pp. 1–20, 2012, doi:10.1145/2078861.2078864.

[7] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-locked encryption and secure deduplication," in Proc. Cryptology—EURO-CRYPT, 2013, pp. 296–312, doi:10.1007/978-3-642-38348-9_18.