ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

Deduplication on Encrypted Big data in Cloud

Sumedh Deshpande, Anupama Murkute, Saurabh Patil, SachinKamble

Prof. Swati Shekapure

Dept. of Computer Engineering, Marathwada Mitra Mandals College of Engineering, Pune, India

ABSTRACT: In existing system data duplication is sufferingfrom security issues and even the security issues. So, to solve these issues with are taking help of hash code generation and double encryption techniques. This helps us to Integrateswith the cloud data deduplication with control access. Because of the authorized data holders who obtain the symmetric the encrypted data can also be securely accessed
Keys used for decryption of data. The results show the superior efficiency andeffectiveness of the scheme for big data deduplication in cloud storage.

KEYWORDS: Cloud Computing, Data Encryption Security and Protection, Access control

I. INTRODUCTION

The important and popular cloud service is data storage service. Cloud users upload personal or confidential content data to a cloud service provider (CSP) data center and allow to keep this data.In addition, the loss of control over Personal data leads to high data security risks, especially data privacy leaks. Due to the rapid Development of data mining and other analysis technologies, the question of privacy becomes serious. Therefore, a good practice is to encrypted data in order to guarantee the security of the data and the privacy of the user.

II. LITERATURE SURVEY

Blowfish algorithm is use to provide encryption to the files Blowfish is a 64-bit block cipher with a variable-length key [1]. External adversaries which aim to extract secrete informationAs much as possible from public and private cloud [2] Now, by taking advantage of the intra-user deduplication performed between clients and their DP, from a client point of view, the visibility of what is stored in the storage system is limited to his/her own account[3]. It is customary to define security in context like ours by using model involving game between two parties[4]. The advantage of deduplicationunfortunately comes with a high cost in terms of new security and privacy challenges. Cloud deduplication is secure and efficient storage service which is as secure as block level deduplication[5]

III. EXISTING SYSTEM

In the existing system Cloud computing is a service which provide user the rearranging data on demands. The cloud computing is internet based computing. It provides shared file and computer resources. The big network resources provided by the cloud computing. The secure and confidentiality is important on cloud computing. Therefore authorize cant able to find the duplication of data and only gives effective result. The client transfers their information to the server form of cloud service provider is not so secure and effective. In existing system there are also many drawbacks which leads to the failure of the system, such as Overwriting of files of same name, space efficiency due to storing multiple same file with different names. In this the comparison of files was lacking and so existing system drawbacks needs to overcome in proposed system.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

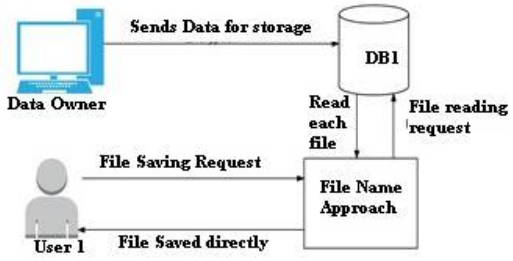

Fig 1. Existing System Architecture

IV. PROPOSEDSYSTEM

ARCHITECTURE

In recent time, there are many problems of storage places in cloud. If data holderstore file in cloudwhich is already available in cloud, so this is waste of memory.
So, in this paper we remove the deduplication. Also, security issues are occurduring the cloud server data storage. We are solving these problems with the help ofdouble encryption technique and hash code techniques for generate hash value

Data Holders

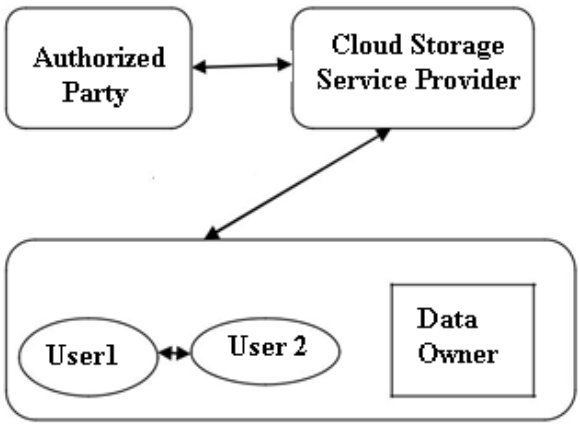

Fig 2: System architecture

1.1 CSP (Cloud Service Provider): The CSP (Cloud Service Provider) is allowing to data owner for storageservices of data. It cannot be fully trusted. Therefore it is curious about the contents of stored data. It should perform honestly on data storage in order to gain profit.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

1.2 Data Holder: The data holder can uploads and savestheir data and files in the CSP. In this system is possible to number of data holders could save their files in encrypted raw data in the CSP. The file is regarded as data owner by the data holder that produces or creates theFile. The data holder is in normal form than the higher priority of owner.

1.3 AP (Authorized Party): An authorized party (AP) that is fully trusted by thedata holders. The data holders to verify data ownership and handle data deduplication. It does not collude with CSP. In this case, CSP should not know the plain user data in its storage. AP cannot know the data stored in CSP.

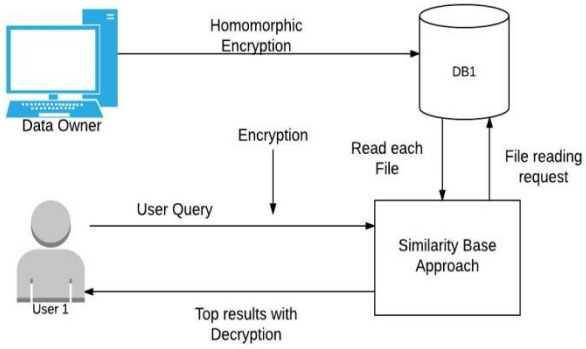

Fig 3. Proposed System Architecture

V. EXPECTED RESULT

For the system performance evaluation, calculate the matrices for accuracy. The system is implemented on java 3-tier MVC architecture framework with INTEL 3.0 GHz i5 processor and 8 GB RAM with public cloud Amazon EC2 consol. System also evaluated the computation.

VI. CONCLUSION

Discuss about to managing encrypted data with deduplication is for achieving a successful cloud storage service, for big data storage. In this paper, proposed a scheme to manage the encrypted big data in cloud with deduplication based on PRE and ownership challenge. This scheme can flexibly support data update and sharing with deduplication even when the data holders are offline. Only authorized data holders can use the symmetric keys used for data decryption by encrypted data can be securely accessed. The performance analysis and graphs showed that there scheme is secure under the described security model for big data deduplication. The results of the computer simulations further showed in the variation of graph. Future work is to include optimizing there design and implementation for big data storage in cloud. To ensure that CSP behaves as expected in deduplication management Studying verifiable computation. The part of cloud computing has brought many researchers from different fields; yet, much effort remains to reach use and the broad acceptance of cloud computing technology. In the future scope propose a big data deduplication. For the future environment system can focus on personalize search on user feedback sessions as well as recommendation based on user point of interest with database security is the interesting part of system

REFERENCES

[1] MsNehaKhatri – Valmik1, Prof. V. K Kshirsagar2 proposed paper on Blowfish algorithm.
[2] Chi Chen at. Al. proposed An Efficient Privacy-Preserving Ranked Keyword Search Method IEEE 2016.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 3, March 2018

[3] Chunhua Su at. al. proposed Analysis and Improvement of Privacy-Preserving Frequent Item Protocol for Accountable Computation Framework IEEE 2012.
[4] Cheng Huang and Rongxing Lu proposed EFPA: Efficient and Flexible Privacy-Preserving Mining of Association Rule in Cloud in IEEE 2015.
[5] Bharath K. Samanthula at. Al. k-Nearest Neighbor Classification over Semantically Secure Encrypted Relational Data MAY 2015.
[6] Lichun Li at. al. Privacy-Preserving-Outsourced Association Rule Mining on Vertically Partitioned Databases in AUGUST 2016.
[7] M. Bellare, S. Keelveedhi, and T. Ristenpart,
[8] "DupLESS: Serveraided encryption for deduplicated storage," in Proc. 22nd USENIXConf. Secur., 2013, pp. 179–194.