JES
Journal of Engineering Sciences

Vol 13, Issue 05, MAY/2022

ISSN NO:0377-9254

AVOIDANCE OF REDUNDANT DATA IN CLOUD WITH SHA-512 SHELTERED APPROACH

1. MOHAMMED SAMEER,2.MOHD ABDUL FAIZAN,
3.MOHAMMAD ALTAF HUSSAIN, 4.Dr. MOHAMMED ABDUL BARI

1. B.E STUDENT,Dept. of Computer Science Engineering,ISL Engineering College,
2. B.E STUDENT,Dept. of Computer Science Engineering,ISL Engineering College,
3. B.E STUDENT,Dept. of Computer Science Engineering,ISL Engineering College,
4. HOD & Associate Professor, Dept. of Computer Science Engineering,
ISL Engineering College

Abstract:

Distributed computing gives another method for administration by offering different assets over Internet. One of the essential administrations given by cloud benefit is information stockpiling. So as to save the protection of clients, these information are put away in cloud in a scrambled shape. Deduplication ends up critical and a testing errand when the information is put away in scrambled shape, which additionally prompts intricacy in putting away extensive information and handling in cloud. A customary deduplication strategy does not take a shot at scrambled information. Existing arrangements accessible for deduplicating encoded information has different security issues. They doesn't give get to control and repudiation as far as capacity. Thus, the deduplication plans are not for the most part conveyed practically speaking. In this paper, we propose a system to deduplicate encoded information put away in cloud dependent on access control along these lines evading repetitive capacity. It coordinates cloud information deduplication with access control. The aftereffect of our plan demonstrates prevalent proficiency and has potential for functional arrangement on account of enormous information stockpiling.

1.INTRODUCTION

Distributed computing gives different administrations by revamping the assets over the Internet. The essential cloud benefit is information storage.In request to safeguard the security of these information, they are regularly put away in an encoded frame. Scrambled information make new difficulties for cloud deduplication which Becomes significant for enormous information stockpiling and handling in cloud. A customary deduplication plot does not take a shot at scrambled information. Along these lines in this undertaking we acquaint a plan with deduplicate encoded information in could dependent on proprietorship to deduplicate numerous duplicates of same information. We plan to unravel the issues in deduplication that are being looked by information holders by giving security to getting to the record. The outcomes indicate unrivaled productivity

and viability of the plan for functional arrangement in cloud. The commitments of this paper can be condensed as tails We propose strategies to spare distributed storage without uncovering the protection of information holders by giving a plan to deduplicate and oversee encoded information. The plan oversees information deduplication with information sharing even without the information holder while saving their security. We join cloud information deduplication with information get to control essentially.

POST-PROCESS DEDUPLICATION

With post-process deduplication, new information is first put away on the capacity gadget and after that a procedure at a later time will break down the information searching for duplication. The advantage is that there is no compelling reason to sit tight for the hash counts and query to be finished before putting away the information in this manner guaranteeing that store execution isn't debased. Usage offering approach based task can enable clients to concede enhancement on "dynamic" records, or to process documents dependent on sort and area. One potential disadvantage is that you may pointlessly store copy information for a brief timeframe which is an issue if the capacity framework is close full limit.

IN-LINE DEDUPLICATION

This is where the deduplication hash counts are made on the objective gadget as the information enters the gadget progressively. In the event that the gadget recognizes a square that it previously put away on the framework it doesn't store the new square, just references to the current square. The advantage of in-line deduplication over postprocess deduplication is that it requires less capacity as information isn't copied. On the negative side, it is as often as possible contended that since hash estimations and queries takes such a long time, it can imply that the information ingestion can be slower subsequently diminishing the reinforcement throughput of the gadget. In any case, certain sellers with in-line deduplication have shown hardware with comparable execution to their post-process deduplication partners. Post-process and in-line deduplication techniques are frequently vigorously discussed.

II. PROPOSED SCHEME

In this paper, we propose an enhanced technique of deduplicating the various kinds of data that can be stored in

cloud. Our proposed scheme consists of the following steps

1) Generate hash for given data
2) Check if file exists, Encrypt hash value with random key.
3) Rehashing and encrypting the hash value for more security.
4) Double encryption and decryption is added.
5) Provide access without uploading.
6) Otherwise, Encrypt and store the given data with hash as key. Store the encrypted hash with individual keys.
7) On deletion, Revoke access by removing the individual key

8) If a file is not used by anyone, then remove the file

Using the SHA-512 algorithm we generate the hash value for the given data. The hash value is encrypted using AES algorithm and the encrypted hash value in stored in the cloud storage.

Whenever a next user upload the same file in to the cloud, first the hash value is encrypted and compared the available lists in the cloud.

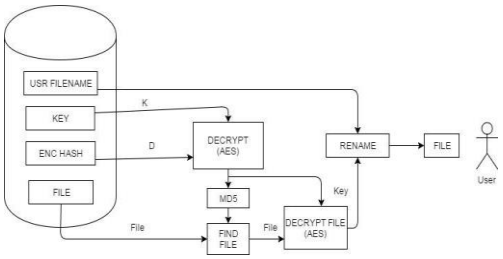

Fig 1: Architecture diagram

III.IMPLEMENTATION

Cloud: The cloud comprises of an open cloud which is responsible for information stockpiling and we can see all information in cloud.

Client: At the client side, User can transfer a document as encode record every client can download the record utilizing decode the record with the hash esteem . Each client can transfer a **SHA-512**

Algorithm:

SHA was published in the year of 1993 by NIST. Then it was reviewed in the year 1995 in order to remove some of its weaknesses. And as a result, a new hashing algorithm has been obtained called SHA-2, which uses a larger message digest. This message-digest were made more resistant to feasible attacks and they were allowed to use the blocks with larger data sizes. This is applicable to SHA

If the value matches, a key is provided to access the available same data in cloud. Otherwise the new data is uploaded. Separate key is provided to multiple users to access the file. Whenever a user removes the file from cloud, his/her access is removed by removing the individual key provided to that user. By this way we can prevent deduplication of data to a large extent with access control document is as of now transferred by somebody, the client get the mapping document id.

512. The SHA 512 consists of 8 words of 64 bits each. The standard hashing algorithm will run individual round for 80 times and the generates 512 bits output. These 512 bits will act as input to the following message block. After processing 1024-bit message blocks the last 512 bits of the message digest will be the message

Fig1: Interface.

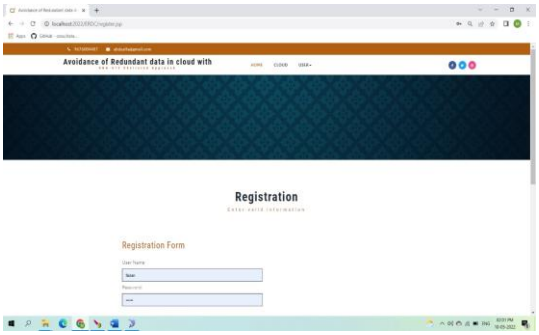

Fig2: Registration Interface.

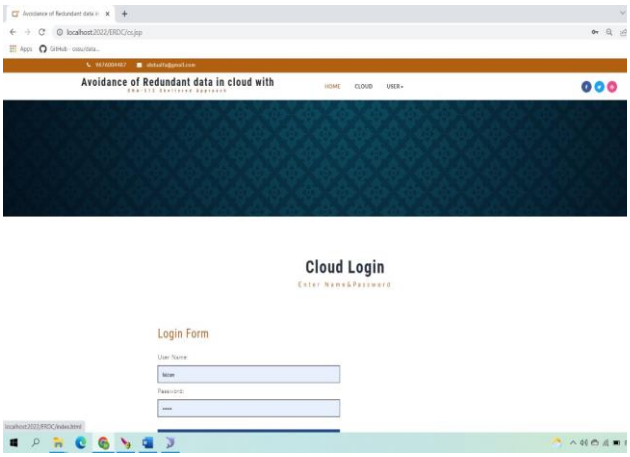

Fig3 : Cloud Interface.

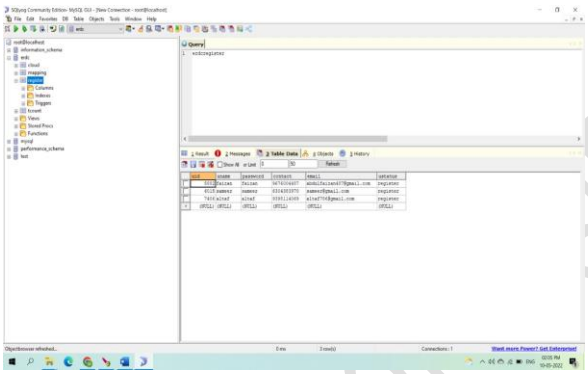

Fig 4: Database.

IV. CONCLUSION

Overseen scrambled information with deduplication is imperative and huge practically speaking for accomplishing a fruitful distributed storage benefit, particularly for enormous information stockpiling. In this paper, we proposed a plan to deal with the encoded documents in a cloud with deduplication dependent on proprietorship. Our plan can adaptably bolster information refresh and offering to deduplication. Encoded information can be safely gotten to just by approved information holders can acquire the symmetric keys utilized for information unscrambling.

REFERENCES

1. A secure data deduplication framework for cloud environments, authors: Fatema Rashid, Ali Miri, Isaac Woungang.
2. Attribute-Based Storage Supporting Secure Deduplication of Encrypted Data in Cloud, authors: Hui Cui, Robert H. Deng, Yingjiu Li.
3. Achieving lightweight, time-specific and secure access control in cloud storage, authors: Yanchao Wang,Fenghua Li, Ben Niu.
4. Design and implementation of various file deduplication schemes on storage devices, authors: Yong-Ting Wu , MinChieh Yu , Jenq-Shiou Leu , Eau-Chung Lee,TianSong.
5. T. T. Wu, W. C. Dou, C. H. Hu, and J. J. Chen, “Service mining for trusted service composition in cross-cloud environment,” IEEE Systems Syst. J., vol. PP, no. 99, pp. 1–12, 2014, doi:10.1109/ JSYST.2014.2361841.
6. Liu, C. Yang, X. Y. Zhang, and J. J. Chen, “External integrity verification for outsourced big data in cloud and iot: A big picture,” Future Generation Comput. Syst., vol. 49, pp. 58–67, 2015.
7. W. Tsai, C. F. Lai, H. C. Chao, and A. V. Vasilakos, “Big data analytics:

Vol 13, Issue 05, MAY/2022
ISSN NO:0377-9254

A survey," J. Big Data, vol. 2, no. 1, pp. 1–32, 2015, doi:10.1186/s40537-015-0030-3.

8. L. F. Wei, et al., "Security and privacy for storage and computation in cloud computing," Inf. Sci., vol. 258, pp. 371– 386, 2014, doi:10.1016/j.ins.2013.04.028.
9. "Deduplication on Encrypted Big Data in Cloud" by Zheng Yan, Senior Member, IEEE, Wenxiu Ding, Xixun Yu, Haiqi Zhu, and Robert H. Deng, Fellow, IEEE .
10. M. Ali, S. U. Khan, and A. V. Vasilakos, "Security in cloud computing: Opportunities and challenges," Inf. Sci., vol. 3doi:10.1016/j.ins.2015.01.025.