IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 7, July 2023

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2023 ||

| DOI:10.15680/IJIRSET.2023.1207186 |

Efficient Re-encryption and Secure Cloud Data Deduplication

Shreyas GN[1], Rajeshwari N[2]

Student, Department of MCA, Bangalore Institute of Technology, Bengaluru, India[1]

Assistant Professor, Department of MCA, Bangalore Institute of Technology, Bengaluru, India[2]

ABSTRACT: To deal with the growing development of data, commercial cloud storage providers have widely adopted the technique of data deduplication, which is both helpful and essential. To further guarantee the security of users' sensitive data in the outsourced storage mode, a number of safe data deduplication algorithms have been developed and are currently being deployed in a variety of scenarios. Dynamic ownership management has been made easier by a number of methods that have been developed, and various researchers have been interested in secure and effective re-encryption for encrypted data deduplication systems. In this study, we concentrate on the re-encryption deduplication storage system and demonstrate that a stub-reserved attack may be used to compromise the recently announced lightweight rekeying-aware encrypted deduplication technique (REED). We also provide an effective and secure data deduplication method.

In light of Our method can survive the stub-reserved attack and preserve the privacy of data owners' sensitive information because of the one-way hash function's intrinsic strength. Additionally, data owners only need to use the CAONT to re-encrypt a small section of the package rather than the entire package, considerably reducing the processing cost of the system. Finally, security research and experimental results show that our re-encryption strategy is effective and secure.

KEYWORDS: User joining, User revocation, Data deduplication, and Re-encryption

I. INTRODUCTION

A growing number of people and companies are outsourcing sensitive data to remote cloud service providers on a pay-per-use basis as a result of the rapid rise of cloud storage. According to a study funded by Dell EMC, the size of the digital world doubles every two years, and in 2020, the total amount of data in the universe is predicted to be 44 zettabytes (ZB), or 44 trillion gigabytes (GB), or more than 5200 GB per person. On the other hand, the growing data volume puts a heavy burden on cloud service providers. One easy way to handle it is to mandate that cloud service providers boost storage capacity on a regular basis in order to meet customers' expectations for high-quality storage services.

The phrase "cloud technology" is used often in several setting types. Data and related files may be accessed over the cloud at any time and from any place using any internet-connected device. It is crucial to give users access to user-related data even when they are not in the office as the working environment moves to more flexible and remote working. The increased data security offered by cloud computing is another benefit. Call it an attack with a stub.

Additionally, we offer a safe data deduplication technique based on the convergent all-or-nothing transform (CAONT) and randomly chosen bits from the Bloom filter, together with efficient re-encryption.

II. LITERATURE SURVEY

Rekeying is the process of creating a new encryption key to replace an old one. In order to avoid key compromise and to offer dynamic access control in cryptographic storage, it renews security protection. Nevertheless, it is difficult to achieve efficient rekeying in encrypted deduplication storage systems that use deterministic content-derived encryption keys to enable deduplication on ciphertexts.

We created and deployed REED, a rekeying-aware encrypted deduplication storage system. REED is based on an all-or-nothing transform (AONT) variation that is deterministic, enabling safe and light rekeying while maintaining deduplication capabilities. We recommend two REED encryption algorithms that compromise between speed and

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2023 ||

| DOI:10.15680/IJIRSET.2023.1207186 |

security, together with a REED expansion for dynamic access control. We implement a REED prototype using a variety of speed-optimization techniques. Our sandbox depends on traces.

To reduce storage space and upload traffic, a method known as data deduplication[2] for deleting duplicate copies of data has gained popularity in cloud storage. Even though a file is owned by many different persons, there is only one copy of it in the cloud. The deduplication technique increases storage while reducing reliability as a result. Furthermore, the privacy concern surfaces when users upload sensitive data to the cloud.

Our work is the first to codify the distributed reliable deduplication system concept in order to overcome the aforementioned security concerns. We recommend brand-new distributed deduplication methods that are more trustworthy and distribute data chunks over a number of cloud servers.

III. PROBLEM PREPOSITION

With the increasing adoption of cloud storage services, data deduplication has become an essential technique for optimizing storage efficiency and reducing redundancy. However, traditional data deduplication methods raise security concerns, as they may expose sensitive information to potential attacks. Additionally, as cloud environments often involve multiple users with varying access rights, ensuring data confidentiality and privacy becomes challenging.

The main challenge addressed in this research is to design and implement an efficient re-encryption and secure data deduplication system for cloud storage. This system should provide a high level of data deduplication while preserving data privacy and confidentiality. The proposed solution must guarantee that even the cloud service provider cannot access the original data or deduplicated content, thereby protecting user data from unauthorized access.

IV. PROPOSED METHODOLOGY

In this work, we show how to design reliable safe deduplication systems for cloud computing. To increase fault tolerance, we combine distributed cloud storage servers with deduplication technologies. Data confidentiality is further ensured by the secret sharing method, which is also compatible with distributed storage systems. To put it more precisely, a file is split into pieces utilizing the secret sharing method rather than encryption mechanisms. These shares will be distributed across several storage servers

.

To enable deduplication, a short cryptographic hash value of the content will also be created and sent to each storage server as the fingerprint of the fragment kept at each server. just the data owner who submits the data first.
Our suggested architectures allow for both block-level and file-level deduplication.

The suggested deduplication systems are secure according to the definitions given in the proposed security model, according to security analysis. We can accomplish secrecy, dependability, and integrity in our suggested system, in greater detail. Our methods take into account two different types of collusion assaults. The collusion attack against servers and the collusion attack against data are these. In particular, even if the enemy only controls a few storage servers, the data is still safe.

V. MODEL

We create two entities in this initial module: User and Secure-Cloud Service Provider.

User: A user is an entity that wishes to outsource data storage to the S-CSP and afterward access the data. To conserve upload bandwidth in a storage system that supports deduplication, the user only uploads unique data and does not upload any duplicate data. Furthermore, users in the system demand fault tolerance to give improved reliability.

S-CSP: An S-CSP is an organization that offers users data storage outsourcing services. When users own and save the same material, the S-CSP will only keep a single copy of these files and maintain only unique data in the deduplication system. A deduplication strategy, on the other hand, can minimize server storage costs while also saving user upload bandwidth. We consider a quorum of S- CSPs, each of which is an autonomous entity, for fault tolerance and data

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2023 ||

| DOI:10.15680/IJIRSET.2023.1207186 |

storage confidentially. The user data is spread among several S-CSPs cutting down on user upload bandwidth.

VI. RESULT

To enable effective duplicate verification, tags for each file will be computed and supplied to S-CSPs.
In order to do deduplication when uploading a file F, the user communicates with SCSPs. To be more precise, for file duplication checking, the user computes and sends the file tag F = TagGen(F) to S-CSPs.If a duplicate is found, the user computes it and sends it via a secure connection to a server. Otherwise, the process continues, the secret sharing mechanism is triggered, and the user uploads a file to CSP if no duplicate is found. To download a file from the SCSPs, the user will use the secret shares. With this technique, fault tolerance is provided, and even if a few storage servers malfunction, the user may still utilize the system.

VII. CONCLUSION

We provided distributed deduplication technologies to improve data integrity without compromising the confidentiality of consumers' outsourced data. To help with file-level deduplication, four alternative structures were offered. Consistency and integrity of the tag were guaranteed.

REFERENCES

[1] X. Chen, J. Li, J. Weng, J. Ma, and W. Lou, “Verifiable computation over large database with incremental updates,” IEEE Trans. Computers, vol. 65, no. 10, pp. 3184–3195, 2016.
[2] M. Gerla, J. Weng, and G. Pau, “Pics-on-wheels: Photo surveillance in the vehicular cloud,” International Conference on Computing, Networking and Communications, pp. 1123–1127, 2013.
[3] X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, “New algorithms for secure outsourcing of modular exponentiations,” IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 9, pp. 2386–2396, 2014.
[4] H. Yuan, X. Chen, T. Jiang, X. Zhang, Z. Yan, and Y. Xiang, “Dedupdum: Secure and scalable data deduplication with dynamic user management,” Inf. Sci., vol. 456, pp. 159–173, 2018.
[5] H. Huang, X. Chen, Q. Wu, X. Huang, and J. Shen, “Bitcoinbased fair payments for outsourcing computations of fog devices,” Future Generation Comp. Syst., vol. 78, pp. 850–858, 2018.
[6] IDC. (2014) The digital universe of opportunities : Rich data and the increasing value of the internet of things. [Online]. Available:https://www.emc.com/leadership/ digitaluniverse/2014iview/index.htm

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.423

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

Scan to save the contact details