e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707105 |

Reliable and Effective Disaster Recovery of Data within Shared Cloud Storage

Keerthana C N, Rajesh N

Student, Department of MCA, AMC Engineering College, Bengaluru, India

Assistant Professor, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: Data deduplication effectively reduces data redundancy within cloud storage systems and decreases users' bandwidth requirements. However, many existing schemes that rely on a reliable key server (KS) are problematic due to vulnerabilities, exposure of information, poor attack resistance, and significant computational overhead. Notably, if the reliable KS fails, the entire system ceases to function, creating a single point of failure. In this paper, we introduce a Secure and Efficient Data Deduplication (SED) scheme for a Joint Cloud storage system. This system takes advantage of collaboration among multiple clouds to provide global services.. SED supports the ability to update and share data dynamically without relying on a trusted KS. Additionally, SED addresses the problem of having a single point of failure common in conventional cloud storage systems. Our theoretical analyses demonstrate that SED guarantees provide semantic security under the random oracle model and demonstrate robust resistance to attacks, including brute-force and collusion attacks. Moreover, SED effectively reduces data redundancies with low computational complexity and minimal communication and storage overhead. The enhanced efficiency and functionality of SED enhance client-side usability. Finally, comparative results show that our scheme outperforms existing solutions.

KEYWORDS: client-side, SED, KS, Deduplication

I. INTRODUCTION

Cloud storage offers large- large-scale data storage and access to services on a "pay-as-you-go" basis. However, repetitive data stored in cloud storage significantly wastes and occupies storage resources. Deduplication is an effective technique for detecting and eliminating redundant data. After deduplication, only one copy of the data is uploaded and stored, reducing client-side bandwidth requirements and enhancing server-side space utilization. This technology is commonly utilized in various cloud computing services to enhance user experience and conserve storage space.

Traditional deduplication methods and their variants, which involve a key management server (KS), and cloud storage providers (CSPs), and users, depend on the security provided by the reliable KS. Unfortunately, these schemes are prone to single-point-of-failure and "platform lock-in" problems arise. If the reliable KS fails, the entire cloud storage system becomes non-functional, preventing the execution of data outsourcing protocols. To tackle these challenges, a novel cloud computing model, known as the Joint-Cloud computing system, has been developed. The Joint Cloud architecture includes users and multiple CSPs providing a range of services. These clouds work together without relying on a trusted KS, allowing users to connect to any cloud for computing services. Joint Cloud provides effective services across multiple clouds meets the demands of global cooperative cloud services by collaborating among different clouds. Additionally, it is possible to implement decentralised systems. Joint Cloud computing has garnered significant attention from both academia and industry.

II. LITERATURE REVIEW

"Disaster Recovery Strategies for Shared Cloud Storage: A Comparative Analysis" **Abstract**: This paper reviews various disaster recovery strategies tailored for shared cloud storage environments. It evaluates the effectiveness of replication, backup, and hybrid approaches in ensuring data reliability and availability during disasters. The study examines the trade-offs between cost, recovery time objectives (RTOs), and objectives for recovery points (RPOs) to provide insights into selecting the most suitable strategy based on specific organizational needs and resource constraints.

"Challenges and Solutions in Disaster Recovery for Multi-tenant Cloud Storage" **Abstract**: Focusing on the unique challenges of multi-tenant cloud storage, this review surveys current disaster recovery solutions. It explores challenges

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707105 |

like data isolation, security, and performance degradation during recovery processes. The paper discusses advancements in virtualization, data deduplication, and distributed consensus protocols to enhance disaster recovery capabilities while maintaining data integrity and confidentiality across multiple tenants.

"Resilient Data Replication Techniques for Disaster Recovery in Shared Cloud Environments" **Abstract**: This literature survey examines resilient data replication techniques designed to mitigate loss of data and guarantee continuity in shared cloud storage environments. It analyses synchronous versus asynchronous replication methods and how they affect recovery time and data consistency. The review evaluates recent advancements in fault tolerance mechanisms and consensus protocols to enhance the reliability and effectiveness of contingency plans for essential data stored in shared cloud infrastructures.

"Backup and Recovery Mechanisms for Reliable Data Protection in Shared Cloud Storage" **Abstract**: Addressing the importance of robust backup and recovery mechanisms, this survey explores different backup strategies and technologies applicable to shared cloud storage.

III. EXISTING SYSTEM

Bellare et al. introduced convergent encryption as a fundamental technique for securing data in deduplication, safeguarding outsourced data from untrusted and malicious CSPs.. formalized this approach as a message-locked encryption (MLE) approach. Following their work, several variants were proposed. However, these schemes based on message-locked encryption (MLE) encounter substantial risks due to the keys being derived directly from the files. Abadi et al. developed both a fully randomized scheme and a deterministic encryption scheme for bounded message distributions using non-degenerate efficiently computable bilinear maps. Li et al. introduced a scheme for reliable key management in deduplication, while Jiang et al. introduced a secure method for deduplication schemes with randomized tags, however, they did not cover data update requirements.

Later, Hur et al. considered management of dynamic ownership for enhanced deduplication, where each client stores keys located along Li et al. proposed using a binary tree path to achieve deduplication. They also introduced secure deduplication with key management based on secret sharing schemes. Following this, Shin et al. developed a decentralized server-aided encryption method deduplication, which required multiple user-KS interactions, providing attackers opportunities to gain valuable information from communications. Miao et al. later introduced a secure method of deduplication approach for multi-server environments. For fine-grained deduplication, Xia et al. developed a rapid and effective content-defined chunking approach. Zhao et al. introduced a deduplication scheme utilizing a Docker registry architecture, optimizing space by deduplicating layers and minimizing overhead during layer restoration.

Regarding security, common attacks on recent deduplication approaches have been discussed. Additionally, emerging technologies like blockchain are becoming more prevalent for deduplication in various applications, drawing significant attention from researchers.

Disadvantages

There is no intra-deduplication technique that only considers data outsourced by a single data owner using the same KS. This approach would be more efficient for backup systems, which are presently not included in existing systems. Additionally, Currently, there is no inter-deduplication technique that handles data outsourced by multiple data owners through various KSs.

IV. PROPOSED SYSTEM

In this paper, we introduce SED, a secure and efficient data deduplication scheme designed for the Joint Cloud storage system, operating independently of a trusted KS. Certain sub-algorithms in SED are inspired by the fully randomized tag generation algorithm, which assists in detecting duplicates and safeguards outsourced data from collusion attacks. Unlike earlier deduplication methods, SED guarantees semantic security for both the ciphertext and the tag, preventing adversaries from gaining valuable information.

SED is the pioneering scheme to securely enable both data updates and sharing. We have developed an encryption algorithm that supports data deduplication, updating, and sharing. To our knowledge, SED is the first scheme to address scenarios where data owners share their outsourced data with authorized users. A master encryption key is

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707105 |

collaboratively generated by participating CSPs, ensuring adaptable and secure key management. The data access control, facilitated by SED authentication, also ensures secure operations for updating and sharing data.

SED combines intra-deduplication and techniques for inter-deduplication to remove redundant data within the Joint Cloud system, enhancing deduplication efficiency. Theoretical analyses demonstrate that SED excels in areas like data confidentiality, integrity, resilience against severe attacks and overall functionality collusion. To experimentally evaluate its complexity, SED was executed and simulated using the Crypto++, GNU, and PBC libraries on Ubuntu. The assessment demonstrates that SED operates efficiently with minimal computational overhead.

Advantages

The proposed SED's encryption method and tag generation technique guarantee semantic security. Additionally, SED can withstand common threats like brute force, tampering, and collusion attacks. It achieves secure deduplication without depending on a trusted key server, while also facilitating data updating and sharing across multiple clouds. Moreover, SED addresses the single-point-of-failure problem and enhances the scalability of traditional deduplication schemes.

System Architecture

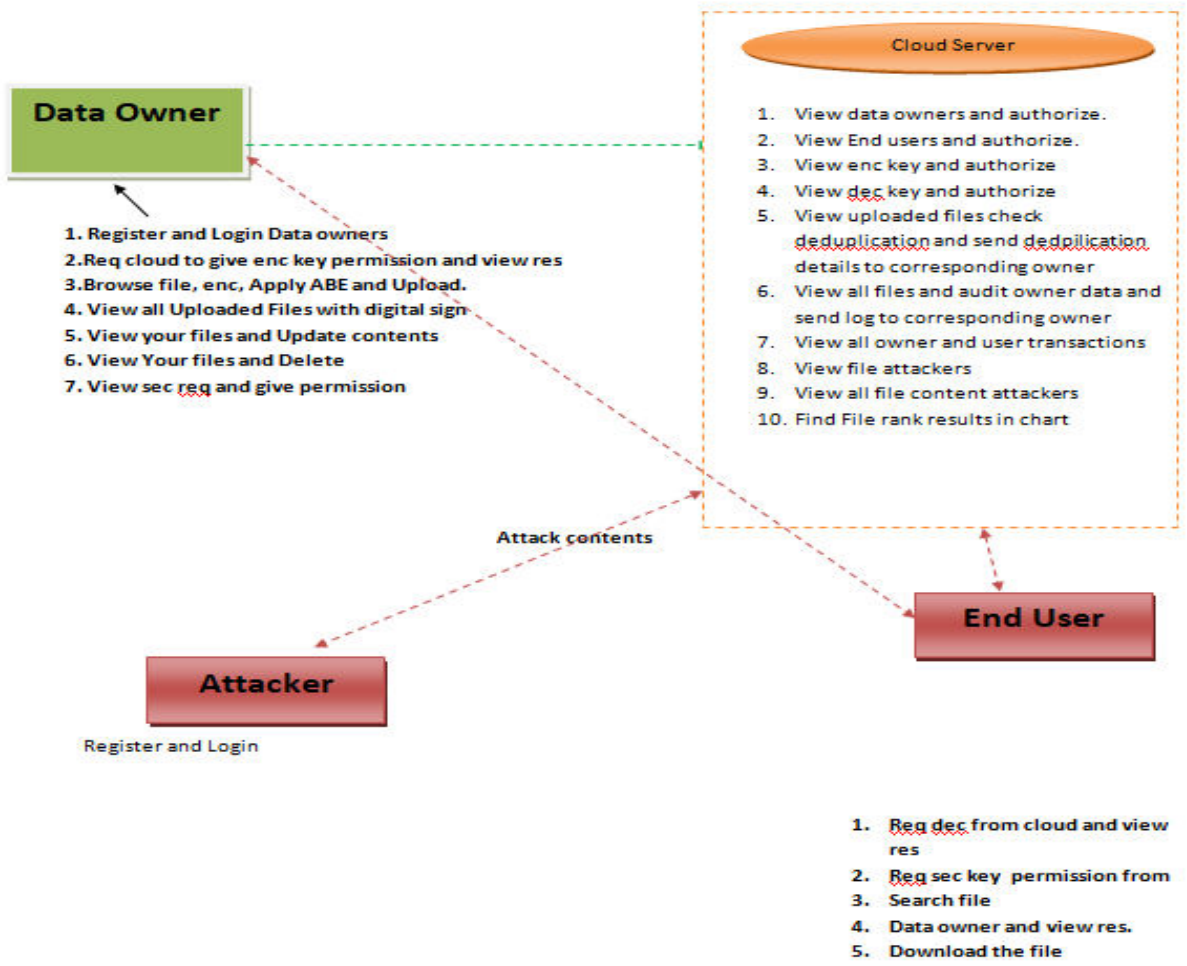

Fig1 system architecture

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707105 |

V. RESULT

In this paper, we present SED, a scheme that is both secure and efficient. for data deduplication. that operates without relying on a reliable key server (KS). SED reduces communication and processing overhead on the client side while enhancing Efficiency in the Joint Cloud storage system is achieved through streamlined encryption and tag generation algorithms, ensuring semantic security and tackling the CDH problem and tag consistency, covering aspects of security and validity. Additionally, SED enhances scalability and addresses SED addresses the vulnerability of traditional KS-based cloud storage systems to single-point-of-failure issues. It also offers a strong defense against various attacks like brute-force attempts and collusion between malicious CSPs and unauthorized users. Moreover, SED enables dynamic data operations such as deletion, modification, and sharing, enhancing its overall functionality and usability.

VI. CONCLUSION

In summary, the SED scheme we propose represents a major improvement in safe and effective data deduplication about Joint Cloud storage systems. By eliminating the need for a trusted Key Server (KS), SED reduces both communication and computation overhead on the client side and boosts overall system efficiency. It utilizes the Computational Diffie-Hellman (CDH) challenge for encryption and tag generation, ensuring semantic security and tag consistency, critical for maintaining security and validity.

SED overcomes several limitations of conventional cloud storage systems such as scalability issues and single points of failure related to the KS. It provides strong protection against common attacks, including brute force and cooperation among malicious Cloud Service Providers (CSPs) Moreover, SED facilitates dynamic data operations such as deletion, modification, and sharing, enhancing both functionality and usability. It also addresses collaboration among unauthorized users. compared to previous schemes.

REFERENCES

[1] P. Christen, "An overview of indexing techniques for scalable record linkage and deduplication," IEEE Transactions on Knowledge and Data Engineering, vol. 24, no. 9, pp. 1537–1555, 2012.

[2] G. Jia, G. Han, J. J. P. C. Rodrigues, J. Lloret, and W. Li, "Integrated memory deduplication and partitioning for improved performance in cloud computing," IEEE Transactions on Cloud Computing, vol. 7, no. 2, pp. 357–368, 2019.

[3] W. Xia, X. Zou, H. Jiang, Y. Zhou, C. Liu, D. Feng, Y. Hua, Y. Hu, and Y. Zhang, "Efficient design of content-based chunking for data deduplication in storage systems," IEEE Transactions on Parallel and Distributed Systems, vol. 31, no. 9, pp. 2017–2031, 2020.

[4] J. Li, J. Li, D. Xie, and Z. Cai, "Ensuring secure auditing and deduplication of data in cloud environments," IEEE Transactions on Computers, vol. 65, no. 8, pp. 2386–2396, 2016.Y.-C. Liang et al., “Sensing-Throughput Trade-off for Cognitive Radio Networks, ”IEEE Trans. Wireless Commun., vol. 7, pp. 1326–37 ,April 2008.

[5] L. Liu, Y. Zhang and X. Li, "KeyD: Secure key-deduplication with identity-based broadcast encryption", IEEE Trans. Cloud Comput..J. Ni, K. Zhang, Y. Yu, X. Lin and X. S. Shen, "Providing task allocation and secure deduplication for mobile crowdsensing via fog computing", IEEE Trans. Dependable Secure Comput., vol. 17, no. 3, pp. 581-594, May/Jun. 2020.Q. Wang, H. Zheng, “Route and spectrum selection in dynamic spectrum networks,” in Proc. IEEE CCNC 2006, pp. 625-629, Feb. 2006.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com