Using Data Partition Techniques to Improve Cloud Data Storage Security

Jinrui Zhang
Tandon School of Engineering
New York University
jz6578@nyu.edu

***Abstract*—Cloud computing has become central to the evolution of distributed systems, providing enhanced capabilities for computation, processing, and communication. Cloud storage offers easy access anytime, anywhere, enabling users to store and retrieve data seamlessly from remote servers. However, security and privacy are key concerns, especially for cloud storage. The proposed work enhances data privacy by encrypting user data with AES and RSA. The Cloud Manager (CM) handles data partitioning and secure storage, achieving both data security and privacy while improving performance through replication techniques.**

***Index Terms*—cloud storage, data partition, AES, RSA, encryption**

I. Introduction

In recent decades, rapid internet expansion has driven cloud computing's popularity, enabling users to access software and services from anywhere without managing technical details. Cloud storage, a key feature, allows developers to store and access data remotely while managing resources and access controls. Cloud services include SaaS, PaaS, and IaaS, with public, private, and hybrid deployment models accessible via the web.

However, since users lack physical control over outsourced data, ensuring data integrity and security is challenging, particularly for those with limited computing resources. Unauthorized access or hacking risks in unsecured cloud networks can lead to data leaks or loss, making cloud security essential.

Recent work focuses on third-party auditing (TPA) and remote integrity checking (RIC) to ensure data integrity. While these methods offer benefits, they also have drawbacks. TPA can lead to increased costs over time and pose privacy risks by exposing sensitive data. RIC can increase bandwidth usage and latency, especially with large datasets, and may still leave data vulnerable to malicious or negligent providers.

In the proposed solution, the Advanced Encryption Standard (AES) algorithm encrypts the file and the RSA algorithm encrypts symmetric keys before handing them to the Cloud Manager (CM), who splits the data across cloud servers. This approach reduces latency in integrity checks, allows sensitive data to remain private during third-party audits, and improves integrity checking efficiency due to smaller data partitions.

This proposed work offers several advantages: it ensures the secrecy of user data on the cloud by fragmenting it; prevents hackers from accessing meaningful information; and protects data from being viewed by cloud service providers, as cryptographic methods maintain data confidentiality.

The rest of the paper is structured as follows: In Section II, it involves an outline of the associated work in the cloud storage field. Section III includes the proposed methods. And Section IV presents evaluations of the specific proposed solution. Finally, Section V discussed the conclusion and the future works.

II. Related Research

This section reviews literature on data integrity verification and current data storage mechanisms. Dynamic data storage methods, such as token pre-computation and cloud storage, are examined for their efficiency [1]. Integrity checking methods focus on detecting server misbehavior and ensuring data correction and error localization [2]. To maintain data quality and reliability, distributed schemes are employed.

Several studies analyze cloud storage data integrity [2], [3], with dynamic data operations and public auditability supporting integrity checks. A third-party auditor model ensures independent evaluation and high service quality, with a storage model devised to handle multiple auditing tasks for efficiency.

Wang et al. [4] proposed a privacy-preserving remote data integrity auditing scheme using random masking. Worku et al. [5] improved efficiency over [3] with a modified random masking technique supporting data privacy protection.

The proposed approach designs a flexible storage scheme ensuring data availability and integrity using a partitioning technique. Files are encrypted with AES and symmetric keys are encrypted by RSA before being split across cloud servers by the cloud manager, maintaining privacy and reducing cost and latency.

III. Hypothesis

The paper proposes a data partitioning technique combined with AES and RSA encryption, inspired by [6]. In the proposed approach, the file is encrypted with AES at the user's end, and the symmetric key used for AES is encrypted with RSA. This method leverages the efficiency of AES and the security benefits of RSA. The encrypted file is then sent to the CM, which partitions the file for further RIC or TPA verification. This approach is expected to significantly improve storage performance while maintaining robust security. The figure 1 shows the process of my approach.

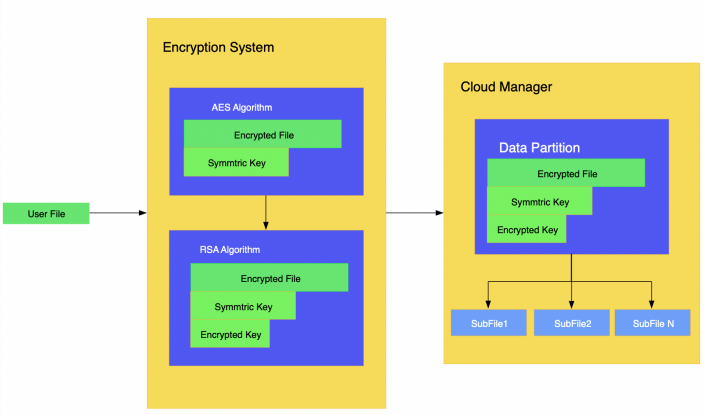

Fig. 1. Example of a figure caption.

To evaluate performance improvements, we will measure the time from encryption to integrity verification completion and assess data throughput under specific bandwidth constraints. The data partition technique is expected to reduce time consumption and increase data throughput.

IV. Empirical Evidence

In the simulation, we apply AES and RSA encryption with RIC for integrity checking. The figure illustrates the total time consumption of the cloud storage process. The figure 2 demonstrate a significant reduction after deploying the data partition technique.

The figure 3 shows the variation in time consumption when applying the data partition technique for a specific data size across different numbers of partitions.

Figure 4 illustrates the bandwidth efficiency achieved after deploying the data partition technique, showing the amount of data that can be processed within a specified time limit.

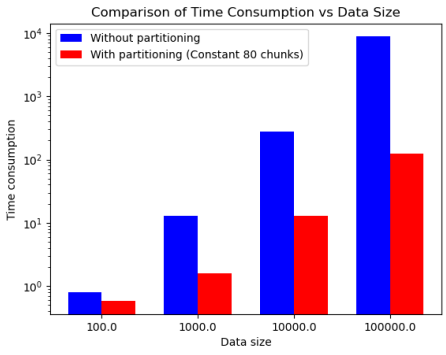

Fig. 2. Experiment on data size

V. Conclusion and Future Work

The proposed work enhances cloud data security by dividing encrypted user data into fragments, preventing unauthorized access to the complete file. Even if accessed, fragments reveal neither content nor sequence due to encryption. Proper validation ensures secure file distribution. Experiments show that data partitioning significantly reduces RIC latency and optimizes bandwidth.

Future work will involve implementing a multi-threaded mechanism for data integrity checks and developing a more advanced simulator to examine the relationship between partition count, file size, and performance. Because a large amount of partition will cause high-frequency auditing to perform badly due to additional processing and data transfers.

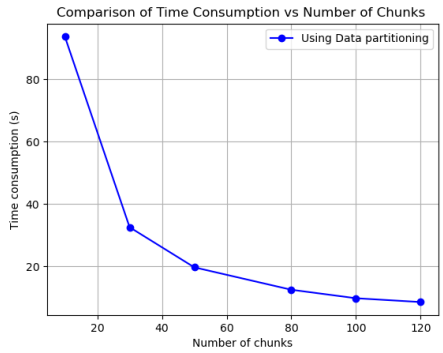

Fig. 3. Experiment on number of partitions

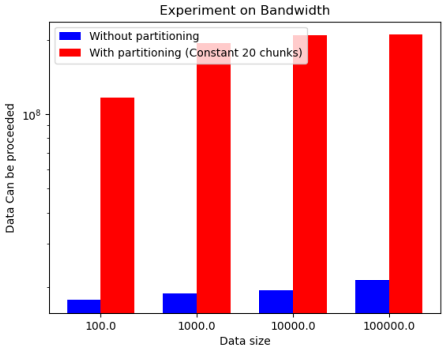

Fig. 4. Experiment on bandwidth efficiency

References

[1] Wang Cong, Wang Qian, Ren Kui, Cao Ning and Lou Wenjing , "Toward Secure and Dependable Storage Services in Cloud Computing," Services Computing, IEEE Transactions on , vol.5, no.2, pp.220-232, April-June 2012.

[2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Ensuring Data Storage Security in Cloud Computing," Proc. 17th Int'l Workshop Quality of Service (IWQoS '09), pp. 1-9, July 2009.

[3] Qian Wang; Cong Wang; Kui Ren; Wenjing Lou; Jin Li; , "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing," Parallel and Distributed Systems, IEEE Transactions on, vol.22, no.5, pp.847-859, May 2011.

[4] C. Wang, S. S. M. Chow, Q. Wang, K. Ren, and W. Lou, "Privacypreserving public auditing for secure cloud storage," IEEE Trans. Comput., vol. 62, no. 2, pp. 362–375, Feb. 2013.

[5] B. Lynn. (2015). The Pairing-Based Cryptographic Library. [Online]. Available: https://crypto.stanford.edu/pbc

[6] Khedkar, Swapnil V., and A. D. Gawande. "Data partitioning technique to improve cloud data storage security." International Journal of Computer Science and Information Technologies 5.3 (2014): 3347- 3350.