International Journal of Management, IT & Engineering
Vol. 14 Issue 12, December 2024,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Data Privacy in Multi-Cloud: A Fragmentation and Distribution Approach

Rajesh Daruvuri
Sravani Chirumamilla
Pravallika Mannem
Suchitha Reddy Aeniga

Abstract

Security and specifically data privacy remains a big concern especially in multi tenancy since it brings resource sharing that exposes environment to more security threats such as hacking. Standard data protection techniques provide security measures but are not efficient in dealing with new and complex threats, such as between inhabitant's data leakage. This paper presents a privacy-preserving framework that uses data fragmentation, multi-cloud distribution, and management control through an intermediary entity known as a trusted proxy. Storing such data in segments across the different CSPs means that none of the CSP has complete access to the dataset, improving on the data privacy, reliability, and solidity. The experimentation proves that the framework optimizes privacy and throughput by outperforming a comparable solution targeting privacy-sensitive, multi-tenant systems.

Keywords:

Cloud Computing;
Data Privacy;
Computer Architecture;
System Design;
Security.

Author correspondence:

Rajesh Daruvuri, Google Inc, USA, Venkatrajesh.d@gmail.com
Sravani Chirumamilla, P2C Technosol LLC, USA. Sravani@p2ctechnosol.com
Pravallika Mannem, ProBPM, Inc, USA, Pravi.sit05@gmail.com
Suchitha Reddy Aeniga, P2C Technosol LLC, USA, Suchitha.bi.dev@gmail.com

1. Introduction

Cloud computing has revolutionized data storage and processing, offering scalable, cost-effective infrastructure that supports organizations in managing and utilizing vast amounts of data [1]. Among the different cloud deployment models, multi-tenant environments have gained popularity due to their ability to host multiple users or tenants on shared resources, maximizing resource efficiency and operational flexibility [6]. However, this shared infrastructure poses unique privacy and security challenges, as data from different tenants coexist within the same virtual and physical spaces. In such settings, data breaches, unauthorized access, and inference attacks become real risks, as data belonging to different users can potentially be accessed or inferred by others [10].

Traditionally, encryption has been a primary method for securing data in the cloud, offering protection against unauthorized access [4]. However, in complex multi-tenant environments, encryption alone is insufficient to address the range of privacy concerns, especially when attackers attempt to reconstruct data from fragments [5]. Moreover, centralized data storage on a single Cloud Service Provider (CSP) introduces vulnerabilities, as a single compromised provider can expose large amounts of sensitive information [9]. This has led researchers to explore advanced privacy-preserving techniques that go beyond traditional encryption, such as data fragmentation and multi-cloud distribution.

Data fragmentation, a technique that divides data into smaller pieces before storing each fragment separately, has shown promise in enhancing data privacy by minimizing the amount of sensitive information accessible through any single fragment [7]. When combined with multi-cloud distribution, in which data fragments are spread across multiple CSPs, this approach further enhances privacy and security by ensuring that no single provider holds the complete dataset [3]. The use of a trusted proxy in this model enables centralized management of data fragmentation, encryption, and retrieval, allowing a more secure and manageable method of access control [2].

This paper presents a novel data privacy framework designed to address the privacy challenges inherent in multi-tenant cloud environments. The framework leverages data fragmentation and multi-cloud distribution, with a trusted proxy managing data access, storage, and retrieval. Key contributions of this research include the development of a fragmentation and distribution model that enhances privacy while ensuring data availability through redundancy. Additionally, we explore

hybrid encryption models to support secure data operations on fragmented data without requiring full reassembly, thereby expanding the potential applications of this framework.

2. Research Method

With the rise of cloud computing, multi-tenant environments have become essential, allowing organizations to share infrastructure efficiently. However, this shared model brings significant privacy risks, including unauthorized access and data inference attacks, which traditional security methods struggle to mitigate [1].

2.1 Encryption and Access Control

Encryption remains a foundational tool in cloud security, with methods like homomorphic encryption enabling data computations without decryption [4]. However, encryption alone cannot address all privacy concerns in multi-tenant setups, as attackers may still infer sensitive data. Advanced access control models, such as Attribute-Based Access Control (ABAC), add layers of security by managing access based on policies [2]. While ABAC enhances control, it does not fully eliminate risks related to shared data storage.

2.2 Data Fragmentation

Data fragmentation splitting data into smaller, independently encrypted parts—has gained traction for minimizing data exposure in multi-tenant settings [9]. Each fragment alone is unintelligible, thus lowering risks even if accessed. When fragments are stored across multiple CSPs, it further reduces the possibility of unauthorized data reconstruction [3].

2.3 Multi-Cloud Distribution and Trusted Proxy Models

Multi-cloud distribution mitigates the risk of relying on a single CSP by spreading data fragments across providers [8].This model enhances fault tolerance and privacy, as no single provider holds the full dataset. However, coordinating this distribution is complex, and trusted proxies have emerged as intermediaries to manage fragmentation, encryption, and secure access. A trusted proxy acts as a central authority, ensuring data is only reassembled under secure, verified conditions [2].

2.4 Limitations of Current Approaches

While encryption, access control, data fragmentation, and multi-cloud architectures each address aspects of cloud privacy, they have limitations. Encryption can be computationally intensive, while multi-cloud setups increase operational complexity and cost [4][3].Trusted proxies, while helpful, can become a single point of failure if compromised, potentially exposing data across CSPs.

3. Methodology

3.1 Data Fragmentation Framework

The proposed data privacy framework involves the following core components:

Data Fragmentation: Data is partitioned into many small and independent parts each of which is encrypted and distributed in the system. This fragmentation process helps in the way that even if one fragment is loaded, the amount of information gained is quite limited.

Primary and Secondary CSPs: Original data shards are stored at the primary CSPs while duplicate copies of these shards are stored at secondary CSPs in order to support fast restore in the event of failure on the part of the primary CSPs.

Trusted Proxy Management: The fragment manager that is recognized as the central trusted proxy is solely in charge with regard to the tasks pertaining fragment storage as well as fragment retrieval and fragment reconstruction. This proxy ensures HTTPS access to the cloud and keeps data confidentiality between CSPs and such requests' negotiator.

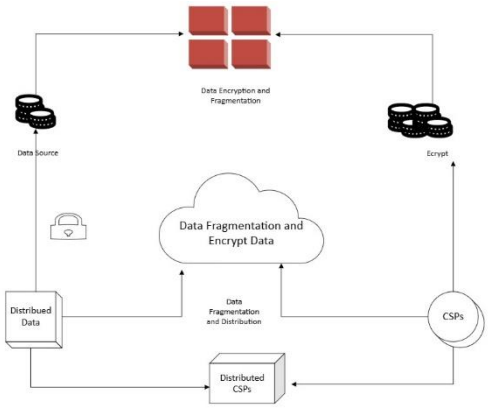

Fig. Data Fragmentation and Distribution

3.2 Data Storage and Retrieval

The trusted proxy manages the lifecycle of each data fragment:

Data Storage: The proxy must evaluate the CSP for each fragment using the security policies, availability and redundancy needed to choose the best CSPs. Before storing, fragments are encrypted with unique keys so that if one of CSPs is compromised, data content remains meaningless to a user.

Data Retrieval and Reconstruction: To perform authorized data retrieval, the proxy collects identified fragments from the related CSPs and recompiles the obtained parts, with secure re-encryption if necessary. When a fragment is not retrieved, the secondary cleaning system CSPs provide the missing fragment**.**

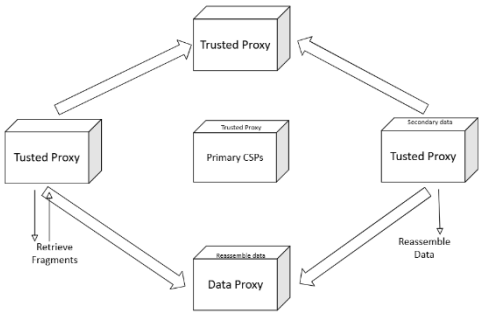

fig. Trusted Proxy and Data Reassembly

3.3 Conflict and Consistency management

Since multi-tenant environments are complex, there must be changes in the data while ensuring that the fragments are synchronized across the CSPs. This principle of trusted proxy simplifies updates by the conflict manager which also checks for every data fragment to make sure that any update is complete at all storage points.

3.4 Security and Privacy Mechanisms

The model combines encryption with fragmentation and assembles set of fragments using corresponding unique keys. Restriction mechanisms and identification techniques are incorporated within the frame work of the trusted proxy to allow only authorized individual to request data reassembly.

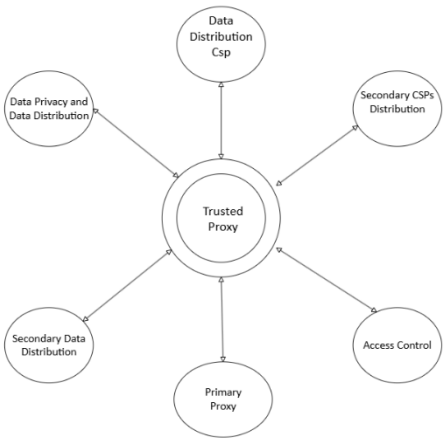

Fig. Trusted Proxy Model

4. Experiment Evaluation

To understand how this data fragmentation approach performs, we have implemented it in a mimic model with different CSPs and tenants in a cloudy setup. We focused on three main areas: Privacy, it's application for fast data retrieval and scalability.

4.1 Setup

The propose primary and backup CSPs to store data fragments with only the fragment identifiers being presented to them while the trusted proxy was the only entity that ran the data requests, we granted access to. The proxy split and encrypted each piece of data before sending that data to the different CSPs we used for the experiment.

4.2 Results

Privacy: Some of the tests proved that no individual CSP was able to obtain a full dataset as each of the providers only stored part of the data. If a person attempted to get the information, he or she would not get enough of it to make any sense of it.

This provided an added security angle to the data as the encryption that was put on each fragment became a protection mechanism even if any fragment was compromised.

Data Retrieval Efficiency: It could restore the broken pieces of collected information and reconstruct it in less than 2 seconds. Small and medium data pieces were returned almost immediately while large data fragments took slightly longer.

This proves that breaking up data and storing them did not slow the rate of data access by much.

Scalability: When we scaled more tenant and data, the system performed good indeed. From one thousand users requesting the data at once, the retrieval time showed that it was still fast.

This means that even with an increased number of users the approach is capable of running without being slowed down, thus it is ideal for large scale systems.

The findings indicate that data fragmentation and this multi-cloud environment gives good privacy and fast access regardless of the number of users. However, the trusted proxy could potentially become a bottle neck or a point of failure. It can be made better by incorporating a backup proxy so as to ensure that data access is always up

5. Discussion

5.1 Strengths

Stronger Data Privacy: What the framework does is divide data into small chunks encrypts each data piece and stores them with different cloud providers, which makes it nearly impossible for anyone including the cloud provider to come into view with the full dataset. This setup assists in protecting information privacy and creating the necessary security in multi-tenant, common cloud server systems, where multiple users are present.

Simplified Management with a Trusted Proxy: The framework intentionally employs only one 'trusted proxy' through which all access to the data is controlled. This proxy decides who can view data and names databases where fragments are

located, as well as methods of accessing this information. This setup makes security and data management much easier because all access goes through this one single point of attack.

Backup for High Reliability: Data fragments are stored with several different "primary" and "secondary" cloud services. A second copy is maintained by the opposite side of the mirror, so it always has fragment backups in case the primary one has failed. This redundancy is especially crucial in such industries as healthcare and finance since data is often needed constantly.

Ideal for High-Privacy Industries: It is very suitable for industries that necessitate high level of data security like healthcare, finance and government industries. For instance, hospitals can apply this frame work to safeguard patient records while companies dealing with monetary activities may apply it to protect transactions information. The government can also protect the data of citizens by ensuring the data is never centralized for one party to access at one instance.

5.2 Weaknesses and Limitations

Trusted Proxy as a Single Point of Failure: By effectively acting as one single point of access to the data, a so-called trusted proxy is created. The entire data access system is at risk if a component is hacked, fails, or becomes overloaded. Potential adversaries are capable of manipulating the proxy to disrupt data acquisition, let alone take control over it. The distinction between both can be conceived as having everything pass through a single point which can be a risk when the element in question is a particularly large high traffic system.

Potential for Slowdowns: Since all the access to the database is via the trusted proxy it may become slower as more people are using the system or more data is being requested. This single proxy is responsible for storage, data retrieval, and access verification, so if it becomes troubled, data retrieval may be slow and even freeze during peak usage.

Complexity and Cost: It also involves additional procedures such as data encryption before sending data fragments, the distribution of data over several providers and operational fragments of backup data. These steps increase the level of complication making the system difficult to install and manage. More so, the nature of having multiple cloud service providers and the option of having a secure trusted proxy is likely to add on costs and can be a big issue for small organizations.

5.3 Future Work

Hybrid Encryption Approaches: Subsequent generations of this framework might employ differentially encrypted schemes in an effort to enhance the security and functionality of the data storage. Such techniques as homomorphic encryption, provide the opportunity to process data in its encrypted form without having to decrypt it at all. Other approach uses in secure multi-party computation where multiple cloud providers can collaborate to perform data processing without getting to see the entire data. These approaches would enhance the flexibility and security of the framework particularly for data analytical purposes better than the current methods being employed.

Real-Time Monitoring and Threat Detection: The trusted proxy could increase the security of the system by incorporating the real-time monitoring feature. Monitoring could identify any harassment such as data requesting, periodical and unusual cloud provider unavailability etc. Such an arrangement would go a long way in addressing any data threats that would be identified because the system would notice a suspicious activity and immediately take appropriate action on the fragments of the data.

6. Conclusion

This paper discusses an approach that adopts data fragmentation, data encryption and multi-cloud distribution as means of mitigating privacy risk within multi-tenant cloud context. Since data is split and its components are stored across several CSPs, it allows none of the providers to have the full data which minimizes privacy concerns. Experimental findings reveal the feature set and privacy preservation capability of the framework, which calls for future large-scale applications with privacy concerns. More development on the timely tracking of threats, such as hacking and intrusions, and better data encryption will continue to improve the data security on cloud environments.

7. References

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., ... & Zaharia, M. (2010). *A view of cloud computing. Communications of the ACM*, 53(4), 50-58.
[2] Jin, X., Krishnan, R., Sandhu, R., & Kant, K. (2018). A Multi-Tenant Attribute-Based Access Control Model with Tenant Trust. *IEEE Transactions on Dependable and Secure Computing*, 15(2), 138-152.
[3] Yuan, D., Yang, Y., Liu, X., & Chen, J. (2011). A data placement strategy in cloud computing to minimize data transfer cost and storage cost. *IEEE Transactions on Cloud Computing*, 1(1), 36-51.
[4] Kamara, S., & Lauter, K. (2010). Cryptographic cloud storage. In Financial Cryptography and Data Security (pp. 136-149). *Springer*, Berlin, Heidelberg.
[5] Wang, C., Wang, Q., Ren, K., Cao, N., & Lou, W. (2010). Toward secure and dependable storage services in cloud computing. *IEEE Transactions on Services Computing*, 5(2), 220-232.

[6] Zissis, D., & Lekkas, D. (2012). Addressing cloud computing security issues. *Future Generation Computer Systems*, 28(3), 583-592.
[7] Subashini, S., & Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. *Journal of Network and Computer Applications*, 34(1), 1-11.
[8] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing. *National Institute of Standards and Technology*.
[9] Ren, Y., & Gu, M. (2012). Data fragmentation for privacy protection in multi-cloud storage. In 2012 *IEEE Asia-Pacific Services Computing Conference* (pp. 201-208). *IEEE*.
[10] Mannem, P., Daruvuri, R., & Patibandla, K. (2024). Leveraging Supervised Learning in Cloud Architectures for Automated Repetitive Tasks. *International Journal of Innovative Research in Science, Engineering and Technology*, 13(11), pp 1-10.
[11] Kasula, V. K., Yadulla, A. R., Konda, B., & Yenugula, M. (2024). Fortifying cloud environments against data breaches: A novel AI-driven security framework. *World Journal of Advanced Research and Reviews* 24 (01), 1613-1626.
[12] Pearson, S., & Benameur, A. (2010). Privacy, security and trust issues arising from cloud computing. In 2010 *IEEE Second International Conference on Cloud Computing Technology and Science* (pp. 693-702). *IEEE*.