A Review for Data Fragmentation in Cloud Computing[1]

***Qays Jabbar Abed, **Dr. Rami Tawil**

**Ph. D. Student, High Institute of Technology*
***Associate Professor, Faculty of Science*
Lebanese University, Lebanon

DOI:10.37648/ijrst.v12i03.004

Received: 20 May 2022; Accepted: 05 July 2022; Published: 21 July 2022

ABSTRACT

The importance of network security is demonstrated by preserving data from being stolen or tampered with by attackers. The computerized cloud supports the protection of both large and small data as well as authorized remote access. Although its protection is secured in the cloud, it needs to use additional methods to enhance data protection from attackers. In this paper, we present the various methods that researchers used to protect data through the computerized cloud, as the diversity in the use of methods led to obtaining good results, albeit to varying degrees. This relieved data fragmentation using multiple technologies and storing it in the cloud, and some used data encryption methods. The aim of this research is to present different types of methods used in the field of data protection across the network via the cloud, which provides an opportunity for researchers to present research that addresses the gaps in order to obtain better results.

INTRODUCTION

Since its inception, the cloud computing technology has been applied widely across a wide range of industries, assisting businesses in lowering maintenance expenses and enhancing the reliability of their data. A cloud storage service relies on the client to store data on a cloud server and to access it whenever necessary due to the availability of services with low charges. Despite all the advantages modern technology has to offer, numerous risks have also surfaced [1]. Data fragmentation aims are useful for keeping the database confidentiality and privacy. In addition, it makes the data more secure which is used in the cloud to be more effective and efficient performance. Figure 1 illustrated the sample of data fragmentation with cloud.

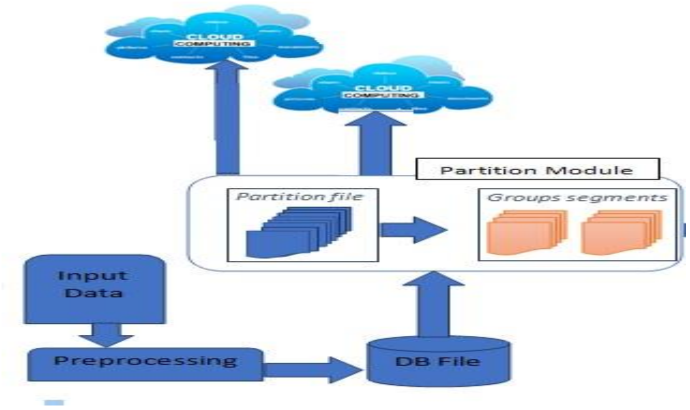

Figure 1: Sample of data fragmentation with cloud

[1] *How to cite the article:*
Qays, J.A, Rami, T., A Review for Data Fragmentation in Cloud Computing, IJRST, Jul-Sep 2022, Vol 12, Issue 3, 19-23, DOI: http://doi.org/10.37648/ijrst.v12i03.004

Cloud data centers are increasingly being attacked by both external attackers and nefarious internal users [2]. Additionally, the cloud provider is in charge of managing and keeping user data secure and, in most circumstances, does not make these practices known to its consumers [3]. There are three types of data fragmentation methods; Horizontal fragmentation, Vertical fragmentation, and Mixed fragmentation.

Horizontal fragmentation refers to the process of separating a table horizontally by locating each row or (a group of rows) of relation to one or more fragments. These separates are then locates to other sides in the distributed system. The table's rows or tuples are divided into different systems, some of which are placed in one system and others in others. A condition on one or more relational properties identifies the rows that make up the horizontal fragments [4].

Vertical fragmentation is the technique of vertically dividing a table into its columnar properties. Some of the attributes in this fragmentation are stored in one system, while the rest are stored in other systems. This is because not every site will require a table's entire set of columns. Each fragment must have the main key field(s) in a database in order to handle restoration

There are two methods for mixed fragmentation; The first approach entails first producing a set or group of horizontal pieces, after which one or more of the horizontal fragments are used to produce vertical fragments. The second approach entails first producing a set or group of vertical pieces, after which one or more of the vertical fragments are used to produce horizontal fragments [4, 5]. This paper will present a review of many researchers who used various data fragmentation methods. In the following paragraphs, research that uses various methods of data segmentation and storing in the cloud will be addressed.

1- Performance analysis of data fragmentation techniques on a cloud server.

Santos, N. et al. [6] investigates various techniques for dividing cloud data to block direct access by unauthorized users. It examines their output in a cloud case, accounting for the total processing time, which includes data upload and download. The results of this experiment show that fragmentation techniques outperform encryption strategies.

The first technique is called predetermined pattern fragmentation, and it operates on pieces that have been introduced into split files with odd or even indexes. The chunks are saved in the split file in accordance with the index they are given after being separated from the original file. Only two split files are produced as a result, and the size of each chunk is determined. In order to reconstruct the file using this technique, the attacker will need to be aware of the chunk's size.

Random pattern fragmentation is the second method, involves splitting the original file into N pieces, as with the other techniques, and inserting each chunk into a split file whose length is equal to the length of the associated pattern. The key benefit of this technique is that an attacker won't be aware of the size of any individual chunk or the arrangement of chunks within each split file.

The original file is first encrypted with AES 256 before being transferred to the cloud using Simple AES encryption which is the last approach. The file is not fragmented, unlike the earlier techniques. This strategy was taken into account in the experiment not only to compare its performance with the other methods but also to examine the effectiveness and acceptability of combining a popular encryption algorithm with data fragmentation.

2-A Comparison of Data Fragmentation Techniques in Cloud Servers.

The year 2018 focuses on data fragmentation strategies and how they can impact the overall performance of a cloud service [7]. Using Amazon Web Services, an analysis of the time taken to execute each algorithm based on pattern fragmentation, cipher, or both. Securing data using pattern fragmenting provides a good level of security.

3-Security Enhancement of Data in Cloud Using Fragmentation And Replication

In this proposed system once [8], the file is uploaded by the client, the cloud manager will encrypt the file, split the file into parts, and place them in featured locations within the cloud. In this proposed system once the file is uploaded by the client, the cloud manager will encrypt the file, split the file into parts, and place them in featured locations within the cloud. Using the T-coloring method, to increase performance, the fragments are replicated across the nodes that generate the highest read/write request. T-coloring increases the effort of an attacker to intrude on the system.

4-Enhancing Data Security in Cloud using Random Pattern Fragmentation and a Distributed NoSQL Database

This research approaches the problem by proposing the use of a fragmentation algorithm [9], combined with a distributed NoSQL (Not only SQL) database to secure data stored in the cloud, using the method of splitting data into multiple parts and aggregating them into divided files. These split files, in turn, are inserted into the Apache Cassandra database, which is distributed across multiple nodes. These nodes are stored in two different clouds to increase performance.

This provides a faster alternative to secure data in the cloud and this distributed approach allows the data to be processed simultaneously, taking advantage of the high resources offered by cloud computing and thus facilitating its adoption in this environment. Scenarios suitable for the proposed method lie mainly in environments where speed is paramount and the client resources are limited.

5-Data Confidentiality using Fragmentation in Cloud Computing

Hudic et al.[10] The proposal focuses on secure and confidential data outsourcing to Cloud environments by using fragmentation techniques and applying only minimal encryption to prevent data exposure. It relies on database normalization, user requirements, and confidentiality levels in order to enforce privacy before distribution.

Fragmentation and storage in the cloud with minimal encryption: Fragmenting is applied to relational databases where tables are treated as independent parts. And the storage is distributed into two areas:

1-The trusted local domain for sensitive data and encryption keys.

2-The public domain for distributing data to the required service providers.

6-Cloud Security Solution Fragmentation and Replication

Chavan, Mrs Radhika, and S. Y. Raut. [11]. they proposed a new solution that presents the Graphical Authentication System with fragmentation and replication techniques. The graphical password authentication provides the security and usability of the proposed system. Here in this system, when users upload any file that the file is fragmented and replicated to provide better security and performance in terms of access time. When any network is not available to access then, the data will be accessed by using replicas in a very short time. The t-coloring method is used to assign the fragments and their replicas to improve security. This system mainly focuses on the data and authentication system with good performance. the extra replication can also result in high storage cost or drops in the system's overall performance due to extreme use of bandwidth. So, here controlled replication is used.

7-A Fast Fragmentation Algorithm for Data Protection In a Multi-Cloud Environment

In this research [12], researchers introduce a novel algorithm for data fragmentation that is particularly well adapted to be used in a multi-cloud environment.

The proposed algorithm fragments initial data d into k fragments f0,...,fk−1 that will be dispersed and stored over c non-colluding clouds. and the fragmentation process uses a combination of secret sharing and data permuting to create dependencies between data inside the fragments. Such k fragments are then dispersed over c clouds. The dispersal follows special rules ensuring that the fragments stored at one provider do not reveal the content of the initial data. A performance comparison with published related works demonstrates it can be more than twice faster than the fastest of the relevant fragmentation techniques while producing reasonable storage overhead.

They purpose to address this issue by fragmenting the user's data, encoding them, and dispersing fragments over several non-colluding sites.

8-Data Security Model Using Artificial Neural Networks and Database Fragmentation in Cloud Environment

The proposed work presents a new cloud data security model with the help of an Artificial Neural Network [13]. This proposed algorithm is implemented using dynamic hashing fragmented components. It is implemented for storing fragmented sensitive secret data. This algorithm is applied to a various number of cloud databases.

To increase the security in the new data security model was proposed using neural networks. It is composed of highly interconnected networks called neurons. Those neurons are trained using supervision and learning. One of the methods for learning neural networks is counter propagation for various applications of the cloud. This leads to a guarantee for improvement in confidentiality

The proposed is to achieve confidentiality for cloud data using fragmentation. The fragmentation is done both horizontally and vertically. It is a technique in which data can be stored for different cloud data centers by fragmenting the databases into several partitions called fragments. Confidentiality is achieved by dividing a database into different fragments for the identification of different locations. First, the data is divided into different fragments. Such as F1, F2, F3, etc. All the fragments are stored in the various data centers of the cloud. The sensitive data M is encrypted separately. The fragments are also stored using a dynamic hashing mechanism in respective data centers.

9-A Method for Text Data Fragmentation to Provide Security in Cloud Computing.

Archana M, and Mallikarjun Shastry P M [14] they are proposed a method to implement a unique approach to provide security in cloud computing, where text files will be fragmented based on random number generation. A linear congruential generator (LCG) but here initial values for the algorithm are changed since the number of fragments for any file is more. Random number generation algorithm lower range and higher range values will be calculated so that the new random number which needs to compute should be within the lower and higher range of random numbers. Fragments can be obtained based on the new random number. With reference to the random number, block size or fragment size will be calculated. Hence text file will be divided into fragments.

Hash value to be calculated and the file will be transmitted to the node. At the node, file fragments will be encrypted and keys will be generated. The file fragments will be retrieved and a file will be reconstructed and the file will be available to the user.

10-Securing Big Data Using Mixed Fragmentation Based on Multi-Cloud Environment

In this research, a new framework for securing Big Data using Multi-Cloud is presented. By splitting the incoming data into three files using a Mixed-Fragmentation technique [15], the proposed technique, Big data Security Fragmentation (BSF) stores and collects data from the three clouds without the need to Encrypt or Decrypt these data.There are two essential fragmentation strategies: Horizontal and Vertical and there is a possibility of Hybrid Fragments.

Horizontal fragmentation divides a single relation R into subsets of rows using query predicates. It reduces query processing costs by selecting the horizontal fragments that are built, and the original relation is reconstructed by the union of the fragments

Vertical fragmentation splits a single relation R into sub relations that are projections of relation R with respect to a subset of attributes. These relations are in a grouping with attributes and are frequently accessed by queries. Projection builds the vertical fragments. By joining the fragments, the original relation is reconstructed

Mixed/Hybrid fragmentation is a Combination of horizontal and vertical fragmentations. This type is the most complex one because both types are used in which horizontal and vertical fragmentation of the DB

CONCLUSION

This research addressed various techniques whose using for data fragmentation, All of these techniques came out with good results despite their differences. Some techniques were used called pre-determined pattern segmentation, and another method is random pattern segmentation. While other methods relied on colors and some of them relied on fragmentation and replication, in addition to some of them used encryption. One of its downsides is that it consumes time, especially when the file size is large, although encryption increases the strength of data protection. We did not notice any processing of files that are called continuously and repeatedly, which requires researchers to stop it to address this thing. It is also worth noting that the type of data and its importance play a major role in choosing the appropriate method, taking into account the importance of time for each case.

REFERENCES

1 M. Bahrami and M. Singhal, "The Role of Cloud Computing Architecture in Big Data", *Information Granularity Big Data and Computational Intelligence*, pp. 275-295, 2015.

2 Z. Yan, R. Deng and V. Varadharajan, "Cryptography and Data Security in Cloud Computing", *Information Sciences*, vol. 387, pp. 53-55, 2017.

3 Cloud Security Alliance, "Top Threats to Cloud Computing", CSA, 2010.

4 P. Kumar, P. Raj and P. Jelciana, "Exploring Data Security Issues and Solutions in Cloud Computing", Procedia Computer Science, vol. 125, pp. 691-697, 2018.

5 R. Hegarty and J. Haggerty, "Extrusion detection of illegal files in cloud-based systems", *International Journal of Space-Based and Situated Computing*, vol. 5, no. 3, pp. 150, 2015.

6 Santos, N., Lentini, S., Grosso, E., Ghita, B.V., & Masala, G.L. Performance analysis of data fragmentation techniques on a cloud server. *Int. J. Grid Util. Comput.*, *10*, 392-401, 2019.

7 Lentini, S., Grosso, E., Masala, G.L. A Comparison of Data Fragmentation Techniques in Cloud Servers. In: Barolli, L., Xhafa, F., Javaid, N., Spaho, E., Kolici, V.(eds) Advances in Internet, Data & Web Technologies. EIDWT 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 17. Springer, Cham. 2018 https://doi.org/10.1007/978-3-319-75928-9_50

8 Patni, P. D., and S. N. Kakarwal. "Security Enhancement of Data in Cloud using Fragmentation and Replication." *International Journal of Engineering and Management Research (IJEMR)* 6.5 (2016): 492-497.

9 N. L. Santos, B. Ghita and G. L. Masala, "Enhancing Data Security in Cloud using Random Pattern Fragmentation and a Distributed NoSQL Database," *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, pp. 3735-3740, doi: 10.1109/SMC.2019.8914454.

10 Hudic, Aleksandar, Shareeful Islam, Peter Kieseberg, Sylvi Rennert, and Edgar R. Weippl. "Data confidentiality using fragmentation in cloud computing." *International Journal of Pervasive Computing and Communications* (2013).

11 Chavan, Mrs Radhika, and S. Y. Raut. "Cloud Security Solution: Fragmentation and Replication" *International Journal of Advance Research and Innovative Ideas in Education* 2, no. 4 (2016).

12 Kapusta, Katarzyna, and Gerard Memmi. "A fast fragmentation algorithm for data protection in a multi-cloud environment." *arXiv preprint arXiv:1804.01886* (2018).

13 D.Suneetha, D.Rathna Kishore, G.G.S.Pradeep. " Data Security Model Using Artificial Neural Networks and Database Fragmentation in Cloud

Environment" International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8 Issue-2, July 2019.

14 Archana M, and Mallikarjun Shastry P M, "A Method for Text Data Fragmentation to Provide Security in Cloud Computing." International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), Volume-9 Issue-2, December, 2019.

15 Rabab M.Nabawy, Heba El Beh, Hamdi M. Mousa," Securing Big Data Using Mixed Fragmentation Based On Multi-Cloud Environment", International Journal of Scientific & Technology Research (IJSTR). Volume 9 - Issue 3, March 2020