International Journal of Research in Engineering, Science and Management
Volume-2, Issue-3, March-2019
www.ijresm.com | ISSN (Online): 2581-5792

Optimizing Cloud Security and Performance using Graph T-Coloring and Fragmentation

R. Padma Priya[1], S. Swethaa[2], J. Harini[3], Konduru Sai Sriyanka[4]
[1]*Assistant Professor, Department of Computer Science, Panimalar Institute of Technology, Chennai, India*
[2,3,4]*Student, Department of Computer Science, Panimalar Institute of Technology, Chennai, India*

***Abstract*: Cloud computing is considered to have a great impact on the way data will be stored during the next years. Cloud computing has been met with security concerns over the past few years. Nevertheless, this IT paradigm has made significant impact to the protocols of different firms and has brought about the emergence of various cloud providers offering high security protocols matched with high performance. It moves the software and databases to the large data centers, where the management of the data may not be trustworthy. In this paper, we propose a design mechanism that not only protects sensitive data but also assures data availability at all times. The mechanism divides a file into fragments, encrypts the data and then replicates the fragmented data over the cloud nodes thereby securing data during attacks.**

***Keywords*: Cloud Security, fragmentation, T-coloring, replication, centrality.**

1. Introduction

The cloud computing paradigm has reformed the management of the information technology infrastructure. Access to shared resources in pay-as-you-go mode cuts the management effort of the user to a minimal level. Cloud computing is characterized by on-demand services, ubiquitous network accesses, elasticity, and measured services. The above mentioned characteristics of cloud computing make it a striking candidate for organizations and users for adoption. However, the benefits of low-cost, negligible management and greater flexibility come with increased security concerns. Security is the most crucial aspects among those prohibiting the wide-spread adoption of cloud computing. Due to the security apprehensions and industry, on occasions rate cloud computing unwise for business-critical organizations. Cloud security issues may decrease due to the core technology's implementation (virtual machine, cloud service offerings and arising from cloud characteristics. A cloud consists of many entities. For a cloud to be secure, all the participating entities must be secure. In any system with multiple units, the highest level of the system's security is equal to the security level of the weak entity. The above mentioned fact can bring the security level of other entities down to the level of the victim entity. The reduced security of the victim entity becomes the gateway for an attacker to enter the system that, in turn, puts the whole system and resources at risk. Therefore, in a cloud, the security of the assets does not completely depend on an individual's security measures. The neighboring entities may provide an opportunity to an attacker to bypass the user defenses. A cloud utility, called off-site data storage, helps the customer from focusing on the data storage systems. The off-site data storage requires users to move data in cloud's virtualized and shared environment that could result in various security concerns. Similarly, the availability of data is also a primary concern because of the fact that data may move in the cloud that is not under the administrative control of a customer.

2. Related work

Mei and Mancini proposed a distributed algorithm for file allocation that guarantees high assurance, and scalability in a large distributed file system. The confidentiality and integrity are preserved, even in the presence of a successful attack. But there were issues of load balancing, job scheduling, and the effects of malicious node. [3] Bilal, Manzano, Calle and Zomaya analysed the robustness to failures and ambiguities to deliver the required Quality-of-Service (QoS) level and satisfy service-level agreement (SLA). But network traffic and throughput analysis were not performed and high level of network robustness leads to higher cost.[8] William K. Hale introduced two generalizations of chromatic number and show that many frequency assignment problems are equivalent to generalized graph coloring problems.[1] Thanasis Loukopoulos and Ishfaq Ahmad evaluated different algorithms w.r.t the storage capacity constraint of each site as well as variations in the popularity of objects, and also examined the trade-off between running time and solution quality.[4] Vasilakos, Keqin and Zomaya proposed a mechanism that provides data confidentiality, protected data sharing without re-encryption, access control for malicious insiders, and forward and backward access control [12].

3. Existing system

In existing system, there are mechanisms for the secure and optimal placement of data objects in a distributed system. However, when the data are to be shared, the cryptographic services need to be flexible enough to handle many users, exercise the access control, and manage the keys to protect data

confidentiality. A single key shared between group members will result in the access of past data to a newly joining member. The above said situation violates the confidentiality and the principle of least privilege. [12] An encryption key is divided into n shares and distributed on different sites within the network. The division of a key into n shares is carried out through the (k, n) threshold secret sharing scheme. The number of replicas and their placement is determined through heuristics.

A primary site is selected in each of the clusters that allocate the replicas within the cluster. the centralization of services allows a reasonable management of complex distributed applications, assuring security, availability, and consistency [3]. Nevertheless, the scheme focuses only on the security of the encryption key. The data files are handled as a single file. It can take long time for uploading and downloading the data.

4. Proposed system

In the proposed system, we propose the Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that for file storage and data security. we divide a file into many fragments, and replicate the fragmented data over the cloud nodes. Once the file is split into many fragments, the DROPS methodology selects the cloud nodes for fragment placement. The selection is made by keeping an equal focus on security and performance in terms of the access time. We choose the nodes that are central to the cloud network to provide better access time. Each node stores only a single fragment of a particular data file that ensures that even in case of an attack, no meaningful information is revealed to the attacker. And the nodes storing the fragments are separated by certain distance by means of graph T-coloring to prevent an attacker of guessing the locations of the fragments. For handling the download request from user, the cloud manager collects all the fragments and re-assemble them into a single file. Later the file is sent to the user. In this methodology, we propose not to store the entire file at a single node. The DROPS methodology fragments the file and it makes use of the cloud for replication. The fragments are distributed such that none of the nodes in a cloud holds more than a single fragment, so that even an attack on the node leaks no significant information.

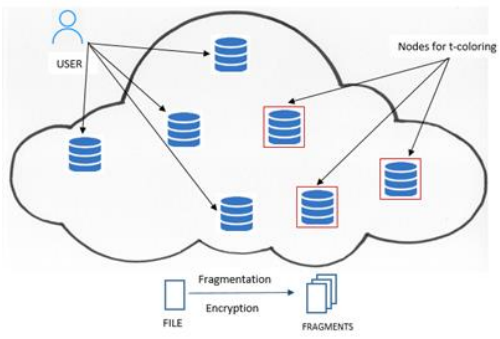

Fig. 1. Architecture diagram

5. Implementation of modules

A. *Registration*

Initially all the data holders are required to register their details in the organization. After registration, the data holder can upload the data to the cloud. The authorized user is provided access to download or view the encrypted file. If the user wants upload the file, then the user must login with their id.

B. *Data Uploading*

After login, the data holder can upload the data into the cloud by sending the uploading file with a request to the cloud server.

C. *Drops*

DROPS methodology utilizes the concept of data fragmentation for securing the user data within the cloud. To further enhance the security, the fragments are not stored on the neighbor nodes. To separate the storage of fragments by certain distance, the concept of T-coloring is used. the fragments are stored on the most central nodes. The selection of central nodes is done by evaluating the centrality measures for the nodes.

1) Data fragmentation

Consider a file with m number of fragments on a cloud with t nodes. Let ss be the number of successful intrusions on distinct nodes, such that ss>m. The probability that ss number of victim nodes contains all of the m sites storing the file fragments. This methodology fragments the file and makes use of the cloud for replication. The fragments are distributed such that none of the nodes in a cloud holds more than a single fragment, so that even an attack on the node leaks no significant information. This methodology uses replication where each of the fragments is replicated only once to improve the security. In this methodology, user sends the data file to cloud. The cloud manager system (a user facing server in the cloud that entertains user's requests) upon receiving the file performs: (a) fragmentation, (b) first cycle of nodes selection and stores one fragment over each selected node, and (c) second cycle of nodes selection for fragment replication. The cloud manager maintains a record of the fragment placement and is assumed to be a secure entity. Once the file is split into fragments, this methodology selects the cloud nodes for fragment placement. The selection is made by keeping an equal focus on security and performance in terms of the access time. We choose the nodes that are central to the cloud network to provide better access time. For the above said purpose, this methodology uses the concept of centrality to reduce access time.

2) Centralities Calculation

There are three centrality measures, namely: (a) betweenness, (b) closeness, and (c) eccentricity centrality. However, if all the fragments are placed on the nodes based on the descending order of centrality, then it is possible that adjacent nodes are selected for fragment placement.

3) T-coloring

Such a placement may provide clues to an attacker as to where other fragments might be present, decreasing the security

level of the data. To deal with the security aspects of placing fragments, we use the concept of T-coloring that was used for the channel assignment problem. We generate a non-negative random number and build the set T starting from zero to the generated random number. The set T is used to restrict the node selection to those nodes that are at hop-distances not belonging to T. For the above said purpose, we assign colors to the nodes, such that, initially, all of the nodes are given the open color. Once a fragment is placed on the node, all the nodes within the neighborhood at some distance belonging to T are assigned a color. Although increase in retrieval time could be observed, the data on cloud is kept highly secure.

D. Data Downloading

The cloud manager collects all the fragments from the nodes and reassemble them into a single file whenever a request is posted. Later, the file is sent to the user.

E. Encryption Module (NTRU Method)

Before placing the fragment of data into the storage node, NTRU encryption scheme executed. It consists of key generation, Data encryption and Decryption.

- KeyGen()→ (pk, sk): let $f \in R$, $g \in R$, while f, g follows the discrete Gaussian distribution, $f = 1 \bmod q$, and f is reversible. Thus, the secret key is denoted by $sk = f$; the public key is denoted by $pk = h = g \cdot f^{-1} \bmod q$.
- Enc($pk = h$, $\mu \in Rp$)→ $c \in Rq$: let $r \in R$,$m \in R$,$m = \mu \bmod p$. Both m' and r follow the discrete Gaussian distribution, and we have $m = p \cdot m' + \mu$, $c = p \cdot r \cdot h + m \bmod q$.
- Dec($sk = f$, $c \in Rq$) → μ: calculate $\_b = f \cdot c \bmod q$, and make it an integer polynomial b, with factors within $[-q/2, q/2)$. Thus, we have $\mu = b \bmod p$.

After Encrypted data, the data will be placed in the storage node.

6. Conclusion

The proposed system is a cloud storage security scheme that deals with the performance and security in terms of retrieval time. The data file was fragmented and the fragments are distributed over multiple nodes. The nodes were separated by the T-coloring. The fragmentation and distribution ensured that no significant information was obtainable by an adversary in case of an attack. No node in the cloud store more than a single fragment of the same file. The performance of this methodology was compared with full-scale replication techniques.

References

[1] W. K. Hale, “Frequency assignment: Theory and applications,” Proc. IEEE, vol. 68, no. 12, pp. 1497–1514, Dec. 1980.
[2] L. Qiu, V. N. Padmanabhan, and G. M. Voelker, “On the placement of web server replicas,” in Proc. IEEE Comput. Commun. Soc. 20th Annu. Joint Conf., 2001, vol. 3, pp. 1587–1596.
[3] A. Mei, L. V. Mancini, and S. Jajodia, “Secure dynamic fragment and replica allocation in large-scale distributed file systems,” IEEE Trans. Parallel Distrib. Syst., vol. 14, no. 9, pp. 885–896, Sep. 2003.
[4] T. Loukopoulos and I. Ahmad, “Static and adaptive distributed data replication using genetic algorithms,” J. Parallel Distrib. Comput., vol. 64, no. 11, pp. 1270–1285, 2004.
[5] M. Tu, P. Li, Q. Ma, I.-L. Yen, and F. B. Bastani, “On the optimal placement of secure data objects over Internet,” in Proc. 19th IEEE Int. Parallel Distrib. Process. Symp., 2005, p. 14.
[6] W. A. Jansen, “Cloud hooks: Security and privacy issues in cloud computing,” in Proc. 44th Hawaii IEEE Int. Conf. Syst. Sci., 2011, pp. 1–10.
[7] Y. Tang, P. P. Lee, J. C. S. Lui, and R. Perlman, “Secure overlay cloud storage with access control and assured deletion,” IEEE Trans. Dependable Secure Comput., vol. 9, no. 6, pp. 903–916, Nov. 2012.
[8] K. Bilal, M. Manzano, S. U. Khan, E. Calle, K. Li, and A. Zomaya, “On the characterization of the structural robustness of data center networks,” IEEE Trans. Cloud Comput., vol. 1, no. 1, pp. 64–77, Jan.–Jun. 2013.
[9] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya, “Energy-efficient data replication in cloud computing datacenters,” in Proc. IEEE Globecom Workshops, 2013, pp. 446– 451.
[10] A. N. Khan, M. L. M. Kiah, S. U. Khan, and S. A. Madani, “Towards secure mobile cloud computing: A survey,” Future Gener. Comput. Syst., vol. 29, no. 5, pp. 1278–1299, 2013.
[11] G. Kappes, A. Hatzieleftheriou and S. V. Anastasiadis, "Virtualization-aware access control for multitenant file systems," *2014 30th Symposium on Mass Storage Systems and Technologies (MSST)*, Santa Clara, CA, 2014, pp. 1-6.
[12] M. Ali *et al.*, "SeDaSC: Secure Data Sharing in Clouds," in *IEEE Systems Journal*, vol. 11, no. 2, pp. 395-404, June 2017.