|| Volume 3 || Issue 2 || February 2018 || ISSN (Online) 2456-0774

INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS

A Review on Ensuring Data Security in Cloud

Asst.Prof.B.K.Patil[1], Ms.Rajguru Ashwini Vasant[2]

Asst Prof.,Computer Science and Engg,.Everest Educational Society's College of Engineering & Technology Aurangabad, India[1]
Student,Computer Science and Engg,. Everest Educational Society's College of Engineering & Technology Aurangabad, India[2]
ashwinirajguru22@gmail.com[2]

Abstract— **In cloud system the data is outsourced on the cloud, this may create security issues. In this dissertation I propose Division and Replication of Data in Cloud (DRDC) which can take care of security issues without compromising the performance. In this system, file uploaded by the client is first encrypted then divided into fragments. Then these fragments are replicated over the cloud nodes. Fragmentation and replication is carried out in such a way that each node contains only a single fragment. Thus if any one of the node is intruded by hacker, no significant information is revealed, and thus security is maintained. To further increase the security, nodes are separated by T-coloring graph method. Due to the T-coloring, the effort needed by an attacker to breach the security is increased multiple times.**

Keywords:- cloud Storage, distributed file system, Division and Replication of data(DRDC),T-coloring.

I INTRODUCTION

Cloud computing is an internet-based computing technology, where shared re-sources such as software, platform, storage and information are provided to customers on demand. Cloud computing is a computing platform for sharing resources that include infrastructures, software, applications, and business processes. Cloud Computing is a virtual pool of computing resources. It provides computing resources in the pool for users through internet. Cloud computing, as an emerging computing paradigm aims to share storage, computation, and services transparently among massive users. The exact definition of cloud computing is a large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet [1].

Current cloud computing systems pose serious limitation to protecting user's data confidentiality. Since user's sensitive data is presented in unencrypted forms to re-mote machines owned and operated by third party service providers, the risks of unauthorized disclosure of the user's sensitive data by service providers may be quite high. There are many techniques for protecting user's data from outside attackers. An approach is presented to protecting the confidentiality of user's data from service providers, and ensures that service providers cannot collect user's confidential data while the data is processed and stored in cloud computing systems. Cloud computing systems provide various internet based data storage and services. Along with the rapid growth of the Internet with the rise of the era of cloud computing, concerns about Internet Security continue to increase. To address this problem we propose the design of a system that will capture the movement of information on the cloud. We will be identifying whether there is a need for some type of security capture device/measure on the cloud, which will allow users to know whether their information is secure and safe without comprising from threats and attack.

II RELATED WORK

M. Tu et al. presented a secure and optimal placement of data objects in a distribution system. The encryption key is divided and division is done through threshold secret sharing scheme. This scheme pays attention to the replication problem with security and access time improvement. In this scheme data files are not fragmented and are handled as a single file. This scheme mainly focuses on encryption key security unlike our methodology [2].Number of requests and the response time are considered as main points for deciding in which site within the region the file has to be placed. Therefore, their strategy increases the data availability and also reduces the number of unnecessary replications [3]. Any weak entity can put the whole cloud at risk. In such a scenario, the security mechanism must substantially increase an attacker's effort to retrieve a reasonable amount of data even after a successful intrusion in the cloud. Moreover, the probable amount of loss (as a result of data leakage) must also be minimized [4]. data enter database (Data enter DB) is used to replicate the most frequently used data items from the central database. Each rack hosts at least one server capable of running local rack-level database (Rack DB), which is used for replication of data from the data enter database [5]. Fragmentation consists in splitting the attributes of a relation R producing different vertical views (fragments) in such a way that these views stored at external providers do not violate confidentiality requirements (neither directly nor indirectly)Note that singleton constraints are correctly enforced only when the corresponding attributes do not appear in any fragment that is stored at a cloud provider [6]. The node separation is ensured by the means of the T-coloring. To improve data retrieval time, the nodes are selected based on the centrality measures that ensure an improved access time. To further improve the retrieval time, we judicially replicate fragments over the nodes that generate the highest read/write requests [7].

A technique presented by Juels et al. involves data migration to the cloud by iris file system. To maintain the integrity and freshness of data, a gateway application uses the merkle tree This system is unable to prevent data loss or access by other VM [8].G. Kappes et al. in their presentation addresses virtualized and multitenancy related problems. DRDC methodology involves fragmentation of data file and multi nodal storage of single file and thus prevents leakage of critical information [9].D. Zissis et al. in their paper advises the use of a third party for security of cloud data. They use public key infrastructure, so that the level of trust is increased in the communication between the involved parties. However tampering and loss due to virtualization and multitenancy are not prevented [10].Although the greedy algorithm depends upon estimates of client distance and load predictions, they find that it is relatively insensitive to errors in these estimates and therefore is a viable algorithm for use in the general Internet environment where workload information will always be imperfect [11]. Each entity is individually untrusted as long as a majority of them can be trusted. Therefore, an intrusion into a part of the system will have no consequence on the system security if only a minority of the security entities is affected by the intrusion this approach is thus intrusion-tolerant[12].

III PROPOSED SYSTEM

In Proposed System Cloud Manager performs following functions:

- Receiving the file
- Encryption of file by AES Algorithm
- Fragmentation and nodal selection
- Second cycle of nodal selection for fragments replication.

As soon as file is divided into fragments, this system assigns the cloud nodes for each fragment. Centrality measures are employed to reduce the retrieval time. Three centrality methods are mainly used; these are between’s, closeness and eccentric centrality. These measures may result in placement of fragments on adjacent nodes, thus compromises security. This is where the concept of T-coloring is justified. In this concept a set T is built starting from zero to random positive number. To make this system work, colors are given to the nodes. Let’s consider there’s an open_color before placing the fragment, as soon as fragment is placed on one of the node, then close_color is given to nodes surrounding the assigned node up to the T distance. This system makes cloud more secure, although somewhat performance is decreased due to less availability of central nodes.

Further in order to increase performance or to decrease the access time, data replication is done in a controlled manner. Again while placing the replicated fragments; concept of T-coloring is used. In data replication, some of the fragments may not be replicated due to T-coloring due to less number of nodes. If client requests to download the uploaded file, then all the fragments are reassembled into a single file by cloud manager and then that file is sent to the client.

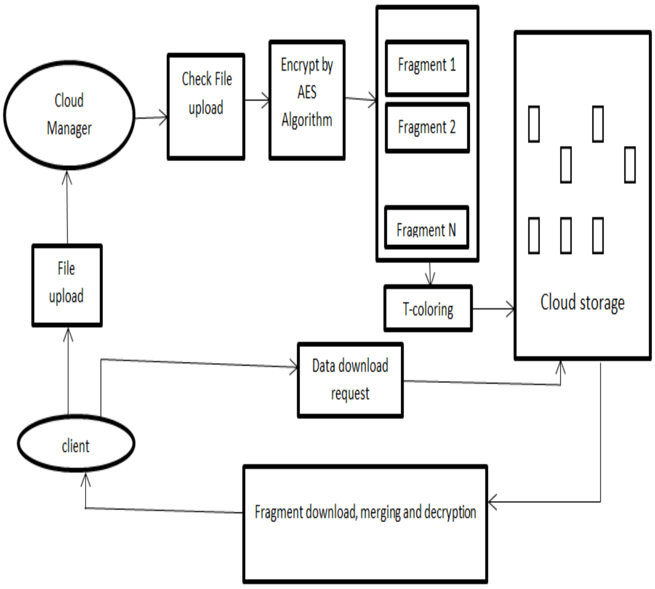

Figure 1: DRDC Framework

IV COMPONANTS OF DRDC

A. Fragmentation engine

The fragmentation engine is an initial component which is located in the cloud storage. This component is mainly used for fragmenting the user files. Initially, data owner will upload a file to cloud storage system. The cloud manager will collect a file and encrypt the file using cryptographic techniques. Convert that encrypted file into ‘n’ number of fragments using fragments engine.

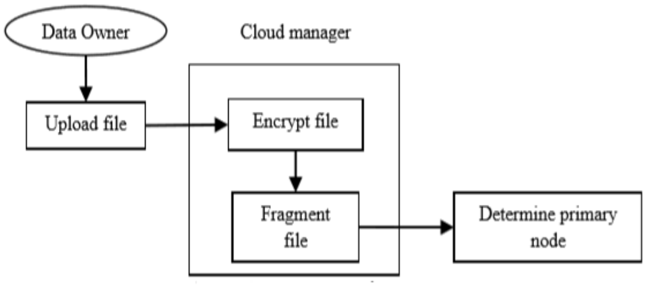

Figure 2: Fragmentation engine

B. T-coloring algorithm

The node identifier component uses the T-coloring concept for determining the available and unavailable nodes for allocation of fragments. Once the file is fragmented into number of pieces, the primary node is determined. Then, a non-negative random number will get generate and build the set T starting from zero to the generated random number. After that, the primary fragment should be placed in the cloud storage, based on centrality measures.

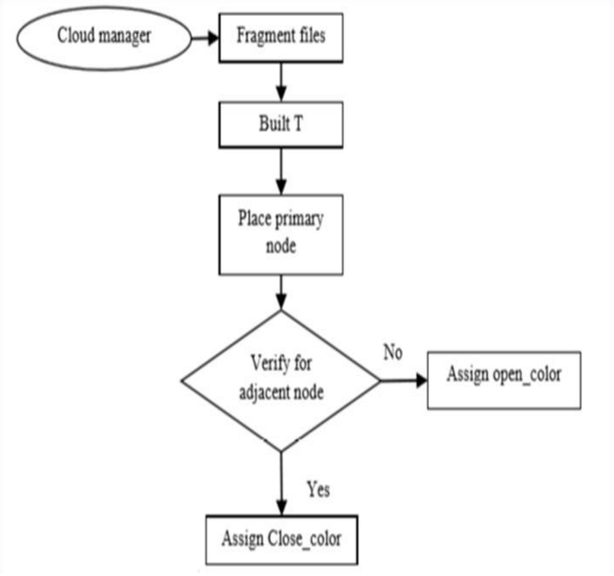

Figure 3: T-coloring algorithms

C. Nearest-Neighbourhood Node Identification

After determining the available nodes and unavailable nodes, the 'n' number of file fragments should be placed in a distinct node. Each and every fragment should have a size. Every single storage node calculate read and write a fragment, same time primary node stores primary copy of fragment. Here, every storage area has two field records, first field is to store primary node for primary fragment using centrality measures, and second field is to identify the nearest neighbour node for storing kth fragment data using T-coloring and centrality measures.

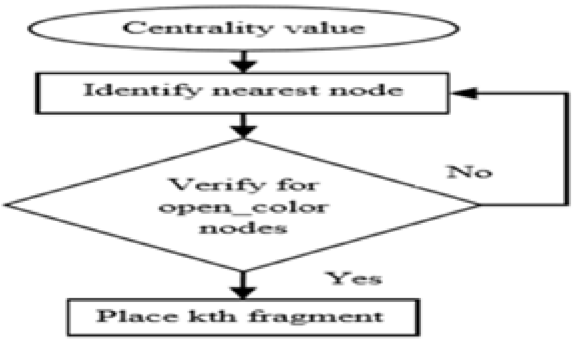

Figure 4:Nearest-neighborhood node identification

D. Cloud management system

Once all the fragments are placed on its appropriate locations, the cloud manager should maintain all the nodes in the cloud. Cloud storage has a different unique storage in different region. This all node should be followed by a single primary node that represent first placement of fragment. Then, T-coloring algorithm is used to plan the remaining nodes and also it uses centrality measures.

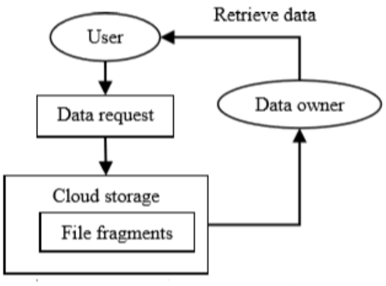

Figure 5: Cloud management systems

V CONCLUSION

In this we proposed a secured system for storage of data in cloud that is also very good in performance. The file which is uploaded on the server first encrypted, and then fragmentation and replication of fragments takes place. Nodes are assigned to the fragments and replicas with the help of T-coloring. Fragments are placed over the nodes in such a way that no node contains more than one fragment. This system of fragmentation and T-coloring increases the effort of an attacker to intrude the system. Even in case of successful attack on a fragment, no significant information is revealed to an attacker. In addition to a high level of security, performance of this system is also very good. There are replicas of fragments which result in fast retrieval of data. This replication is also in a controlled manner, so that performance is increased without compromising security.

REFERENCES

[1] K. Bilal, S. U. Khan, L. Zhang, H. Li, K. Hayat, S. A. Madani, N. Min-Allah, L. Wang, D. Chen, M. Iqbal, C. Z. Xu, and A. Y. Zomaya, "DROPS: Division and Replication of Data in Cloud for Optimal Performance and Security," Concurrency and Computation: Practice and Experience, Vol. 25, No. 12, 2015, pp. 1771-1783.

[2] T. Loukopoulos and I. Ahmad, "Static and adaptive distributed data replication using genetic algorithms," Journal of Parallel and Distributed Computing, Vol. 64, No. 11, 2004, pp. 1270-1285.

[3] A. Juels and A. Opera, "New approaches to security and availability for cloud data," Communications of the ACM, Vol. 56, No. 2, 2013, pp. 64-73.

[4] G. Kappes, A. Hatzieleftheriou, and S. V. Anastasiadis,

"Dike: Virtualization-aware Access Control for Multitenant Filesystems," University of Ioannina, Greece, Technical Report No. DCS2013-1, 2013.

[5] L. M. Kaufman, "Data security in the world of cloud computing," IEEE Security and Privacy, Vol. 7, No. 4, 2009, pp. 61-64.

[6] S. U. Khan, and I. Ahmad, "Comparison and analysis of ten static heuristics-based Internet data replication techniques," Journal of Parallel and Distributed Computing, Vol. 68, No. 2, 2008, pp. 113-136.

[7] A. N. Khan, M. L. M. Kiah, S. U. Khan, and S. A. Madani, "Towards Secure Mobile Cloud Computing: A Survey," Future Generation Computer Systems, Vol. 29, No. 5, 2013, pp. 1278-1299.

[8] A. N. Khan, M.L. M. Kiah, S. A. Madani, and M. Ali, "Enhanced dynamic credential generation scheme for protection of user identity in mobile-cloud computing, The Journal of Supercomputing, Vol. 66, No. 3, 2013, pp. 1687-1706 .

[9] A. Mei, L. V. Mancini, and S. Jajodia, "Secure dynamic fragment and replica allocation in large-scale distributed file systems," IEEE Transactions on Parallel and Distributed Systems, Vol. 14, No. 9, 2003, pp. 885-896.

[10] L. Qiu, V. N. Padmanabhan, and G. M. Voelker, "On the placement of web server replicas," In Proceedings of INFOCOM 2001, Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 3, pp. 1587-1596, 2001.

[11] D. Sun, G. Chang, L. Sun, and X. Wang, "Surveying and analyzing security, privacy and trust issues in cloud computing environments," Procedia Engineering, Vol. 15, 2011, pp. 2852 2856.

[12] Y. Tang, P. P. Lee, J. C. S. Lui, and R. Perlman, "Secure overlay cloud storage with access control and assured deletion," IEEE Transactions on Dependable and Secure Computing, Vol. 9, No. 6, Nov. 2012, pp. 903-916.