International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 12 | Dec 2018 www.irjet.net p-ISSN: 2395-0072

Detecting Data Leakage and Implementing Security Measures in Cloud Computing

Syeda Hajira Tabassum[1], Soumya Naik[2]

[1]MTech, Dept. of Computer Science & Engineering, City Engineering College, Karnataka, India
[2]Assitant Professor, Dept. of Computer Science & Engineering, City Engineering College, Karnataka, India

---

Abstract – *Most of the data has been proceed through third-party application (TPA) and user are unaware about the security essentials in cloud computing. The data owner will store the data in the cloud. Every user must have registered in the cloud. Cloud provider must verify the authorized user. If someone try to access the account, data will get leaked. This leaked data will present in an unapproved site as in on the internet or someone's computer. In this paper, to address performance and security issue. We propose DROPS methodology, abbreviated as Division and Replication of Data in the Cloud for Optimal Performance and Security. In this methodology, we have to select the file and then store the particular file in the cloud account. In order to provide security, we are going to implement DROPS concepts. Now we divide the file into various fragments based on the threshold value. Each and every fragments are stored in the node using T-Coloring. After the placement of fragments in node, it is necessary to replicate each fragments for one time in cloud.*

***Key Words*: DROPS, leaked, cloud computing, data leakage, T-Colouring,**

1. INTRODUCTION

Every company focus on security issues of securing the data from different third parties form being out sourced. Every company follows a different strategy which does not match with any other company. The employees in the organizations are being trained to support the data secrecy and the fundamental structure of the organization. The safety must be beyond the employees' knowledge so that the employee has no idea of cracking it by covering logical and physical security. Information security is frequently subjected to metaphors. The information security must be targeted at global level by not letting the user know the problematic issues faced by the security department and also the logical security i.e. the sensitive data, applications and also the operating system used in a particular institute. The security must also be extended to telecommunication department of the institute so that they have a network security also.

Data Leakage may take place at any time, there is no designated time. Data leakage rely only on the significance of sharing information by the agent. The information distributed is considered as sensitive data when it consists of information about the client, budget, code and any design specification. If the leakage occurs, it leaves the institute in unprotected state. This data leakage places the institute or organizations in the state of uncertainty resulting in the decry of the business and eventually defeat of the company.

Cloud computing security is growing in information security, computer security. In most cases it is recommended that based on the segment of risks security of data control to be choose and executed, by evaluating the vulnerabilities threats, and influences. The security concerns of the cloud can then be organized in different ways; Gartner [9] discover 14 regions of distress during the alliance of Cloud Security.

Traditionally, Watermarking technique is used to handle data leakage detection. For instance, in every distributed copies a unique code is embedded using different types of watermarking algorithm. If this copy at later time is found by any unauthorized person, the person can't be identified. This result in the major drawback in watermarking technique. Though Watermarking technique is advantageous in real-time domain, yet again, few changes in the original information. In no time, if the recipient of the data is malicious, watermarks will be destroyed. Security is one of the most crucial aspects in cloud computing. Hence this prohibit the adoption of cloud computing. Hence, in this paper, we addressed the issue of security and performance. We proposed Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) this will break the owner's files into various sections and replace them in the cloud space. The owner of the data decides the threshold value, based on this value fragments of a file is performed. So that the meaning full information is not held by the fragments.

Any attack that is successful on any node should not hint the position of other fragments present within the cloud space. This in turn make the intruder cannot predict locations of file fragments so that it improves the security in the cloud. By using T-Coloring concept is used to do the node separation in order to select nodes in such a way that nodes

are not adjacent and are at obvious distance from each other. Furthermore, the fragments storing in the nodes are placed with an obvious distance with the aid of graph T-Coloring in order to block an intruder from accessing the site of the fragments.

2. LITERATURE SURVERY

Prashant Khobragade [8] [10] proposed that to detect the region where the attack has been taken place it is required to inspect data in the browser history. The inspected data is then brought together and put into database as a proof. The suggested method and forensic toolkit helps to analyze the data for the law enforcement.

Sandip A. Kale, S.V. Kulkarni [2] focuses on robust watermarking technique, this technique will be useful in few cases, yet again, few modification of the original data is required. Moreover, watermarks will be destroyed if the data recipient is vengeful. The agent can add fake items to the spread data to improve the potency in finding criminal agents.

Sushilkumar, B.Shinde, Archana. Bhosale [3], shows that in different institutes and organizations data leakage is spotted as the very serious and big challenge. They have further explained that sellers add fake information that is shared among the consumers in an attempt to increase the chance to detect criminal agents. This is detection of data leakage is then handled by an algorithm.

Data leakage [7] [12] takes place in everyday life this happens when private business information like information of consumer, technical code or design description, amount lists, intellectual property and secrets of trade and spreadsheets of budgets are leaked. When this information or data is dripped out it puts the company in the state of unprotected. This leads business in an unsafe position, Putting the company in serious risk due to data leakage.

3. EXISTING SYSTEM

Water marking Technique: is a type of security technique which deals with the idea of embedding a particular code or encryption on the information that is to be distributed. The information can be image or a video or any official file. This encryption helps the company to claim the ownership on any particular data. Water marking is a technique where a bit pattern is added to the data at a particular position on the tuples and subset of the data. The tuple and subset and their attributes are algorithmically coded in such a way that they are controlled by a key which can be accessed only by the owner.

Fake Objects Method: this method causes less problems to the real objects by introducing fake objects in some applications. Consider an example, where hospitals are agents and medical records are distributed data items. Here in the records even the small change is undesirable. Still, the addition of fake information to the records will not affect as this data would not be matching any patient record and hence no patient would be treated on the basis of fake records injected. In this case, company A Company X sells products to company Y a list to send advertisements be used once. Company X then injects trace data which contain addresses possess by company X. Thus, every time company Y purchases the product through list, company will receive a copies of the mails. This trace data are fake items that enables to identify the misuse of data.

4. PROPOSED SYSTEM

Detecting data leakage: The data owner first registered into the cloud account. Each and every user has to registered into the cloud. Now the data owner and user will become the authorized person. The data owner will upload the file into the cloud. Now the data owner login into the account, at that time the cloud provider verifies the already registered owner or not. If they are registered owner and then they will transfer the file to the users who have registered. The user can now login into their account, cloud provider again verifies the user. The user will download the file sent by the data owner.

If someone try to copy the URL, the data get leaked in someone's laptop. Now the details about the unauthorized person will be tracked. This tracked information sent as a mobile intimation to the data owner. The mobile intimation will hold information like IP address, MAC address and GPS location.

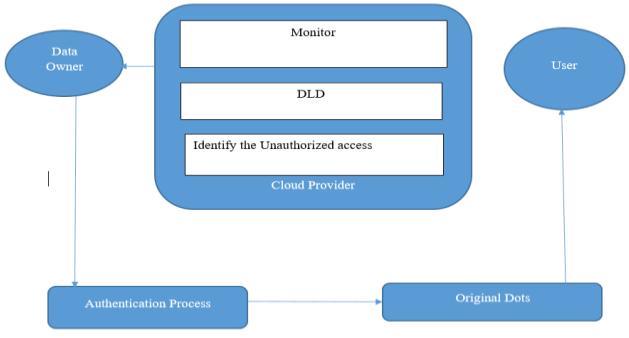

Fig -1: Data Leakage Architecture

Drops Methodology: In DROPS methodology, we are not storing the entire file in cloud space. Now we are splitting the entire file into various fragments. These fragments have to distribute in the cloud space. Each and every fragment has to place in a particular node. So that each node contains only a single fragment. In each successful attack the node will not reveal the significant information. After the fragmentation process replication will takes place. In replication process, each fragment has to replicate its content once in the cloud space. In this way in the cloud computing we can achieve security. In DROPS methodology, data file is sent to the cloud space by the user. Cloud manager performs the following on file receiving: (1) File Fragmentation (2) Selecting Nodes (3) Fragment stores (4) Nodes selection for fragments replication.

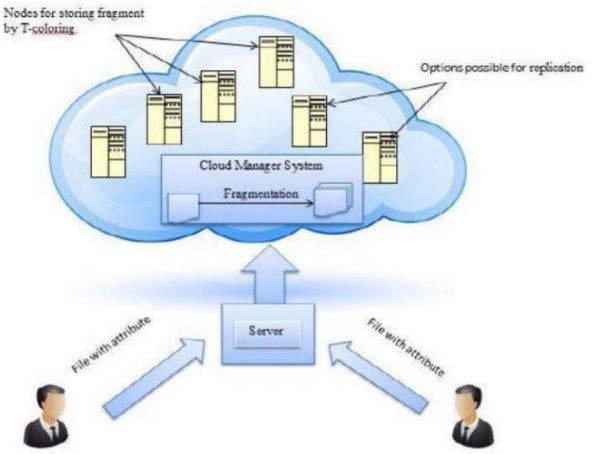

Fig -2: DROPS Methodology

Drops Implementation

File Fragmentation: In cloud, security is the major aspects for a large-scale system. This provides the system security as individual nodes and whole. On every successful action of intruding into single node it provides more result or effect for data and other nodes. A successful intrusion may lead to software failure. It may also lead to administration defenseless. File fragmentation is a term that describes a group of files that are scattered throughout the cloud. The data owner splits the file into various pieces called fragments. The size of fragmentation is decided by data owner using threshold value. In various aspects the threshold value can be fixed, they are, 1. Decide percentage based on file size. 2. Numbering of each fragments. For example, first fragment the file by 10% of total size or split the file by using various size like fragment 404MB, fragment 1000MB. It is also possible to number the fragments like fragment1, fragment2 etc., The numbering of each fragments only known to the data owner. There are three types of fragmentation: 1. Horizontal. 2. Vertical. 3. Mixed (Hybrid).

Fragment Placement: In order to give security when putting the fragments in to the cloud, the T-coloring concept is essential. Channel task problems mainly uses T-Coloring. This will generate a positive number, random number, to build the set T by generating random number from zero as start. It assigns colors to the each and every node, in the beginning, all nodes will be in open color. The fragment is then placed on particular node, all the nodes in the neighborhood are at certain distance which will be belonging to set T and these are allocate to close color. In this series of actions, this will lose few central nodes in cloud so that may rise time to retrieve. If anyhow the intruders try to track the node position and take the fragment, user cannot make out the precise location of other fragments. The intruder ends up just predicting the location of other fragments in cloud. Therefore, T-coloring concepts are used to separate the nodes.

Fragment Copying/Reproducing: In replication process, unique copy exists and the same copy will exist once again. The replication process is used to elevate the availability of data and their by improve the time to retrieve. This carry out a controlled reproducing. In replication process, copies of the same data item have the same value. There are three types of replication, 1. No replication. 2. Fully replicated. 3. Partial replication. It positions the fragment on the node that gives the limited cost access. Hence improve the time to retrieve the access of the fragments and rebuilding of the original file. While reproducing the fragment, by T-coloring fragment separation is achieved, it also takes care of node placement. In instance of a large counts of fragments or small count of nodes, there is a possibility that few of the fragments could be left without being reproducing due to the use of T-coloring concepts.

T-Coloring: T-coloring forbid keeping the fragment and also avoid storing a fragment in the node of a neighborhood, this results in eliminating the nodes that is used to retain. In such scenarios, where only the remaining fragments, those nodes are selected which are not possessing any fragments for random storage. T-coloring concept is applied to separate nodes. The process of fragmentation made sure no successful information is obtained in successful attack. In the cloud, no node stored the instances of same file of single fragment. The DROPS methodology execution is compared with full-scale reproducing techniques. This outcome draws the attention on the performance and security. This rise in the level of data security.

5. CONCLUSION

DROPS methodology is propped security scheme in the cloud storage that deals with the performance and security that increase the time to retrieve. Data in the files were fragmented and these fragments are spread over multiple or many nodes. The fragmented file will be replicated their by increasing data availability and provide security to the cloud. The fragmented file will be replicated their by increasing data availability and provide security to the cloud.

REFERENCES

[1] Chandu Vaidya and Prashant Khobragade,2015, "Data Security in Cloud Computing", ISSN: 2321-8169 .

[2] Sandip A. Kale, Prof. S.V.Kulkarni," Data Leakage Detection", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 1, Issue 9, November 2012.

[3] Prof. Sushilkumar N. Holambe, Dr.Ulhas B.Shinde, Archana U. Bhosale,"data Leakage Detection Using Cloud Computing ", International Journal Of Scientific & Engineering Research, Volume 6, Issue 4,(April-2015).

[4] P. Papadimitriou and H. GarciaMolina, "Data Leakage Detection," technical report, Stanford Univ., 2008.

[5] Hartung and kutter,.Watermarking technique for multimedia data.2003.

[6] Priyanka Barge, Pratibha Dhawale, Namrata Kolashetti," A Novel Data Leakage Detection", International Journal of Modern Engineering Research (IJMER) ISSN: 2249-6645, Vol.3, Issue.1, Jan-Feb. 2013.

[7] Archana Vaidya, Prakash Lahange, Kiran More, Shefali Kachroo and Nivedita Pandey," Data Leakage Detection", International Journal of Advances in Engineering & Technology, ISSN: 2231-1963, March 2012.

[8] Prashant Khobragade, Latesh G. Malik,"A Review on Data Generation for Digital Forensic Investigation using Datamining", IJCAT International Journal of Computing and Technology, Volume 1, Issue 3, April 2014.

[9] Jon Brodkin, "Gartner: Seven cloud-computing security risks". InfoWorld. 2008-07-02. Retrieved 2010-01-25.

[10] Khobragade, P. K., & Malik, L. G., "Data Generation and Analysis for Digital Forensic Application Using Data Mining". In Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on (pp. 458-462). IEEE. April, 2014.

[11] Iqbal, C. Z. Xu, and A. Y. Zomaya, "Quantitative comparisons of the state of the art data center architectures," Concurrency and Computation: Practice and Experience, Vol. 25, No. 12, 2013, pp. 1771-1783.

[12] Chandu Vaidya etl. & BE scholars "Data leakage Detection and Dependable Storage Service in cloud Computing" IJSTE volume 2 issues 10 April 2016 ISSN online 2349-784X