A Secure Data in Cloud Storage using Lightweight Approaches

O.V. Narayanan , M.E
CSE Department
SRR Engineering College,Chennai, India

Mrs. R. Ramya, M.E.,
Assistant Professor of CSE Department
SRR Engineering college, Chennai, India,

Abstract: **Cloud computing is an emerging technology which attracts customers by giving offers like reduced cost, space and virtually unlimited dynamic resources for storage, computation etc. User shared the sensitive data over the cloud which gives rise to security issues in cloud computing Outsourcing data to a third-party administrative control as is done in cloud computing gives rise to security concerns. The data compromise may occur due to attacks by other users and nodes within the cloud. Therefore, high security measures are required to protect data within the cloud. However, the employed security strategy must also take into account the optimization of the data retrieval time. In this project, protect user's data a secure methodology fragmentation and replication of data is used in this paper. The data is fragmented into pieces and then replicate them over the cloud nodes for maintaining the availability, performance level and backing up the data. T-coloring term is used here which is not giving any idea about locations of the fragments to an attacker. Centrality is used for node selection process.**

Key Terms: Fragmentations, Replications, T – Coloring

I. INTRODUCTION

Outsourcing of data on the cloud computing system is new trend in the IT field to overcome from issues of Data security, Data maintenance and to reduce the cost of Data storage equipment. Also the cloud data can be remotely accessible to every authorized user from any geographic area. Once the data is placed by data owner on cloud system he /she lost their control from data , hear is the new data security and data confidentiality problems comes in front of Data owner and cloud service provider. Cloud service provider should provide the maximum data security. As much as burden increased on the cloud provider so that we purposed this system called Dynamic Data Possession In cloud Computing System Using Fragmentation and Replication, in this system main focus is on the cloud data security and providing the assurance to the cloud users that their data is secured at cloud system.

Security is one of the most crucial aspects among those the wide-spread adoption of cloud computing. Cloud security issues sustained due to the core technology implementation as like virtual machine (VM) escape or session riding, etc. The service offerings by cloud as structured query language injection or weak authentication schemes and cloud characteristics like data recovery vulnerability and Internet protocol vulnerability, etc.) For each of the cloud nodes use the term node to represent computing, storage, physical, and virtual machines) contains a distinct fragment to increase the data security. A successful attack on a single node must not reveal the locations of other fragments within the cloud. The selection of the nodes is performed in two phases. In the first phase, the nodes are selected for the initial placement of the fragments based on the centrality measures. In Second phase, the nodes are selected for replication. The working of the DROPS methodology is shown as a high-level work flow and comparative techniques to the DROPS methodology. The various implemented replication strategies are: (a) A-star based searching technique for data replication problem (DRPA-star) (b) Weighted A-star (WA-star), (c) A-star, (d) Suboptimal A-star1 (SA1) (e) suboptimal A-star2 (SA2), (f) Suboptimal A-star3 (SA3) (g) Local Min-Min, (h) Global Min-Min, (i) Greedy algorithm, and (j) Genetic Replication Algorithm (GRA). Here three Data Center Network (DCN) architectures, namely: (a) Three tier, (b) Fat tree, and (c) D Cell. We use the aforesaid architectures because they constitute the modern cloud infrastructures and the DROPS methodology is proposed to work for the cloud computing paradigm.

II. LITERATURE SURVEY

A. Replication for Improving Availability & Balancing Load in Cloud Data Centers:

A large number of replication strategies for management of replicas have been proposed in literature. As a result of replication, data replicas are stored on different data nodes for high reliability and availability. Replication factor for each data block and replica placement sites need to be decided at first. A replication strategy for region based framework based on the demand of files over a geographically distributed Grid environment. Access frequency of each file is calculated and on that basis it is determined that in which region the replicas need to be placed and the number of replicas need to be placed. When a file is created, the access frequency is calculated for each region and replicas are placed in the regions with the large order of the access frequency. 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.

B. Quantitative comparisons of the state of the art data center architectures, Concurrency and Computation: Practice and Experience:

The data outsourced to a public cloud must be secured. Unauthorized data access by other users and processes (whether accidental or deliberate) must be prevented. 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.

C. Encryption and fragmentation for data confidentiality in the cloud:

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). Intuitively, fragmentation protects the sensitive association represented by an association constraint c when the attributes in c do not appear all in the same (publicly available) fragment, and fragments cannot be joined by non-authorized users. 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.

D. Division and Replication of Data in Cloud for Optimal Performance and Security:

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.

III. PRELIMINARIES

1. Cloud Security

With the increasing popularity of cloud computing, technology experts along with security specialists are always trying different standards to secure their infrastructure in cloud locking them from outside networks. As of today there are no universal perfect solutions for cloud security. Cloud devices are likely to attacks from cyber criminals.

Data privacy and data protection are major concerns for any security expert in an organization regarding their infrastructure in the cloud. Data may not get stored in the same system within a public or community cloud, resulting in multiple concerns legally. As of today, there are no safety standards and regulations stated by the providers for the customers to ensure sufficient security. Virtualization security, identity and access management, threat management, content security, and data privacy need to be given priority and require more focus.

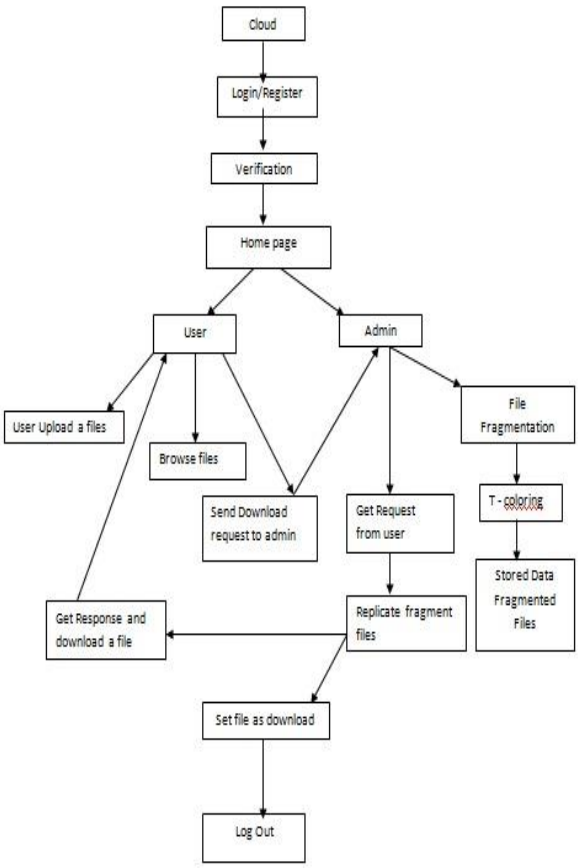

Fig.1. Overview Diagram

Data encryption all through the lifecycle can be one method of data protection. In cloud we are not known where our data resides, what we know is these are shared servers. Our information is shared among cloud nodes. So the chances are that the data might be leaked. It is important to have a strong security strategy for giving relief. This can curb data leakage and protect your valuable data.

2. Data Fragmentation

In large –scale system the security depends upon the whole system as well as the single node of a system. If the file is attacked by an attacker then there will be single point failure. So to avoid this data division or fragmentation technology is used. Fragmentation can increase an attacker's effort. In a fragmentation schema file f is split into n fragments, all fragments are signed and distributed to n remote servers, one fragment per server. The user can reconstruct file f by accessing m fragments arbitrarily chosen.

3. Replication

The tremendous growth of cloud computing enabled the deployment of immense IT services that are built on top of geographically distributed platforms and offered globally. For better reliability and performance, resources are replicated at the redundant locations and using redundant infrastructures. , Number of data replication methods have been proposed to address an exponential increase in Internet data traffic and optimize energy and bandwidth in datacentre systems.

Availability is assured by replication, without encryption, with the idea that files can be encrypted by the client before storing when confidentiality is an issue. Data replication means maintaining multiple copies of same data on same server or on different servers. If data is present at one site only, then it will be single point failure. Server will face a heavy load balancing condition and system performance. Also if that site fails, all that data will be lost, this is also a serious concern. Replication is necessary for maintaining the availability, performance level, backing up the data and also for balancing load.

4. T-coloring

T-coloring is basically used for channel assignment, such that the channels are separated by a distance to avoid interference. Suppose we have a graph G= (V, E) and a set T containing non-negative integers including 0. The T-coloring is a mapping function f from the vertices of V to the set of non-negative integers, such that | f(x) - f(y) | ∉T, where(x, y) ∈E. The mapping function f assigns a color to a vertex. In simple words, the distance between the colors of the adjacent vertices must not belong to T.

5. Centrality Measures

The centrality of a node in a graph provides the measure of the relative importance of a node in the network. The objective of improved retrieval time in replication makes the centrality measures more important.

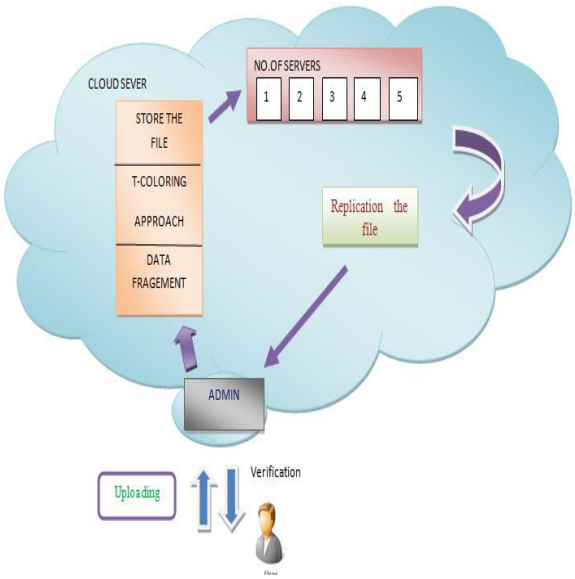

Fig 2. Proposed Work

I. Proposed System Work Flow

So here in proposed System the data is fragmented and replicate to achieve both secrecy and ideal performance. The Centrality and T-coloring method is used for node selection so it prohibits the single point failure. In this methodology, The propose to store the fragments on different nodes. After the fragmentation of file, it will be used for replication. The fragments are distributed such that no node in a cloud holds more than a single fragment, so that even a successful attack on the node then also no valuable information leaks.

This methodology uses controlled replication where each of the fragments is replicated only once in the cloud to improve the security. Although, the controlled replication does not improve the retrieval time to the level of full-scale replication, it significantly improves the security.

Firstly, in this methodology user sends the data file to cloud. The cloud manager system upon receiving the file performs: (a) fragmentation, (b) first cycle of nodes selection and stores one fragment over each of the selected node, and (c) second cycle of nodes selection for fragments replication. The cloud manager maintains record of the fragment placement and is assumed to be a secure entity. The fragmentation threshold of the data file is specified to be generated by the file owner. The file owner can specify the fragmentation threshold in terms of either percentage or the number and size of different fragments. By using this method, we can secure the data and also increases the performance level.

IV CONCLUSION

The propose a methodology which deals with cloud storage security and optimal performance in terms of retrieval time. Before uploading file we are fragmenting that file into multiple fragments and allocate that fragments using T-coloring technique in cloud. This provides security at client level as well as in network layer. The fragmentation and dispersal ensured that no significant information was obtainable by an adversary in case of a successful attack. No node in the cloud, stored more than a single fragment of the same file. To protect the original data privacy against the TPA, we randomize the coefficients in the beginning rather than applying the blind technique during the auditing process. Considering that the data owner cannot always stay online in practice, in order to keep the storage available and verifiable after a malicious corruption, we introduce a semi-trusted proxy into the system model and provide a privilege for the proxy to handle the reparation of the coded blocks and authenticators.

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, "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.

[2] 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 Transactions on Cloud Computing, Vol. 1, No. 1, 2013, pp. 64-77.

[3] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya, "Energy-efficient data replication in cloud computing datacenters," In IEEE Globecom Workshops, 2013, pp. 446-451. .

[4] K. Hashizume, D. G. Rosado, E. Fernndez-Medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," Journal of Internet Services and Applications, Vol. 4, No. 1, 2013, pp.