ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

A Review on Achieving Security by Fragmentation and Replication of Data

Ashlesh S. Patole[1]; Shripadrao Biradar[2]

M.E., Dept. of Computer , RMD Sinhgad School of Engineering, Pune., India[1]

Assistant Professor, Dept. of Computer, RMD Sinhgad School of Engineering, Pune., India[2]

ABSTRACT: Cloud Computing (CC) is an emerging trend that offers number of important advantages. One of the fundamental advantages of CC is pay-as-per-use, where customers will pay only according to their usage of the services. Currently data generation is improving users storage availability. There is need to outsource such big amount of data. There are many Cloud Service Providers(CSP). CSP is growing trend for numbers of customers and organizations reduces the burden of local data storage and maintenance. 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 paper, we propose Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that collectively approaches the security and performance issues. In the DROPS methodology, we divide a file into fragments, and replicate the fragmented data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker.

Moreover, the nodes storing the fragments, are separated with certain distance by means of graph T-coloring to prohibit an attacker of guessing the locations of the fragments. Furthermore, the DROPS methodology does not rely on the traditional cryptographic techniques for the data security; thereby relieving the system of computationally expensive methodologies. We show that the probability to locate and compromise all of the nodes storing the fragments of a single file is extremely low. We also compare the performance of the DROPS methodology with ten other schemes. The higher level of security with slight performance overhead was observed.

KEYWORDS: Centerality, cloud security, fragmentation, replication, performance, T-coloring.

I. INTRODUCTION

The employed security strategy must also take into account the optimization of the data retrieval time. Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that collectively approaches the security and performance issues. In the DROPS methodology, we divide a file into fragments, and replicate the fragmented data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker. Moreover, the nodes storing the fragments are separated with certain distance by means of graph T-colouring to prohibit an attacker of guessing the locations of the fragments. Data replication is one of the most important technique used for removing theidentical copies of repeating data and it is used in the cloud storage for the purpose of reduce the storage space.However, there is only one copy for each file stored in cloud even if such file is owned by huge number of users. Keeping the multiple data copies with similar content replicationeliminates redundant data by keeping only one physical copy and refer other redundant data to that copy. Data replicationcan be file level or block level. The duplicate copies of identical file eliminates by file level de-duplication. And block level replicationeliminates duplicate blocks of data that occur in non-identical files. This system allocates the file fragments using T-colouring graph technique. To maintain integrity we are providing the Third Party Auditor scheme which makes the audit of file stored at cloud and notifies the data owner about file status stored at cloud server.This system supports security challenges such as authorized duplicate check, integrity, data confidentiality and reliability.

II. RELATED WORK

The employed security strategy must also take into account the optimization of the data retrieval time. Division and

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that collectively approaches the security and performance issues. The DROPS methodology divides a file into fragments, and replicate the fragmented data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker. Moreover, the nodes storing the fragments are separated with certain distance by means of graph T-colouring to prohibit an attacker of guessing the locations of the fragments.

1. A HybridCloud Approach for Secure Authorized Replication[1] From This Paper we Referred-

In the proposed system we are achieving the data replicationby providing the proof of data by the data owner. This proof is used at the time of uploading of the file. Each file uploaded to the cloud is also bounded by a set of privileges to specify which kind of users is allowed to perform the duplicate check and access the files.New replicationconstructions supporting authorized duplicate check in hybrid cloud architecture in which the duplicate-check tokens of files are generated by the private cloud server with private keys. Proposed system includes proof of data owner so it will help to implement better security issues in cloud computing.

2. Secured Authorized De-duplicationBased Hybrid Cloud Approach [2] From This Paper we Referred-

Convergent encryption provides data confidentiality in de-duplication. A userderives a convergent key from each original data copy and encrypts the data copywith the convergent key. In addition, the user also derives a tag for the data copy,such that the tag will be used to detect duplicates. To detect duplicates, the user firstsends the tag to the server side to check if the identical copy has been already stored.Both the convergent key and the tag are independently derived, and the tag cannotbe used to deduce the convergent key and compromise data confidentiality. Authorized data replicationwas proposed to protect the data security by includingdifferential privileges of users in the duplicate check several new de-duplicationconstructions that support in authorized duplicate check in hybrid cloud architecture.

3. Implementation ReplicationSystem with Authorized Users [3] From This Paper we Referred-

This paper represents that, many techniques are using for the elimination ofduplicate copies of repeating data, from those techniques, one of the important datacompression technique is data duplication. Many advantages with this data duplication,mainly it will reduce the amount of storage space and save the bandwidthwhen using in cloud storage. To protect confidentiality of the sensitive data whilesupporting replicationdata is encrypted by the proposed convergent encryptiontechnique before out sourcing. Problems authorized data duplication formally addressed by the first attempt of this paper for better protection of data security. This is different from the traditional duplication systems. The differential privileges of users are further considered in duplicate check besides the data itself. In hybrid cloud architecture authorized duplicate check supported by several new duplication constructions. Based on the definitions specified in the proposed security model, our scheme is secure. Proof of the concept implemented in this paper by conducting test-bed experiments. A Client program is used to model the data users to carry out the file upload process. A Private Server program is used to model the private cloud which manages the private key and handles the file token computation. A Storage Server program issued to store and de-duplicates files. The Client provides the function calls to support token generation and replicationalong the file upload process. We observed that the information to Check replicationand upload the files, Fetching the Signs using Hashing Algorithm, Checking for Duplication, file uploading, file downloading and attacker trying to attack(block) the cloud.

4. Location-aware type ahead search on spatial databases: emetics and efficiency [4] From This Paper we Referred-

Users often search spatial databases like yellow page data using keywords to and businesses near their current location. Such searches are increasingly being performed from mobile devices. Typing the entire

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

query is cumbersome and prone to errors, especially from mobile phones. We address this problem by introducing type-ahead search functionality on spatial databases. Like keyword search on spatial data, type-ahead search needs to be location-aware, i.e., with every letter being typed, it needs to return spatial objects whose names (or descriptions) are valid completions of the query string typed so far, and which rank highest in terms of proximity to the user's location and other static scores. Existing solutions for type-ahead search cannot be used directly as they are not location-aware. We show that a straight-forward combination of existing techniques for performing type-ahead search with those for performing proximity search perform poorly. We propose a formal model for query processing cost and develop novel techniques that optimize that cost. Our empirical evaluations on real and synthetic datasets demonstrate the effectiveness of our techniques. To the best of our knowledge, this is the rst work on location-aware type-ahead search.

5. A Secured and Authorized Data Replicationwith Public Auditing [5] From This Paper we Referred-

This paper studies private data replication technique for cloud storages. Intuitively, a private data replication protocol allows a client who holds a private data proves to a server who holds a summary string of the data that he/she is the owner of that data without revealing further information to the server. The proposed private data replication protocol is provably secure in the simulation based framework assuming that the underlying hash function is collision resilient ,the discrete logarithm is hard and the erasure coding algorithm E can erasure up to fraction of the bits in the presence of malicious adversaries.

III.GOALS AND OBJECTIVE

Goals: The division of a file into fragments is performed based on a given user criteria such that the individual fragments do not contain any meaningful information. Each of the cloud nodes; we 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. To keep an attacker uncertain about the locations of the file fragments and to further improve the security, we select the nodes in a manner that they are not adjacent and are at certain distance from each other. The node separation is ensured by the means of the T-coloring.

Scope: The data owners lose the control over their sensitive data once the latter is outsourced to a remote CSP which may not be trustworthy. This lack of control raises new formidable and challenging tasks related to data confidentiality and in- tegrity protection in cloud computing. Customers require that their data remain secure over the CSP. Also, they need to have a strong evidence that the cloud servers still 2 possess the data and it is not being tampered with or partially deleted over time, especially because the internal operation details of the CSP may not be known to cloud customers. Encrypting sensitive data before outsourcing to remote servers can handle. The employed security strategy must also take into account the optimization of the data retrieval time. Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that collectively approaches the security and performance issues. The DROPS method- ology divide a file into fragments, and replicate the fragmented data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker. Moreover, the nodes storing the fragments are separated with certain distance by means of graph T-colouring to prohibit an attacker of guessing the locations of the fragments.

IV.PROPOSED ALGORITHM

1. ALGORITHM FOR FRAGMENT PLACEMENT

O = {O1;O2;....;ON}
O = {SIZEOF(O1); SIZEOF(O2);.....; SIZEOF(ON)}
COL = {OPEN COLOR; CLOSE COLOR}
CEN = {CEN1; CEN2; .....;CENM}
COL ← OPEN COLOR FOR ALL I

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

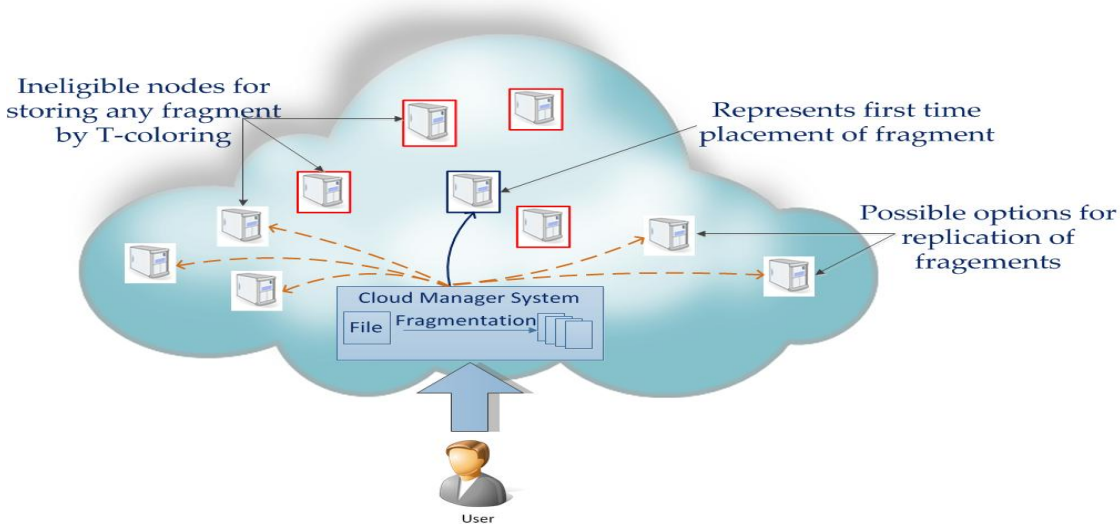
Vol. 3, Issue 12, December 2015

CEN ←CENI¦ I

COMPUTE:
FOR EACH OK € O DO
SELECT SI | SI ← INDEXOF(MAX(CENI))
IF COLSI = OPEN COLOR AND SI >= OK THEN
SI← OK
SI ← SI − OK
COLSI ← CLOSE COLOR
SI'←DISTANCE(SI; T) P /*RETURNS ALL NODES AT
DISTANCE T FROM SI AND STORES IN TEMPORARY SET SI'*/

COLSI. ← CLOSE COLOR
END IF
END FOR

V.ARCHITECTURE



The cloud manager keeps record of the fragment placement and is assumed to be a secure entity.Suppose a graph G = (V;E) and a set T containing non-negative integers including 0. The Tcoloring 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. Formulated by Hale, the T-coloring problem for channel assignment assigns channels to the nodes, such that the channels are separated by a distance to avoid interference.

VI. CONCLUSIONS

Outsourcing data to remote servers has become a growing trend for many organizations because it remove the burden of local data storage and maintenance. In this work consists the problem of creating multiple copies of dynamic data file and verifying those copies stored on untrusted cloud servers. The DROPS methodology, a cloud storage security scheme that collectively deals with the security and performance in terms of retrieval time. The data file was fragmented and the fragments are dispersed over multiple nodes. The nodes were separated by means of T-coloring.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

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. The performance of the DROPS methodology was compared with full-scale replication techniques. The results of the simulations revealed that the simultaneous focus on the security and performance resulted in increased security level of data accompanied by a slight performance drop.

REFERENCES

1. Jin Li, Yan Kit Li, Xiaofeng Chen, Patrick P. C. Lee, Wenjing Lou "A HybridCloud Approach for Secure Authorized De-duplication" IEEE Transactions on Paralleland Distributed Systems:PP Year 2014

2. MrVinod B JadhavProf Vinod SWadne Secured Authorized De-duplicationBased Hybrid Cloud Approach International Journal of Advanced Research in ComputerScience and Software Engineering

.**3.** A.Abdul Samadhu, J.Rambabu, R.Pradeep Kumar, R.Santhya DetailedInvestigation on a Hybrid Cloud Approach for Secure Authorized De-duplication InternationalJournal for Research in Applied Science and Engineering Technology(IJRASET)

4. S. B. Roy and K. Chakrabarti, "Location-aware type ahead search on spatial databases: Emantics and efficiency," in Proc.ACM SIGMOD Int. Conf. Manage. Data, 2011, pp. 361–37..

5. JadapalliNandini, RamireddyNavateja Reddy Implementation De-duplication System with Authorized Users International Research Journal of Engineering and Technology (IRJET).

6. Mazhar Ali, Student Member, IEEE, Kashif Bilal," Division and Replication of Data in Cloud for Optimal Performance and Security" IEEE Transactions on Cloud Computing