International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 4, Issue 2, No 3, 2016. ISSN: 2348-6600
Reference ID: IJCS-130 PAGE NO: 766-771

Secure Dynamic Fragment and Replica Allocation in Large-Scale Distributed File Systems

[1,] S.Shakila Banu, [2,] Dr.V.Jayaraj
[1] Research Scholar , Bharathidasan University ,Trichy
[2] Assistant Professor, School of CSE & Applications, Bharathidasan University ,Trichy

***Abstract*—** 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 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.

Keywords: Centrality, cloud security, fragmentation, replication, performance.

1 INTRODUCTION

Security is one of the most crucial aspects among those the wide-spread adoption of cloud computing [14, 19]. 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.) To secure cloud all of the participating entities must be secure. In a cloud the security of the assets does not solely depend on an individual's security measures because In any given system with multiple units, the highest level of the systems security is equal to the security level of the weakest entity [12] [5] and so the neighboring entities may provide an opportunity to an attacker .The off-site data storage cloud utility requires users to move data in cloud's virtualized and shared environment that may result in various security concerns. The Pooling and elasticity of a cloud allows the physical resources to be shared among many users [22]. Shared resources may be reassigned to other users at some instance of time that may result in data compromise through data recovery methodologies [2].The data [9]. Similarly, cross-tenant virtualized network access may also compromise data privacy and integrity. Improper media sanitization can also leak customer's private

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 4, Issue 2, No 3, 2016. ISSN: 2348-6600
Reference ID: IJCS-130 PAGE NO: 766-771

data [5]. The Unauthorized data access by users and processes must be prevented [4]. An 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. The probable amount of loss (as a result of data leakage) present Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that judicially fragments user files into pieces and replicates them at strategic locations within the cloud.

2 LITERATURE SURVEY

Computing Datacenters.

1.Cloud computing is an emerging paradigm that provides computing resources as a service over a network. Communication resources often become a bottleneck in service provisioning for many cloud applications. Therefore, data replication, which brings data closer to data consumers is seen as a promising solution. It allows minimizing network delays and bandwidth usage. In this paper we study data replication in cloud computing data centers. Unlike other approaches available in the literature, we consider both energy efficiency and bandwidth consumption of the system, in addition to the improved Quality of Service as a result of the reduced communication delays. The evaluation results obtained during extensive simulations help to unveil performance and energy efficiency tradeoffs and guide the design of future data replication solutions.

2. Data Security Issues in Cloud Computing. Cloud computing is an enticing technology which is a combination of many existing technologies such as parallel computing, grid computing, distributed computing and others. It offers services like data storage, computing power, shared resources at low cost to its users over internet at anytime from anywhere. Costing model on cloud computing is based on pay as you go method, hence companies are saving millions by adopting this technology. As more and more individuals and companies are relying on cloud for their data, the question arises here is how secure cloud environment Though cloud computing has many advantages, it also have some security problems.

3. On the Characterization of the Structural Robustness of Data center Networks. A Data Center Network (DCN) constitutes the communicational backbone of a data center, ascertaining the performance boundaries for cloud infrastructure. The DCN needs to be robust to failures and uncertainties to deliver the required Quality of Service (QoS) level and satisfy Service Level Agreement (SLA). In this paper, analyze robustness of the state-of-the-art DCNs. Our major contributions are: (a) we present multi-layered graph modeling of various DCNs; (b) we study the classical robustness metrics considering various failure scenarios to perform a comparative analysis; (c) The present the inadequacy of the classical network robustness metrics to appropriately evaluate the DCN robustness; and (d) The propose new procedures to quantify the DCN robustnessCurrently, there is no detailed study available centering the DCN robustness. Therefore, we believe that this study will lay a firm foundation

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 4, Issue 2, No 3, 2016. ISSN: 2348-6600
Reference ID: IJCS-130 PAGE NO: 766-771

for the future DCN robustness research . Motivated by the question of access control in cloud storage, we consider the problem using Attribute-Based Encryption (ABE) in a setting where users' credentials may change and cipher may be stored by a third party.

4. Secure Overlay Cloud Storage with Access Control and Assured Deletion This paper describes outsource data backups off-site to thirdparty cloud storage services so as to reduce data management costs. However, we must provide security guarantees for the outsourced data, which is now maintained by third parties. We design and implement FADE, a secure overlay cloud storage system that achieves fine-grained, policy-based access control and file assured deletion. It associates outsourced files with file access policies, and assuredly deletes files to make them unrecoverable to anyone upon revocations of file access policies. To achieve such security goals, FADE is built upon a set of cryptographic key operations that are self-maintained by a quorum of key managers that are independent of third-party clouds. In particular, FADE acts as an overlay system that works seamlessly atop today's cloud storage services. We implement a proof-of-concept prototype of FADE atop Amazon S3, one of today's cloud storage services. We conduct extensive empirical studies, and demonstrate that FADE provides security protection for outsourced data, while introducing only minimal performance and monetary cost overhead. Our work provides insights of how to incorporate value-added security features into today's cloud storage services.

5. Security and Privacy Issues in Cloud Computing Environment Cloud computing is emerging as a powerful architecture to perform large-scale and complex computing. It extends the information technology (IT) capability by providing ondemand access to computer resources for dedicated use. The information security and privacy are the major concerns over the cloud from user perspective. This paper surveys and evaluates the architecture, data security and privacy issues in cloud computing like data confidentiality, integrity, authentication, trust, service level agreements and regulatory issues. The objective of this paper is to review comprehensively the current challenges of data security and privacy being faced by cloud computing and critically analyse these issues.

3. Existing System

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 . As discussed above, 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.

A cloud must ensure throughput, reliability, and security . A key factor determining the throughput of a cloud that stores data is the data retrieval time. In large-scale systems, the problems of data reliability, data availability, and response time are dealt with data replication strategies .

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 4, Issue 2, No 3, 2016. ISSN: 2348-6600
Reference ID: IJCS-130 PAGE NO: 766-771

Drawbacks of Existing System

Placing replicas data over a number of nodes increases the attack surface for that particular data. For instance, storing m replicas of a file in a cloud instead of one replica increases the probability of a node holding file to be chosen as attack victim, from 1/ n to m /n, where n is the total number of nodes.

Proposed System

In this paper, we collectively approach the issue of security and performance as a secure data replication problem. We present Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that judicially fragments user files into pieces and replicates them at strategic locations within the cloud. 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.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.

Advantages of Proposed System

- The proposed DROPS scheme ensures that even in the case of a successful attack, no meaningful information is revealed to the attacker.
- We do not rely on traditional cryptographic techniques for data security. The non-cryptographic nature of the proposed scheme makes it faster to perform the required operations (placement and retrieval) on the data.

T-COLORING ALGORITHM FOR FRAGMENTS ALLOCATION:

T-coloring technique selects nodes in cloud for fragment placement by keeping focus on performance and security. The central node in cloud network gives improved access time. The DROPS method utilize centrality concept to decrease the access time. Centrality concludes central node based on various measures. T-coloring restricts node selection at hop distance. T-coloring gives more security performance in the cloud. The fragments are stored in different node so location of fragments can't able to determine.

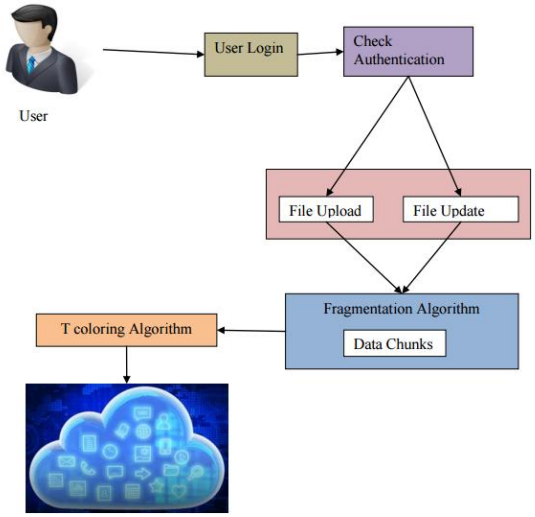
ALGORITHM: FRAGMENTATION

Step 1: I = {I1;I2,...,IN}
Step 2: i = {sizeof(i1); sizeof(i2),...,sizeof(iN)}
Step 3: data = {open_data, close_data}
Step 4: cen = {cen1; cen2,....,cenM}
Step 5: data ←open_data ∀ i
Step 6: cen ←ceni ∀ i
Step 7: for each Ik > I do
Step 8: choose C^i |C^i indexof(max(ceni))
Step 9: if $data_{C^i}$ = open_data andc_i>= ik then
Step 10: S^i←Ik
Step 11: Ci ←Ci – Ik
Step 12: $data_{C^i}$ ←close_ data
Step 13: $C^{i''}$←distance(C^i; T
Step 14: $data_{C^i}$ ← close_data
Step 15: end if
Step 16: end for

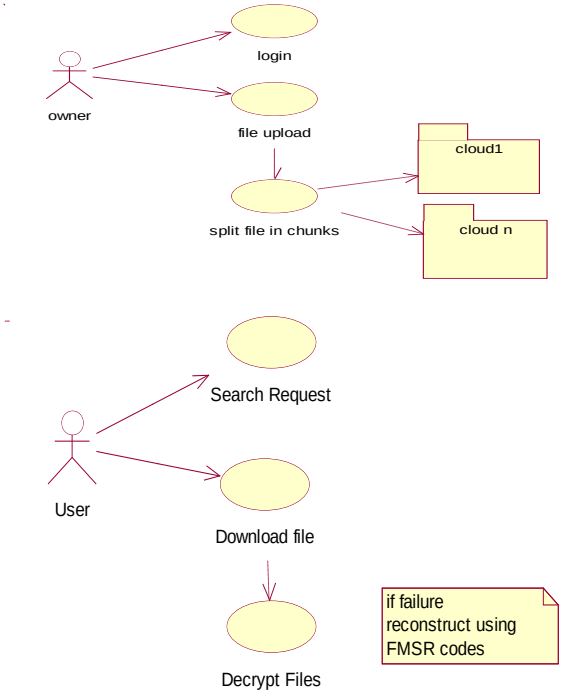
International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 4, Issue 2, No 3, 2016. ISSN: 2348-6600
Reference ID: IJCS-130 PAGE NO: 766-771

4. SYSTEM MODEL

A cloud that consists of M nodes, each with its own storage capacity. Let Si represents the name of i -th node and si denotes total storage capacity of Si Communication time between Si and Sj is the total time of all of the links within a selected path from Si to Sj represented by c(i, j) . We consider N number of file fragments such that Ok denotes k -th fragment of a file while ok represents the size of k-th fragment. Pk denote the primary node that stores the primary copy of Ok, replication scheme for Ok denoted by Rk is also stored at Pk and Whenever there is an update in not as an independent document. Please do not revise any of the current designations Ok , the updated version is sent to Pk that broadcasts the updated version to all of the nodes in Rk . Let colSi store the value of assigned color to Si . The colSi can have one out of two values, namely: open color and close color. The value open color represents that the node is available for storing the file fragment. The value close color shows that the node cannot store the file fragment The set T is used to restrict the node selection to those nodes that are at hop-distances not belonging to T. In the DROPS methodology, we propose not to store the entire file at a single node. The DROPS methodology fragments the file and makes use of the cloud for replication. The fragments are distributed such that no node in a cloud holds more than single fragment, so that even a successful attack on the node leaks no significant information. In the DROPS 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 of the selected node, and (c) second cycle of nodes selection for fragments replication. The cloud manager keeps record of the fragment placement and is assumed to be a secure entity



6.DATAFLOW DIAGRAM



International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 4, Issue 2, No 3, 2016. ISSN: 2348-6600
Reference ID: IJCS-130 PAGE NO: 766-771

7. CONCLUSION AND FUTURE WORK

The DROPS methodology, a user has to download the file, update the contents, and upload it again. It is strategic to develop an automatic update mechanism that can identify and update the required fragments only. The aforesaid future work will save the time and resources utilized in downloading, updating, and uploading the file again. Moreover, the implications of TCP in cast over the DROPS methodology need to be studied that is relevant to distributed data storage and access.

8. REFERENCES:

[1] 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.

[2] Y. Deswarte, L. Blain, and J-C. Fabre, "Intrusion tolerance in distributed computing systems," In Proceedings of IEEE Computer Society Symposium on Research in Security and Privacy, Oakland CA,110- 121, 1991

[3] B.Grobauer, T.Walloschek, and E. Stocker, "Understanding cloud computing vulnerabilities," IEEE Security and Privacy, Vol.9, No. 2, 2011.

[4] W. K. Hale, "Frequency assignment: Theory and applications," Proceedings of the IEEE, Vol. 68, No. 12, 1980,

[5] 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,

[6] W. A. Jansen, "Cloud hooks: Security and privacy issues in cloud computing," In 44th Hawaii IEEE International Conference on System Sciences (HICSS), 2011,

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

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

[9] 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.

[10] 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.