SRICSO: Section and Replication of Information in Cloud for Security and Optimum Operation

Varun A H[1], S. Pramela Devi[2]

M.Tech(CSE) student, Dept of CSE, MVJCE, Bangalore.
Assistant Professor, Dept of CSE, MVJCE, Bangalore.

***Abstract*— The data is being outsourcing to third party into the cloud computing requires the security to be done while accessing. But while providing the security there comes the performance issues which needs to be taken to be consideration, to increase the performance of the cloud along with secure data accessing the new SRICSO methodology has introduced. SRICSO methodology deals with the dividing the data among the nodes of the cloud as, the main data is being distributed over several nodes while accessing the user must provide the tag value from that tag value the sub data divided among nodes is being tracked by creating T-path and is being retrieved from main node. But in this concept the data stored early may be stored again this can be advanced by avoiding the data which is already present in the cloud hence performance is increased while retrieving or downloading the required data by the user.**

***Index Terms*— Centrality, cloud security, fragmentation, de-duplication, performance**

I. INTRODUCTION

Cloud is mainly used for the public use and is said to be said to public cloud, where at public cloud the data being outsourced to third party. So, there comes the issue of the security which is to be concentrated. To provide security there are many techniques like the old methodology of cryptographic security providing the public key and private key encryption while uploading and downloading the data which leads to many of the performance issues such as hacking of data and even the time consumed to download and upload the data to cloud through Network.

So, a new technique or the methodology was provided by this SRICSO where in this technique, we divide the data that is being uploaded to the cloud in the threshold specified from main node to the sub nodes and a T-Path is created while the downloader specifies the tag value while downloading.

Motivated by the old methodology of DROPS Methodology were only the division of the data is being taken among nodes and the duplication of the data is not avoided which makes the cloud consume more space and even it takes more cost.

Hence, SRICSO Methodology overcomes the issue created by providing the de-duplication of the data to be stored on the cloud by the help of two techniques such as file level and node level de-duplication techniques.

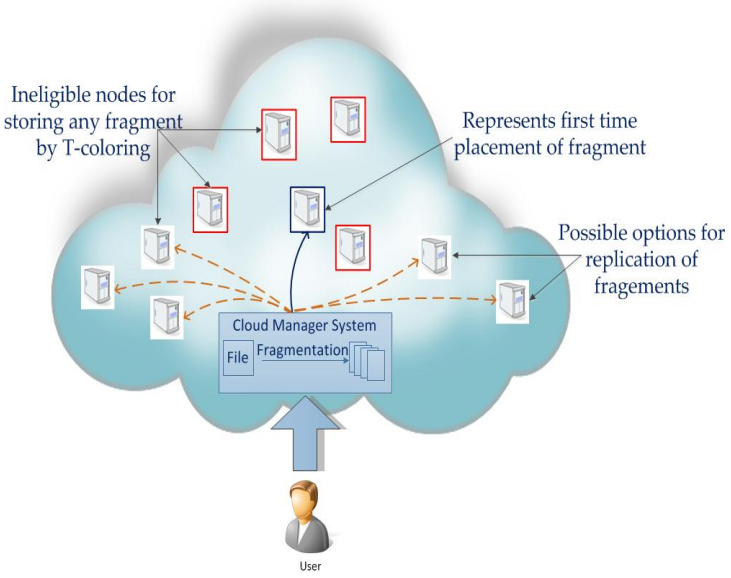


Figure 1: Architecture

The rest of the paper is organized as follows section II deals with related work , section III speaks about Sectioning and Replication Of Data With De-Duplication, section IV and section V speaks about the results regarding performance improvement and future work.

II. LITERATURE SURVEY

Mazhar AliKashif Bilal, Student Member, Samee U. Khan, Bharadwaj Veeravalli, Keqin Li, Senior Member, and Albert Y. Zomaya [1] were the first to introduce DROPS Division and Replication of Data in Cloud for Optimal Performance and Security Methodology, where the data to be stored into the cloud and to be accessed by the other user is being divided and replicated over the cloud from main node to the sub nodes and while downloading the data the path creation is made from sub nodes to main node which

overcomes the cryptographic way of accessing the data and an optimal performance is gained as data is divided.

K. Bilal, M. Manzano, S. U. Khan, E. Calle, K. Li, and A. Zomaya,[2] "On the characterization of the structural robustness of data center networks," introduced about the data centred networks for communicational backbone of data and performance boundries of cloud. Analyze the state of art of DCN for Robustness research, Multilayered graph modeling for various DCN, new techniques of quantifying the DCN and finally the comparision of new techniques with old technique.

D.Boru, D.Kliazovich, F.Granelli, P.Bouvry, and A.Y.Zomaya,[3] "Energy-efficient data replication in cloud computing datacenters," data replication brings data closer to the consumers which was the advantage to the users to use data easily. Mining the network delay and bandwidth usage is taken to analyze the delay caused on the network, which considers both Energy ad bandwidth consumption of system in addition to Quality of services to reduce communication delays.

III. Sectioning And Replication Of Data With De-duplication

The data before uploading, user who is trying to upload the data should be register with the cloud then he gets the key to his registered mail id which is needed to be entered while uploading the data. While downloading the data user need to enter the same key to download the data, so that based on the key entered the data divided while uploading gets centralized to main node and gets downloaded.

A. Data Sectioning and Replication:

While uploading the data the data being divided into mainly five nodes based on the MD5 algorithm for division and the SHA1 algorithm for replication, here we combine the MD5 and SHA1 to generate the Hash value based on which a key is generated.

MD5 algorithm as follows:

Step1 Append padding bits

The input message is "padded" (extended) so that its length (in bits) equals to 448 mod 512. Padding is always performed, even if the length of the message is already 448 mod 512.

Padding is performed as follows: a single "1" bit is appended to the message, and then "0" bits are appended so that the length in bits of the padded message becomes congruent to 448 mod 512. At least one bit and at most 512 bits are appended.

Step2. Append length

A 64-bit representation of the length of the message is appended to the result of step1. If the length of the message is greater than 2^64, only the low-order 64 bits will be used.

The resulting message (after padding with bits and with b) has a length that is an exact multiple of 512 bits. The input message will have a length that is an exact multiple of 16 (32-bit) words.

Step3. Initialize MD buffer

A four-word buffer (A, B, C, D) is used to compute the message digest. Each of A, B, C, D is a 32-bit register. These registers are initialized to the following values in hexadecimal, low-order bytes first):
word A: 01 23 45 67
word B: 89 ab cd ef
word C: fe dc ba 98
word D: 76 54 32 10

Step4. Process message in 16-word blocks

Four functions will be defined such that each function takes an input of three 32-bit words and produces a 32-bit word output.

F (X, Y, Z) = XY or not (X) Z
G (X, Y, Z) = XZ or Y not (Z)
H (X, Y, Z) = X xor Y xor Z
I (X, Y, Z) = Y xor (X or not (Z))

B. De-Duplication

De-Duplication refers to the avoidance of the duplicate data into the cloud while uploading of the file only. In this section of De-Duplication implementation, we are implementing based on the two techniques

File level De-Duplication and Node-Level De-Duplication Where at File Level as generated hash value matches with old one then that file will not be uploaded internally just a success message is displayed to the user where for the Downloader the file that was previously uploaded is pointed and the file gets downloaded.

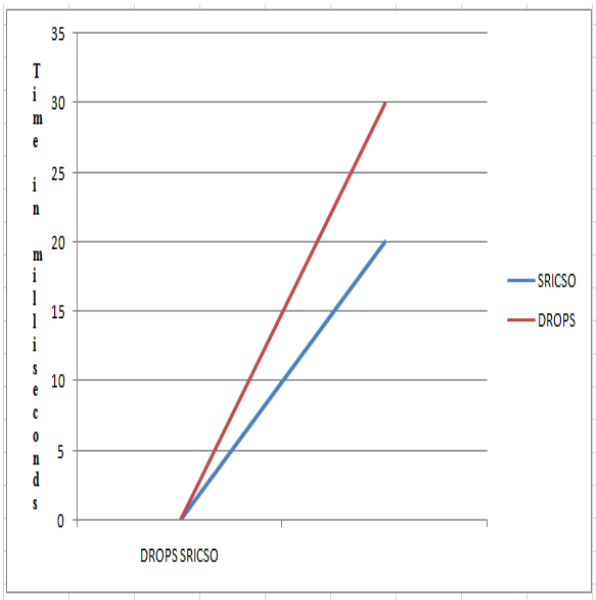
Node-Level De-Duplication where the data being distributed among the nodes if gets any of the data being to be uploaded with the same hash coded data then it will not be uploaded instead just it shows successful message to Data up loader.

Hence by the De-Duplication the more amount of Space and even Time for uploading the data is being saved which leads to more performance to Cloud.

IV. Comparision Results

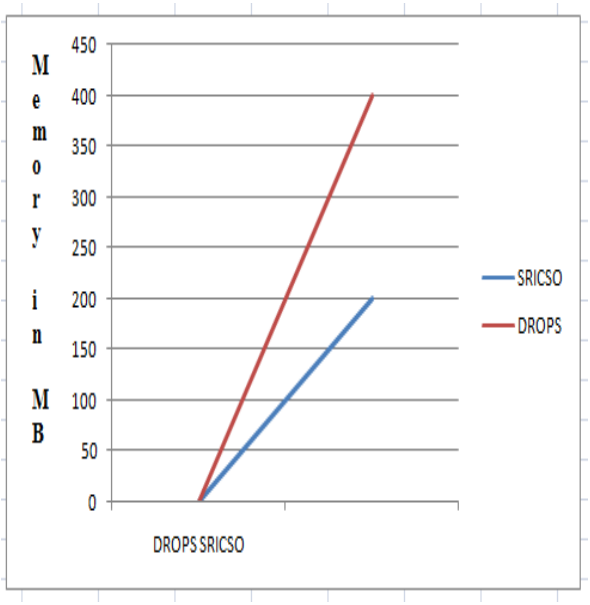
Compared to the old Methodology of the DROPS, which lacks in the unwanted and repeated storage of the data into the cloud, our SRICSO methodology introduces the De-Duplication technique which saves much of the space to storage and the time take to upload the data file to the cloud, hence a gradual increase in the Performance of the Cloud.

We compare the performance of the SRICSO with the DROPS based on the space and the time constraints through the graphs as follows:



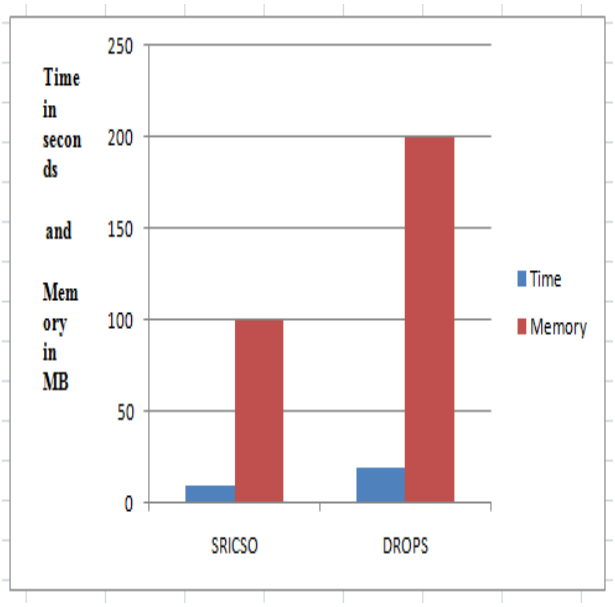
Graph 1: Line graph comparison based on Time used to upload data.

Graph 1 Compares the both DROPS and SRICSO methodology based on time consumption. As graph shows SRICSO takes less amount of time.



Graph 2: Comparison based on Memory Consumed by DROPS and SRICSO.

Comparing the DROPS and SRICSO methodology with the amount of space consumed on the cloud storage in MB.



Graph 3: Overall comparision for Performance Evaulation

Comparing the Over All performance of the SRICSO methodology with DROPS and concluding SRICSO is more performance Oriented than DROPS methodology.

V. Conclusioin And Future work

This paper proposes the novel approach of the division and the de-duplication of the data in the Cloud where much amount of the space to be taken is saved and even the user time and the Cloud usage time is also reduced. As the data being distributed among the nodes of the cloud if any hacker hacks the cloud also there also the hacker may not get full data and even the data is in the form of hash codes which might not be able to read to the hacker. Hence this paper proposes a new methodology to improve the Performance and the Security of the Cloud.

In SRICSO approach there requires a bit more amount of time for division of the data and even downloading as a path need to be created from sub nodes to main node which can be advanced as a Future work.

VI. References

[1] DROPS: Division and Replication of Data in Cloud for Optimal Performance and Security Mazhar Ali, Student Member, IEEE, Kashif Bilal, Student Member, IEEE, Samee U. Khan, Senior Member, IEEE, Bharadwaj Veeravalli, Senior Member, IEEE, Keqin Li, Senior Member, IEEE, and Albert Y. Zomaya, Fellow, IEEE DOI 10.1109/TCC.2015.2400460, IEEE Transactions on Cloud Computing

[2] 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.

[3] 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

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

[5] "An analysis of security issues for cloud computing," Journal of Internet Services and Applications, Vol. 4, No. 1, 2013, pp. 1-13. by K. Hashizume, D. G. Rosado, E. Fernndez-Medina, and E. B. Fernandez.

[6] "NIST cloud computing standards roadmap," NIST Special Publication, July 2011. by M. Hogan, F. Liu, A.Sokol, and J. Tong

[7] "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. by Y. Tang, P. P. Lee, J. C. S. Lui, and R. Perlman,

[8] "On the optimal placement of secure data objects over Internet," In Proceedings of 19th IEEE International Parallel and Distributed Processing Symposium, pp. 14-14, 2005. by M. Tu, P. Li, Q. Ma, I-L. Yen, and F. B. Bastani

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

[10] "Addressing cloud computing security issues," Future Generation Computer Systems, Vol. 28, No. 3, 2012, pp. 583-592. by D. Zissis and D. Lekkas,