Multi-Cloud Hosting for Performance Optimization and Security

Gade Pooja V.[1] Mate Shweta A.[2] Janjalkar Priti S.[3]
[1,2,3]B.E. Student
[1,2,3]Department of Computer Engineering
[1,2,3]Chakan, Pune

Abstract—Many organizations and enterprises stores data into the cloud to reduce maintenance cost and also to increase reliability of data. While hosting their data into cloud, customer faces many cloud vendors as well as their pricing policies so that the customer may get confused with which cloud(s) are suitable for storing data. CHARM propose a data hosting scheme which combine two key functions- first function is selecting a cloud to store data in minimum cost and guaranteed availability and second is triggering transition process to re-distribute data. It can do according to pricing of clouds and variations of data patterns. While sending data to Third-Party administrative control in cloud, it also becomes an issue related to security. We need high security measures to protect data in a cloud so we use DROPs concept that solves security and performance issue. In this we divide data into fragments and replicate that fragments into cloud nodes and each node contain only a single data. All these nodes are separated with a certain distance using graph T-colouring. Using this security and performance optimization is done.

Key words: Multi Cloud, Reliability, Cloud Security, Fragmentation, Replication

I. INTRODUCTION

Selecting suitable clouds and an appropriate redundancy strategy to store data with minimized monetary cost and guaranteed availability [1] [5]. The second is triggering a transition process to re-distribute data according to the variations of data access pattern and pricing of clouds. Multi-cloud data hosting has received wide attention from researchers, customers, and startups. The basic principle of multi-cloud (data hosting) is to distribute data across multiple clouds to gain enhanced redundancy and prevent the vendor lock-in risk [2].

The cloud computing has performed the usage and management of the information technology infrastructure. Cloud computing is characterized by ubiquitous network accesses, on-demand self-services, elasticity, resource pooling, and measured services. The characteristics of cloud computing make it a striking candidate for individual users for adoption businesses, organizations. However, the benefits of low-cost from a user's perspective and security concerns with greater flexibility. Security [9] [10] is one of the most important factor in cloud computing [11].

II. LITERATURE SURVEY

Z. Li, C. Jin, T. Xu, C. Wilson, Y. Liu, L. Cheng, Y. Liu, Y. Dai,and Z.-L. Zhang [1], in this paper the author introduced the various advantages of Cloud storage services like Google drive, Drop box and many other. Also Microsoft is a one drive which allows user to store, share data very convenient and reliable. User can access there data from anywhere, on any device, at any time as there need. In this paper it points to a simple question like is the current data sync by cloud efficiently? Then it define a novel metric named TUE i.e. Traffic Usage Efficiency to check the synchronization of a data.

A.Li, X. Yang, S. Kandula, and M. Zhang [2], in this paper introduces the, Cloudcmp to help customer to spick a cloud that fits by their needs. The Cloudcmp also provide a performance and cost of cloud. Cloudcmp measures the different services like, networking, computing, storage and many more.

A. Mei, L. V. Mancini, and S. Jajodia [3], this survey represents a distributed algorithm. This distributed algorithm is used for file allocation that guarantees scalability, availability, reliability in a large distributed file system. The distributed algorithm can also use other schemas to allocate files over multiple servers i.e. replication and fragmentation.

III. PROPOSED SYSTEM

We used the DROPS methodology which is a cloud storage security scheme that deals with the security of data and performance in retrieval time. In this the data was fragmented and fragments are distributed over multiple nodes. These nodes are separated by T-coloring graph. Only one node can be saved in the single fragment of that same file. The performance of DROPS is compared with replication techniques; it means it checks for multiple data. With the DROPS methodology user has to update the contents of file, upload the file, and download the file [4]. DROPS use the distributed data storage with TCP in cast.

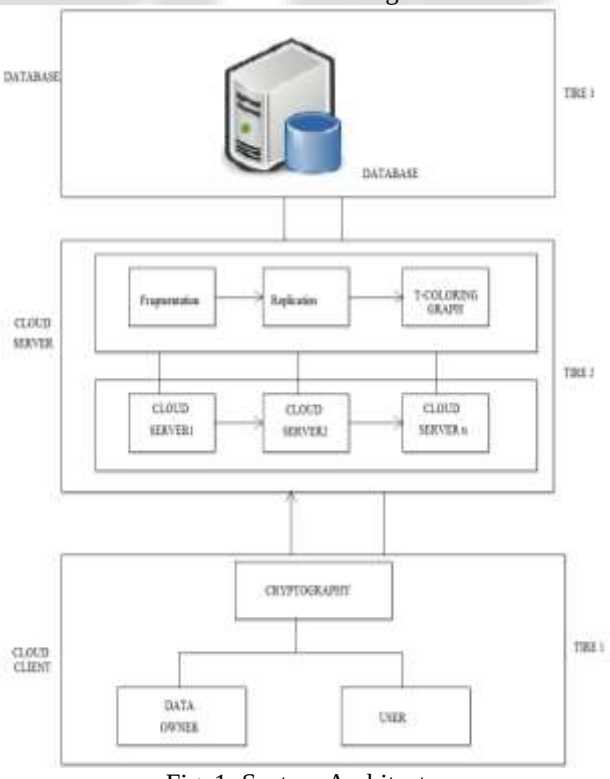

Fig. 1: System Architecture

We also propose a cost-efficient data hosting scheme with high availability in heterogeneous multi-cloud,

named "CHARM" [3]. It puts data into multiple clouds with minimum cost and guaranteed availability. We combine the two widely used redundancy mechanisms, i.e., replication and erasure coding, into a uniform model to meet the required availability in the presence of different data access patterns. Next, we design an efficient heuristic-based algorithm to choose proper data storage modes (involving both clouds and redundancy mechanisms). Moreover, we implement the necessary procedure for storage mode transition (for efficiently re-distributing data) by monitoring the variations of data access patterns and pricing policies [6] [7]. The Figure 1 shows the normal system architecture.

A. Data owner

There are mainly two types of data that are stored on cloud.

- The data which is created by user before uploading the file into cloud.
- The data which is created on cloud it.

The owner has their own services like register, login, file upload, view file, select cloud, download file and logout.

B. User

User will be any person who will use cloud. The users have their own services which are provided by data owner like register, login, view file, select cloud, download file and logout.

The Figure 2 describes the overall architecture i.e. the fragmentation, replication and T-coloring graph. The web application is to maintain the network through the internet.

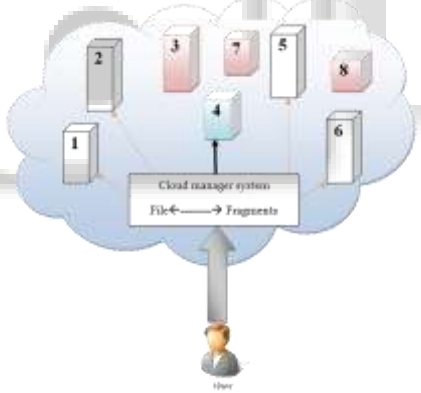

Fig. 2: Block diagram of fragmentation, replication and T-coloring graph

Where, Block 1,2,5,6 represents possible blocks for replication.

Block 3, 7 and 8 represents the storing nodes fragmented by T-coloring.

Block 4 represents the first time placement of fragment.

Fragmentation [12] means splitting the data into smaller parts. In this cloud the fragmentation in means splitting data into small fragments and distributing them across various nodes on cloud. It is a procedure in cloud which is performed by cloud server which automatically breaks the data into fragments for storage in different or various locations based on T-coloring graph technique. If an attacker attack on any user node and if the attacker gets the information successfully then at that time our system proved a security. Because the data is divided into fragments. So it increases the level of security.

Once the file is split into different fragments, the DROPS methodology is used i.e. Division and Replication of Data in Cloud for Optimal Performance and Security. The DROPS methodology selects the cloud nodes for fragment. The selection is done by keeping equal focus on security and performance, in terms of access time. We choose the most central nodes which are provided in cloud storage for better access time. This DROPS method uses the centrality to reduce the access time.

However, if all the fragments data are placed on nodes based on centrality on descending order, then there is possibility to get the clues for an attacker, reducing the security level of user data. To deal with this security issue the concept of T-coloring graph is used.

The T-coloring graph [8] technique generates a non-negative random number and builds the set T which is starting from zero to generated non-negative random number. Where the set T is used to restrict the selection of node that are at hop distance but not belonging to set T. For this it assigns a color to the nodes. Initially, all the nodes are given in open color. Then once the fragments are placed then all the neighborhood nodes belonging to set T are marks as close nodes. Replication [13] means copies of multiple data.

IV. CONCLUSION AND FUTURE WORK

Cloud services are rapidly used. They experiencing the services based on the multi cloud. So, in this system we deign a scheme CHARM, which guide customer to distribute there data on different clouds which is cost effective. We also used the DROPS concept for storage of data. The CHARM helps user to a decision about storage and cloud to place data in. This system can be enhanced by developing an automatic update methodology that updates the required fragments only. In future this system will save the time for downloading and uploading the file again.

ACKNOWLEDGMENT

It gives us great pleasure in presenting the preliminary project report on 'Multi- Cloud Hosting for Performance Optimization and Security'. I express my sincere thanks to my parents, friends and all those who helped us directly or indirectly in many ways in completion of this dissertation work.

REFERENCES

[1] Z. Li, C. Jin, T. Xu, C.Wilson, Y. Liu, L. Cheng, Y. Liu, Y. Dai, and Z.-L. Zhang, "Towards network-level efficiency for cloud storage services," in Proceedings of the 2014 Conference on Internet Measurement Conference, pp. 115–128, ACM, 2014.

[2] A. Li, X. Yang, S. Kandula, and M. Zhang, "Cloudcmp: comparing public cloud providers," in Proceedings of the 10th ACM SIGCOMM conference on Internet measurement, pp. 1–14, ACM, 2010.

[3] A. Mei, L. V. Mancini, and S. Jajodia, "Secure dynamic fragment and replica allocation in large-scale distributed file systems," Parallel and Distributed

Systems, IEEE Transactions on, vol. 14, no. 9, pp. 885–896, 2003.
[4] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya, "Energyefficient data replication in cloud computing datacenters," in Globecom Workshops (GC Wkshps), 2013 IEEE, pp. 446–451, IEEE, 2013.
[5] C. M. M. Erin Allen, "Library of Congress and DuraCloud Launch Pilot Program Using Cloud Technologies to Test Perpetual Access to DigitalContent," The Library of Congress, News Releases, http://www:loc:gov/today/pr/2009/09-140:html.
[6] "Windows Azure pricing updates," http://azure:microsoft:com/en-us/updates/azure-pricing-updates/.
[7] "Google Cloud Platform pricing updates," http://googlecloudplatform:blogspot:com/2014/03/google-cloud-platformlive-blending-iaas-and-paas-moores-law-for-the-cloud:html.
[8] W. K. Hale, "Frequency assignment: Theory and applications,"Proceedings of the IEEE, Vol. 68, No. 12, 1980, pp. 1497-1514.
[9] B. Grobauer, T.Walloschek, and E. Stocker, "Understanding cloud computing vulnerabilities," IEEE Security and Privacy, Vol. 9, No. 2, 2011, pp. 50-57.
[10] A. N. Khan, M. L. M. Kiah, S. U. Khan, and S. A. Madani, "Towards Secure Mobile Cloud Computing: A Survey," Future Generation Computer Systems, Vol. 29, No. 5, 2013, pp. 1278-1299.
[11] 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.
[12] A. Mei, L. V. Mancini, and S. Jajodia, "Secure dynamic fragment and replica allocation in large-scale distributed file systems," IEEE Transactions on Parallel and Distributed Systems, Vol.14, No. 9, 2003, pp. 885-896.
[13] M. Tu, P. Li, Q. Ma, I-L. Yen, and F. B. Bastani, "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.

