International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

Multi-Cloud Data Storing Strategy with Cost Efficiency and High Availability

Shilpa B Kodli[1], Pramila D Reshmi[2]

[1]Assistant Professor, Department of MCA, Visvesvaraya Technological University Post Graduate Centre,Kalaburagi, India

[2]PG Scholar, Department of MCA, Visvesvaraya Technological University Post Graduate Centre, Kalaburagi, India

Abstract: ***Cloud computing has been envisioned as the de-facto solution to the increasing storage cost of IT enterprises. The cloud vendors, customers and different pricing policies are confused to choose suitable cloud(s) for storing their data and to choose the cheaper methodology. Apart from the reduction in storage moves the user's data outsourcing to the cloud also helps in reducing the maintenance. This provides a scheme which gives a proof of data integrity in the cloud which the customer has employ to check the correctness of his data in the cloud. This strategy ensures that the storage at the client side is minimal which will be beneficial for clients.***

Keywords: CHARM, cost efficient, multi-cloud, suitable cloud data hosting.

1. Introduction

Cloud computing get its name as a metaphor for today's internet world. Existing clouds such as Amazon S3, Windows Azure, Google cloud storage belongs to great differences in terms of working performances and pricing policies. Selecting suitable clouds and appropriate redundancy strategy to store data with minimum cost and ensured availability plays a major role is storing of data.

Many Cloud vendors develop their infrastructure and keep on upgrading them with newly emerging technologies. The cloud services can be accessed by using different models like Single service provider and multiple services providers. The problem with these models is easy for hacking and not guaranteed availability.

Data outsourcing to cloud storage severs is raising trend among many firms and users owing to its economic advantages. This essentially means that the owner (client) of the data moves its data to a third party cloud storage server which is supposed to - presumably for a fee - faithfully store the data with it and provide it back to the owner whenever required. As data generation is far out spacing data storage it proves costly for small firms to frequently update their hardware whenever additional data is created. Also maintaining the storages can be a difficult task.

Storing of user data in the cloud despite its advantages has many interesting security concerns which need to be extensively investigated for making it a reliable solution to the problem of avoiding local storage of data.

Many problems like data authentication and integrity outsourcing encrypted data and associated difficult problems dealing with querying over encrypted domain were discussed in research literature. In this paper we deal with the problem of implementing a protocol for obtaining a proof of data possession in the cloud sometimes referred to as Proof of irretrievability (POR).

Any such proofs of data possession schemes do not, by itself, protect the data from corruption by the archive. It just allows detection of tampering or deletion of a remotely located file at an unreliable cloud storage server. To ensure file robustness other kind of techniques like data redundancy across multiple systems can be maintained.

In this paper, we focus on efficient and heuristicbased data hosting scheme for heterogeneous multi-cloud environment and flexible transaction scheme for CHARM (Cost Efficient Multi Cloud Data Hosting Scheme with High Availability). This CHARM scheme intelligently puts data into multiple clouds with minimized monetary cost and guaranteed availability. In the second section we discuss our related work and what server has been used by the different authors and in the third section about the system working steps. In the fourth section we discuss about the modules which has been used in this paper and in the next section about the algorithm and last conclusion.

2. Related Work

Data Storage and data integration has received a lot of attention at the data management and application level. Mansouri Y, Toosi, A.N., Buyya [1] Authors dealt with the problem of multi-cloud storage with a focus on availability and cost factors. Because of more cost the customers are unable to choose suitable cloud.

There is a concern about moving large amount of data into a single cloud is similar to vendor lock-in risk. Depsky[2] author deals with stores data, even critical data into multiple clouds assuming data availability and security. Shaik. AafreenNaaz [3] author reviewed the cloud computing features provides more benefits to the users in terms of low cost and availability of data, providing security to the cloud computing is a main factor. The single cloud service provider for outstanding is not trusted because of failure in service availability and possibility of attacker like malicious virus which corrupts the stored data.

Here a multi-cloud is emerged by inter clouds or cloud of ouds where research related to single cloud problems can be addressed by using multi-cloud.

Many new tools like Apache library cloud which provides a unique interface on different clouds for convenient deployment of multi-cloud services information given in [4]. This methodology helps in communication between different clouds.

The advantages and disadvantages of erasure coding and replication in peer-to-peer system is given in paper[5] and [6].Here the mechanism in multi cloud environment cannot be compared because it is proved very different from the results in two works.

A research is done on data hosting in Grid/peer-to-peer storage systems has stated in [9], [10], [11], [12]. Here the authors deal with the prominent feature of storage system is that storage nodes are unstable.

A similar work on storage system in [13],which erasure coding and replication in multiple data centers are discussed here the author deal with cache in the primary data centers. The heterogeneity of multi cloud and the selection of clouds are not considered .the cache helps in storing back of file when accessed by erasure coding frequent data swap inevitably induces additional cost which makes long competitive when compare to other data hosting.

M.P.Papazogloneetc[14] has reviewed that cloud computing technology has main drawback vendor lock-in. The cloud service developers will not allow to get service for free and does not allow to mix and match applications and services .hence they introduced cloud blueprint so that developers to mix and match services for free of cost. By this it is facilitate to mix and match the configuration, application and stacking the resources into cloud.

This approach provides simplified method for provisioning and automating cloud services and also applications run dynamically on fully virtualized clouds.

R.Thandeeswaran et al.[15] has reviewed that security need tobe addressed as major concern for handling critical application and sensitive information. The use of multiple cloud has following advantages

1. Exchange of data from multiple clouds.
2. Selection of clouds based on price and services.

S.Oritz Jr. [16] has reviewed that many of the industries are lack in extension of adopting cloud computing technology. The implementation leads to instability in area such as security and interoperability in turn this leads to vendor lock-in. hence the standardization is introduced which involves virtualization which play very important role in cloud computing. The data hosting schemes mentioned in our paper focus on different aspects like vendor lock-in, selecting suitable data hosting strategy, optimization of performance guaranteeing flexible availability and security.

3. Architecture

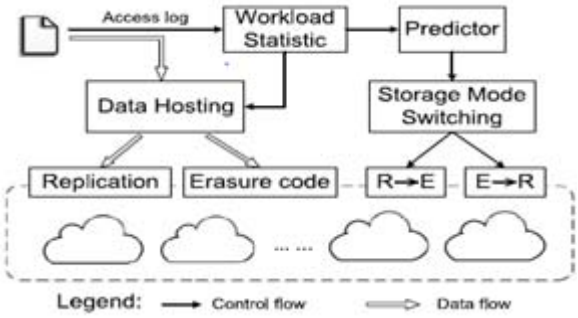

Figure 1: System architecture

By considering the above fig, there are two main factors and the all data are stored in the clouds, first the replication indicates no repetition of files in the cloud it avoids duplicate files so that more storage space is available in the cloud. And there will never be a work load for the admin and no complexity in the database. The following steps are:

Step 1: In the step1 the application interacts with the cloud and sends all the data to the server, if the user is interacting with the cloud1 or requesting the cloud for example the cloud1 is busy it is forwarded to the cloud2 no time is wasted here shifting of files takes place.

Step 2: The predictor is used to gathering the file frequency the given time is for only one month and in the next month its changes and many algorithms supports for the collecting the right data in the clouds.

Step 3: In the third step the storage modes shift towards the storing capacity, that decides which data is suitable for cloud because storing process must not be complexity in the cloud.

4. Module Description

1. Multi-cloud:
Around the world many different data centers have been situated. There may be a same or different data centers contains a similar or distinguished cloud suppliers. Data can be viewed by a user's in some region but he may be satisfied with the data or may not be depending on own views. He may get different outcomes some of them may be high and some of them may be low. But in our project it chooses a particular cloud for a particular data which is suitable according to the file and information which gives a good output and overcomes the limitation from the present system without effecting the other information. When server has a less load of a work we may slow down the increasing of operations and can also implement other data.

2. Cloud Storage:
As the cloud has become popular and getting benefits day by day, its importance is it provides security and protects the data complexity, the information which is saved is full of safe. some may try to take the information present in the cloud due to encryption, if sometimes hacked there may be no use, In our project keys are generated and forwarded to their respective accounts, if the user mail is sometimes

hacked there will be no use because the data will be in the form of cipher text, clients information is protected and maintained safety.

3. Owner module

Owner modules is about to accepting the clients and their pricing models, when accepted they take their permission of admin and using secret key he uploads the files for a particular file particular key is generated and each owner gets a unique code.

4. User module:

This module helps the clients to download the files, first he register and logins with a given id and select the clouds he gets the generation of a secret key takes place to his mail and request to the clouds for a particular file with the id when admin gives the permission the session key is generated so that by using that key he can view the file and make in the form of decrypted file and also can be download.

5. Algorithms

RSA Algorithm

To encrypt the files RSA Algorithm is used and the data is provided with high security. It was developed by Rivest Shamir Adelman and used under cryptosystem. It is one of the most used algorithms, it contains two keys one is encryption key which can be use by anyone and is called public, the other is decryption key which is not given to anyone and kept secret. We use prime numbers and an auxiliary value and are consider as one of the slowest algorithm.

KPABE Algorithm

In the proposed system we use this algorithm for the policy attributes, to encrypt and termed as key policy attribute based encryption. In this algorithm the time is less and the text which is cipher is small when compared with CPABE. It provides more advantages in implementing the hardware requirements and maintains high complexity. We take input as a security parameter and message and gives output as a public and cipher text and is considered as one of the important algorithm because it encrypt the files of pricing models.

6. Analysis and Results

Figure 1: Home page

Home page which includes Data owner, Proxy server and User.

Figure 2: File Symmetric key

The different symmetric key is generated for a different file when uploaded by the dataowner.

Figure 3: File Decryption

The process done by users when downloading, they first decrypt the file and then downloaded.

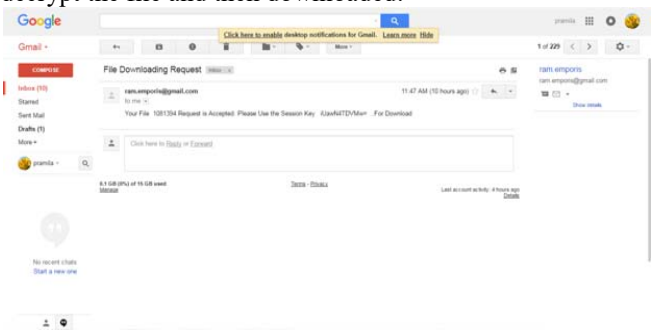

Figure 4: Generation of session key

Session key is generated to the user accounts to download the file; if user enters the invalid session key their account is blocked.

7. Conclusion

The scheme provided in this paper facilitates the client in getting a proof of integrity of the data, a cloud hosting and storage security that collectively deals with security and performance. It is done by two functions, the bit generator function g and the function h which is used for encrypting the data. Hence the storage at the client is very much minimal compared to all other schemes that were developed.

Cloud services are nowadays more expensive and vastly developing service based on the several clouds and is a prevalent. One of the most limitations in the clouds is price. So we developed this project which helps the clients to choose the clouds with efficient price and assurance data which helps the user in cost and time complexity is maintained because we are maintaining and generating the

session keys and private keys for high security and data confidentiality.

Cloud services are experiencing rapid development and the services based on multi-cloud also become prevailing. One of the most concerns, when moving services into clouds, is capital expenditure. So, in this paper, we design a novel storage scheme CHARM, which guides customers to distribute data among clouds cost-effectively. CHARM makes fine-grained decisions about which storage mode to use and which clouds to place data in. The evaluation proves the efficiency of CHARM.

References

[1] Mansouri, Y,Toosi, A.N., Buyya, R.: Brokering algorithms for optimizing the availability and cost of cloud storage services. In: Proceedings of the 2013 IEEE International Conference on Cloud Computing Technology and Science, Washington, DC, USA, vol. 01, pp. 581–589 (2013).
[2] H.Shacham and B. Waters, "Compact Proofs of Retrievability," In Proceedings. of Asiacrypt "08,Dec. 2008.
[3] Shaik. AafreenNaaz, Pothireddygari. Ramya, P.VishunuVardhanReddy.,"Cloud Computing: Use of Multi-Clouds" pp. 295-304s.
[4] "ApacheLibcloud"http://libcloud:apache:org/.
[5] H. Weatherspoon and J. D. Kubiatowicz, "Erasure Coding vs. Replication. A Quantitative Comparison," in IPTPS. Springer, 2002.
[6] R. Rodrigues and B. Liskov, "High Availability in DHTs: Erasure Coding vs. Replication," in IPTPS. Springer, 2005.
[7] Y. Ma, T. Nandagopal, K. P. Puttaswamy, and S. Banerjee, "An Ensemble of Replication and Erasure Codes for Cloud File Systems," in INFOCOM. IEEE, 2013.
[8] H. H. Liu, Y. Wang, Y. R. Yang, H. Wang, and C. Tian, "Optimizing Cost and Performance for ContentMultihoming," 2012.
[9] M. Pitkanen, R. Moussa, M. Swany, and T. Niemi, "Erasure Codes for Increasing the Availability of Grid Data Storage," in AICTICIW. IEEE, 2006.
[10] H. B. Ribeiro and E. Anceaume, "Datacube: A P2P Persistent Data Storage Architecture Based on Hybrid Redundancy Schema," in PDP. IEEE, 2010.
[11] A. Duminuco and E. W. Biersack, "Hierarchical Codes: A Flexible Trade-off for Erasure Codes in Peer-to Peer Storage Systems," Peerto- Peer Networking and Applications, vol. 3, no. 1, pp. 52–66, 2010.
[12] W. Lin, D. Chiu, and Y. Lee, "Erasure Code Replication Revisited," in P2P. IEEE, 2004.
[13] Y. Ma, T. Nandagopal, K. P. Puttaswamy, and S. Banerjee, "An Ensemble of Replication and Erasure Codes for Cloud File Systems," in INFOCOM. IEEE, 2013.
[14] M.P. Papazoglou and W. van den Heuvel, "Blueprinting the Cloud," IEEE Internet Computing, Nov. /Dec 2011, pp. 74-79
[15] R. Thandeeswaran, S. Subhashini, N. Jeyanthi1, M. A. SaleemDurai, "Secured Multi- Cloud Virtual Infrastructure with Improved Performance", cybernetics and information technologies XII, (2), pp. 11-22, 2012 Society IEEE, 2013
[16] S. Ortiz Jr., "The Problem with Cloud Computing Standardization," Computer, July 2011, pp. 13-16.
[17] DuraCloud. [Online]. Available: http://www.duracloud.org/, 2014.
[18] Cloud Foundry. [Online]. Available: http://www.cloudfoundry. org/, 2014.
[19] ApacheLibcloud. [Online]. Available: http://libcloud.apache. org/, 2014.
[20] AmazingStore. [Online]. Available: http://cn.amazingstore.org/, 2014.
[21] S. Liu, X. Huang, H. Fu, and G. Yang, "Understanding data characteristics and access patterns in a cloud storage system," in Proc. 13th IEEE/ACM Int. Symp. Cluster, Cloud, Grid Comput., 2013, pp. 327–334.
[22] P. Wendell, J. W. Jiang, M. J. Freedman, and J. Rexford, "Donar: Decentralized server selection for cloud services," in Proc. ACM SIGCOMM Conf., 2010, pp. 231–242.
[23] H. H. Liu, Y. Wang, Y. R. Yang, H. Wang, and C. Tian, "Optimizing cost and performance for content multihoming," in Proc. ACM SIGCOMM Conf. Appl., Technol., Archit., Protocols Comput. Commun., 2012, pp. 371–382.
[24] T. G. Papaioannou, N. Bonvin, and K. Aberer, "Scalia: An adaptive scheme for efficient multi-cloud storage," in Proc. Int. Conf. High Perform. Comput., Netw., Storage, Anal., 2012, p. 20.
[25] J. S. Plank, "Erasure codes for storage systems: A brief primer," Usenix Mag., vol. 38, no. 6, pp. 44–50, 2013.
[26] J. S. Plank, K. M. Greenan, and E. L. Miller, "Screaming fast galois field arithmetic using Intel SIMD instructions," in Proc. 11th USENIX Conf. File Storage Technol., 2013, pp. 299–306.
[27] H. Weatherspoon and J. D. Kubiatowicz, "Erasure coding vs. replication: A quantitative comparison," in Proc. 1st Int. Workshop Peer-to-Peer Syst., 2002, pp. 328–338.
[28] R. Rodrigues and B. Liskov, "High availability in DHTs: Erasure coding vs. replication," in Proc. 4th Int. Conf. Peer-to-Peer Syst., 2005, pp. 226–239.