

# A Survey on Cost-Efficient Multi-Cloud Data Hosting Scheme with High Availability

Shruthi L<sup>1</sup>, D Khasim Vali<sup>2</sup>

PG Scholar, Department of CSE, Vidya Vardhaka College of Engineering, Mysuru, India<sup>1</sup>

Associate Professor, Department of CSE, Vidya Vardhaka College of Engineering, Mysuru, India<sup>2</sup>

**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 paper 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 thin clients.

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

## I. 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 servers 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 outpacing 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 survey paper, we focus on efficient and heuristic-based data hosting scheme for heterogeneous multi-cloud environment and flexible transaction scheme for CHARM. This CHARM scheme intelligently puts data into multiple clouds with minimized monetary cost and guaranteed availability.

## II. 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. Aafreen Naaz [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 clouds 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 can not 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.Papazoglone etc[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 to be 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.Ortiz 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.

### III. 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.

### 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 Asiacypt '08, Dec. 2008.
- [3]. Shaik. Aafreen Naaz, Pothireddygari. Ramya, P. Vishnu Vardhan Reddy., "Cloud Computing: Use of Multi-Clouds" pp. 295-304s.
- [4]. "Apache Libcloud" <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 Content Multihoming," 2012.
- [9]. M. Pitkanen, R. Moussa, M. Swamy, and T. Niemi, "Erasure Codes for Increasing the Availability of Grid Data Storage," in AICT-ICIW. 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," Peer-to-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. Jeyanthi, M. A. Saleem Durai, "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.