

A Survey on Guaranteed Availability of Cloud Data with Efficient Cost

Vignesh M¹ Dr. Rajesh T M²

²Associate Professor

^{1,2}Department of Computer Science & Information Technology

^{1,2}Dayananda Sagar University, Bangalore, India

Abstract— Cloud computing is an innovative proficiency in the field of information and expertise. It provides so many things in terms of “As-A-Service” basis. Cloud Computing is the basically visualization of computing as a utility, where customers can tenuously store up their data into the cloud so as to benefit from the on-request services from a collective pool of configurable computing assets. In recent day with increasing the population of data in whole world, the cloud storage space requirement is also increasing rapidly. The main advantage of using cloud storage is, customer can reduce their expenditure for purchasing and maintain the storage area, but they have to use the cloud in optimized manner. In this work, we propose a cost proficient multi-cloud for the data with high data availability. The proposed work contains two principle works; the first is selecting a few reasonable cloud storage providers and a fitting repetition procedure to store information with minimalized monetary cost and guaranteed data availability. The second is setting up a redundant storage method to re-distribute the data among different available data access patterns and cloud pricing models. The proposed technique not just saves around 20 percent of monetary cost additionally shows comprehensive flexibility to information and value changes.

Key words: Multi-Cloud, Data Availability, Cloud Storage

I. INTRODUCTION

Cloud computing will be the core of information infrastructure in today's technology. It offers all kinds of services for the users. The considerable service happening by the cloud is nothing but storage competence. Storage outsourcing is fetching more and more attractiveness to both industry and academic outstanding to the advantages of low cost, ease of access, and easy distribution. As one of the storage outsourcing form, cloud storage has gained broad interest in recent years. Many companies, such as Amazon, Google, and Microsoft provide their own cloud storage space services, where the users can upload their files to the servers, access them from a variety of devices, and share them with the others even though cloud storage services are extensively adopted in current days, there still stay behind many security issues and potential threats. Cloud computing provides different types of services to the users over the network. It enables organizations to consume resources as utility just like electricity. Data hosting services provide users with an efficient and reliable way to store data and this stored data can be accessed from anywhere, on any device, and at any time. Cloud computing is Internet based technology that provides on demand access to shared pool of computing resources and on pay per use basis. Cloud computing provides distributed environment which is important to develop large scale applications rapidly.

In recent years data hosting services has become more popular so that there are many cloud service providers offering data hosting services. In most of the cases

companies moving towards hosting their data into a single cloud. However, there are several choices available from various cloud vendors.

II. LITERATURE SURVEY

There are many benefits of cloud computing, it also brings new challenges in maintaining data integrity and high reliable data availability. Cost is also a major concern when switching to multi-cloud system. The author Nitesh Shrivastava et al [1] made an research related to single and multi-cloud cost, security and availability based scenario. The work purposes to sustenance the usage of multi-cloud environment over single cloud in order to reduce the risk.

Prof.V.N.Dhawas et al [2] proposed a secured cost-effective multi-cloud storage prototype in cloud computing which holds an efficient distribution of data amongst the available cloud Service Providers, to deliver customers with high data availability as well as secure storage. Providing improved privacy and guaranteeing data availability can be achieved by fragmenting the customer's data blocks into data pieces and putting them among the various available Service Providers.

High data availability in peer-to-peer data hosting scheme demands data redundancy. In this comparison, the author HakimWeatherspoon et al [3] compares two common redundancy schemes: replication and erasure coding. In contrast to previous comparisons, the authors have taken the features of the nodes that encompass the overlap into consideration, and investigate that in some cases the advantages from coding are restricted, and possibly will not be worth its disadvantages.

Zhenhua Li et al [4], Cloud storage services such as Dropbox, Google Drive, and Microsoft OneDrive provide users with a convenient and reliable method to store and share data from anyplace, on any device, and at any time. The keystone of these services is the data synchronization operation which automatically maps the changes in users' local filesystems to the cloud through a series of network communications in a appropriate timely manner. If not structured properly, the tremendous amount of data synchronization congestion can potentially cause troubles to both cloud service providers and cloud consumers.

Understanding the integral system characteristics is crucial in designing and optimization of cloud storage system, and few studies have systematically explored the data features and access patterns. Authors S.Liu, et al [5] [6] have found that there are much diversity amongst cloud storage system and traditional file systems: the cloud storage system has high file sizes, smaller read and write ratio, and low set of active files than that of a typical traditional file system. With a trace driven mockup, they found that the cache effectiveness can be enhanced by 5 times using the inputs from authors observations. To tackle the challenges in migrating the enterprise application services into multi-

cloud environment -based deployments, organizations processes are partially put on-premise and partly in the multi-cloud. The use multi-cloud architectures enable organizations to benefit from cloud-based architectures, though accepting application performance rations, and privacy limitations on which all services could be migrated to the multi-cloud.

With the increasing popularity of cloud storage is leading many enterprises to consider moving data out of their own data centers and into the cloud. However, achievement for cloud storage providers can be a significant risk to clients, as it becomes very expensive to shift storage providers. In the work presented by author Hussam Abu-Libdeh et al [7], achieves a instance for using RAID-like techniques using disks and file systems in the cloud storage, writer claims that by method of striping the user's data into multiple cloud service providers can consent customers to prevent from vendor lock-in risk, reduce the cost of switching cloud providers, and improved fault-tolerate provider failures. With introduce Redundant Array of Cloud Storage, a proxy technique that transparently blowouts the storage load across providers. The author evaluates a model of the system and evaluates the costs incurred and benefits reaped.

The study on storage system Manorama et al [8] the author implements erasure coding and replication across multiple data centers are conferred that with the cache is primary for data centers. The heterogeneity of multi cloud and the identifying of clouds are not worth the advantage, the cache benefits in storing back of data when accessed by erasure coding frequent data swap certainly induces additional cost which is longer competitive when compare to other data hosting schemes.

An extensive research is done on data hosting in peer-to-peer storage systems the authors has stated in [9],[10],[11],[12] deals with the major feature of storage system is that storage nodes are not stable. There are many merits and demerits of peer-to-peer system mechanism in multi cloud environment cannot be compared since it is proved very different from the results in other method.

III. PRESENT METHODOLOGY

An extensive number of procedures have been proposed by various analysts for hosting their data in the cloud, Numerous Cloud vendors develop their own infrastructure and upgrade them with recently emerging technologies. The cloud services can be accessed by using Single service provider and multiple services providers. The problem with these models Is that it doesn't provide guaranteed availability. By the use of single cloud data hosting scheme, the cloud customers host their data into single based hosting scheme and then just trust to luck, leading to vendor lock-in risk. In addition, the loss of service availability and data integrity are the major problems for the customer. To tackle the existing scenario, a collaborative cost-effective solution for multi-cloud data hosting scheme should be found. Also, to meet the availability requirement of different services, an economical redundancy mechanism (i.e., replication or erasure coding) based on specific data access patterns can be found. Much research has been done on the single cloud and more research is going in the area of multi-cloud to overcome these cost and security issue as well

IV. ISSUES AND CHALLENGES

In the existing methodology the major concern for cloud users, is how to efficiently select suitable clouds providers to minimize monetary cost in the existence of heterogeneous cloud storage pricing policies, to can we achieve different availability requirements for different services. As to decrease the cost, it primarily depends on the level of data usage, mostly storage capacity consumption and consumption of network bandwidth. The Existing clouds models exhibit excessive heterogeneities in terms of both working performances and pricing strategies. The Customers usually host their data into a single cloud and then simply trust to luck. This may lead to a problem so-called vendor lock-in risk. The foremost major concern is data availability. It is not wise for any enterprise or an organization to host all data in a single cloud for security and data availability problems

V. CONCLUSION AND FUTURE SCOPE

The present days, the cloud services are quite popular among enterprise and are developing a numerous techniques to overcome the problem of vendor-lock in and single cloud data hosting scheme, in which if the failure of a particular cloud we cannot retrieve the data. Hence, we can implement a multi-cloud data hosting system, where we can store the data across multiple storage spaces and can retrieve data from any one of the storage cloud without having to concern about cloud failures.

REFERENCES

- [1] NiteshShrivastava, Ganesh Kumar "A survey on cost effective multi-cloud storage in cloud computing" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013
- [2] Prof.V.N.Dhawas,PranaliJuikar "A Secured Cost Effective Multi-Cloud Storage in Cloud Computing" International Journal of Scientific & Engineering Research, Volume 4, Issue 5, May-2013 1017 ISSN 2229-5518.
- [3] HakimWeatherspoon and John D. Kubiatowicz "Erasure Coding Vs. Replication: A Quantitative Comparison" 2012
- [4] Zhenhua Li 1;2, Christo Wilson 3, Zhefu Jiang 4, Yao Liu 5, Ben Y. Zhao 6, Cheng Jin 7, Zhi-Li Zhang 7, and Yafei Dai. Efficient Batched Synchronization in Dropbox-like Cloud Storage Services
- [5] S.Liu, X.Huang, H.Fu, and G.Yang, "Understanding Data Characteristics and Access Patterns in a Cloud Storage System" in CCGrid IEEE, 2013
- [6] Mohammad Hajjat, Xin Sun, Yu-Wei Eric Sung, David Maltz, Sanjay Rao Kunwadee Sripandkulchai and Mohit Tawarmalani,Cloudward Bound: Planning for Beneficial Migration ofEnterprise Applications to the Cloud,2010
- [7] Hussam Abu-Libdeh, Lonnie Princehouse, Hakim Weatherspoon RACS: A Case for Cloud Storage Diversity
- [8] 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.
- [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.

