

CHARM: Intelligently Cost and Bandwidth Detection for FTP Servers using Heuristic Algorithm

Shiva Urolagin¹

Student, Department of CSE, KLE Dr. M S Sheshgiri College of Engg & Tech., Belgaum, India¹

Abstract: CHARM - In this sense it is a data hosting schema and it helps to reduce the cost of hosting data and also helps to detect the bandwidth and speed. Now a day more and more companies, organizations and enterprises are hosting their data into the cloud to reduce maintenance, cost and data reliability. Many cloud storage companies are providing cloud spaces to store the data, but vendors are getting problems with which cloud space is better to store the data and also getting problem with costs. Customers usually store their data into single cloud but some clouds are high price and low bandwidth in this situation customer may face problems with uploading their data. Sometimes customer may struck with lock in risk, In this sense the customer suffer from price adjustments of cloud vendors because price is not fixed according to market. According to their need they bind up with one company and company polices. This proposed CHARM system helps to analysis the cost and bandwidth of different clouds, In this system customers possible to analysis the speed and cost of different clouds before they putting data. This system integrates two key functions. First is selecting a severable suitable clouds and appropriate redundancy to store the data with low cost and guaranteed availability. Second is putting a transmission process to re-distribute the data according to variations of data axis and pricing. Performance of the CHARM can be analyse using trace-driven simulation and prototype experiments.

Keywords: Cloud computing, Multi-cloud, Data hosting, Bandwidth analyser.

I. INTRODUCTION

Cloud storage services such as Amazon S3, Windows Azure and Google cloud storage are providing cloud services with reliable, scalable and low cost. Different Cloud vendors build their respective infrastructure and they apply various techniques to make their service competitive for example, Amazon charges more for storage and Google cloud storage charges more for bandwidth. In this system multi cloud data hosting has introduced. Distributing data into many clouds is known as multi cloud data hosting. Now a days it's receiving wide attention from customers, researchers and start-ups.

Now a day's many companies, organization and enterprises are hosting their data in clouds to reduce IT maintenance and costs but sometimes they may get struck with cost and speed. And they also face with lock in risk. In the sense according to customer need they put their in a particular company and they make agreement with particular company but after some days they face problem with cost and speed. Particular customer data is not possible to transfer data from one cloud to companies. This proposed system helps to analyses the cost, bandwidth and speed. User is possible to analyses the different cloud through this system.

In this proposed system they are helped to analysis the cost and bandwidth of different clouds, in this system customer possible to analysis the speed and cost of different clouds before them putting data. This system

integrates two key functions. First is selecting a severable suitable clouds and appropriate redundancy to store the data with low cost and guaranteed availability. Second is putting a transmission process to re-distribute the data according to variations of data axis and pricing. We analyses the performance of the CHARM using trace-driven simulation and prototype experiments. So a result shows that CHARM saves around 20% of price adjustments.

A. Vendor lock in risk:

In this customers suffer from price adjustments and company policies because according to market requirement price may vary. Customers usually put their data into single cloud and simply trust to luck, the another risk for customer is if cloud services suddenly stop their services then customer has to spend lot of money to transfer data. For example if the customer switch to particular cloud then after sum days he may want to more bandwidth and speed then particular customer cant possible to transfer the whole data, this time vender may struck with lock-in-risk.

Customer according to their need they bind up with one particular company and company polices, later if customer may want to change the services then customer has to pay huge amount of money to transfer the data. If any customer makes agreement to particular company then customer will not possible to switch to another company.

The vendor lock in risk makes users to suffer from price adjustment of cloud providers which are not uncommon. For example, fluctuation of electricity bills will affect the prices of the cloud services. Another vendor lock in risk is quality of services, guaranteed services, level of agreements and failure do occur etc.

B. Heterogeneous clouds:

Different cloud providers build the different infrastructures and they keep upgrading according to market. They also design different architectures and apply various techniques to make service competitive. Such cloud providers keep their clouds upgrading according to market requirements. All cloud providers set their prices as per as services and customer needs. For instance Rack space does not charge for web operations and Google cloud storage charges more for bandwidth and Amazon S3 charge more for storage.

C. Multi Cloud data hosting:

The basic principle of multiple cloud data hosting is distributing the data across the multiple clouds to prevent vendor lock in risk. Now a days multi cloud data hosting getting wide attention from customers, researchers and starters.

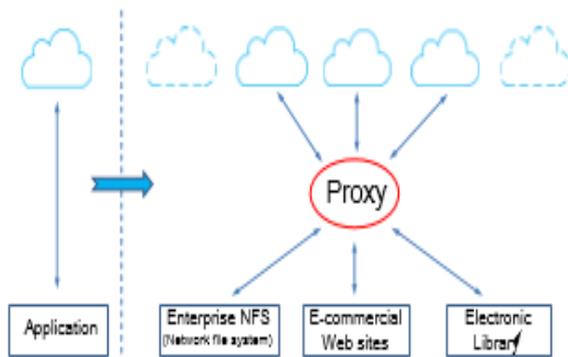


Figure 1: Multi Cloud data hosting

The proxy work as a mediator it collects the user requirements and access from the cloud to provide services to users. Proxy play as a key rule, it collects the user requirements and spreads across different clouds. The Application server handles the applications and operations. The Multi cloud is divided in three folds, first is “Depsky” it guarantees that data availability and security on multiple clouds and also stores the data trustfully. Example, financial and medical data. “RACS” In this system it connects between the different clouds to deploy the vendor lock in risk and it helps to reduce the cost. Second is new type of cloud vendors grow up rapidly to provide services on multi cloud. Third is a new deployment tool like Apache Libecloud provides a interface on different clouds.

Multi cloud People still encounters two critical problems they think how to choose appropriate clouds and minimize the clouds and how differentiate availability requirements of different services, to avoid this and to maintain proper costs and achieve multi clouds this propose system helps.

II. THE PROPOSED CHARM SCHEME

Proposed CHARM technique helps to reduce the cost with high availability in heterogeneous multi cloud. This technique intelligently puts data in multiple cloud with low costs and guarantee availability it combines to redundancy mechanism either replication and ensure coding. Heuristic-based algorithm to choose proper data storage. CHARM helps to reduce the 20% of cost and also helps to find out one good cloud storage. Multi cloud helps to distribute the data over the Different clouds.

This proposed system helps to reduce the cost and also help to choose a proper cloud according to user need. It is a heuristic based data hosting scheme for multi clouds. CHARM help to find out different clouds, different pricing strategies, requirement of availability and data accessing pattern. It chooses the appropriate redundancy strategy to store data with minimized cost and guaranteed availability. It displays each cloud price and available bandwidth. CHARM helps to monitoring the variations of pricing policies of different clouds and data access system. It also starts a data migration process among different clouds if necessary. Performance of the CHARM is possible to calculate using both trace driven experiments and simulation.

III. LITERATURE SURVEY

In this paper [1] CloudCmp-comparing public cloud provider (2010), in this paper they are concentrating on cost of per clouds and services of per clouds. These peoples are referring and comparing different cloud users. This technique is helps to analyse the cost and service of different clouds using CloudCmp. CloudCmp is a tool it intelligently helps to analyse the services of clouds. These system keywords are Cloud computing, cost, performance and comparison. In this paper they are creating a tool for analyse the behaviour of different clouds. Because of this system user may not struck with problems, intelligently user can possible to analyse cost and behaviour of clouds.

In this paper [2] Meeting Service Level Agreement Cost-Effectively for Video-on-Demand Applications in the Cloud (2014) in this paper they used VOD-Video-on-demand application. VOD Applications are high demand and high performance applications. In this paper used heuristic algorithm, this algorithm helps to analyse the performance of the cloud services. VOD applications are contain more space so these applications need more bandwidth, high performance and low cost.

In this paper [3] Optimizing Cost and Performance for Content Multi homing (2014)-Many large content publishers use multiple content distribution networks to deliver their content, and many commercial systems have become available to help a broader set of content publishers to benefit from using multiple distribution networks, which we refer to as content multi homing.

In this paper [3] Also design a novel, lightweight client adaptation algorithm executing at individual content viewers to achieve scalable, fine-grained, fast online adaptation to optimize the quality of experience (QoE) for individual viewers. We prove the optimality of our optimization algorithms and conduct systematic, extensive evaluations, using real charging data, content viewer demands, and performance data, to demonstrate the effectiveness of our algorithms. We show that our content multi homing algorithms reduce publishing cost by up to 40%. Our client algorithm executing in browsers reduces viewer QoE degradation by 51%.

In this paper [4] Redundancy is the basic technique to provide reliability in storage systems consisting of multiple components. A redundancy scheme defines how the redundant data are produced and maintained. The simplest redundancy scheme is replication, which however suffers from storage inefficiency. Another approach is erasure coding, which provides the same level of reliability as replication using a significantly smaller amount of storage. When redundant data are lost, they need to be replaced. While replacing replicated data consists in a simple copy, it becomes a complex operation with erasure codes: new data are produced performing a coding over some other available data.

In this paper [4] The amount of data to be read and coded is d times larger than the amount of data produced, where d , called repair degree, is larger than 1 and depends on the structure of the code. This implies that coding has a larger computational and I/O cost, which, for distributed storage systems, translates into increased network traffic. Participants of Peer-to-Peer systems often have ample storage and CPU power, but their network bandwidth may be limited. For these reasons existing coding techniques are not suitable for P2P storage.

This work explores the design space between replication and the existing erasure codes. We propose and evaluate a new class of erasure codes, called Hierarchical Codes, which allows reducing the network traffic due to maintenance without losing the benefits given by traditional erasure codes.

III. SYSTEM ARCHITECTURE

Most popular cloud storages are Amazon S3, Windows Azure, Google cloud storage, Rack space and Aliyun OSS. According to market all the companies pricing system differs. The Amazon cloud services charges more for storage like that Windows Azure charge more for bandwidth. These companies are upgrading those platforms according to market and user requirements.

In this system Heuristic algorithm has help to analyse the bandwidth of clouds. Because of this system customer may not struck with lock in risk. Bandwidth can be analyse throw uploading and downloading speed. The bandwidth analyser pop up the graph to analyse the speed

of cloud. Using this system customer can come to know about price and speed and also customer can possible to choose proper cloud.

Multi-cloud data hosting helps to spread the data throw different clouds and also helps to gain the data from different clouds. Conceder for example if any user upload his data then it store in different clouds. In multi-cloud data hosting operation the proxy server works likes a mediator. Proxy server helps to extract the data to different clouds and also helps to access the data from different clouds.

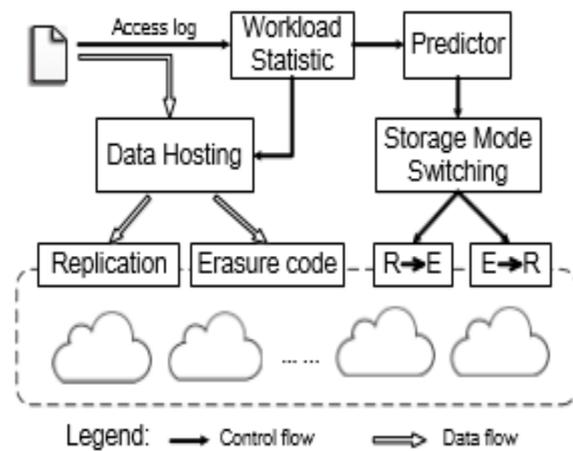


Figure 2: System architecture

In this system there are four main components those are data hosting, storage mode switching, work load statistic and predictor. This whole model located in proxy. It helps in reduction of cost and data hosting with high availability.

- Replication: Replication is nothing but data replacement. It helps to put the data into different clouds. Data hosting stores the data using replication.
- Work load statistics: It helps to balance the load according to data, it keeps connecting and access the data to guide the placement of the data and it also sends the static information to the predictor. Predictor guides the action of SMS.
- Data hosting: Data hosting helps to put the data into the cloud and also stores the data using the replication or ensure coding, According to the size and axis frequency of the data.
- Predictor: Predictor is a outcome of result and it runs in the background. It is used to predict the future access frequency of the files. One month time interval for prediction of data. There are lot of algorithms for the prediction. It takes data from work load statics and data hosting.
- Storage mode switching: It's nothing but switching from one cloud to another cloud. Storage mode switching decides whether the storage mode and certain data should be changed from replication to ensure loading or reserve according to the output of predictor.

IV. RESULTS

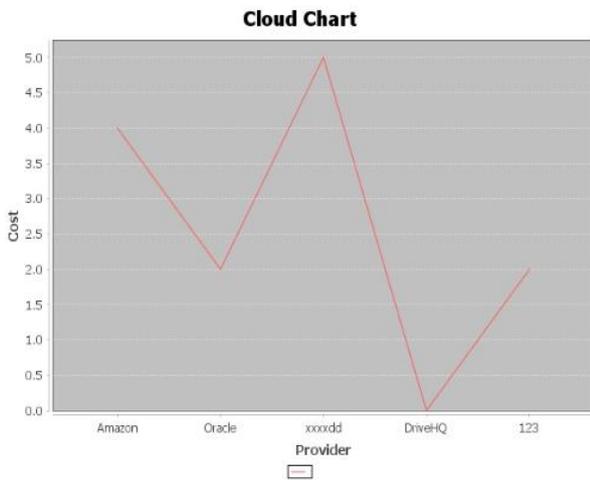


Figure 3: Cost analyser graph.

The figure 3 shows the result of cost variation between different clouds. In graph one side is cost representation is there and another side cloud providers are there. This graph helps to analyse the different clouds cost. For example oracle make 2.0 cost for data hosting, same as amazon charges 4.0. Because of this graph the customer can analyse the cost of different clouds.

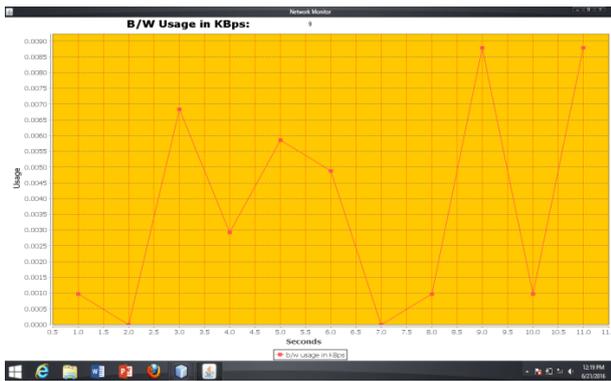


Figure 4: Bandwidth analyser graph.

The figure 4 represent bandwidth analyser graph. It helps to analyse the bandwidth of different clouds. According to user need the user can possible to buy the clouds.

V. CONCLUSION

This proposed system helps to analyse the cost and bandwidth of different clouds. Usually cloud users may struck with lock-in-risk. Users sum times struck with those needs, so this CHARM system helps to analyse cost and bandwidth intelligently. This system gives a clear idea about cloud providers and company polices. User after register with this system user can get a clear idea about cost and bandwidth of different clouds and user possible to select clouds according to his needs. User after connecting with CHARM system user can check the rate description and bandwidth variation. This system displays the graph of

bandwidth analyser and cost analyser. Thrown this system user can get brief idea about clouds and according to his need user can possible to choose particular data. In this system user can directly possible to upload and download the data. User may not need to go on clouds and upload the data, throw a CHARM system user can upload the data.

REFERENCES

- [1] A. Li, X. Yang, S. Kandula, and M. Zhang, "CloudCmp: Comparing Public Cloud Providers," in IMC. ACM, 2010.
- [2] Y. Zhao, H. Jiang, K. Zhou, Z. Huang, and P. Huang, "Meeting Service Level Agreement Cost-Effectively for Video-on-Demand Applications in the Cloud," in INFOCOM. IEEE, 2014.
- [3] H. H. Liu, Y. Wang, Y. R. Yang, H. Wang, and C. Tian, "Optimizing Cost and Performance for Content Multihoming," 2012.
- [4] 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.
- [5] C. M. M. Erin Allen, "Library of Congress and DuraCloud Launch Pilot Program Using Cloud Technologies to Test Perpetual Access to Digital Content," The Library of Congress, News Releases, <http://www.loc.gov/today/pr/2009/09-140.html>.
- [6] "Summary of the Amazon EC2 and Amazon RDS Service Disruption in the US East Region," <http://aws.amazon.com/cn/message/65648/>.
- [7] P. Wendell, J. W. Jiang, M. J. Freedman, and J. Rexford, "Donar: Decentralized Server Selection for Cloud Services," 2010.
- [8] C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li, S. Yekhanin et al., "Erasure Coding in Windows Azure Storage," in ATC. ACM, 2012.
- [9] M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A. G. Dimakis, R. Vadali, S. Chen, and D. Borthakur, "XORing Elephants: Novel Erasure Codes for Big Data," The VLDB Endowment, vol. 6, no. 5, pp. 325–336, 2013.
- [10] Y. Hu, H. C. Chen, P. P. Lee, and Y. Tang, "NCCloud: Applying Network Coding for the Storage Repair in a Cloud-of-Clouds," in FAST. ACM, 2012.
- [11] W. Lin, D. Chiu, and Y. Lee, "Erasure Code Replication Revisited," in P2P. IEEE, 2004.
- [12] Y. Zhao, H. Jiang, K. Zhou, Z. Huang, and P. Huang, "Meeting Service Level Agreement Cost-Effectively for Video-on-Demand Applications in the Cloud," in INFOCOM. IEEE, 2014.
- [13] Y. Hu, H. C. Chen, P. P. Lee, and Y. Tang, "NCCloud: Applying Network Coding for the Storage Repair in a Cloud-of-Clouds," in FAST. ACM, 2012.
- [14] M. Pitkanen, R. Moussa, M. Swamy, and T. Niemi, "Erasure Codes for Increasing the Availability of Grid Data Storage," in AICT-ICIW. IEEE, 2006.
- [15] H. B. Ribeiro and E. Anceaume, "Datacube: A P2P Persistent Data Storage Architecture Based on Hybrid Redundancy Schema," in PDP. IEEE, 2010.
- [16] 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.

BIOGRAPHY



Shiva M Urolagin is an M.Tech student in the Department of Computer Science & Engineering, KLE DR. M S Sheshgiri College of Engineering Belgaum. He completed his Bachelor of Engineering in Computer science & Engineering from S.G. Balekundri Engineering College, Belagavi, in the year 2014.