|| Volume 3 || Issue 9 || September 2018 || ISSN (Online) 2456-0774

INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS

Survey on Minimizing Payment Cost of Cloud Service Provider

Tarun Sharma[1], Prof. Nagaraju Bogiri[2]

Department of Computer Engineering K.J. College of Engineering & Management Research Pune

tphsharma16@gmail.com[1],mail2nagaraju@gmail.com[2]

Abstract— **Many industries and research center using a cloud service provider (CSP) provider for storing data on that and CSP used many web applications such as web portal, online social network providing services to the clients all over the world. These types of datacenters provide the different unit prices and get/put latencies for resources reservation and allocations. Selection of different CSPs datacenters and cloud customers facing two challenges (I.) How to allocating data to the datacenters in worldwide to satisfy application Service Level Objectives (SLO) requirement which includes both data availability and retrieval latency. (II.) How to allocate reserve resources and data in the datacenters, which belongs to different CSP to minimizing payment cost. Find out the solution of these challenges firstly we used integer programming techniques for handles cost minimization problems. We propose three techniques for reducing the service latency and payment cost 1. Multicast Based Data Transferring, 2. Coefficient Based Data Reallocation and 3. Request Redirection Based Congestion.**

Keywords: *Cloud Service Provider, Service Level Objectives, Payment Cost Minimization, and Data Availability.*

I INTRODUCTION

In recent year worldwide cloud storage such as Microsoft Azure, Google Cloud storage and Amazon S3, has become more popular. Cloud Service Provider (CSP) provides various data storage services such as Get and Put services help of worldwide geographically distributed datacenters. Fig. 1 Shows that the cloud storage service provider's overview. Most of enterprisers shift data workload on the cloud storage to save capital expenditure to maintain and build the hardware infrastructures and avoid the complexity of managing datacenters. CSP used many web applications such as web portal, online social networking for the providing services to clients all worldwide. Such type of web applications availability and data access delay are critical, which is affected on the cloud customer's incomes. In order to reduce data access latency, the data requested by clients needs to be handled by datacenters near the clients, which requires worldwide distribution of data replicas.

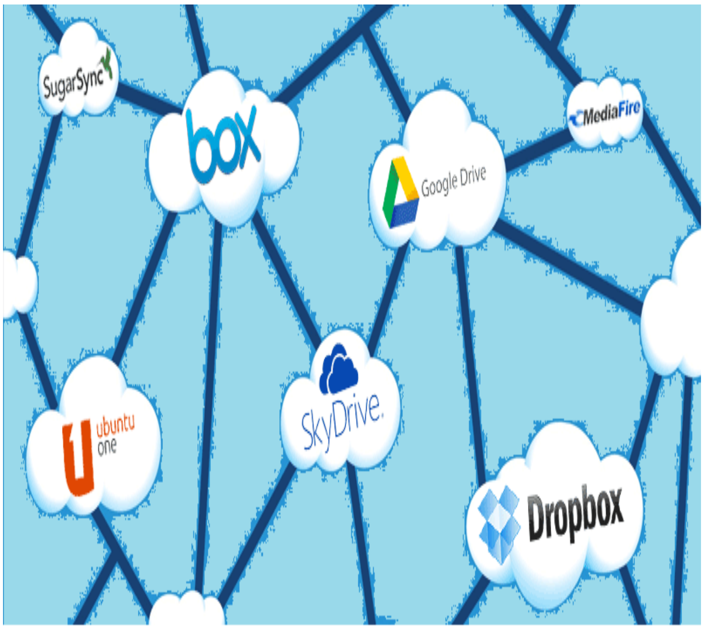

Figure 1 : Overview of Cloud Storage Service Providers

The payment cost of CSP depends on the cost of data GET/PUT, Storage and Transfer of Data. Different types of DSP have different cost. For examples Amazon provides less data storage cost (($0.01/GB and $0.005/1,000 request) and Windows Azure data Get/Put price ($0.024/GB and $0.005/100,000 requests). Table 1 Shows that the examples of CSS offered by cloud service providers.

Table1. Cloud Storage Service Provider Examples

Provider Name	Storage Services
Amazon	Queue, File, Blob, Tables, Redis Cache
Google	EBS, S3, EFS, Glacier, DynamoDB
Microsoft	Data store, Big-table, Storage

II LITERATURE SURVEY

In this section discuss the literature review in detail about the multiple cloud storage providers:

In this paper [1] they propose SPANStore to export a unified view of geographically distributed storage services to applications and to automate the process of trading off cost and latency, while satisfying consistency and fault-tolerance requirements. SPANStore achieves this goal by spanning data centers of multiple cloud providers, by judiciously determining replication policies based on workload properties, and by minimizing the use of computer resources. SPANStore minimizes the use of computer resources to implement tasks such as two-phase locking and data propagation, which are necessary to offer a global view of the storage services that it builds upon.

MINERVA combines a number of different elements like a constraint-based representation of each design problem and a set of optimization strategies and heuristics to search for solutions, guided by predictions of the likely effects of each choice made. MINERVA addresses all of these issues: it designs storage systems that are as good as or better than the ones produced by people; it does so in minutes rather than days; and it accompanies them with predictions of the resulting system's performance [2].

Cloud services make use of Volley by submitting logs of datacenter requests. Volley analyzes the logs using an iterative optimization algorithm based on data access patterns and client locations, and outputs migration recommendations back to the cloud service. In this paper [3] they have focused on using Volley to optimize data placement in existing datacenters, service operators could also use Volley to explore future sites for datacenters that would improve performance. By including candidate locations for datacenters in Volley's input, the operator can identify which combination of additional sites improve latency at modest costs in greater inter-datacenter traffic.

This paper propose [4] SafeStore distributed storage system designed to maintain long-term data durability despite conventional hardware and software faults, environmental disruptions, and administrative failures caused by human error or malice. The architecture of SafeStore is based on fault isolation, which Safe-Store applies aggressively along administrative, physical, and temporal dimensions by spreading data across autonomous storage service providers (SSPs).

Cloud storage service provider provides the platform for storing large scale data on it with the minimum prize. Massive Parallel Processing (MPP) is used to bridge the gap between the traditional data warehouse and modern cloud storage. In this paper [5] they implement Open Source Prototype GPCloud for Load/Upload data on the cloud storage this technique is based on cloud storage Amezon S3 and MPP data warehouse Greenplum. An experimental result shows that performance of this technique better than the existing system. This system is supported for INSERT and SELECT operation.

Resource provisioning to the computational task is the major challenge for cloud computing. In a paper [6] they propose prediction of cloud resources utilization on the pre-resource level and pre-task. In this they used machine learning techniques for prediction of resource utilization on the cloud storage. For this experiment they used dataset on the GitHub Travis CI and Travis CI.

The performance of this system is compared with the simple learning regression approach and results shows that the increase the accuracy of this system. In this [7] firstly study the problems of jointly optimizing resource allocation and video management for the large scale Video-on-Demand (VoD) cloud. They propose RAVO (Resource Allocation and Video Management Optimization) model for jointly mange videos to achieve low cost and allocate system resources. For managing large video pool they used clustering algorithm. The performance of this system is compares with the other state-of-art techniques such as iGreedy, IPTV-RAM (internet protocol television-resources allocation and management) and super-optimum.

Cloud computing is a powerful technology for performing operation such as complex and massive-scale computing. At this time size of data increasing and also verity of data generated and expanding every day. The use of cloud service provider is that to process, store and analyze data. In this paper [8] their present classification techniques for big data, cloud service model and conceptual view of data. They also review on the challenges in big data processing. The review conducted on scalability, volume, availability, data protection, data transportation, data heterogeneity/quality, regularity/ legal issues, privacy, governance and data access.

In this paper [9] they present techniques such as "Supercloud", it is deployed using resources from the several cloud service providers including Rackspace, Amazon EC2, and HP Clouds. A supercloud allows organizations, companies and individuals move on a cloud computing environment. In particular, cloud users control live migration and location of their computation, storage and networking without owning all of underlying infrastructure.

In this paper [10] they propose privacy in frequent itemset mining on the encrypted cloud. In this they use three protocols such as Protocol 1 achieves extremely higher mining performance and Protocol 2 provides the strong privacy guarantee and Protocol 3 for improved efficiency. Mining performance is achieved separate form protocol 1 and protocol 3. Performance of system compared with the association rule mining and they also used a chess database for this experiment, it has totally 3196 transactions and 74 possible attributes. Performance of comparison carried out into two different security levels: Database privacy and Item Privacy.

In this paper [11] they present DEPSKY techniques for improving confidentiality, integrity and availability of information stored in the cloud using an encoding, encryption and replication of data on the cloud. They achieved these objectives by building Clouds of Clouds on the top of a set of storage clouds combining Cryptographic Secret Sharing with the Byzantine Quorum System Protocols.

|| Volume 3 || Issue 9 || September 2018 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

III COMPARATIVE ANANLYSIS

Table 1: **Survey Table**

Sr. No	Paper Name	Author	Method Proposed	Limitations
1.	SPANStore: Cost-effective geo-replicated storage spanning multiple cloud services	Z. Wu, M. Butkiewicz, D. Perkins, E. Katz-Bassett, and H. V. Madhyastha	SPANStore, a key-value store that exports a unified view of storage services in geographically distributed data centers.	Applications seek to tolerate failures of data centers and Internet paths.
2.	Minerva: An automated resource provisioning tool for large-scale storage systems	G. A. Alvarez et al	MINERVA suite of tools for designing storage systems automatically.	The maximum number of disks per LUN and the maximum number of LUNs per array
3.	Volley: Automated data placement for geo-distributed cloud services	S. Agarwal	Volley is designed to work in SCOPE.	Latency at modest costs in greater inter-datacenter traffic.
4.	SafeStore: A durable and practical storage system	R. Kotla, L. Alvisi, and M. Dahlin	SafeStore, to maintain long-term data durability despite conventional hardware and software faults, environmental disruptions, and administrative failures caused by human error or malice.	SSFS's throughput is within 12% of NFS.
5.	On Storing and Retrieving Geospatial Big-Bata in Cloud	Y. Yao, Haozhou Wang and K. Liu	Open-Source Prototype GP Cloud for fill the gap between modern cloud storage and traditional Data warehouse.	Proposed application dose not support file operations such as MODIFY and EXPAND, which is directly related to DELETE and UPDATE operations.
6.	Predicting cloud resource utilization	Stefan Schulte, Michael Borkowski and Christoph Hochreiner	Machine learning techniques to prediction models of resources in clouds.	Training dataset small size is used and the current approach lacks renormalization techniques to balance out the neuron weights.
7.	Video management and resource allocation for a large-scale VoD cloud	S.H. Gary Chan and Zhangyu Chang	Resource Allocation And Video Management Optimization (RAVO) techniques for jointly mange videos to achieve low cost and allocate system resources.	Maximum delay due to links and servers.
8.	The rise of big data on cloud computing: Review and open research issues	Ibrar Yaqoob, Samee Ullah Khan, S. Mokhtar and A. Gani	They present classification techniques for big data, cloud service model and conceptual view of data.	The issue with integrity is that previously Developed hashing schemes are no longer applicable to large amounts of data.
9.	Supercloud: Opportunities and challenges	Zhiming Shen, Qin Jia, Weijia Song	"Supercloud", it is deployed using resources from the several cloud service providers including	Low bandwidth for VMs that run in clouds and regions other than the one

		And Cornell University	Rackspace, Amazon EC2, and HP Clouds.	Where the NFS server is hosted.
10.	Toward Practical Privacy-Preserving Frequent Itemset Mining on Encrypted Cloud Data	Boyang Wang, Jiqiang Liu, Ming Li, and Shuo Qiu	Privacy in frequent itemset mining on the encrypted cloud.	Experiment conduct on small dataset.
11.	DepSky: Dependable and secure storage in a cloud-of-clouds	Miguel Correia, Alysson Bessani, and Bruno Quaresma	DEPSKY techniques for improving confidentiality, integrity and availability of information stored on the cloud.	All these works on the a single cloud (not a cloud-of-clouds), and require a cloud with the ability to run code, and have limited support for cloud unavailability,

IV CONCLUSION

The aim of this paper is minimizing the cost of cloud storage service providers with the guarantee of services for this we use worldwide distributed datacenters, which is belonging to the different CSPs with different resource unit prices. Firstly, we are using an integer programming for cost minimizing problems. It is NP-hard problem; we introduce DAR techniques for the dominant cost based data allocation algorithm to the reducing cost of each datacenters. Reducing service latency and payment cost datacenters we introduce three methods 1.Multicast Based Data Transfer, 2.Coefficient Based Data Allocation and 3.Request Redirection Based Congestion Control.

REFERENCES

1. Z. Wu, M. Butkiewicz, D. Perkins, E. Katz-Bassett, and H. V. Madhyastha, “SPANStore: Cost-effective geo-replicated storage spanning multiple cloud services,” in Proc. SOSP, Nov. 2013, pp. 292–308.
2. G. A. Alvarez et al., “Minerva: An automated resource provisioning tool for large-scale storage systems,” ACM Trans. Comput. Syst., vol. 19, no. 4, pp. 483–518, Nov. 2001.
3. S. Agarwal et al., “Volley: Automated data placement for geo-distribute cloud services,” in Proc. NSDI, 2010, p. 28
4. R. Kotla, L. Alvisi, and M. Dahlin, “SafeStore: A durable and practical storage system,” in Proc. ATC, Jun. 2007, pp. 129–142.
5. Y. Yao, Haozhou Wang and K. Liu, “On Storing and Retrieving Geospatial Big-Data in Cloud”, SIGSPTIAL International workshop on use of GIS in the emergency management, ACM, 31- October-2016.
6. Christoph Hochreiner, Stefan Schulte and Michael Borkowski, “Predicting cloud resource utilization”, 9th International conference on utility and cloud computing, 6-December-2016.
7. S.H. Gary Chan and Zhangyu Chang, “Video management and resource allocation for a large-scale VoD cloud”, Nature Communication, 24 January 2016 ACM Transaction on Multimedia Computing Communication Application, Vol. 12, Article 72, September 2016.
8. Ibrar Yaqoob, Samee Ullah Khan, S. Mokhtar and A. Gani, “The rise of big data on cloud computing: Review and open research issues, 2015-Elsevier.
9. Zhiming Shen, Qin Jia, Weijia Song, “Supercloud: opportunities and challenges.”, SIGOPS Oper. Syst. Rev, Jan-2015.
10. Boyang Wang, Jiqiang Liu, Ming Li, and Shuo Qiu, “Toward Practical Privacy-Preserving Frequent Itemset Mining on Encrypted Cloud Data”, IEEE Transactions on Cloud Computing, 2017.
11. Miguel Correia, Alysson Bessani, and Bruno Quaresma, “DepSky: Dependable and Secure Storage in a Cloud-of-Clouds”, ACM Transaction Storage, November-2017.