Journal of Engineering Sciences
Vol 11, Issue 11, NOV/2020
ISSN NO:0377-9254

TO ACCESS MULTIPLE CLOUD SERVERS BY USING MINIMUM -COST CLOUD STORAGE SERVICE

[1]D. MURALI, [2]S. KAVITHA

[1]Asst. Professor, Dept. of CSE, Brahmaiah College of Engineering, Nellore, Andhra Pradesh, India.

[2]PG Scholar, Dept. of CSE, Brahmaiah College of Engineering, Nellore, Andhra Pradesh, India.

Abstract **–** Cloud storage services are a valuable and widely used resource used in high volume by many different applications. There are currently numerous cloud service providers which offer various types of unique cloud storage services with a range of competitive prices and performance to attract application developers. Majority of cloud service providers (CSPs) provide data storage services and data access services from around the worldwide. These data centers provide different get/put latencies and reasonable prices for resource utilization. Thus, when selecting different CSPs' data centers, Cloud computing providers delivers application via internet, which are access from desktop and mobile apps, cloud customers of globally distributed applications (e.g., online social networks) face two challenges: 1) how to allocate data to worldwide data centers to satisfy application service level objective (SLO)requirements, including data access latency, security and availability and 2) how to allocate data and reserve resources in data centers belonging to different CSPs to minimize the payment cost. To handle these challenges, we first model the cost minimization problem under SLO constraints using the integer programming [1]. Due to its NP-hardness, we then introduce heuristic solution, including a dominant cost-based data allocation algorithm and an optimal resource reservation algorithm. Our proposed schema has three enhancement methods to reduce the cost of services and service latency: 1) coefficient-based data reallocation; 2) multicast-based data transferring and (3) Request redirection-based congestion control. In this proposed schema goal is to provide the multi-cloud environment over single cloud to reduce the risk and more security.

Keywords – Cloud storage, data availability and security, payment cost minimization.

I. INTRODUCTION

Cloud data storage, such as Hostinger, Godaddy and Google cloud storage, has become a popular commercial service. a cloud service provider (CSP)provides data storage service (including get and put functions) using its worldwide distributed data centers. Nowadays, more and more enterprisers shift their data workloads to the cloud storage in order to save capital expenditures to build and maintain the hardware infrastructures and avoid the risk of managing the data centers [2].

Cloud computing is an emerging technology that allows users to use and share resource, storage and data access services from around the world. However, Cloud service providers charge something from users for using these services. Specifically, to access data from their globally distributed storage edge servers, they charge users depending on the user's location and the amount of data transferred.

Cloud Storage Services:

A cloud storage service is a business that maintains and manages its customers' data and makes that data accessible over a network, usually the internet [3].Cloud storage service is a cloud computing service offering in which data can be stored, edited and retrieved from a remote cloud storage server over the Internet under a utility computing model. Cloud storage service, an Infrastructure as a Service (IaaS) service model, delivers scalable, flexible and redundant storage capacity through Web services API, online interfaces and thin client applications. Cloud storage service is also known as utility storage service and storage as a service. Cloud storage service is provided, hosted and managed by the storage service provider (SSP), and it works on the combination of storage servers, which are designed on storage virtualization architecture. This technique allows a single storage server to create multiple logical an virtual drives with scalable capacity and tight coupling. End users and applications access the logical storage by the online management interface or integrating vendor APIs with the application and are only billed for the storage capacity metered.

II. RELATED WORK

In the existing system an application execution platform across multiple CSPs was proposed. COPS and Volley automatically allocate user data among data centres in order to minimize user latency. Blizzard is a high-performance block storage for clouds, which enables cloud-unaware applications to fast access any remote disk in clouds [4]. Unlike these systems, DAR additionally considers both SLO guarantee and cost minimization for customers across multiple cloud storage systems. Cloud storage service is used by many web applications, such as online social networks and web portals, to provide services to clients all over the world. In the web applications portals, data access delay and availability are very critical, which affect cloud customers' incomes. Experiments at the ecommerce portal demonstrated that a small increase of100ms in webpage load time significantly reduces user satisfaction, and degrades sales by one percent. For a request of data retrieval in the web presentation process, the typical latency budget inside a storage system is only 50-100ms. In order to reduce data access latency, the data requested by client's needs to be handled by data centers near the clients, which requires worldwide distribution of data replicas [5].Also, data replication between data centers enhances data availability since it avoids a high risk of service failures due to data center failure, which may be caused by disasters or power shortages. However, a single CSP may not have data centers in all locations needed by a worldwide web application.

Besides, using a single CSP may introduce a data storage vendor lock-in problem, in which a customer may not be free to switch to the optimal vendor due to prohibitively high switching costs. Storage providers charge clients for bandwidth (Transfer), data requests(Get/Put), and Storage. Thus, a client moving from one CSP to another pays for Transfer cost twice, in addition to the Storage cost. The clients are vulnerable to price hikes by vendors, and will not be able to freely move to new and better options [6]. The quickly evolving cloud storage marketplace may leave a customer trapped with an obsolete provider later. The vendor lock-in problem can be addressed by allocating data to data centers belonging to different CSPs.

Building such geo-distributed cloud storage is faced with a challenge: how to allocate data to world wide data centers to satisfy application SLO (service level objective) requirements including both data retrieval latency and availability? The data allocation in this paper means the allocation of both data storage and Get requests to data centers. These computers can be dynamically provisioned as per users' requirements. Thus, to achieve better performance and scalability, applications could be managed using Commercial services provided by Clouds, such as Amazon AWS, Google App Engine, and Microsoft Azure. However, the cost of computing, storage and communication over these resources could be overwhelming for compute intensive and data intensive applications. It's needless to mention that cloud computing has made it way easier to run a small or medium scale business. With the exponential growth of the data that needs to be stored, the requirement of a data centre has become evident. Since most organizations are not able to afford the infrastructure and the technical expertise that is required to maintain a server, cloud hosting seems to be a boon to them. Cloud hosting service also makes it easier for the remote employees to access the important information. One of the best features of cloud hosting is that you can access the data any time and from any corner of the globe.

Since security is a huge concern with cloud hosting. The location of a data centre determines how safe and secure all your data will be. Keeping this in mind we propose both its data centres in low cost and security in very high. Moreover, none of the centres is easily accessible and each of these has been protected with proper security measures. The payment cost of a cloud storage service consists of the costs for Storage, data Gets/Puts and Transfers. Different data centers of a CSP or different CSPs offer different prices for Storage, data Gets/Puts and Transfers. , the second challenge is introduced: how to allocate data to data centers belonging to different CSPs and make resource reservation to minimize the service payment cost? Since website visit frequency varies over time, unexpected events may introduce a burst of requests in a short time. It may affect the accuracy of predicting the visit frequency. Thus, we need to dynamically adjust the data center Get serving rate to save the payment cost [7]. In summary, to use a cloud storage service, a customer needs to determine data allocation and resource reservation among data centers worldwide belonging to different CSPs to satisfy their application requirements on data retrieval latency and availability minimize cost, and dynamically adjust the allocation to adapt to request variation. Many previous works focus on finding the minimum amount of resources to support the application workload to reduce cloud storage cost in a single CSP. However, there have been only a few works that studied cloud storage cost minimization for a storage service across multiple CSPs with different prices. Within our knowledge, SPAN Store is the only work that handles this problem.

III. PROPOSED WORK

It aims to minimize the cloud storage cost while satisfying the latency and failure requirement across multiple CSPs. However, it neglects both the resource reservation pricing model and the data center capacity limits for serving Get/Put requests. A data center's Get/Put capacity is represented by the Get/Put rate(i.e., the number of Gets/Puts in a unit time period) it can handle [8]. Reserving resources in advance can save significant payment cost for customers and capacity limit is critical for guaranteeing SLOs. For example, Amazon DynamoDB shows a capacity limitation of 360,000 reads per hour. Also, data center network overload occurs frequently, which leads to packet loss. If the integer program used to create a data allocation schedule is modified to be capacity-limit aware, it becomes NP-hard, which cannot be easily resolved. Therefore [9], we need to consider the resource reservation pricing model and data center capacity limits when building a minimum-cost cloud storage service across multiple CSPs. To handle the above-stated two challenges, we propose a geo-distributed cloud storage system for Data storage; request Allocation and resource reservation across multiple CSPs. It transparently helps customers to minimize their payment cost while guaranteeing their SLOs. Building the geo-distributed cloud storage across multiple CSPs can avoid the vendor lock-in problem since a customer will not be constrained to an obsolete provider and can always choose the optimal CSPs for the cloud storage service. In the proposed we introduced the DAR system as a heuristic solution to this problem, which includes a dominant-cost based data allocation algorithm among storage data centres and an optimal resource reservation algorithm to reduce the cost of each storage data centre.

We also proposed several enhancement methods for DAR to further reduce the payment cost and service latency including 1) coefficient-based data reallocation; which aims to balance the workloads among all billing periods in order to minimize the payment cost by maximizing the reservation benefit. 2)multicast-based data transferring, which builds a minimum spanning tree to create new data replicas in order to minimize the Transfer cost for replica creation in a new data allocation deployment? (3) Request redirection-based congestion control, which redirects Get requests from overloaded data centers to under loaded data centers that have received Gets more than(or less than) their expected number of Gets after data allocation to minimize the payment cost, respectively.

DAR also incorporates an infrastructure to conduct the algorithms. Cluster storage automate configuration methods are proposed to use the minimum resources needed to support the desired workload. Adya et al. proposed a file system with high availability and scalability and low cost, named as Farsite. It depend on randomized replication to achieve data availability, and minimizes the cost by lazily propagating file updates. SPAN Store is key-value storage over multiple CSPs' data centres to minimize cost and guarantee SLOs. However, it does not consider the capacity limitation of data centres. We have modelled the cost minimization problem under multiple constraints using the integer programming.

We introduce a heuristic solution including: (1) A dominant-cost based data allocation

Vol 11, Issue 11, NOV/2020

ISSN NO:0377-9254

algorithm, which finds the dominant cost (Storage, Get or Put) of each data item and allocates it to the data center with the minimum unit price of this dominant cost to reduce cost in the pay-as-you-go manner. (2) An optimal resource reservation algorithm, which maximizes the saved payment cost by reservation from the pay-as you-go payment while avoiding over reservation. We conduct extensive trace-driven experiments on a supercomputing cluster and real clouds (i.e., Hostinger, Godaddy and Google Cloud Storage) to show the effectiveness and efficiency of our system in cost minimization, SLO compliance and system overhead in comparison with previous systems. Our dominant-cost based data allocation algorithm finds the dominant cost (Storage, Get or Put) of each data item.

For example, if a data item is storage dominant, it means its storage cost is much higher than the sum of other costs. Therefore, DAR is suitable for the scenarios in which most customer data items have dominant cost. To accessed these cloud services security and reliability we are using different models like: i) Using single service provider. ii) Using multiple service providers. The weakness of single service provider is that it can be easily be hacked by intruders and if the service provider fails or down for some technical reasons than client will not at all access his/her data. The problem in multiple service provider models is to compromise the security because there is lack of security technique used here but availability is up to the mark.

Cloud Computing Service Models:

Private and Public clouds serve as the backbone for a variety of different cloud computing service models. Currently the industry has been successfully adopting three common types of cloud computing service models [10]. Infrastructure as a Service (IaaS), is a service model around servers (compute power),storage capacity, and network bandwidth. Examples include Amazon S3, Hostinger, Rackspace, AT&T, and Verizon. Platform-as-a-Service (PaaS) provides an externally managed platform for building and deploying applications and services. This model typically provides development tools such as databases and development studios for working with the supplied frameworks, as well as the infrastructure to host the built application. Examples include Force.com, Microsoft Azure, and Google App Engine. Software as-a-Service (SaaS) is simply having a software system running on a computer that doesn't belong to the customer and isn't on the customer’s premises [11]. It is based on the concept of renting an application from a service provider rather than buying, installing and running software yourself.

CONCLUSION:

This work aims to minimize the payment cost of customers while guarantee their SLOs by using the worldwide distributed data centers belonging to different CSPs with different resource unit prices. We first modeled this cost minimization problem using integer programming. Due to its NP-hardness, we then introduced the DAR system as a heuristic solution to this problem, which includes a dominant-cost based data allocation algorithm among storage data centers and an optimal resource reservation algorithm to reduce the cost of each storage data center. We also proposed several enhancement methods for DAR to further reduce the payment cost and service latency including i) coefficient-based data reallocation, ii)multicast-based data transferring, and iii) request redirection-

Journal of Engineering Sciences

Vol 11, Issue 11, NOV/2020

ISSN NO:0377-9254

based congestion control [12]. DAR also incorporates an infrastructure to conduct the algorithms. Our trace-driven experiments on a test bed and real CSPs show the superior performance of DAR for SLO guaranteed services and payment cost minimization in comparison with other systems. Since more replicas of a more popular data item can help relieve more loads from overloaded data centers, in our future work, we will study how to adjust the number of replicas of each data item to further improve the performance of SLO conformance.

Further, we will conduct experiments against varying workload conditions and using other traces.

REFERENCES:

[1] Minimum-cost Cloud Storage Service Across Multiple Cloud Providers. Guoxin Liu and Haiying Shen. Department of Electrical and Computer Engineering. Clemson University, Clemson, SC 29631, USA. {guoxinl, shenh}@clemson.edu.

[2] Amazon S3, accessed on Jul. 2015. [Online].Available: http://aws. amazon.com/s3/.

[3] Microsoft Azure, accessed on Jul. 2015. [Online].Available: http://www. windowsazure.com/

[4] Goolge Cloud Storage, accessed on Jul. 2015.[Online]. Available:https://cloud.google.com/products/cloud-storage/

[5] R. Kohavl and R. Longbotham. (2007). OnlineExperiments: Lessons Learned, accessed on Jul. 2015.Available:http://expplatform.com/ Documents/IEEEComputer2007OnlineExperiments.pdf.

[6] B. F. Cooper et al., "PNUTS: Yahoo!'s hosted dataserving platform," Proc. VLDB Endowment, vol. 1,no. 2, pp. 1277–1288, Aug. 2008.

[7] A. Hussam, P. Lonnie, and W. Hakim, "RACS: Acase for cloud storage diversity," in Proc. SoCC, Jun.2010, pp. 229–240.

[8] Z. Wu, M. Butkiewicz, D. Perkins, E. Katz-Bassett,and H. V. Madhyastha, "SPANStore: Cost-effectivegeo-replicated storage spanning multiple cloud services," in Proc. SOSP, Nov. 2013, pp. 292–308.

AUTHORS

Mr. **D. MURALI** has received his B.Tech degree in Computer Science and Engineering and M.Tech degree in Computer Science and Engineering from JNTUA in 2011 and 2014 respectively. He is dedicated to teaching field from the last 7 years. He has guided 5 PG and 10 UG Students. At present he is working as Assistant Professor in Brahmaiah College of Engineering, North Rajupalem, Nellore, Andhra Pradesh, India.

Mrs. **S. KAVITHA** has received her M.Sc degree in Information Technology at Acharya Nagarjuna University in 2010 and pursuing M.Tech degree in Computer Science and Engineering at Brahmaiah College of Engineering affiliated to JNTUA in 2018-2020.