

# An optimized cost estimator for multi cloud storage environment

M.J JEYASHEELA RAKKINI,  
AP/CSE, Computer Science Department,  
Moogambigai College of Engineering<sup>2</sup>

## Abstract

The growing need of internet usage of many enterprises and organizations has lead to demand of significant storage requirements. The most recent breakthroughs include the integration of various storage devices in the cloud environment. To minimize the cost of storage and to improve the data availability, data can be hosted in the cloud. Many cloud vendors offer various heterogeneous pricing policy. Since the customers may be perplexed with various options, providing a consultancy sort of service to recommend them with a cheaper hosting strategy is the objective of this work. The users need not rely on single cloud service provider as they may end up in vendor lock in risk. They can meet out the storage requirement through multi cloud providers. This work proposes a cost estimator that can estimate the cost involved in storing the client data across multiple clouds, in multiple formats and hybrid storage devices. The outcome of this work is to suggest an optimal strategy to ensure guaranteed availability at a minimized monetary cost.

## INTRODUCTION

Cloud computing domain has earned a significant attention and a huge “gold rush” as a new computing paradigm and has become enormously popular. Public cloud service providers offer several pricing policies for different specification and performance requirements. It has now become a challenging task to select more appropriate service that can map the user requirements satisfying the customer budget constraints. There is a need to develop a comprehensive cloud service selection mechanism among various service providers such as Amazon S3, Google Cloud Storage,

Windows Azure. These provide services to their customers with scalable, reliable storage at low-cost data hosting functionality. Enterprises are storing all or part of their data into the cloud, to trim down the operation and acquaintance cost with enhanced data availability and reliability [1], [2], [3]. The available cloud services and their related information such as the cloud users can monitor important aspects of the cloud service or platform which they are using, are provided by the cloud monitoring Application Program Interfaces. APIs provide information to decide which cloud service providers can be selected by comparing the available cloud services. The above comparison cannot be done in a meaningful manner using single criterion. A comprehensive selection mechanism must be taken into account since there are multiple-criteria decision-making for which various number of techniques are available which can be used for IaaS cloud service selection.

## MOTIVATION

*Vendor lock-in risk.* Customers get perplexed as there are numerous cloud vendors and they have varied pricing policies. These policies may confuse the cloud users in taking proper decision to choose the optimal hosting strategy that can meet their requirements. The prevailing state of affair is that customers usually store their data, into a single cloud. This situation may lead to “vendor-lock-in-risk”. The customers may be in a dilemma to either migrate to other cloud vendors or to remain on the same cloud provider. This work tries to address this issue by estimating the cost involved in storage either in single cloud or in multicloud. The fluctuations in electricity bills have a great influence in the

pricing policy of the cloud vendors like Windows Azure and Google Cloud Storage as mentioned in [4],[5]. *Multiple-cloud data hosting*. Recently, multi-cloud data hosting has attracted wide attention from researchers and enterprises. The basic principle is to intensify the

redundancy and distribute the data among multiple cloud to prevent the vendor lock-in risk. Fig. 1. shows the concept of multi cloud hosting where the “proxy” redirects requests from client applications and coordinates data storage in dispensed manner across multiple clouds

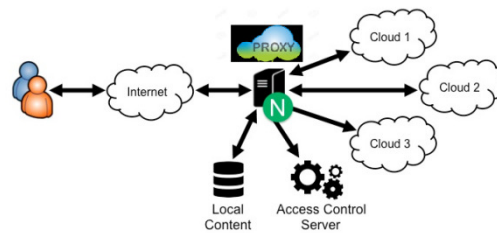


Fig. 1. Multi-cloud data hosting.

The efficiency analysis of multi-cloud is illustrated in three folds. First, there have been a few researches conducted on multi-cloud. DepSky guarantees availability of data and security based on multiple clouds, thus allowing critical data (e.g., medical and financial data) to be stored in a trust worthy manner as mentioned in [6]. RACS deploys erasure coding among different clouds in order to prevent vendor lock-in risk and reduce monetary cost [7]. Second, new types of cloud vendors (e.g., Dura-Cloud [8] and Cloud Foundry [9]) have emerged to provide real services based on multiple clouds. Third, new development tools like Apache libcloud [10] provide a unified interface above different clouds, which facilitates migrating services among clouds.

#### RELATED WORK

As cloud computing provides the needs of researchers and industry professional such as on-demand self-service, rapid elasticity, broad network access, large pool of compute resources, and measured services [11] there has been an increase in its adoption over recent years. On measured services characteristic of cloud computing, several tools have been developed, utilizing various billing models, to provide cloud services providers the ability to bill users on for resources used on monthly basis. The most widely used model is the “Pay-as-you-go”, which is adopted by all major cloud services providers [12]. On the other hand, cloud services providers have developed tools such as GCP Pricing Calculator, AWS Simple Monthly Calculator that allow users to estimate cost of adopting services on cloud environments such as EC2, GCE, AWS S3 etc.

Table 1. Summary of Related Work

Pricing Model	Economic Model	Provider Orientation	Consumer Orientation	Remarks
Buyya [13]	Commodity Market Model	Considered	Considered	Providers can recover their investment before the age of resources
Dhawas [14]	Posted Price Model	Not considered	Considered	Security issues are easily handled
Sururah [15]	Commodity Market Model	Not considered	Considered	Rewards are offered to consumer
Shifeng [16]	Double Side Auction Model	Not considered	Considered	More centric towards consumers but helpful for providers also

Kento [17]	Double Side Auction Model using K pricing	Not considered	Considered	Consumer specific, can order its requirement from spot/forward market
Pourebrahimi [18]	Double Auction Approach with continuous aspect	Considered	Not considered	Beneficial for the providers due to decision making approach.
Michael [19]	Double Auction Approach(English Auction)	Not considered	Considered	It is beneficial for the user because they purposed a scarcity pricing approach

### PROPOSED WORK

#### *Theoretical concepts*

#### *Cloud computing characteristics:*

Key characteristics of cloud computing are  
i) rapid elasticity – capability to provision computing resources to the customers on demand  
ii) on-demand self-service – capability that enables user to avail the services and resources without intervention of services providers  
ii) Broad network access – ability to

### DEPLOYMENT MODELS

Cloud computing has three deployment models namely public cloud, private cloud, community cloud, and hybrid cloud.

- *Public Cloud* – In this model, a single virtual machine is used by multiple tenants and they will pay for each as they consume.
- *Private Cloud* – In this model, a company manages its own data center that provides scalability, provisioning, automation and management. The company gains the benefits of cloud computing with private cloud, but incur the cost of maintaining it as well.
- *Hybrid* – This model, allows a company to maintain an internal private cloud and at the same time they can use public cloud as and when needed.

Cloud computing has three service models, which includes Infrastructure-as-a-Service (IaaS),

### CONCEPTUAL FRAMEWORK

#### *A. Problem Formalization*

In this section cloud service selection problem is being formulated in a generalized and abstract-mathematical form.

access resources and services over network  
iv) Resources pooling – To suffice multiple customers computing resources are pooled and dynamically assigned  
v) Measured Services – resources and services usage are measured, controlled, and monitored to provide transparency for both provider and consumer of the utilized service.

#### *Service models:*

Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). In IaaS, the customers are provided with the underlying infrastructure where they could deploy arbitrary software including operating systems and applications. In PaaS, customers do not manage the underlying cloud infrastructure such as operating systems, servers, network but rather manage applications developed using tools and libraries supported by the providers. In SaaS, the capability provided to the customer is to use providers' applications running on cloud infrastructure such as email services, spreadsheets and word processor applications. The proposed framework is designed to allow users to estimate cost of application running in the IaaS service model. The figure 3.1 shows the the services models and associated providers.

Let the users requests and workload profiling can be characterized as  
1) Requests set: Let  $R = \{r_1, r_2, r_3, \dots, r_l\}$  where  $l \geq 2$ . This set contains all the service schemes offered for the customer by the service provider.

- 2) Performance metric (set): Let  $P = \{p_1, p_2, p_3, \dots, p_p\}$  where  $p_1, p_2, p_3, \dots, p_p$  where  $p \geq 2$  and each  $p_i \in P$  represents a criterion that can be used for choosing the service.
- 3) Decision matrix: It contains descriptor vector that can be combined to take the decision and it is represented as  $l \times p$  matrix.
- 4) User priority vector: It represents the user priority weights vector where each weight  $w_i$

represents the weight associated with each performance metric  $p_i$  and a value 1 indicates the default weight.

The objective of this work is to map the user request with the performance metric and formulate the decision matrix. Since users may have different preferences, each user requirements can be still refined by including the priority vector.

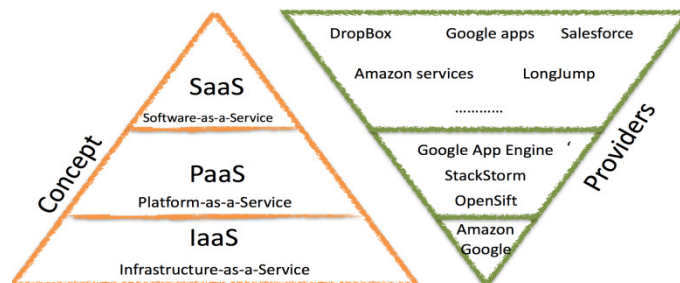


Fig.2. Cloud computing service models and few service providers

## ARCHITECTURE

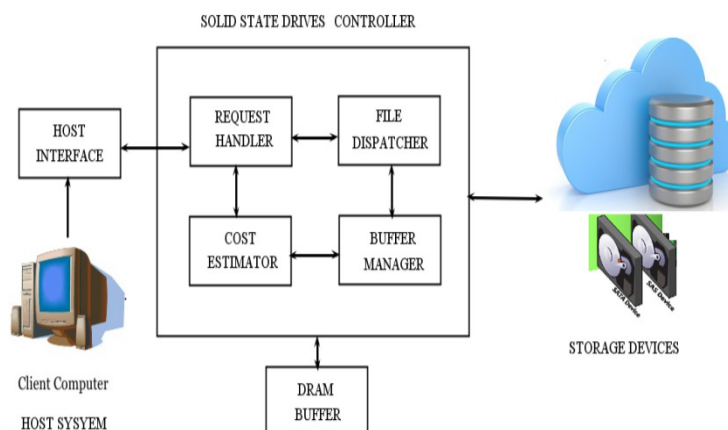


Fig. 3: Price estimator structure

Figure 3 shows the price estimator structure where the customer will submit the request through the host interface and the request

handler will try to map the request with the service providers specification. The cost estimator will estimate the cost by considering the request set, performance metric set and the

user priority vector. The decision matrix will provide the details about the matching user specification with service providers availability. The cost estimator can now recommend the user lists down various service providers and their billing pattern, so that the cost estimator will use these information in arriving the cost involved in storage. User enters their requirements related to the storage capacity and performance. All their requirements are mapped to the storage devices using mapper. Request Handler has the mapper to process the user request and to allocate the storage devices according to the file format and file types. The users request is classified on the basis of the user's need for example, Central Processing Unit (CPU) power, storage space, speed of Random Access Memory (RAM). Customer submits the request to the admin validates the user request and after passing the payment gateway, the processed order is dispatched. The context handler analysis the context and feeds information to the next phase and examines the required storage structure in a cost and performance effective manner. The work load profiler gets the input from the user. Buffer manager distributes the user data in various storage devices namely HDD, SSD FLASH memory based on the nature and frequency of operation.

### CONCLUSION

In this paper we have formalized the cloud service selection problem into a rigorous mathematical form and have presented a multiple criteria cloud service selection methodology. Since cloud services are experiencing fast development and the services based on multi-cloud also become prevailing and has become one of the important storage units. An analysis have been made to reduce the initial cost and to increase the performance involving storing, processing, and retrieval process. Multistage analyzer has been built to analyze the performance of hybrid storage structure. The proposed cost estimator estimates the cost involved to fix different storage strategies by interpreting the memory usage pattern in terms

with various options with providers and the associated cost, so that based on the application requirement and the affordability the customer can settle down with the desired setup. of size, speed and frequency. Cloud resource examiner monitors the requirements from the cloud users.

### References

- [1] Gartner: Top 10 cloud storage providers. [Online]. Available: <http://www.networkworld.com/news/2013/010313-gartnercloud-storage-265459.html?page=1>, 2013.
- [2] Z. Li, C. Jin, T. Xu, C. Wilson, Y. Liu, L. Cheng, Y. Liu, Y. Dai, and Z.-L. Zhang, "Towards network-level efficiency for cloud storage services," in Proc. ACM SIGCOMM Internet Meas. Conf., 2014, pp. 115–128.
- [3] Z. Li, C. Wilson, Z. Jiang, Y. Liu, B. Y. Zhao, C. Jin, Z.-L. Zhang, and Y. Dai, "Efficient batched synchronization in dropbox-like cloud storage services," in Proc. ACM/IFIP/USENIX 14th Int. Middleware Conf., 2013, pp. 307–327.
- [4] Windows Azure Pricing Updates. [Online]. Available: <http://azure.microsoft.com/en-us/updates/azure-pricing-updates/>, 2014.
- [5] Google Cloud Platform Pricing Updates. [Online]. Available: <http://googlecloudplatform.blogspot.com/2014/03/googlecloud-platform-live-blending-iaas-and-paas-moores-law-for-thecloud.html>, 2014
- [6] A. Bessani, M. Correia, B. Quaresma, F. Andre, and P. Sousa, "DepSky: Dependable and secure storage in a cloud-of-clouds," in Proc. 6th Conf. Comput. Syst., 2013, pp. 31–46.
- [7] H. Abu-Libdeh, L. Princehouse, and H. Weatherspoon, "RACS: A case for cloud storage diversity," in Proc. 1st ACM Symp. Cloud Comput., 2010, pp. 229–240.
- [8] DuraCloud. [Online]. Available: <http://www.duracloud.org/>, 2014.
- [9] Cloud Foundry. [Online]. Available: <http://www.cloudfoundry.org/>, 2014.
- [10] Apache Libcloud. [Online]. Available: <http://libcloud.apache.org/>, 2014.
- [11] "http://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-

- 145.pdf,” 2016/4/26/. [Online]. Available: <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>
- [12] *An Approach to Estimating Cost of Running Cloud Applications Based on AWS*, ser.201219<sup>th</sup> Asia-Pacific Software Engineering Conference.
- [13] B. Sharma, R. K. Thulasiram, P. Thulasiraman, S. K. Garg, and R. Buyya, “Pricing cloud compute commodities: A novel financial economic model,” in 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. IEEE Computer Society, 2012.
- [14] V. N. Dhawas, P. Juikar, N. Patekar, N. Lendghar, and S. Vartak, “A Secured Cost Effective Multi-Cloud Storage in Cloud Computing,” *International Journal of Scientific & Engineering Research*, vol. 4, Issue 5, pp. 1017-1021, 2013.
- [15] S. A. Bello, and G. A. Wakil, “Flexible Pricing Models for Cloud Services,” *Transactions on Networks and Communications*, vol. 2, no. 5, pp. 15-28, 2014.
- [16] Shang, S., J. Jiang, Y. Wu, Z. Huang, G. Yang, and W. Zheng, “DABGPM: A Double Auction Bayesian Game-Based Pricing Model in Cloud Market,” *Network and Parallel Computing*, pp. 155-164, 2010.
- [17] I. Fujiwara, K. Aida, and I. Ono, “Applying Double-sided Combinational Auctions to Resource Allocation in Cloud Computing”, *Proceedings of the 10th Annual International Symposium on Applications and the Internet*, IEEE, July 19- 23, pp. 7-14, 2010.
- [18] B. Pourebrahimi, K. Bertels, G. Kandru, and S. Vassiliadis, “Market-based resource allocation in grid,” In 2nd IEEE International conference on eScience and grid Computing, 2006.