Efficient Schemes for Resource Allocation In the Cloud for Media Streaming Applications

Shreya Kulkarni, PragnaPrakash S, Shaila S, Mrs.Janaki.K
Department of Computer Science, RRCE,Bangalore-74
kulkarnis642@gmail.com, karur.janaki@gmail.com

ABSTRACT-A large number of users in the Internet have recently attracted by the media streaming applications. Due to advent of this bandwidth-intensive applications, with guaranteed QoS relying only on central resources at a media content provider it is economically inefficient to provide streaming distribution. media content providers(e.g., Video on Demand (VoD) providers) can use to obtain streaming resources that match the demand by the elastic infrastructure offered by the cloud computing. The amount of resources allocated (reserved) in the cloud are charged by the media content providers. Based on non-linear time-discount tariffs (e.g., Amazon Cloud Front and Amazon EC2) the existing cloud providers employ a pricing model for the reserved resources. This type of pricing scheme offers discount rates depending non-linearly on the period of time during which the resources are reserved in the cloud. In this case, the right amount of resources reserved in the cloud, and their reservation time are used to decided by the open problem such that the financial cost on the media content provider is minimized. A simple - easy to implement- algorithm are proposed for resource reservation that maximally exploits discounted rates offered in the tariffs, while ensuring that as much as resources are reserved in the cloud. Our algorithm is carefully designed based on the prediction of demand for streaming capacity to reduce the risk of making wrong resource allocation decisions. The results show that the proposed algorithm significantly reduces the monetary cost of resource allocations in the cloud as compared to other conventional schemes according to our numerical evaluation and simulation.

Index Terms—Media streaming, Cloud Computing, Non-linear pricing models, Network economics.

1 INTRODUCTION

Huge number of users in the Internet is attracted by the media streaming application. As compared to the past few years as the number of video streams served increased 38.8% to 24.92 billion. Due to this huge demand it creates a burden on centralized datacenters at media content providers such as Video on-Demand (VoD) providers to sustain the required QoS guarantees [2]. The problem becomes critical with the increasing demand for higher bit rates required for the growing number of higher-definition video quality desired by consumers. In this paper, by using cloud computing we explore new approaches that mitigate the cost of streaming distribution on media content provider. A non-linear pricing scheme offers discount rates depending non linearly on the period of time during which the resources are reserved in the cloud. In this case, an open problem is to decide on both the right amount of resources reserved in the cloud, and their reservation time such that the financial cost of the media content provider is minimized. Based on the prediction of this (PBRA), of demand for streaming capacity, our proposed algorithm is carefully designed to reduce the risk of making resource allocation decisions. Cloud computing (IT infrastructure provided over the internet and priced over the usage).cloud computing here is used to avoid the unnecessary capital investment by the media content for unused data. Allocate memory on demand and optimize the resources. To target the expected usage peak, media content provider can make long term investment in infrastructure. This causes problem in view of flash crowd events. Expenses charged by cloud providers convert the upfront infrastructure investment for media content providers. Cloud service providers provide services over the internet and it is priced as it is used.Paas, Iaas, Saas are the services provided by cloud service providers. Where Iaas includes processor and can be scaled up as demand increases and Paas provides operation services from available system of operating system. Saas is used to select required software of your choice like Amazon, Microsoft and Google. These provide cloud services over the internet. Cloud enhances the auto scaling ability to reserve amount of resources to match the demand. Media content provider provides on demand plan and avoids unnecessary amount to be paid and also eliminate the waste by optimizing the space allocated in the memory. The on demand plan is the pay-per-use plan.Our main aim is to do in this paper is to reduce the monetary cost to reserved amount of resources in the cloud by implementing prediction based resources allocation algorithm which gives predictions and discounted rates in the tariffs by ensuring amount of resources reserved in the cloud.

2 RELATED WORK

The web based applications has been widely studied in the literature for prediction of CPU utilization and user access demand.The prediction method based on Radial basis function(RBF)networks has been proposed by Y.Lee.et.al to predict the user access demand request for web type of services in web based applications.Prediction method is to improve the capability for making informed decisions by providing reliable functions based on neural networking and linear regrecision.This were Radial basis functions which is proposed by Y.Lee.et.al for predicting the user access demand request for web types of services in web based applications and output of network is a linear combination of Radial basis function of inputs.Web application and demand prediction for CPU utilization has been studied for long time to gain the popularity recently. Cloud providers offers a

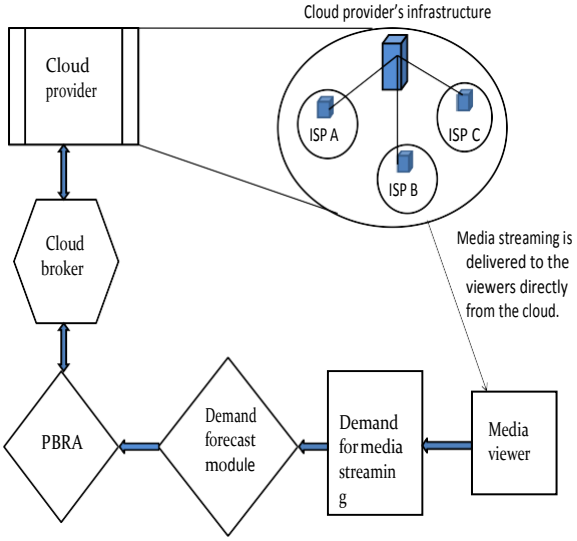
streaming resources to content providers with intensive bandwidth demand.

3 PROPOSED SYSTEM

Here we implement an innovative algorithm based on prediction based resource allocation algorithm in which resource allocation is the process and strategy involving a company deciding where scarce resources should be used in the production of goods and services. BRA gives predictions and minimizes the cost of resources in the cloud by giving discounted rates in the tariffs. Here predictions are made in such a way that it considered decision from cloud service provider. The main view of cloud service provider is it should optimize the memory resource allocation so that it can provide to other content provider. Cloud service provider provides two services to customers they are linear tariff plans and non-linear tariff plans. The main aim is to reduce our expenses by choosing a right plan at right time by ensuring sufficient resources which considerably gives more profit.

The content provider reserves resources in the cloud according to the predicted demand. The proposed algorithm is based on time-slots with varied durations (sizes). In every time-slot, the media content provider makes a decision to reserve amount of resources in the cloud. Both the amount of resources to be reserved and the period of time over which the reservation is made (duration of time-slots) vary from one time-slot to another, and are determined in our algorithm to yield the minimum overall monetary cost (Fig. 2). We alternatively call a time-slot a window, and denote the window size (duration of the time-slot) by w. Since the actual demand varies during a window size, while allocating the resource in the cloud remain the same for the entire window size (according to the third assumption above), the algorithm needs to reserve resources in every window j that are sufficient to handle the maximum predicted demand for streaming capacity during that window with some probabilistic level of confidence η.

We denote the monetary cost of the reserved resources during window j by Cost (wj, Allocj), and can be computed as Cost (wj, Allocj) = tariff (wj, Allocj)×wj, (2) where tariff(wj,Allocj) represents the price (in $ per time unit) charged by the cloud provider for amount of resources Allocj reserved for period of time (window size) wj. More specifically, the demand forecast module predicts streaming capacity demand in the upcoming period of time L and feeds this information to our algorithm .The algorithm upon receiving the demand prediction, computes the right size of window j (i.e., w∗ j), and the right amount of reserved resources in window j (i.e., Alloc∗ j), , such that the cost of the reserved resources during window j (i.e., Cost(wj,Allocj) in (2)) is minimized; or equivalently, the discounted rates offered in the tariffs are maximally utilized. Hence, the objective of our algorithm is to minimize Cost(wj,Allocj) ∀j, subject to Probability(D(t) ≤ Alloc(t)) ≥ η, ∀ t ∈ L. In other words, our objective is to minimize the monetary cost of reserved resources such that the amount of reserved resources at any instant of time is guaranteed to meet the actual demand with probabilistic confidence equals to η.



SYSTEM DESIGN

Fig1 :System Architecture

Table 1. Summary of results for iterations executed for window j = 1

iteration(h)	1	2	3	4	5	6
wh	1.0	2.0	3.0	4.0	5.0	6.0
_max	0.6	1.0	1.184	1.26	1.3	1.37
Alloch	3	1.0	2.0	2.0	2.0	2.0
Xh=tariff(wh,Alloch)	1.0 11	10.85	12.25	12.0	11.75	11.50

Table 2.Media streaming cost given different probability of the demand (in $)

Distribution	Long-normal(σ=3)	Long-normal(σ=6)	Long-normal(σ=8)
cost	34,457	41,543	48,393

Table 3.Media streaming cost using two resource allocation plans provided by the cloud (hybrid resource provisioning approach) (in $)

n	Cost of reservation plan	Cost 0f on demand plan	Total cost
0.75	34,457	12,213	46.670
0.8	36,979	8,854	45,833
0.9	44,033	2,821	46,854
0.95	46,324	2,741	49,065

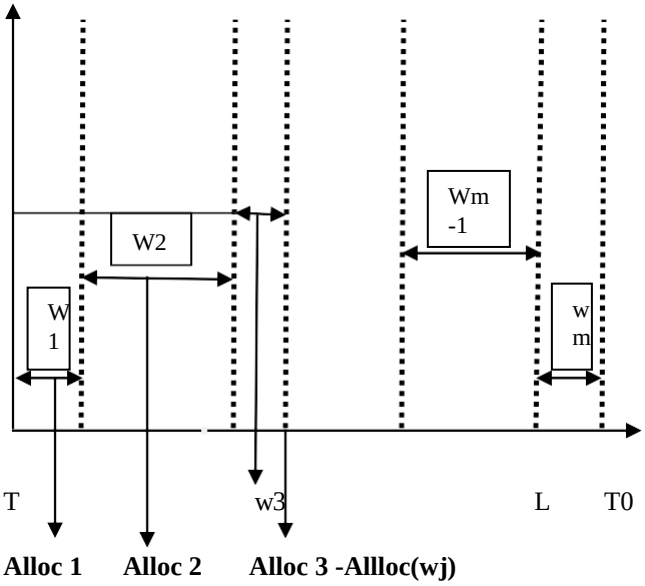


Fig2: PBRA algorithm design

wj: the jth window size M: number of windows Allocj: amount of allocated resources in window j

4 RESULTS AND DISCUSSION

To reduce our expenses by choosing a right plan at right time by ensuring sufficient resources in the cloud which comfily gives more profit? We have considered Non-Linear tariff plans because it is less costlier, more flexible, gives more discounts from customer perspective and non-fixed with respect to customer‘s perspective. The proposed algorithm based on time slots with varied durations. In every time slot the media content provider makes a decision to reserve amount of resource in the cloud. Both the amount of resources to be reserved and period of time over which reservation is made vary from one time slot to another and it is determined in our algorithm to yield the minimum overall monetary cost. Media content provider can predict the demand for streaming capacity of a video channel over a future period of time and it reserves resources in the cloud according to the predicted demand.The proposed algorithm PBRA gives predictions and minimizes the cost of resources in the cloud by giving discounted rates in the tariffs and PBRA provides demand to avoid unnecesaary amount to be paid to eliminate the wastage. The results show that our algorithm adjust the trade-off between the resources in the cloud and support well defined plan such that as much as the demand is coming the trade must be optimized.

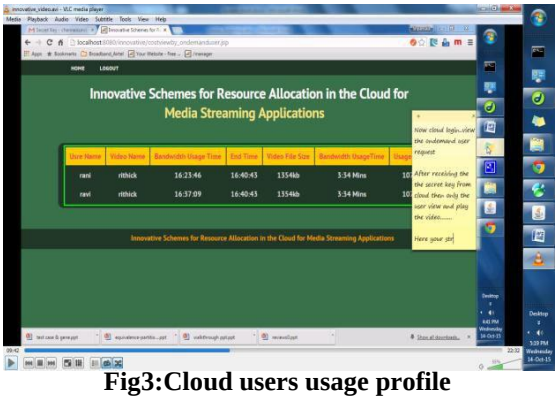


Fig3:Cloud users usage profile

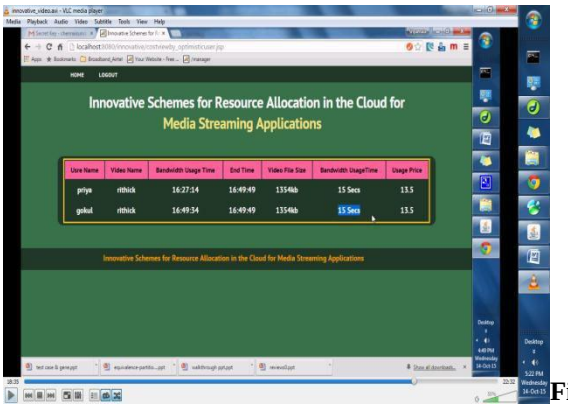


Fig4:Cloud users validity profile

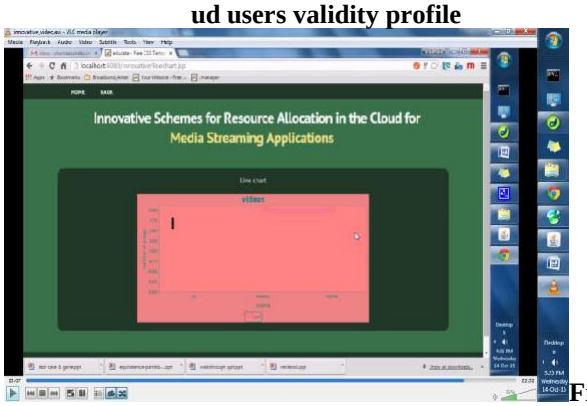


Fig5:Cloud users on demand video usage

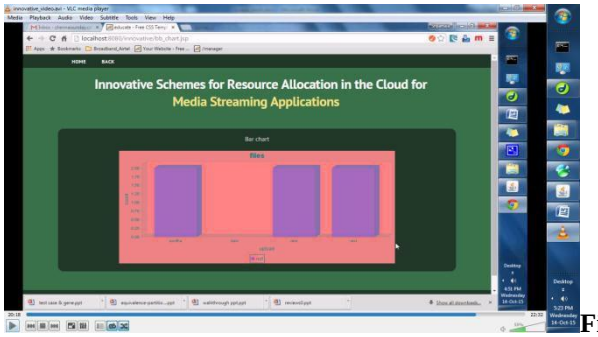


Fig6: Cloud users increment over the years

5 CONCLUSION

This paper studies the problem of resource allocations in the cloud for media streaming applications. We have considered non-linear time-discount tariffs that a cloud provider charges for resources reserved in the cloud. We have proposed algorithms that optimally determine both the amount of reserved resources in the cloud and their reservation time - based on prediction of future demand for streaming capacity - such that the financial cost on the media content provides minimized. The proposed algorithms exploit the time discounted rates in the tariffs, while ensuring that sufficient resources are reserved in the cloud without incurring wastage. We have evaluated the performance of our algorithms numerically and using simulations. The results show that our algorithms adjust the trade-off between resources reserved on the cloud and resources allocated on-demand. In future work, we shall perform experimental measurements to characterize the streaming demand in the Internet and develop our own demand forecasting module. We shall also investigate the case of multiple cloud providers and consider the market competition when allocating resources in the cloud.

6 REFERENCES

[1] Cisco Systems Inc., ―Cisco visual networking index: Forecast and methodology, 2010-2015,‖ white paper, 2010.
[2] Y. Liu, Y. Guo, and C. Liang, ―A survey on peer-to-peer video streaming systems,‖ in Peer-to-Peer Networking and Applications,vol. 18, no. 1, pp. 18–28, 2008.
[3] Cisco Systems Inc., ―Data center virtualization and orchestration: Business and financial justification,‖ white paper, 2007.
[4] Four Reasons We Choose Amazons Cloud as Our Computing Platform, The Netflix Tech Blog, Dec., 2010.
[5] http://www.octoshape.com/
[6] Amazon EC2 Reserved Instances, http://aws.amazon.com/ec2/reserved-instances, 2012.
[7]AmazonCloudFront, http://aws.amazon.com/cloudfront/, 2012.
[8] G. Chuanxiong, G. Lu, H. Wang, S. Yang, C. Kong, P. Sun,H. Ballani, P. Costa, T. Karagiannis, and A. Rowstron, ―Towards predictable datacenter networks,‖ in Proc. of the ACM SIGCOMM conference, pp. 242–253, 2011.
[9] E. White, M. O'Gara, P. Romanski, P. Whitney, ―Cloud Pricing Models,‖ in Cloud Expo: Article, white paper, 2012. http://java.sys-con.com/node/2409759?page=0,1.