



## International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: [www.ijircce.com](http://www.ijircce.com)

Vol. 6, Issue 10, October 2018

# Enhanced and Efficient Scheme for Resource Allocation in the Cloud for Media Streaming Applications

Deepak Zambare<sup>1</sup>, Rahul Gaikwad<sup>2</sup>

P.G. Student, Department of Computer Engineering, Godavari College of Engineering, Jalgaon, Maharashtra, India<sup>1</sup>

Assistant Professor, Department of Computer Engineering, Godavari College of Engineering, Jalgaon,  
Maharashtra, India<sup>2</sup>

**ABSTRACT:** Media streaming applications have recently attracted a large number of users in the Internet. With the advent of these bandwidth-intensive applications, it is economically inefficient to provide streaming distribution with guaranteed QoS relying only on central resources at a media content provider. Cloud computing offers an elastic infrastructure that media content providers (e.g., Video on Demand (VoD) providers) can use to obtain streaming resources that match the demand. Media content providers are charged for the amount of resources allocated (reserved) in the cloud. Most of the existing cloud providers employ a pricing model for the reserved resources that is based on non-linear time-discount tariffs (e.g., Amazon Cloud Front and Amazon EC2). Such a 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 on the media content provider is minimized. We propose a simple - easy to implement - algorithm for resource reservation that maximally exploits discounted rates offered in the tariffs, while ensuring that sufficient resources are reserved in the cloud. Based on the prediction of demand for streaming capacity, our algorithm is carefully designed to reduce the risk of making wrong resource allocation decisions. The results of our numerical evaluations and simulations show that the proposed algorithm significantly reduces the monetary cost of resource allocations in the cloud as compared to other conventional schemes.

**KEYWORDS:** Video-on-Demand. Quality-of-Service, Prediction Based Resource Allocation Algorithm

### I. 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 nonlinear 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



# International Journal of Innovative Research in Computer and Communication Engineering

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

Website: [www.ijircce.com](http://www.ijircce.com)

Vol. 6, Issue 10, October 2018

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.

## II. 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 regression. 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 offer a streaming resource to content providers with intensive bandwidth demand.

## EXISTING SYSTEM

The prediction of CPU utilization and user access demand for web-based applications has been extensively studied in the literature. A prediction method has been proposed with respect to upcoming CPU utilization pattern demands based on neural networking and linear regression that is of interest in e-commerce applications. Y. Lee et al. proposed a prediction method based on Radial Basis Function (RBF) networks to predict the user access demand request for web type of services in web-based applications

## III. 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  $\eta$ .

# International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

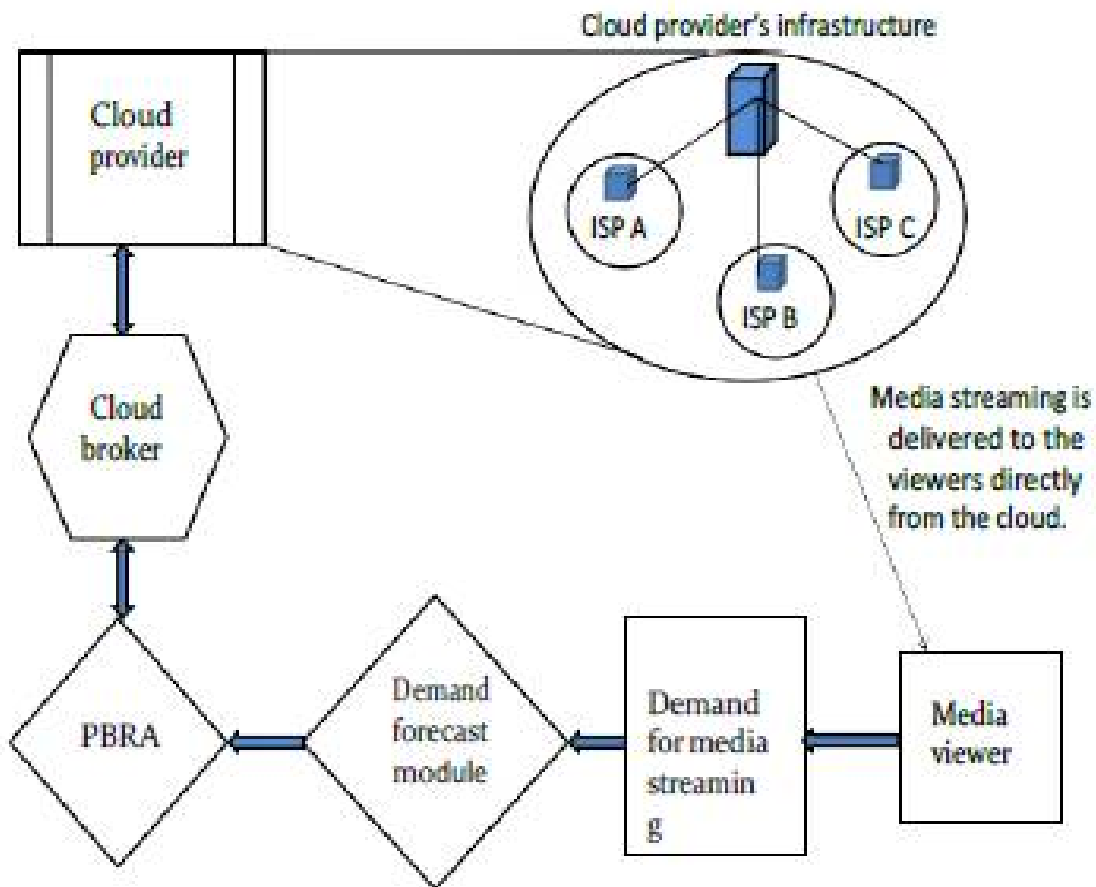
Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 6, Issue 10, October 2018

We denote the monetary cost of the reserved resources during window  $j$  by  $Cost(w_j, Alloc_j)$ , and can be computed as  $Cost(w_j, Alloc_j) = tariff(w_j, Alloc_j) \times w_j$ , (2) where  $tariff(w_j, Alloc_j)$  represents the price (in \$ per time unit) charged by the cloud provider for amount of resources

$Alloc_j$  reserved for period of time (window size)  $w_j$ . 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.,  $Cost(w_j, Alloc_j)$  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  $Cost(w_j, Alloc_j) \forall j$ , subject to  $Probability(D(t) \leq Alloc(t)) \geq \eta, \forall t \in L$ . 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  $\eta$ .

The proposed system is a practical - easy to implement - Prediction-Based Resource Allocation algorithm (PBRA) that minimizes the monetary cost of resource reservation in the cloud by maximally exploiting discounted rates offered in the tariffs, while ensuring that sufficient resources are reserved in the cloud with some level of confidence in probabilistic sense. We first describe the system model



**SYSTEM DESIGN**



# International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: [www.ijircce.com](http://www.ijircce.com)

Vol. 6, Issue 10, October 2018

## IV. CONCLUSION AND FUTURE WORK

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 provider is 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 clouds.

## 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] Four Reasons We Choose Amazons Cloud as Our Computing Platform, The Netflix Tech Blog, Dec., 2010.
- [4] <http://www.octoshape.com>
- [5] Amazon EC2 Reserved Instances, <http://aws.amazon.com/ec2/reserved-instances>, 2012.
- [6] Amazon CloudFront, <http://aws.amazon.com/cloudfront/>, 2012.
- [7] G. Chuanxiong, G. Lu, H. Wang, S. Yang, C. Kong, P. Sun, W. Wu, and Y. Zhang, “SecondNet: a data center network virtualization architecture with bandwidth guarantees,” in *Proc. of the ACM 6th International Conference on emerging Networking EXperiments and Technologies (Co-NEXT)*, 2010.
- [8] 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>.
- [10] D. Niu, Z. Liu, B. Li, and S. Zhao, “Demand forecast and performance prediction in peer-assisted on-demand streaming systems,” in *Proc. of IEEE Infocom conference*, pp 421–425, 2011.
- [11] S. Peichang, W. Huaimin, Y. Gang, L. Fengshun, and W. Tianzuo, “Prediction-based Federated Management of Multi-scale Resources in Cloud,” in *AISS: Advances in Information Sciences and Service Sciences*, vol. 4, no. 6, pp. 324–334, 2012.
- [12] G. Gursun, M. Crovella, and I. Matta, “Describing and Forecasting Video Access Patterns,” in *Proc. IEEE Infocom Mini-Conference*, pp. 16–20, 2011.
- [13] D. Niu, H. Xu, B. Li, and S. Zhao, “Quality-Assured Cloud Bandwidth Auto-Scaling for Video-on-Demand Applications,” in *Proc. of IEEE Infocom Conference*, pp. 421–425, 2012.
- [14] S. Islam, J. Keung, K. Lee, and A. Liu, “Empirical Prediction Models for Adaptive Resource Provisioning in the Cloud,” in *Future Generation Computer Systems*, vol. 28, no. 1, pp. 155–162, 2012.