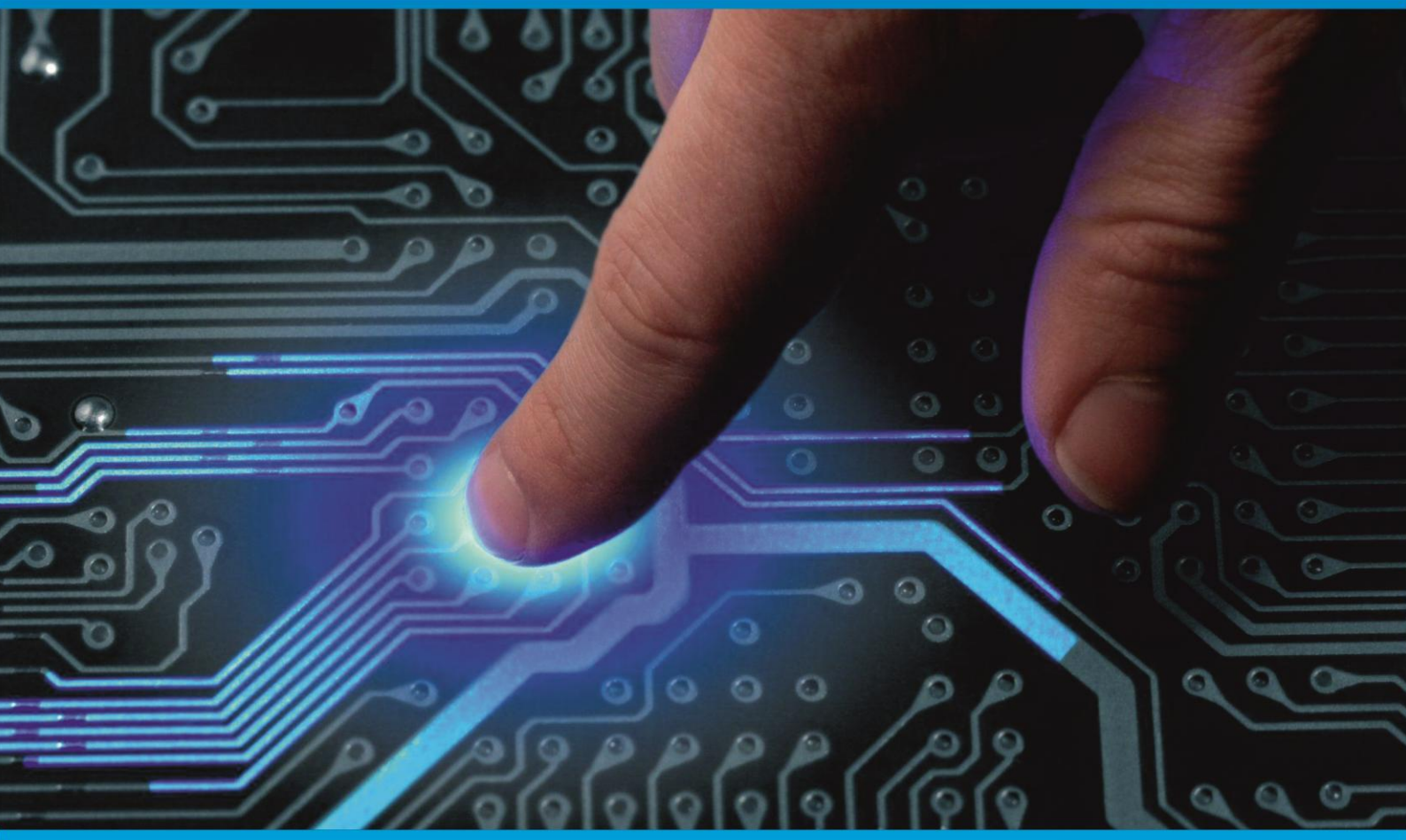
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 1, January 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901031 |

Efficient Scheme for Resource Allocation in the Cloud for Media Streaming Applications

Zambare Deepak[1*], Gaikwad Rahul[2], Gosavi Pramod[3]

Department of Computer Sciences, Godavari COE Jalgaon, MS India[1]

Department of Computer Sciences, Godavari COE Jalgaon, MS India[2]

Department of Computer Sciences, Godavari COE Jalgaon, MS India[3]

*Corresponding Author

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.

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 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

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901031 |

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 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 [15]. 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. Although the demand prediction for CPU utilization and web applications has been studied for a relatively long period of time, the prediction of demand for media streaming has gained popularity more recently [11]-[14]. The access behavior of users in Peer-to-Peer (P2P) streaming with time-series analysis techniques using non-stationary time-series models was predicted in [11]. The method of time-series prediction based on wavelet analysis was studied in [12],[13], Principle component analysis is employed by the authors to extract the access pattern of streaming users. Although most of the above studies predict the average streaming capacity demands, few papers have also studied the volatility of the capacity demand, i.e., the demand variance at any future point in time, which yields more accurate risk factors [14].

The prediction of streaming bandwidth demand is outside the scope of this paper. In this work, we formulate the problem considering a given probability distribution function of prediction of future demand for streaming bandwidth. In addition to demand prediction for resource reservation, other relevant studies have addressed the appropriate joint reservation of bandwidth resources on multiple cloud service providers with the purpose of maximizing bandwidth utilization [12], [14]. In an adaptive resource provisioning scheme is presented that optimizes the bandwidth utilization while satisfying the required levels of QoS. Maximization of bandwidth utilization in turn helps cloud service providers reduce their expenses and maximize their revenues. In an optimization framework for making dynamic resource allocation decisions under risky and uncertain operating environments was developed to maximize revenue while reducing operating costs. This framework considered multiple client QoS classes under uncertainty of workloads. Recently, streaming resources (e.g., bandwidth) have become a feature offered by many cloud providers to content providers with intensive bandwidth demand. The streaming of media content to content viewers located at different geographical regions at guaranteed data-rate is a part of the service offered by the cloud provider. The common way of implementing this service in the cloud is by having multiple data-centers inside the networks of the access connection providers (e.g., Internet Service Providers, ISPs) located at appropriate geographical locations.

Cloud service providers may need to negotiate contracts with a number of ISPs to locate their servers into the networks of those ISPs. In this regard, another group of papers have focused on studying different types of contracts between cloud service providers and ISPs with the purpose of minimizing the expenses of cloud providers. However, an interesting design approach is to look at the resource reservation problem from the viewpoint of content providers. Obviously, content providers are more interested in minimizing their costs, i.e., the amount of money that they are charged directly by cloud providers.

To the best of our knowledge, very few studies have investigated the problem of optimizing resource reservation with the objective of minimizing the monetary costs for content providers. A good example is presented in, wherein a resource reservation optimization problem was formulated to minimize the costs of content providers, so-called cloud consumers, using a stochastic programming model. In the process of problem formulation, uncertain demand and uncertain cloud providers' resource prices are considered. In contrast, the optimization problem formulated in our work takes into account a given probability distribution function obtained from a fore mentioned studies for the prediction of media streaming demands.

Furthermore, the problem of cost minimization is addressed by utilizing the discounted rates offered in the non-linear tariffs. To the best of our knowledge, none of the previous papers has investigated the problem of cost minimization for media content providers in terms of monetary expenses by taking into account both the penalties caused by the over-provisioned
or under-provisioned reserved resources, and the advance purchase of resources at cloud providers for just the right period of time.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901031 |

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 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.
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.

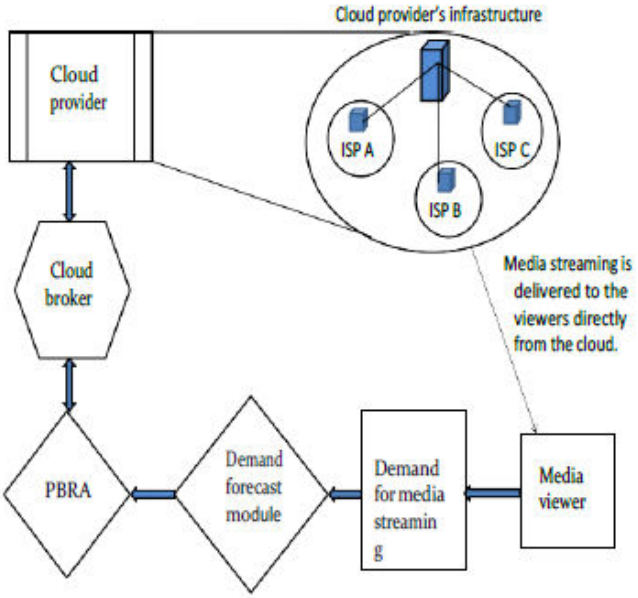


Fig. 1 : System Design

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901031 |

IV. ALGORITHM DESIGN

We summarize the assumptions that we use in our analysis as follows.
1) We assume that upon receiving the resource allocation request by the cloud provider from the media content provider, the resources required are immediately allocated in the cloud, i.e., updating the cloud configuration and launching instances in cloud data-centers incurs no delay.

2) Since the only resource that we consider in this work is bandwidth, it would be important to delve into the relation between the cloud provider and Content Delivery Networks (CDN). However, we assume that the provisioning of media content to media viewers (clients of the media content provider) located at different geographical regions at guaranteed data-rate is a part of the service offered by the cloud provider. The common way of implementing this service in the cloud is by having multiple data-centers inside the networks of the access connection providers (e.g., ISPs) located at appropriate geographical locations.

3) We assume that the media content provider is charged for the reserved resources in the cloud upon making the request for resource reservation (i.e., prepaid resources); and therefore, the media content provider cannot revoke, cancel, or change a request for resource reservation previously submitted to the cloud.

4) In clouds, tariffs (prices of different amount of reserved resources in $ per unit of reservation time) are often given in a tabular form.
Therefore, the cloud service provider requires a minimum reservation time for any allocated resources, and only allows discrete levels (categories)of the amount of allocated resources in in the Amazon Cloud Front resource provisioning plans [7].
We take into account the a aforementioned constraints and propose a practical - easy to implement -algorithm for resource reservation in the cloud, such that the financial cost on the media content provider is minimized.

V. 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-Line a 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 unnecessary amount to be paid to eliminate the wastage. The results show that our algorithm adjusts 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.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901031 |

EXPERIMENTAL RESULTS

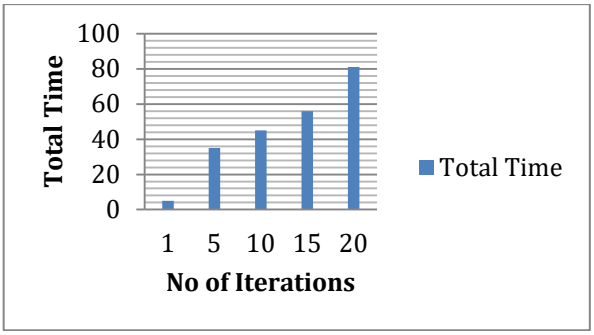


Fig: Total Time Vs No of iterations

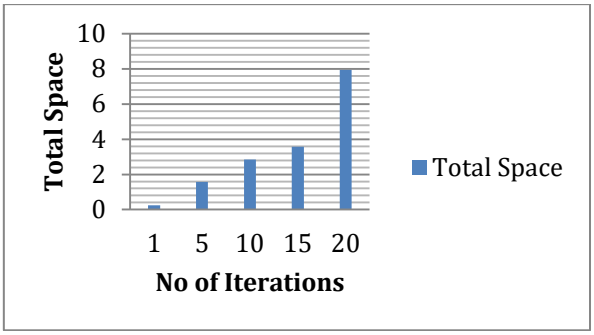


Fig: Total space required Vs No of iterations

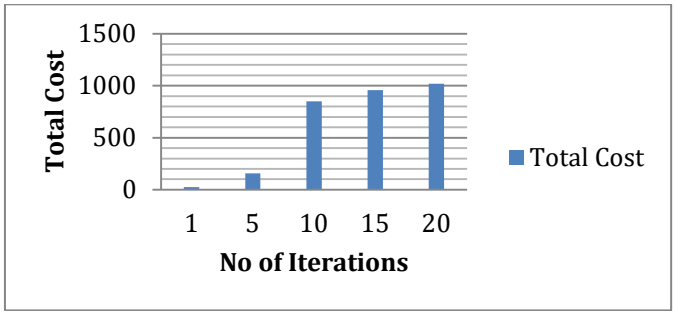


Fig: Total Cost Vs No of iterations

VI. 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] 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] Amazon CloudFront, http://aws.amazon.com/cloudfront/, 2012.
[8] 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.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 1, January 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0901031 |

[9] H. Ballani, P. Costa, T. Karagiannis, and A. Rowstron, "Towards predictable datacenter networks," in Proc. of the ACM SIGCOMM conference, pp. 242–253, 2011.
[10] 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.
[11] 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.
[12] 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.
[13] G. Gursun, M. Crovella, and I. Matta, "Describing and Forecasting Video Access Patterns," in Proc. IEEE Infocom Mini-Conference, pp. 16–20, 2011.
[14] 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.
[15] 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.

INNO SPACE

SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details