ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

CLOUD SYSTEM FOR MEDIA STREAMING APPLICATIONS BASED ON DYNAMIC PREDICTION SCHEDULING

Mythily.S [1], Mohanapriya.M [2]

1. P.G. Student, Dept. of MCA, VSB Engineering College, Karur, Tamilnadu, India
2. HoD, Dept. of MCA, VSB Engineering College, Karur, Tamilnadu, India

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 CloudFront 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.

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

I. INTRODUCTION

Media streaming applications have recently attracted large number of users in the Internet. The increase in the number of users has constantly increased the demand for HD videos, higher bit rates with no compromise on QoS (Quality of Service). To handle the same cloud service providers have to equip themselves with maximum resources to supply the demand. However it is possible to anticipate the usage peaks and thereby allocate resources behind it accordingly thereby increasing efficiency and lowering costs. Studies have shown that most servers in a typical data-center of a media content provider are only used at about 30% of their capacity [1]. The following Prediction Based Resource Allocation as understood from [2] has been studied, simulated and a an appropriate pricing model has been suggested. Cloud computing is combination of three major areas called Distributed processing, parallel processing and grid computing. Using clouds we can do online transactions, access software and use applications from anyplace. End user demand on VoD is equal to 3 billion DVDs per month. The objective of cloud computing is to provide trusty, custom-make and QOS(Quality of service) which ensure efficient incentive environment to the Media content viewer. Media streaming application are viewed by most of the internet user very frequently. Secure to the data is need because data is store in data center through

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

internet connection. Using Application programming interface any cloud user can get access to data center from cloud provider at anyplace and any time. Types of services provide to end user are platform as service or software as service. Positive side of cloud is low-cost, re-provisioning of resources and pervasive. Cloud able to provide high quality of service to end user with less capital investment to their physical infrastructure and maintains. End user of cloud can access data from data center through internet at any place and at any time. The resource allocation is major task where it improves high quality of service with minimum cost in pervasive environment. Over internet many cloud applications required the resources to perform their task. Resources should be distributed among cloud applicants so that there should not be starvation. To solve such problem we should allocate request amount of resource for individual module. Utilizing and allocation of hardly resources in bounded cloud environment required resource allocation strategy. In order to finish the end user task, RAS required type and amount of resources need by cloud application. Time and order is important input to allocate resource allocation. Type of resources provision plan the cloud provider provides based on daily, weekly, monthly and yearly plan up on needs is called On-demand plan. Another resource provision plan is reservation plan where the amount is paid before the utilization of resource. Christo Ananth et al. [5] discussed about a method, Sensor network consists of low cost battery powered nodes which is limited in power. Hence power efficient methods are needed for data gathering and aggregation in order to achieve prolonged network life. However, there are several energy efficient routing protocols in the literature; quiet of them are centralized approaches, that is low energy conservation. This paper presents a new energy efficient routing scheme for data gathering that combine the property of minimum spanning tree and shortest path tree-based on routing schemes. The efficient routing approach used here is Localized Power-Efficient Data Aggregation Protocols (L-PEDAPs) which is robust and localized. This is based on powerful localized structure, local minimum spanning tree (LMST). The actual routing tree is constructed over this topology. There is also a solution involved for route maintenance procedures that will be executed when a sensor node fails or a new node is added to the network.

We design the system model .we modify the secure algorithm by adding more user credentials and used prediction algorithm for finding future demands for resource allocation. Section II gives survey, section III and IV address existing and proposed methodology, section V system modules, section VI address our screenshots and performance. It shows the proposed system model assured security and significantly reduces the cost on resource allocation in the cloud as compare to conventional model.

II. LITERATURE SURVEY

Young Sil Lee et al.[16] [17]address how to provide security by using OTP (one time password) .OTP is used only once and it is generated randomly. For creating of OTP the used Pin- pong 128 chipper which already implements but we have used additional user credential for more security. Final OTP of four digits are generated by with two random digits. RichaChowdhary,SatyakshmaRawat [18]

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

address issues in multi cloud environment and how to provide the multi high level security. Multi-level security by combine OTP and session password and also hardware level security for multi-cloud platform. Shikharesh et al. [20]has given description on basic fundamental point of resource management and what are the problems in cloud while allocating of resources for multiuser. Resource allocation strategies based on execution time and pre-emptible nature job are schedule by estimating time. But predicting time for executing a job is difficult for user and more wrong assumption are made. This problem is solved by Virtualization concept in distributed environment called as infrastructure as a service.

To work on VM concept, Jose et al proposed new idea called matchmaking .Paper talks about advance reservation for resource scheduling polices. In paper Kuo-Chan et al. [10], used most fit processor policy for resource allocation in cluster environment. It checks for immediate subsequent jobs allocation available clusters and assign this policy which gives efficient throughput. But it again leads a problem for searching appropriate cluster to assign policy in distributed environment contain homogenous and heterogeneous clustered distributed geographically . One more problem is process migration, when load increases in one particular VM it automatic migration of job is needed based on location. A system which automatically scale it resource is design in this paper [6] in dynamic environment and dynamic resource allocation can be assigned. Paper focus on a live migration in distributed environment and how it able to relocate across multi infrastructure physical environment and scale automatically scale it resources talks about only considering non-preemptive job scheduling only. Researcher"s studies lead to develop a new efficient resource allocation for real time task in disturbed environment. [9], [8], [6] Power management while VM selection and performing real time job scheduling considering aspect like selecting appropriate policies to increases speed and to minimize cost. Resource allocation in IAAS (Infrastructure as-a service) is different for Resource allocation in SAAS (software as-a service).Among the selfish VMs Zhen Kong et al. proposed a new design for non-cooperative environment. Design use stochastic approximation to improve the quality of service and measure, analyze resource allocation under various circumstances. Proposed method given efficient result but it can"t be implemented in real time work load share environment. Four Reasons We Choose Amazons Cloud as Our Computing Platform [15] address why to choose Amazon web services as our cloud platform. What is the future built and deployed search engines in AWS. This Netflix blog focus on technology. Every one discusses on new technology issues and gives their own perceptive on the issues. Gossip-based protocol [4] is proposed for huge environment for resource"s allocation. In the thesis, performance efficiency is measured using dynamic set model contain set of node having specific cup capacity and memory capacity. It allocates resources in cloud environment in where application requested resources are allocated with I time-dependent memory demand and actively enlarge utilization of resource"s in cloud. Based on time it request and allocated time slot it assign and reutilizing of resource is talked in paper [13].The simulation result

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

shows that when the memory demand is small then available memory in cloud provide good quality of service allocation even number of applicants increases up to threshold value. Still need improvement in field of reliability in case of huge cluster and datacenter. Paul et al. [7] address the three different polices to avoid over provision and under provision when there is change in change in user demand. To leverage the automatic scaling to or from remote node, assuming remote node is set of resource that provides resources for multiple cloud providers. But it is consider being on preemption concept. yang et al.[13] have proposed a methodology for automatic scaling by gathering the scientific database knowledge of scaling application server is called as profile based approach. This approach gives realistic result in utilization of resources and also talks about platform as a services.

There are few works [2], [11] that dynamically allocate CPU resources to meet QoS objectives by first allocating requests to high priority applications. Hence the authors‟ Dorian et al. proposed Utility (profit) based resource allocation for VMs which use live VM migration (one physical machine to other) as a resource allocation mechanism [3]. This controls the cost-performance trade-off by changing VM utilities or node costs. This work mainly focus on CPU resources scaling in IaaS. A few works [1],[8] that use live migration as a resource provisioning mechanism but all of them use policy based heuristic algorithm to live migrate VM which is difficult in the presence of White paper on Forecast methodology, 2010-2015 [14] document analysis cisco VNI (Visual Network Index) global ip traffic forecast and the methodology. It gives How many user are using streaming VOD over internet from 2010-2015.What are the traffic problems and how the traffic increased from 2010 to 2015.Paper pointing about web based file sharing that include traffic in p2p applications like BitTorrent and eDonkey .It also talks about Internet game downloads, online gaming, voice over ip, video communication, mobile data traffic and business ip traffic.

2.1 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.

2.1.1 Disadvantages:

1. Applications are economical inefficient.
2. Servers will be idle most of the time. So, it is the waste of time.

3. PROPOSED SYSTEM

We propose a simple and 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.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

Advantages:

1. Reduce the cost for resource allocation.

Objective: Reserving sufficient resources in the cloud by the media content provider – based on the prediction of future streaming demand – such that no resource wastage is incurred, while QoS for the actual (real) streaming traffic is maintained with some level of confidence (η).

- **D(t) :** Actual demand for streaming capacity of the video channel at an instant of time t, and measured as the number of users that stream the channel at instant of time t multiplied by the data rate required for every downloading user to meet QoS guarantees. D(t) follows a log normal distribution with mean E(D(t)) and variance σ.
- **Alloc(t):** Amount of streaming bandwidth that the media content provider allocates in the cloud at any time instant t.
- **η** : pre-determined threshold (level of confidence)
- **wj**: period of time over which resources are reserved
- **Tariff(wj , Allocj) :** Price (in Rs. per time unit) charged by the cloud provider for amount of resources Allocj reserved for period of time (window size) wj.
- **Cost(wj , Allocj) =** Tariff(wj , Allocj) × wj.

Hence Objective is **Minimizing Cost(wj , Allocj) subject to Probability(D(t) ≤ Alloc(t)) ≥ η**

FIGf

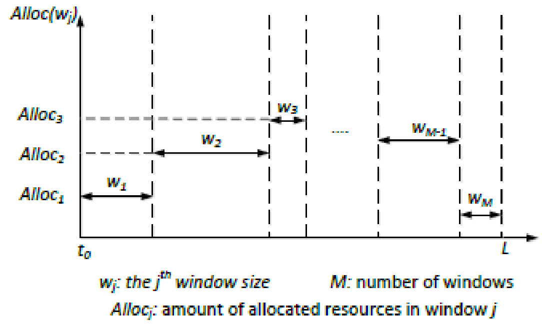

Fig 1. PDRS algorithm design

4.SYSTEM OVERVIEW

The system model that we advocate in this paper for media streaming using cloud computing consists of the following components (Fig. 1). _ Demand forecasting module, which predicts the demand of streaming capacity for every video channel during future period of time. _ Cloud broker, which is responsible on behalf of the media content provider for both allocating the appropriate amount of resources in the cloud, and reserving the time over which the required resources are allocated. Given the demand prediction, the broker implements our proposed algorithm to make decision on resource allocations in the cloud. Both the demand forecasting module and the cloud broker are located in the media content provider site. Cloud provider, which provides the streaming resources and delivers streaming traffic directly to media viewers.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

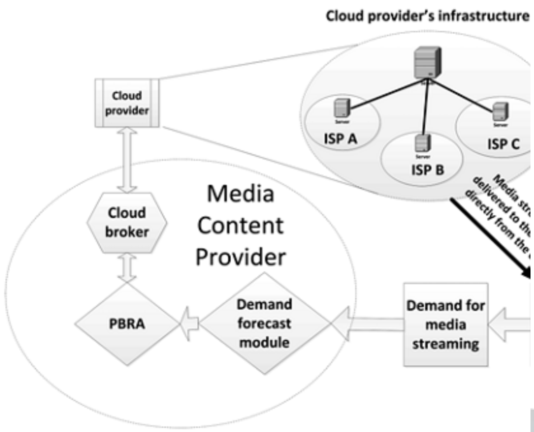

Fig 1. System Design

Algorithm 1: Prediction-based Dynamic Resource Scheduling (PDRS) Algorithm

Input: *PM:* all the physical machines
VM: all the virtual machines
Output: *MigrationSchedule*

% Resource Confliction Prediction Phase;
Let *Busy* be the set of PMs which is migrating VMs;
Let *Available* be the set of PMs which holds *vms* but has no migration action;
Let *Idle* be the set of the rest PMs
for *all pms* in *Available* **do**
PM Conflict Predict
T_c =*Conflict_Predict*(pm_i , *k_step*) //Algorithm 2;
if T_c □□□1 **then**
% Algorithm 3
vm_x =*MoveOutVMDetermination*(pm_i , $T_{conflict}$)
Add vm_x to *ToMigrateList*;
Remove vm_x from *Available*;
end
end
% Resource Consolidation Phase;
for *all* vm_i in *ToMigrateList* **do**
Pick up vm_i with the smallest T_c ;
Let pm_{sour} be the pm which hold vm_i ;%Algorithm 4
pm_{des} =*handleConflict*(vm_i) //Algorithm 4
if pm_{des} □ *NULL* **then**
Add (vm_i □ pm_{des}) into *MigrationSchedule*;
Move pm_{des} and pm_i to *Busy*;
Delete vm_i from *ToMigrationList*;
end
end
Sort the *pm* in *Available* in descending order by their predicted load situation *state*;
while *Available.size*() □ 0 **do**
Pick the last pm_i in *Available* which is also the lightest loaded;
Pick the lightest loaded vm_i in pm_i ;
while pm_t ! □ *Available.end*() **do**
if pm_t has enough spare space to hold vm_i without conflicts **then**
Add (vm_i □ pm_t) into *MigrationSchedule*;
Move pm_t from *Available* to *Busy;*
Add pm_t to *Busy;*
break;
else
Let pm_t be the next one in *Available;*
end
end
Remove pm_i from *Available*;
end
return *MigrationSchedule;*

1)We assume that upon receiving the resource allocation request by the cloud provider from the

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

media content provider, the resources required are immediately allocated in the cloud, i.e., updating the cloud configuration and launching instances in cloud data-centres 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-centres 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 Rupees per unit of reservation time) are often given in a tabular form. See for example the reservation phase in the Amazon CloudFront resource provisioning plans [7].

4.1 Modules Implementation

1. Demand forecast module
2. Cloud broker module
3. Cloud provider module

Demand forecast:

This module predicts the demand of streaming capacity for every video channel during future period of time.

Cloud broker:

This module is responsible on behalf of the media content provider for both allocating the appropriate amount of resources in the cloud, and reserving the time over which the required resources are allocated.

Cloud provider:

This module provides the streaming resources and delivers streaming traffic directly to media viewers.

5 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] Q. Huang, S. Su, S. Xu, J. Li, P. Xu, and K. Shuang, "Migration-based elastic consolidation scheduling in cloud data center," *in Proceedings of IEEE ICDCSW* 2013.

[2] S. Srirama, O. Batrashev, and E. Vainikko, "Scicloud: scientific computing on the cloud," in Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing. *IEEE Computer Society*, 2010, pp. 579–580.

[3] F. W. Li and R. W. Lau, "Emerging technologies and applications on interactive entertainments," *Journal of Multimedia*, vol. 6, no. 2, pp. 107–114, 2011.

[4] S. Ma, "A review on cloud computing development," *Journal of Networks*, vol. 7, no. 2, pp. 305–310, 2012.

[5] Christo Ananth, S.Mathu Muhila, N.Priyadharshini, G.Sudha, P.Venkateswari, H.Vishali, "A New Energy Efficient Routing Scheme for Data Gathering ",International Journal Of Advanced Research Trends In Engineering And Technology (IJARTET), Vol. 2, Issue 10, October 2015), pp: 1-4

[6] "Amazon elastic compute cloud." [Online]. Available: http://aws. amazon. com/ec2/

[7] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," *ACM SIGOPS Operating Systems Review*, vol. 37, no. 5, pp. 164–177, 2003.

[8] J. Sahoo, S. Mohapatra, and R. Lath, "Virtualization: A survey on concepts, taxonomy and associated security issues," *in Computer and Network Technology (ICCNT), 2010 Second International Conference on. IEEE*, 2010, pp. 222–226.

[9] A. Kivity, Y. Kamay, D. Laor, U. Lublin, and A. Liguori, "kvm: the linux virtual machine monitor," *in Proceedings of the Linux Symposium*, vol. 1, 2007, pp. 225–230.

[10] Y. Hashimoto and K. Aida, "Evaluation of performance degradation in hpc applications with vm consolidation," *in Networking and Computing (ICNC), 2012 Third International Conference on. IEEE*, 2012, pp. 273–277.

[11] J. G. F. Hermenier, X. Lorca and J. Lawall, "Entropy: a consolidation manager for clusters," *In Proc. VEE*, 2009.

[12] P. A. A. Verrna and A. Neogi, "pmapper: Power and migration cost aware application placement in virtualized systems," *Proceeding of ACM/IFIP/USENIX 9th International Middleware Conference*, 2008.

[13] A. T. Wood, P. J. Shenoy and M. S. Yousif, "Black-box and gray-box strategies for virtual machine migration," *In Proc. NSDI*, 2007.

[14] A. Corradi, M. Fanelli, and L. Foschini, "VM consolidation: a real case based on openstack cloud," *Future Generation Computer Systems*, 2012.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 12, March 2016

[15] D. N. B. J. M. Armbrust, A. Fox, "Scads: Scale-independent storage for social computing applications, " *In Proc. CIDR*, 2009

[16] P. B. Urgaonkar, M. S. G. Pacifici and A. N., "An analytical model for multi-tier internet services and its applications," *In Proc SIGMETRICS*, 2005.

[17] P. B. Urgaonkar, M. S. G. Pacifici and A. N, "Automated control for elastic storage," *In Proc. ICAC*, 2010.

[18] S. J. E. C. I. C. Clark, K. Fraser and A. Warfield, "Live migration of virtual machines, " *In Proc. NSDI*, 2005.

[19] N. Bobroff, A. Kochut, and K. Beaty, "Dynamic placement of virtual machines for managing SLA violations," *in Integrated Network Management, 2007. IM'07. 10th IFIP/IEEE International Symposium on. IEEE*, 2007, pp. 119–128.

[20] Z. Gong and X. Gu, "Pac: Pattern-driven application consolidation for efficient cloud computing," *Signature*, vol. 1, no. 2, pp. 2–2, 2010.

[21] G. E. Box, G. M. Jenkins, and G. C. Reinsel, Time series analysis: forecasting and control. *Wiley*, 2011, vol. 734.

[22] "Rubis online auction system," *http://rubis. ow2. org/*.