

Resource Allocation with Profit and Reputation Maximization in Cloud Computing: P and P Model.

Abhay Kumar Mishra¹, Nagendra Kumar²M.Tech. Scholar, Shri Ram Institute of Science & Technology, Jabalpur, M.P, India¹Asst. Professor, Shri Ram Institute of Science & Technology, Jabalpur, M.P, India²

ABSTRACT: Cloud computing has introduced a major shift in the IT delivery model by offering computing resources for hosting applications as a utility. This helps businesses and organizations to access advanced IT facilities offered by cloud providers without the expensive up-front investments necessary to establish their own infrastructure. In this context, significant research efforts have already been made that aim to minimize costs for cloud customers; less attention however, has been given to challenges and opportunities that cloud providers face when striving for profit maximization. This thesis presents a set of novel market and economics-inspired policies, mechanisms, algorithms, and software designed to address the profit maximization problem of cloud providers. Our solutions are proposed for two main types of providers: Those that rely solely on their own resources to serve customers and those that participate in a cloud federation and benefit from resource sharing.

KEYWORDS: Cloud Computing, Service Level Agreement, SLA Management, SLA Pricing, Resource provisioning.

I. INTRODUCTION

Cloud computing is originally developed from distributed computing; it can be defined as a type of parallel and distributed system which has many interconnected computers or servers [1]. It is a promising technology which attracts researchers, academicians and computing industries in great extent because of its computing capability to deliver shared cloud objects dynamically. Ever since its conception, cloud computing has been revolutionizing the way data storage and processing mechanisms are envisioned and implemented. It enabled the on-demand availability of services such as Software, Platform, Infrastructure (through SaaS, PaaS, IaaS respectively) and thus formed an economic solution to meet the ever-fluctuating demand for storage and computational resources by growing businesses.

A. Cloud Computing Pricing Model: Cloud Computing Providers offer numerous online services based on SLA (Service Level Agreement) between the provider and the customer. However an important role between providers and customers relationship has pricing model for which they must agree. Each provider has his scheme for calculating the price (has services based on SLA (Service Level Agreement) between the provider and the customer offered for clients. The provider's goal is to have a greater benefit, while each client's goal is to have the maximum service for low price. Therefore, satisfying both parties requires an optimal pricing methodology. The price charged is one of the most important metrics that a service provider can control to encourage the usage of its services [2]. Price is an important factor for the company which provides cloud services because it affects the clients directly and organization profit. The price also has a major impact in economic aspect, where key concepts such as fairness and competitive pricing in a multi-provider marketplace affect the actual pricing. Pricing for competition and fairness affects choices in the design of user applications and system infrastructures. In fact pricing fairness balances user cost and cloud service provider

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

profit. Pricing model in Cloud Computing is more flexible than traditional models. Every cloud provider has its own pricing scheme.

Main focus of Cloud Computing is to fulfil and guarantee quality of service (QoS) for customers. The price in Cloud Computing and value chain is based on business models and framework. The value chain from the traditional IT services is changing as a result of cloud computing. The key issue is how price / costs are measured, accounted, and distributed between different service layers and organizational units responsible of them. There are many business models based on different service models that determine the price of services in the cloud. For example below we see an illustration of a model for Cloud Computing cost accounting [3] which addresses cost accounting issues in production of cloud services.

Price wars have gotten a lot of press, with providers lowering prices dozens of times over the last few years, nominally due to reductions in cost structure, but also in an attempt to gain share, receive publicity, and signal their aggressiveness in the market to competitors. However, although cloud *prices* are of interest, so is cloud pricing models, such as reserved instances, sustained-use pricing, and spot instances. It would be a mistake to consider pricing models to be an afterthought; they're a means of competitive differentiation as well as for creating value for both customers and providers.

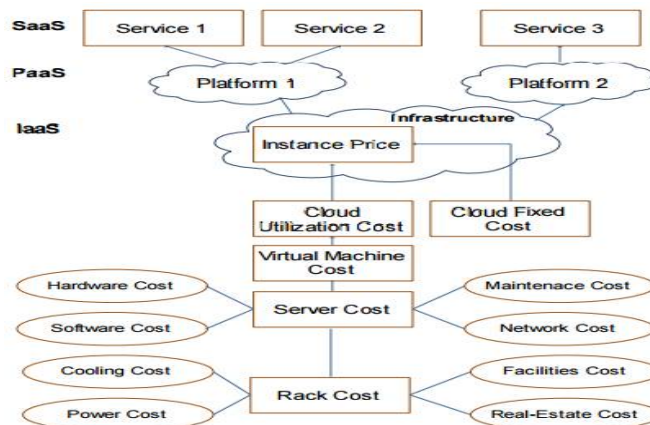


Figure 1: Cloud computing cost accounting model.

CSPs (Cloud Service providers) aim to increase their profits by trying to accept the highest possible number of jobs and increase their customer satisfaction by abiding to their SLAs (Service Level Agreements). As the cloud market gets mature, each CSP tries to maintain her market share through the establishment of long-term business relationships with her customers. The success of these relationships critically depends on the extent to which the SLAs are being respected [4]. "Should this incoming request be accepted?" is a critical question raised when the studied CSP is experiencing saturation and cannot offer its incoming requests the same SLA levels that used to ensure customer satisfaction. Rejecting an incoming job decreases profit while accepting it and failing to respect the SLA severely affects customer satisfaction and CSP's reputation. This question is very complex and hard to be answered due to the variety in available cloud services, where above question for interactive and non-interactive computation services only.

II. RELATED WORK

A) Pricing Models: When providers understand how much it costs to provide the services he or she is offering, how much competitors are charging for the same service, and how customers discern the value of services, it is time to find out what type of pricing model they should utilize. The most commonly used pricing models in cloud computing markets, especially in infrastructure as-a-service, are usage-based, subscription-based, and demand-oriented pricing models.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

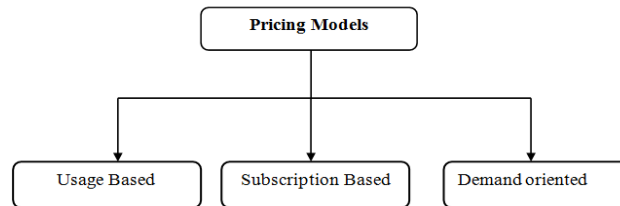


Figure 2: Pricing Models.

B) Usage-based pricing model: Basically, cloud computing can be defined as delivery of on-demand access to computing services on a pay-as-you-go basis. A usage-based pricing model relies on the scheme that customers pay according to the amounts of services that they use. Usage-based pricing model is the most common pricing model used by IaaS cloud service providers. In this model, the provider evaluates the services that they provide, and charge customers accordingly. In the usage-based pricing model, cloud providers often charge for services only on a fixed-rate basis. There is a large area of literature on price analysis of running applications on clouds considering the usage-based pricing model in cloud computing [5, 6].

A related work by Sharma et al. [7] developed a cloud resources pricing model that uses financial option model to give a lower bound on the prices and age of resource to give an upper bound on prices for what they call cloud compute commodities.

C) Subscription-based pricing model: Subscription-based pricing model is a pricing model that allows customers to pay a subscription fee to use the service for a particular time period. This is often popular among Software-as-a-Service (SaaS) cloud providers, where vendors deliver cloud services over the Internet. The idea behind subscription-based pricing is that customers pay a fee to subscribe to a service over a predefined time period and they can regularly use the service during the subscription period. Subscription-based pricing models with more or less modifications are used by IaaS cloud provider as well while it is called with different terms such as reservation contract or prepaid scheme [8].

D) Demand-oriented pricing model: Demand-oriented pricing model is the process of establishing a price for a service based on the level of demand. The service price is changed according to its demand in a way that when the demand is high the price goes up high and when it is low the price goes down.

Dynamic Pricing

The problem of cost maximization with dynamically allocating resources to spot markets has been investigated by Zhang et al. [9]. Supply adjustment and dynamic pricing are used as a means to maximize revenue and meet customer demand. Wang et al. [42] proposed an optimal method for a dynamic pricing in cloud market based on the uniform price auction. In order to maximize the cloud provider's revenue, a dynamic pricing model based on genetic algorithms has been proposed by Mac'ias and Guitart [10]. As shown in Figure 3 below, there are various types of dynamic pricing for setting the price of cloud resources. These dynamic pricing strategies can be categorized into two main groups: price-discovery and price-posted models. In the former, the provider sets the price based on the communication with the customers, e.g., asking them to report their bid. Auction-based and negotiation-based techniques fall into this group. The latter, the price-posted model, does not necessarily require communication with the customers and the provider posts the pre-determined price which dynamically varies during the time based on some external factors such as demand or time of use. In the following figure, we show about various type of dynamic pricing suitable for cloud computing.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017



Figure 3: Dynamic Price Model.

Economic challenges and enabling approaches Cloud interoperability and community clouds raise many more challenges than cloud computing both in terms of functional and non-functional aspects. To overcome these challenges, vast efforts are required to research and develop tools and techniques in this regard. These challenges cover Service Level Agreement (SLA), monitoring, virtualization, portability, economy, provisioning. We review four main economic-related challenges and identify possible enabling approaches. These four main challenges of cloud environments from the economic point of view are shown in Figure 4.

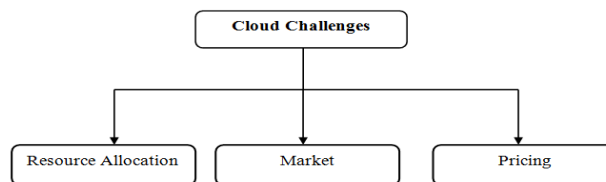


Figure 4: Cloud SLA Challenges.

Resource Allocation Strategy: Resource allocation is a challenging issue from the cloud provider's perspective. Cloud providers usually offer their virtualized resources based on different QoS levels, e.g., best effort and reserved. As the number of resource consumers is increasing, clouds need to share their resources with each other to improve their quality of service. In general, such a collaborative cloud computing system is prone to argument between user requests for accessing resources [11]. Resource utilization is not a new issue in cloud computing environments. Various solutions have been proposed for the resource utilization problem in cloud environments. Gomes et al. [12] propose and investigate the application of market-oriented mechanisms based on the General Equilibrium Theory to coordinate the sharing of resources between clouds. In a Cloud Computing, providers can mutually conspire to share their resources and fulfil the demand among each others. Similarly, a provider that has underutilized resources could lease part of them to other providers in the federation [13]. Therefore, cloud providers require having a comprehensible understanding of the consequences of every decision they make.

Resource Provisioning Policies to Increase Profit: One of the key motivations for Cloud Service providers is the possibility of making profit by outsourcing their available data center resources to serve thousands of users. Therefore, Cloud providers crave to accept as many new requests as possible with the objective of maximizing profit; they must guarantee Quality of Service (QoS) based on the agreed Service Level Agreement (SLA) with customers. Achieving this goal requires efficient resource management strategies. To be able to offer QoS guarantees without limiting the number of accepted requests, providers must be able to dynamically increase the available resources to serve requests. One possible source for additional resources is idle resources from other providers. In order to enable such a scenario, coordination between providers has to be achieved, possibly through establishment of a Cloud federation [14]. A challenging condition for providers occurs when they commit part of their capacity in the form of spot VMs. Spot VMs are VMs that can be terminated by providers whenever the current value for running such VMs exceeds the value that the client is willing to pay for using such resources. Despite several recently proposed platforms for Cloud computing, with different motivations and stimulus for parties to join it, many fundamental problems and questions about cloud computing resource utilization remain unanswered.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

A user satisfaction-oriented scheduling algorithm for service requests is proposed by Lee et al. Such an algorithm tries to maximize cloud providers' profit by accepting as many service requests as possible, as long as QoS is kept at a certain level. In this regard, contracting with other service providers was taken into account as a method to avoid rejection of user requests. The problem of how to value resources and how price may impact utilization is not trivial. Current public cloud service providers like Amazon, GoGrid and Rackspace usually adopt fixed pricing strategies for the infrastructure services they provide. However, fixed pricing models are not suitable for cloud computing environments as a policy to be applied between its participants, because it neither reflects current market price of resources due to dynamism in supply and demand nor generates any incentives for providers to join the cloud markets. Moreover, the problem of dynamic allocation of data center resources to different spot markets to maximize cloud provider's total revenue has been investigated by Zhang et al. [15].

III. PROPOSED METHOD

A) Proposed Resource Utilization Method: Proposed system divides all available resources into two sets where first set is of the slow resources and second set is of the fast resources. This partitioning is done according to the MIPS speed of each resource. Suppose that there are 10 available resources. The first step is to sort these resources into ascending order by considering their MIPS speed and then add first 5 resources into the first set and remaining 5 resources into the second set. When tasks arrived for execution, choose the resource which takes less time for execution among all available resources before assigning any task. If chosen resource is from the first set, then assign average length task to it and if chosen resource is from the second set, then assign the longest task to it. The technique used in proposed algorithm for scheduling tasks is the combination of both Max-Min as well as Enhanced Max-Min. The proposed algorithm minimizes the chances of scheduling a large task to the slow resource with making completion time shorter. Proposed algorithm helps to utilize resources more efficiently and achieve good performance in terms of makespan as compared to existing algorithms. Algorithm is given below:

Input: T is task set, R is Resource set with MIPS speed, and n is total no of Resources

Algorithm ProposedRP (T, R, n)

```
{
    For  $t_i \in T$  do
    {
        Arrange elements of set R in ascending order of execution speed in MIPS.
    }
     $R_{lazy} = R_i$  form 1 to  $n/2$ ;
     $R_{speedy} = R_i$  form  $n/2$  to n;
    While  $t_i \in T$  and T is not empty do
    {
        Find  $R_i$  with minimum execution time
        If  $R_i \in R_{lazy}$  than
            Assign medium length task to  $R_i$ 
        Else
            Assign long length task to  $R_i$ 
        Delete  $t_i$  from T
    }
}
```

B) Proposed Profit Maximization Method The main computing capacity is provided into speedy and lazy resources so that we optimize available resources due to which no need to take them on rent. It reduces investment cost and increases the profit. The lazy computing resources provide the extra capacity in peak period. The proposed max min queuing scheme will be used to assign resources for requests. For each service request entering the system, the system record sits waiting time. The requests are assigned and executed on the speedy computing resources in the order of

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

maximum execution time. It is apparent that all service requests can guarantee their deadline and are charged based on the workload according to the SLA. Hence, the revenue of the service provider increases. Since the revenue has been maximized using our scheme, minimizing the cost is the key issue for profit maximization. Algorithm is explained below:

Input: m is number of servers; E is set of events, Q service queue

Algorithm Proposed Profit (m, E)

```

{
    A queue Q is initialized as empty
    ProposedRP(T,R,n);
    if  $E_i \in E$  and any server is available then
        Assign the service request to one available server
    else
        Put it at the end of queue Q
    end if
    if  $E_i \in E$  and a server becomes idle then
        Search if the queue Q is empty
        if true then
            Wait for a new service request
        else
            Take the service request from queue Q according to maxmin method
        end if
    end if
}

```

IV. EXPERIMENTAL SETTINGS AND RESULTS

The metrics used to measure the performance obtained by resource allocation strategies are mean execution time of all resources utilized. This section contains the implementation part of this research that means development and simulation process of the proposed algorithm using WorkflowSim tool as well as software and hardware requirements for this experiment. The software used for this experiments are Netbeans, JDK 1.8 and WorkowSim 1.0. We have developed the proposed algorithm in Java language. Then WorkowSim simulator is used to test the developed algorithm. To test the proposed algorithm, one data center, one scheduler, and 20 virtual machines are used first. Then proposed algorithm tested by using workflows named Montage50, to compare the mean execution time. We have also developed FCFS and MaxMin algorithms in Java using their pseudo codes for the simulation purpose. Because we compare our proposed algorithm with FCFS and MaxMin to check whose performance is better? Table below shows the comparisons of three algorithm with increased virtual machine size i.e. VM=20 and VM=30.

ALGORITHM No. Of VM	FCFS ALGORITHM(MS)	MAX-MIN ALGORITHM(MS)
VM=10.	41.78	14.94
VM=20.	61.94	24.79
VM=30.	72.79	27.11

Table 3: Execution Time Comparison of Algorithms..

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

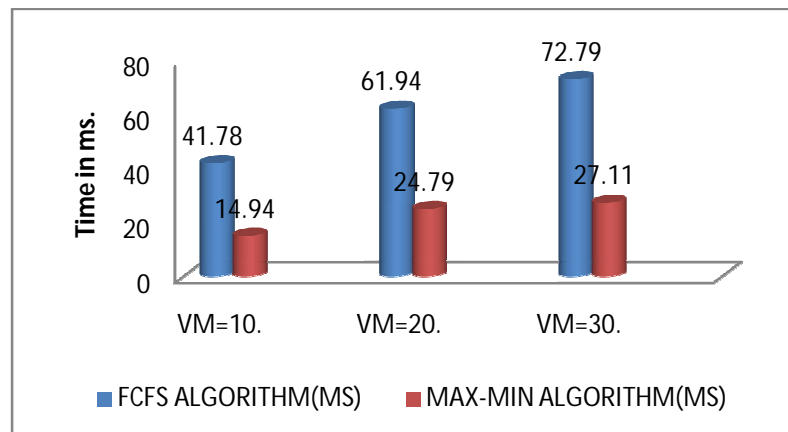


Figure 5: Comparison of Execution time.

Results of above evaluations show that proposed algorithm completes tasks execution with lower execution time and higher performance as compared to scheduling algorithms FCFS. Performance of proposed algorithm is better than FCFS for 10, 20 and 30 virtual machines. Results shows that proposed algorithm behaves better in terms of Execution time after testing it with Montage50 workflows. As we increased a number of virtual machines which are used for simulation, still the total length of scheduling for the proposed algorithm is less than the FCFS.

V. CONCLUSION AND FUTURE WORK

The proposed work presents new resource provisioning algorithm based on proper resource utilization aim to profit maximization. This algorithm is implemented using WorkFlowSim simulation tool with Netbeans. Large task is assigned to the fast resource and smaller ones to lazy resources for proper utilizations. In proposed algorithm possibility of scheduling long length tasks to the slow resource is reduced. Solution used in this paper is division of resources into two groups according to MIPS speed. If the fastest available resource is from the second group which is the group of fast resources, then the largest task is scheduled to it. If the fastest available resource is from the first group which is the group of slow resources, then the average length task is mapped to it.

Resource-provisioning policies are designed to act based on burning time profit calculation without any knowledge about future requests. Developing algorithms that consider the prediction of future demand and resource availability to drive outsourcing decisions can be explored as an extension to our study. In future resource optimization can be performed using extension of proposed algorithm by combining various algorithms to provide hybrid algorithm.

References

- [1]. Peter Mell and Tim Grace, "The NIST definition of cloud computing", NIST Special Publication 800-145 (SP800-145), National Institute of Standards and Technology, Gaithersburg, January 2011, pp. 1-52.
- [2]. M. Al-Roomi, Sh. Al-Ebrahim, S. Buqrais and I. Ahmad, Cloud Computing Pricing Models: A Survey, International Journal of Grid and Distributed Computing Vol.6, No.5, pp.93-106, 2013.
- [3]. J. Jäätmaa, Financial aspects of cloud computing business models, Aalto University, 2010.
- [4]. Í. Goiri, F. Julià, J. Fitó, M. Macías, and J. Guitart, "Supporting CPUbased guarantees in cloud SLAs via resource-level QoS metrics", Future Generation Computer Systems, 28.8, pp. 1295-1302.
- [5]. U. Sharma, P. Shenoy, S. Sahu, and A. Shaikh, "A cost-aware elasticity provisioning system for the cloud," in Proceedings of 31st International Conference on Distributed Computing Systems (ICDCS'11), Minneapolis, U.S.A, June 2011, pp. 559-570.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

- [6]. H.Wang, Q. Jing, R. Chen, B. He, Z. Qian, and L. Zhou, "Distributed systems meet economics: pricing in the cloud," in Proceedings of the 2nd USENIX conference on Hot topics in cloud computing (HotCloud'10), Boston, Massachusetts, USA, June 2010.
- [7]. B. Sharma, R. K. Thulasiram, P. Thulasiraman, S. K. Garg, and R. Buyya, "Pricing cloud compute commodities: A novel financial economic model," in Proceedings of the 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid'12), Ottawa, Canada, May 2012, pp. 451–457.
- [8]. D. Niu, C. Feng, and B. Li, "Pricing cloud bandwidth reservations under demand uncertainty," SIGMETRICS Performance Evaluation Review, vol. 40, no. 1, pp. 151–162, June 2012.
- [9]. Q. Zhang, Q. Zhu, and R. Boutaba, "Dynamic resource allocation for spot markets in cloud computing environments," in Proceedings of the Fourth IEEE International Conference on Utility and Cloud Computing (UCC'11), Melbourne, Australia, Dec. 2011, pp. 178–185.
- [10]. M. A. Salehi, B. Javadi, and R. Buyya, "QoS and preemption aware scheduling in federated and virtualized grid computing environments," Journal of Parallel and Distributed Computing (JPDC), vol. 72, no. 2, pp. 231–245, 2012.
- [11]. "On economic and computational-efficient resource pricing in large distributed systems," in Proceedings of the 10th IEEE/ACM International Conference on- Cluster, Cloud and Grid Computing (CCGrid'10), Melbourne, Australia, May 2010, pp. 838–843.
- [12]. E. R. Gomes, Q. B. Vo, and R. Kowalczyk, "Pure exchange markets for resource sharing in federated clouds," Concurrency and Computation: Practice and Experience, vol. 23, no. 9, pp. 977–991, 2012.
- [13]. I. Goiri, J. Guitart, and J. Torres, "Economic model of a cloud provider operating in a federated cloud," Information Systems Frontiers, vol. 14, no. 4, pp. 827–843, 2011.
- [14]. Hao Hao. Zhou, Su. Deng, HongBin. Huang, and YaHui Wu, "Resource Allocation in Cloud Computing Based on Clustering Method", IEEE 2015.
- [15]. Q. Zhang, Q. Zhu, and R. Boutaba, "Dynamic resource allocation for spot markets in cloud computing environments," in Proceedings of the Fourth IEEE International Conference on Utility and Cloud Computing (UCC'11), Melbourne, Australia, Dec. 2011, pp. 178–185.