International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 10, October - 2013

A Model for an Optimal Approach for Job Scheduling in Cloud Computing

K.Thyagarajan[1], S.Vasu[2], S.Sri Harsha[3]

[1]*Assoc Prof, Department of CSE, SVCET, Chittoor.*

[2]*Assoc Prof, Department of CSE, SVCET, Chittoor.*

[3]*Asst Prof, Department of CSE, SVCET, Chittoor.*

.

***Abstract*- Job scheduling is one of the core research areas in cloud computing. Ultimate goal of the cloud service provider is to maximize the profit by utilizing the infrastructure resources effectively. Job scheduling policy must be effective for gaining the maximum profit. This paper proposes a model for an optimal approach in job scheduling in cloud computing for effective gain of the profits to the cloud computing service provider.**

Keywords: cloud computing, Service Level Agreement (SLA),Infrastructure as a Service (IaaS), Cloudsim 2.1, priority check controller, job scheduler, Activity Based Costing(ABC) .

1. INTRODUCTION

Cloud computing is rising as the next generation platform for computation. It can be a combination of distributed and parallel computing [5]. Main advantages of the Cloud computing are that, it is used for on-demand gathering of information, technology services and products. In general we can define cloud computing as internet-based computing, whereby shared resources, software and information are provided to computers and other services on demand based on the service level agreement between user and service provider. It is one of the emerging technologies and it can replace the large infrastructures with its infrastructure as a service provider (IaaS) service. So for effective utilization of resources there should be an effective scheduling policy which helps in gaining maximum profits. In this paper a model has been proposed for scheduling the requests which obtains maximum profits to the provider.

The remainder of the paper is organized as follows: Section 2 gives related work carried out section 3 describes the overall problem statement. Section 4 gives a brief idea of the proposed method. Section 5 provides the information of behavior of Cloudsim toolkit Section 6 presents the experimental results and Section 7 ends with conclusion and the future work to be done.

2. RELATED WORK

Job scheduling algorithms for cloud computing service provider is very rare in the literature in terms of gaining maximum profits. Several algorithms [2] [7] in the literature follows queuing policy with some priority based parameters.

Costing in cloud computing has several overheads where there are chances of over-costing and over-priced in some cases and under-costing and under-priced in some cases. This problem is overcome using a method called Activity Based Costing (ABC) where it helps in accurate costing [1]. ABC helps in reducing the extra charges; it tries to eliminate unwanted charges which take in costing the cloud resources. Accurate costing helps the cloud computing users for better service and also helps the service provider for estimating the profits.

3. PROBLEM DEFINITION

Cloud computing categorizes the services into three types namely 1. Infrastructure as a service (IaaS) 2.Platform as a service (PaaS) 3. Software as a service (SaaS). Job scheduling algorithms concentrate on request handling and providing services. Efficient usage of resources may be helpful in handling large number of requests. In this paper we aim to utilize the infrastructure service in efficient manner, and to obtain maximum profits to the cloud computing service provider maximum resource utilization helps in obtaining maximum profits in terms of cost for the cloud computing service provider. The proposed model gives an optimal solution by considering all the requests and categorized to a level (low, medium and high) using a prioritization parameter and serving each request which gives maximum profit and makes the resources not to be idle for long time.

4. A NEW MODEL FOR SCHEDULING REQUESTS

For effective handling of requests a new scheduling model has been proposed. Our proposed model demonstrates how the requests are allocated to the cloud resources. In the figure applications describe the number of requests to the cloud service provider. These requests may be of different types, in terms of (like CPU, memory etc.) which are sorted out based on the preference or the criteria by the cloud service provider. Priority check control [3] helps in determining the profit and to categorize in which level the request falls as show in fig 1. A mathematical equation has been used for sorting out the requests, where the equation has different parameters. Based on the number of requests we calculate the prioritization of

each single request and categorize to certain levels and submit it to cloud resources. Only under the idle situation of resources the lower profit jobs are served, which is the condition to be specified in the Service Level Agreement (SLA) between the cloud- service provider and the customer. The model would help in utilizing the maximum resources as we are considering the low prioritized jobs also.

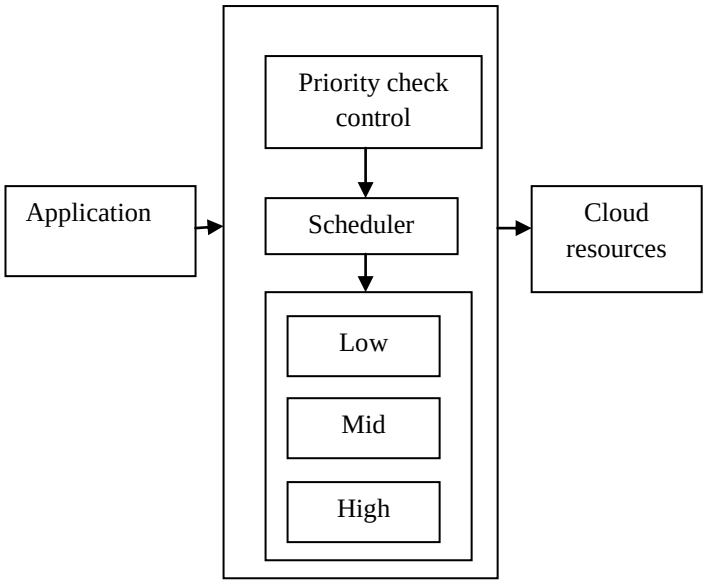

Figure 1. Job scheduler model for allocation of requests

Steps followed in request handling in this proposed approach are:

1. Calculate the profit ratios of each request.
2. These requests are categorized into certain levels
   i. High level is chosen for the job request with maximum priority.
   ii. Mid level follows the next level in priority
   iii. Low is given the last option in serving the requests
3. Submit the jobs to the cloud resources.

The allocation model helps in utilizing the resources at maximum level as we are reducing the idle state of the resources. In this scenario, we consider maximum requests from the customers and based on the prioritizations, we serve the request of each customer.

4.1 Proposed algorithm used for scheduling the requests

The proposed algorithm follows activity based costing model. Prioritization parameters are defined as follows:

(1) $R_{i,k}$: The i^{th} individual use of resources by the k^{th} task.

(2) $C_{i,k}$: The cost of the i^{th} individual use of resources by the k^{th} task.

(3) P_k: The profit earned from the k^{th} task.

(4) C_u: customer urgency

(5) L_k: The priority level of the k^{th} task.

$R_{i,k}$ is the resource used by the request for executing its instructions, $C_{i,k}$ cost of the resource used for executing the request, P_k profit earned from each request for the usage of cloud services, C_u represents the customers urgency which helps in increasing or decreasing the profits which is a key factor, it also helps in sorting the jobs. To the current parameters this could be an added advantage and L_k represents the priority level of each individual request. The priority level of each task can be calculated as in formula [1] which has been modified. The total individual resources use is supposed to be n. so the priority level of the k_{th} task is:

$$L_k = \sum_{i=0}^{n} R_{i,k} \times C_{i,k} / P_k + C_u \qquad (1)$$

This priority check controller sets the request or sort out the request and passes to the broker where the scheduling of request takes place.

Algorithm of preprocess:

For all available tasks **do**

Calculate their priority levels L_k

If Customer urgency exist

$$L_k = \sum_{i=0}^{n} R_{i,k} \times C_{i,k}/P_k + C_u$$

Else

$$L_k = \sum_{i=0}^{n} R_{i,k} \times C_{i,k}/P_k$$

End for

For every L_k **do**

Sort them and then put them into an appropriate list

End for

While the system is running do

If there is new task coming **do**

Calculate its priority and then put it into an appropriate list

End if

End while

Algorithm of process:

Do preprocess as a thread

While the system is running **do**

If every list is not empty **do**

Process the task which has the highest priority

Scan every list to modify the priority base on the restrictive conditions

End if

End while

Activity Based Costing (ABC) helps in accurate costing and urgency parameter followed in the algorithm helps in maximizing the profits to the cloud service provider.

5. IMPLEMENTATION USING CLOUDSIM TOOLKIT

Cloudsim toolkit [5] helps for simulating cloud environment, released by The Cloud Computing and Distributed Systems Laboratory, University of Melbourne. It provides the features of modeling and simulation of cloud computing environment. The behavior of cloudsim is discussed below. The diagram represents how the cloudsim works. User tries to submit his requests in the form of cloudlets. Each cloudlet has the properties of file size, number of instructions to be executed etc. These cloudlets can be considered as user's request, and these properties can be set by user of cloud service. User tries to pass these cloudlets to the broker; it follows the scheduling policy as discussed above in the algorithm and based on the approach it handles the cloudlets and serves each request.

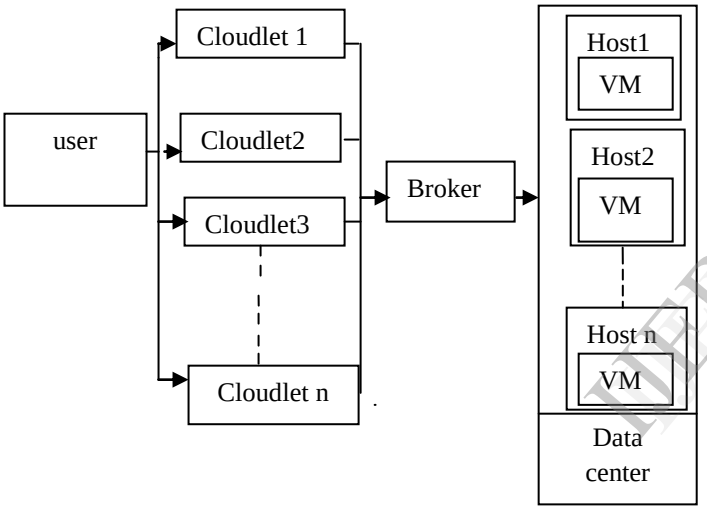

Fig 2. Behavior of Cloudsim

Cloudsim has an advantage of building of broker driven policies. VM's represents the virtual machine to be created on the hosts. Creation of hosts depends on the broker where it allocates each VM's to the different host. Datacenter has the capability to hold maximum number of hosts and the broker can dynamically change the setup of hosts and VM's.

6. EXPERIMENTAL RESULTS

Cloudsim toolkit has been used to create the simulation environment. The input to the simulation environment is considered to be, the number of Machine Instructions (MI), the average machine instructions of each task is 100 and the urgency parameter provided by the user. The resources which are Virtual Machines (VM), used are of same configurations. Comparisons are made between the algorithm with urgency factor and without urgency factor. The results are tabulated below

Table 1 .Profit gained from cloudlets with urgency without urgency

No.of Cloudlets	cost in (Rs) without urgency	with urgency
25	1425	2257
50	2925	4565
75	4120	6900
100	5100	8120

From the figure 3 it is clearly shown that as the urgency parameter is added, the profits been raised when compared with no urgency parameter.

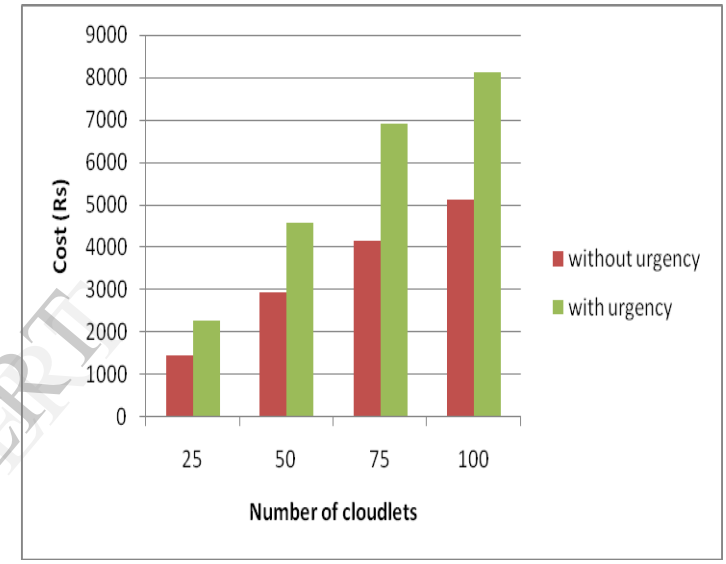

Figure 3.Comparison of algorithm with urgency and without urgency parameter

The algorithm is also compared with some traditional algorithm like First come first serve which show less efficiency compared to the proposed algorithm. The results are show in the fig 4.

Table2 .Comparison between FCFS and proposed algorithm

No.of Cloudlets	cost in (Rs) proposed algorithm	FCFS
25	2257	630
50	4565	1075
75	6900	1980
100	8120	2570

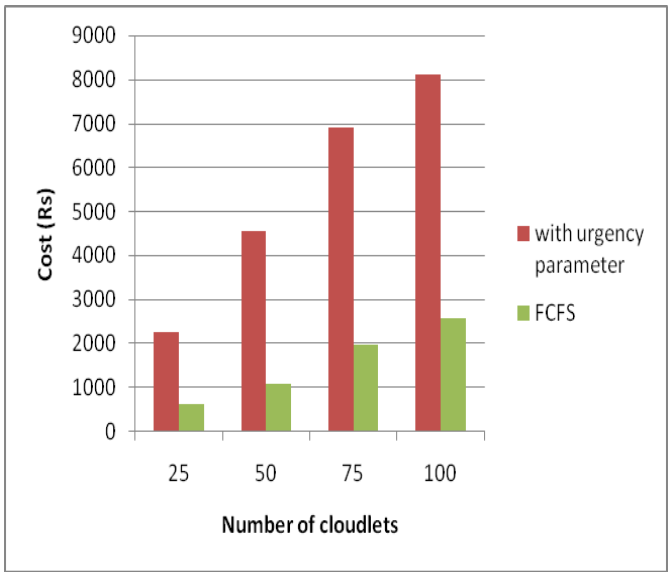

Figure4.comparison of proposed algorithm with FCFS

From the above results it is shown that urgency parameter helps in increasing the profits to the cloud service provider.

7. CONCLUSION & FUTURE WORK

This paper proposes an approach, that models the utilization of the resources efficiently and to gain the maximum profit for the cloud service provider. This work concentrates on certain parameter like customer urgency which helps in maximizing the profits to cloud computing service provider. For future work it can be extended to add few more parameters based on the requirement.

REFERENCES

[1]. Qi Cao, Zhi-Bo Wei, Wen-Mao Gong "*An Optimized Algorithm for Task Scheduling Based On Activity Based Costing in Cloud Computing* "in 3rd international conference held on bio-informatics and biomedical engineering in 2009.

[2]. Luqun Li "*An Optimistic Differentiated Service Job Scheduling System for Cloud Computing Service Users and Providers* "in 2009 Third International Conference on Multimedia and Ubiquitous Engineering.

[3]. Libra — *An Economy-Driven Cluster Scheduling and Service Level Agreements (SLA)-based Resource Allocation System* by R.Buyya.

[4].Mohsen Amini Salehi and Rajkumar Buyya "*Adapting Market-Oriented Scheduling Policies for Cloud Computing*" in International Conference on Algorithms and Architectures for Parallel Processing in 2010.

[5]. Rodrigo N. Calheiros, Rajiv Ranjan, Cesar A. F. De Rose and Rajkumar Buyya "*CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services*".

[6].YeeMingChen,Shin-YingTsai"*Optimal rovisioning of Resource in a Cloud Service*" in International Journal of Computer Science Issues, Vol. 7, Issue 6, November 2010.

[7]. Young Choon Lee , Chen Wang , Albert Y. Zomaya, and Bing Bing Zhou "*Profit-driven Service Request Scheduling in Clouds*" in 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing.