Scientific Journal of Impact Factor (SJIF): 3.134

E-ISSN (O): 2348-4470
P-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 2,Issue 12,December -2015

Load Balancing using PBTLB algorithm in Cloud Computing

Mittal Patel[1],Prof Chaita Jani[2]

[1,2]Computer engineering & Kalol institute of technology & research centre.

Abstract—*Day by day increasing Internet traffic introduces the requirement of the concept of load balancing to maximize utilization of available resources in the cloud. There are many complex calculations and concepts used to make the best use of resources and increasing performance. Out of all this complexity of the algorithm Throttle*d provide the simplest solution *for load balancing in the cloud environment. So we're going to use strangled as a basic algorithm and propose PBTLB (Priority) Based Load Balancing Throttled) an improved version of load balancing algorithm on paper.*

Keywords--*Throttled, PBTLB, Load balancing, Virtual machine balancing, Cloud Analyst.*

I. INTRODUCTION

Cloud refers to a different IT environment that is scheduled for the end remote provisioning scalable IT resources and calculated [1]. It is a type of computing where resources are shared rather than owning personal legacies or local personal servers that can be used to manage system applications. Cloud computing word cloud is used as a symbol for the Internet so we can define a cloud as Internet-based computing where different services such as storage, servers and applications are provided to teams the organization and device via Internet [2]. As compared to conventional "own use" method if you use cloud computing, the cost of purchasing and maintaining infrastructure is removed. It allows users to use resources in accordance with the coming of their needs in real time. Therefore, we can say that cloud computing allows the user to have appropriate access and sharing on-demand computing resources, such as storage, networks, applications and services, etc based on pay as a whole.

A. Cloud Computing Architecture:

Cloud computing is growing in the environment of real-time information about the cloud and the services they provide and deployment models are discussed. Figure 1 illustrates the three layers of basic services that make cloud computing. Provides three basic services that are software as a service, platform as a service and infrastructure as a service [2]. The rest of the paper is organized as follows. In section 2 we discuss virtualization cloud. In section 3, load balancing and "s will oblige. Is discussed in section 4. various load balancing techniques are discussed. In section 5 the challenges of load balancing cloud computing are explained.

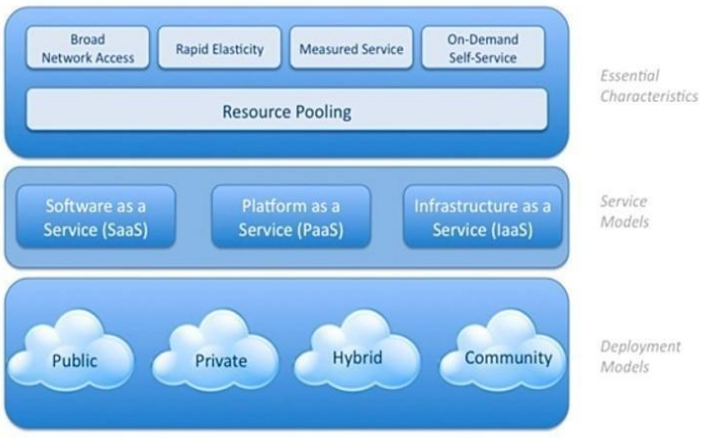


Fig. 1 : Cloud Computing Architecture [1]

II. LOAD BALANCING

Load balancing is one of the main issues related to cloud computing. The load may be a memory, CPU power, network or delay charge. It is always necessary to share the workload between the different nodes of the distributed to improve the

utilization of resources and to improve performance of the system. This may help prevent the condition where nodes are either heavily loaded or low load on the network. Load balancing is the process of ensuring the equal participation of the workload in the group node system or processor so without disturbing the running task is complete. The objectives of load balancing [7] are: -

- Improve the performance
- Maintain system stability
- make fault tolerance system
- Accommodate future modification

There are mainly two types of load balancing algorithms:

A. STATIC ALGORITHM

Static traffic algorithm is divided equally among the servers. This algorithm requires a prior knowledge of system resources, so the decision of displacement of the load is not dependent on the current status of the system. Static algorithm is correct in the system that has a low variation in the load.

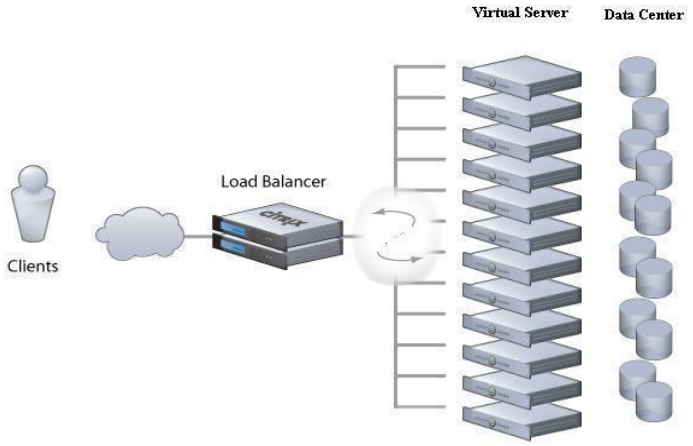


Fig. 2 **:** Load Balancing In Cloud Computing [3]

B. DYNAMIC ALGORITHM

Dynamic algorithm in the lighter server across the network or system is chosen to find and load balancing. For this communication it requires real-time network traffic that can improve the system. This current state of the system is to use the construction of decisions to handle the load.

III. LITERATURE SURVEY

A. Round Robin Algorithm:

It would use the method of interval, as its name means working in a round approach in which each node in the cloud environment is set with a slice of time and each node has to wait their turn to perform their work . In other words, use the random sampling method means that the main driver balancer arbitrarily choose to allocate the burden in case of a balancer is heavily loaded .When compared with other algorithm complexity round robin algorithm is lower. [8]

B. Throttled Load balancing

The Throttled algorithm will discover the specific node for assigning the new job. The job manager will keep a list of node detail using index list; with that it assign a particular job to particular node. If the node is ready to accept the particular job means it will accept and process otherwise it will wait for the other node requesting for processing. [8]

C. Min-Min Algorithm

It starts with a set of all unassigned tasks. First of all, least completion time for all tasks is found. Then among these least times the minimum value is selected which are the least time among the entire the tasks on any resources. Then according to that least time, the task is schedule on the related machine. Then the execution time for all other jobs is updated on that machine by adding the execution time of the assigned job to the execution times of other jobs on that machine and assigned task is eliminate from

the list of the jobs that are to be assigned to the machines. Then once more the same process is followed until all the jobs are assigned on the resources. But this approach has a major disadvantage that it can lead to starvation [10].

D. Max-Min Algorithm

Max-Min is approximately same as the min-min algorithm apart from the following: after finding out minimum execution times, the maximum value is chosen which is the maximum time along with all the jobs on any resources. After that according to that maximum time, the job is scheduled on the related machine. Then the execution time for every other jobs is updated on that machine by adding the execution time of the assigned job to the execution times of other jobs on that machine and assigned job is removed from the list of the tasks that are to be assigned to the machines[10].

E. Biased Random Sampling

M. Randles et al. [11] proposed this distributed load balancing algorithm. Load balancing can be accomplished efficiently across the nodes in this approach, by using random sampling method. Here servers are worked as nodes. In this method a virtual graph is developed representing the load on the nodes and with every in-degree directed to the particular resources to the server. While the server starts executing the job it decreases the in-degree which specifies the decrease in availability of free resources. Likewise while the server completes the job the incoming degree gets increment which in turn indicates the raise in availability of resources. This process is called random sampling. The execution begins at any node and the random neighbouring node will be choosing for the next job to be executed. Thus the load balancing technique used here is fully decentralized and select apt for many cloud networks.[12]

F. Honey Bee Foraging

This algorithm was proposed by Dhinesh B.L.D, P.V.Krishna. This algorithm was originated from the behaviour of honey bees in finding their food [13]. Among the classes of bees the hunter bees hunt for food sources. Once the food source has been found the hunter bees come back to the bee hive and announce the food source by a dance called "waggle dance". The kind of dance demonstrates the quality and quantity of the food and the distance of the source from the bee hive. This makes the survey bees to race for the food. In case of load balancing the servers are grouped into virtual servers. Each Virtual servers will has its own Virtual Server (VS) request queue. A server provide a request, compute its profit and evaluate it with the colony profit, if profit was high, then the server live at the existing virtual server and on the other hand if profit was low, then the server returns to the hunt or survey behaviour, thus balancing the load with the server.

V. PROPOSED WORK

A. Priority Based Throttled Load Balancing Algorithm.(PBTLB):

Proposed Algorithm:

1) Initialize all Request list with no entries.
2) Initialize VM status list with entries for all VM status as available.
3) Create Group of Requests with Priority.
4) Sort Requests as per Priority.
5) Data Center Controller receives a Group of new requests.
6) Allocate a request from starting of queue to load a balancer.
7) If VM state list> allocated request list,

   Then Allocate the VM.

   Else wait for the VM to get free.
8) When the VM finishes processing the request and the Data Center

   Controller receives user response; it notices the load balancer of the VM de-allocation.
9) The load balancer updates the status of VM in VMs state list and allocated request list. Balancer will take next request from Map (waiting Map) and repeat 5 through 10.

VI. EXPERIMENTAL WORK

This section describes the current status of implementation along with appropriate screen shots.

Here, I used Cloud Analyst tools for simulation results according to our requirements.

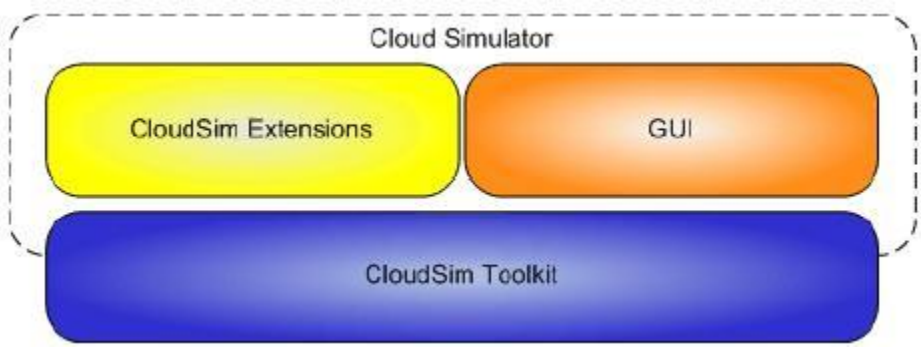


Figure 3 CloudAnalyst Architecture [12]

It is an easy to use tool with a level of visualization capability. It also enables the modeller to execute simulations repeatedly with modifications to the parameters quickly and easily. A Graphical output of the simulation results enables the results to be analysed more easily and more efficiently and accuracy of the simulation logic.
In order to analyze various load balancing policies configuration of the various components of the cloud analyst tool need to be set. We have set the parameters for the user base configuration, application deployment configuration, and data center as shown in figure 6, figure 7 and figure 8 respectively. As shown in figure the location of user bases has been defined in six different regions of the world. We have taken three data centers to handle the request of these users.one data center is located in region 0,second in region 1 and third in region 2.On DC1,DC2 and DC3 number of VM 2 are 50. In order to analyze various load balancing policies configuration of the various component of the cloud analyst tool need to be done. We have set the parameters for the user base configuration, application deployment configuration, and data center configuration as shown in figure 3, figure 4 and figure 5 respectively. As shown in figure the location of user bases has been defined in six different regions of the world. We have taken three data center to handle the request of these users. One data center is located in is located in region 0, second in 1, third in 2. On DC1, DC2 and DC3 number of VM allocated are 50. The duration of simulation is 60hrs.

Cloud analyst enables the modeler to execute the simulation repeatedly with the modifications to the parameters quickly and easily. The graphical output of the simulation results can be analyzed more easily and efficiently.

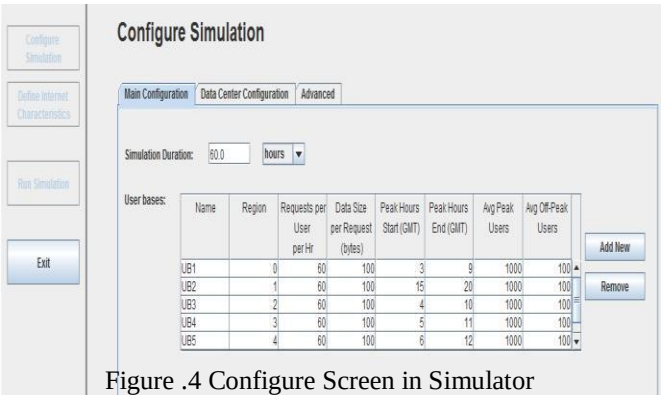


Figure .4 Configure Screen in Simulator

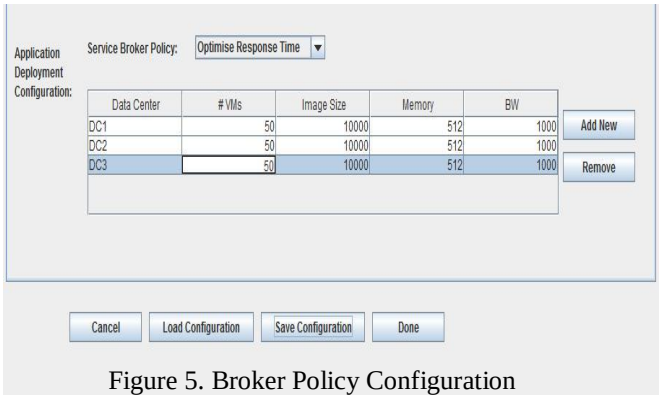


Figure 5. Broker Policy Configuration

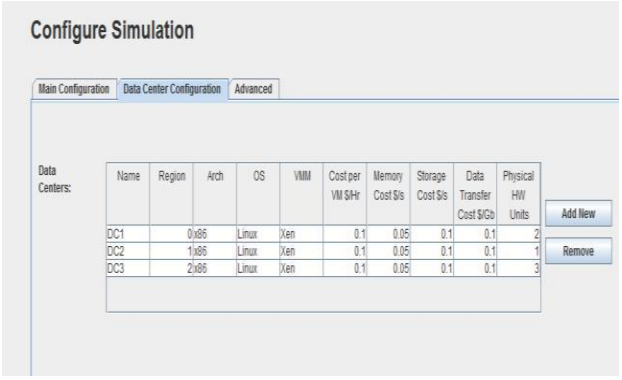
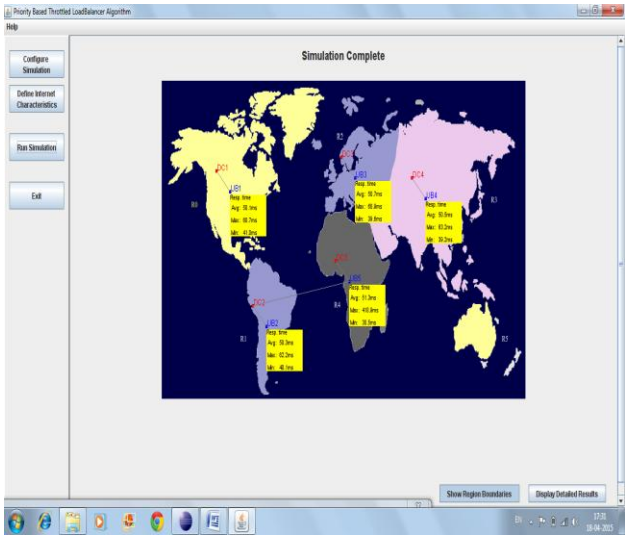


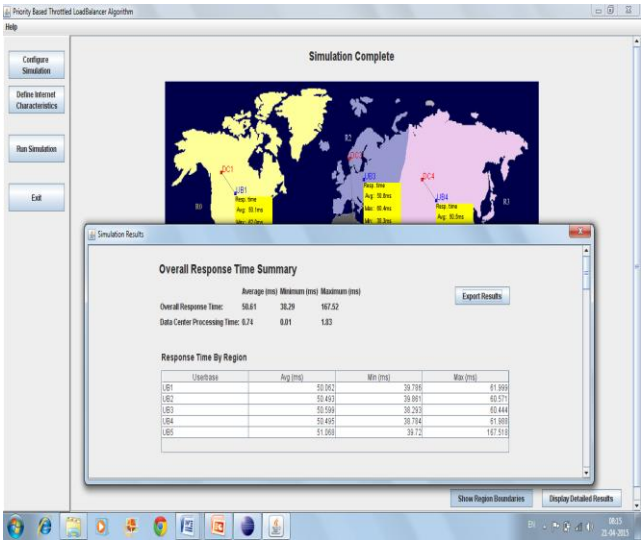
Figure 6. Data centre Configuration

After performing the simulation the result computed by cloud analyst is as shown in the following figures. We have used the above defined configuration for each load balancing policy one by one and depending on that the result calculated for the metrics like response time, request processing time and cost in fulfilling the request has been shown in Figures 5,6,7,8.

A. Response Time:

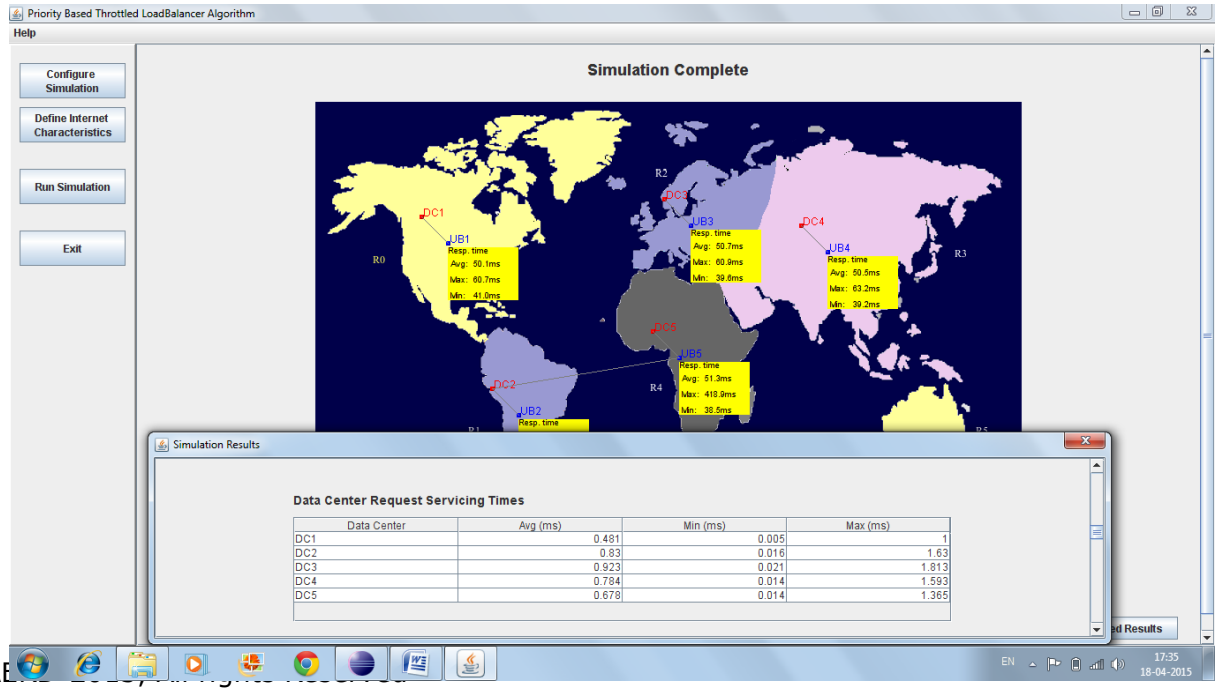
Response time for each user base and overall response time calculated by the cloud analyst for each loading policy has been shown in the figure 9, 10 ,11 and 12 respectively.





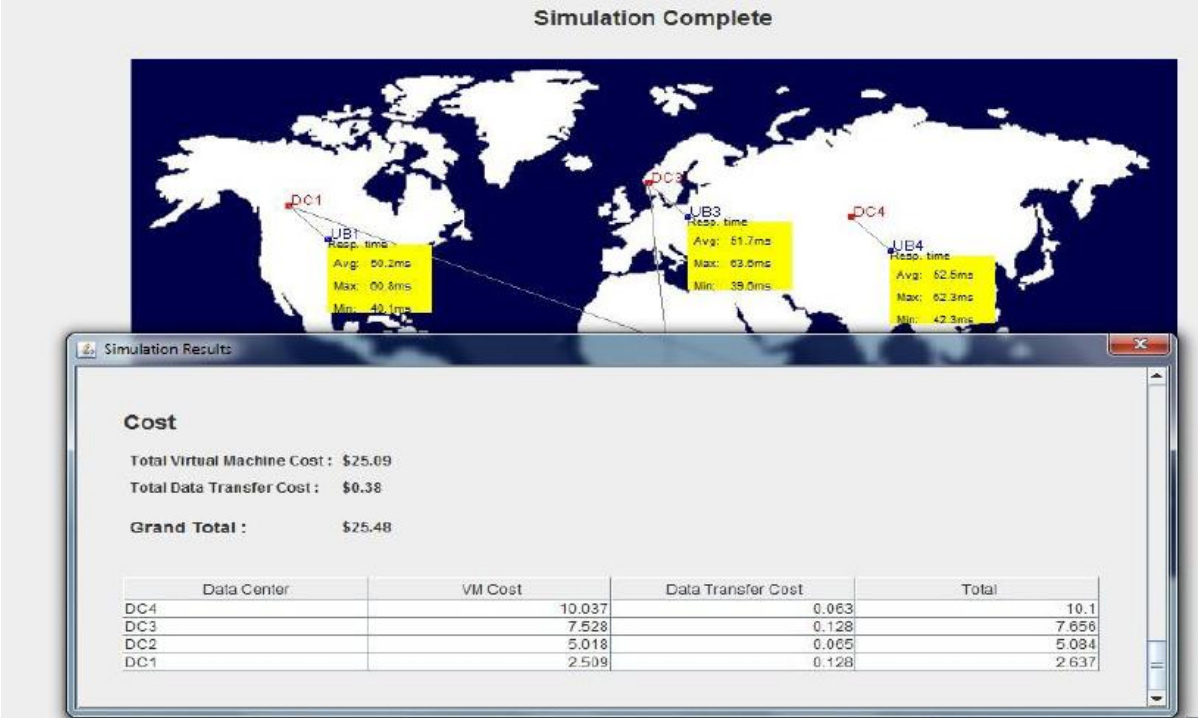
A. Data Center Request Servicing Time:

Data center request servicing time for each data center calculated by the cloud analyst for each loading policy has been shown in the figure.



B. **Processing Cost:**

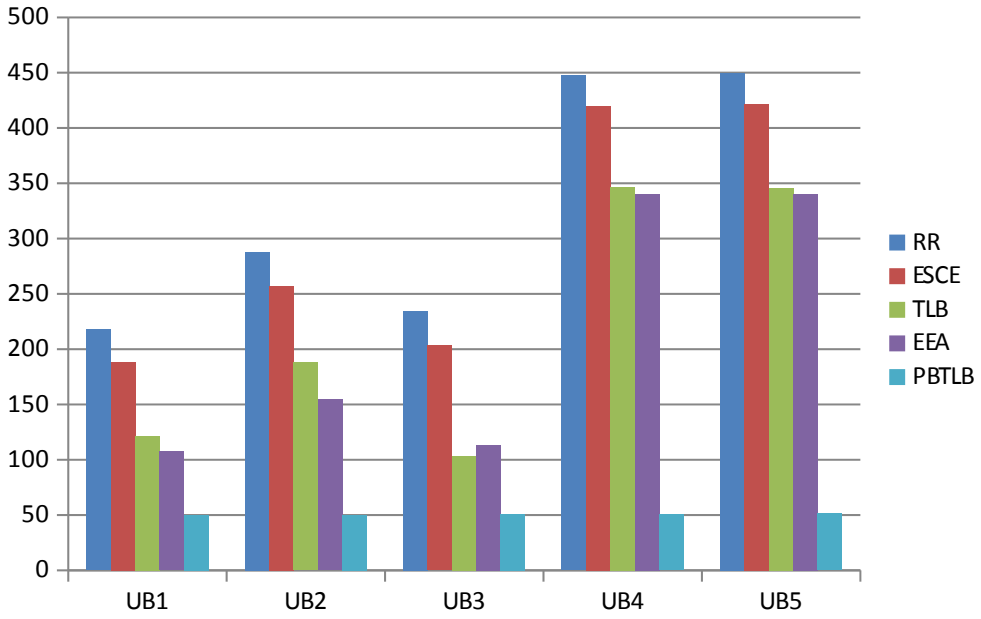
The processing cost for each load balancing policy computed by the cloud analyst as can be seen from the figures



Comparison with other algorithms:

- Comparison among load balancing policies(average response time):

UBs	**RR[9]**	**ESCE[9]**	**TLB[9]**	**EEA[9]**	**PBTLB**
UB1(1000)	217.847	188.268	121.645	107.437	50.103
UB2(2000)	287.676	256.852	187.827	154.423	50.299
UB3(3000)	234.324	203.323	103.607	113.178	50.681
UB4(4000)	447.466	420.094	346.787	340.524	50.477
UB5(5000)	449.673	421.572	345.466	340.498	51.272

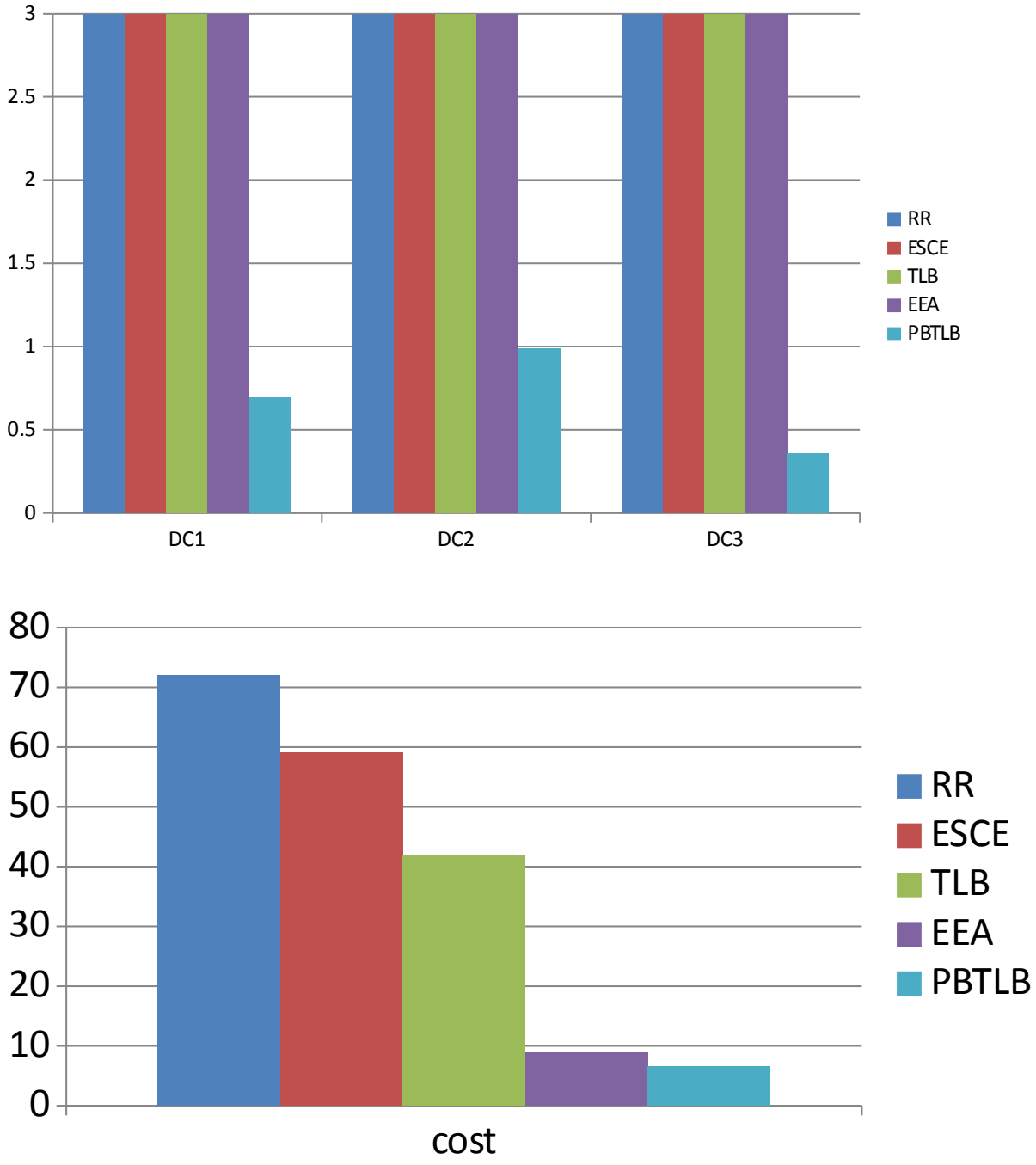


[comparison among different algorithm with averege response time]

- Comparison among load balancing policies(Average data center request servicing time):

DCs	**RR[9]**	**ESCE[9]**	**TLB[9]**	**EEA[9]**	**PBTLB**
121.655	68.436	121.655	68.436	55.335	0.697
DC2	218.021	182.407	124.64	90.149	0.988
DC3	141.005	116.897	42.564	46.672	0.358

[comparison among different algorithm with average data center request servicing time]



[Comparison among different algorithm with data transfer cost]

VII. CONCLUSION AND FUTURE SCOPE:

As cloud computing is new research area for research and development load balancing will be important challenge for service provider. For this various methods are used to solve load balancing in cloud system. In existing method the average response time, processing time and cost are more and tasks have to wait more time for some resource allocation. In existing method job allocation is based on prior knowledge of the system .So I can establish new algorithm DWARR algorithm which is based on Throttled, but some modification on Throttled algorithm. PBTLB algorithm overcomes the problem of existing algorithm

like reduce the average response time, processing time and cost also. Tasks don‟t have to wait more time as compare to existing algorithm. Thus, system throughput increase and also resource utilize efficiently.

REFRENCES

[1] Basic concept and terminology of cloud computing- http://whatiscloud.com
[2] L. Wang, J. Tao, M. Kunze,"Scientific Cloud Computing: Early Definition and Experience", the 10th IEEE International Conference Computing and Communications 2008.
[3] Sotomayor, B., RS. Montero, IM. Llorente, and I. Foster, "Virtual infrastructure management in private and hybrid clouds," in IEEE Internet Computing, Vol. 13, No. 5, pp: 14-22, 2009.
[4] D.Escalnte,Andrew J.Korty,"Cloud Services:Policy and Assessment",Educause review July/August 2011.
[5] RIMA P. LINGAWAR , MILIND V. SRODE ,MANGESH M. GHONGE ,"Survey on Load-Balancing Techniques in Cloud Computing ",International Journal of Advent Research in Computer & Electronics, Vol.1, No.3, E-ISSN: 2348-5523, May 2014.
[6] P.Warstein,H.Situ and Z.Huang(2010),"Load balancing in a cluster computer" In preceeding of the seventh International Conference on Parallel and Distributed Computing,Applications and Technologies,IEEE.
[7] T.Kokilavani J.J. Collage of Engineering and Technology and Research Scholar,Bharathiar University,Tamilnadu,India "Load Balanced Min-Min algorithm for Static Meta Task Scheduling in Grid Computing"International Journal of Computer Applications(0975-8887)Volume 20-No.2,April 2011.
[8] Randles M., Lamb D. and Taleb-Bendiab A.,"A Comparative Study into Distributed Load Balancing Algorithms for Cloud Computing ",(2010) 24th International Conference on Advanced Information Networking and Applications Workshops, 551-556
[9] O. Abu- Rahmeh, P. Johnson and A. Taleb-Bendiab,"A Dynamic Biased Random Sampling Scheme or Scalable and Reliable Grid Networks", INFOCOMP - Journal of Computer Science, ISSN 1807-4545, 2008, VOL.7, N.4, December, 2008, pp. 01-10.
[10] Randles M., Lamb D. and Taleb-Bendiab A. (2010) 24th International Conference on Advanced Information Networking and Applications Workshops, 551-556.

