**Performance Analysis of Load Balancing Algorithms in Cloud Computing**

**Author:Tanisha Agrawal**

**ABSTRACT**

Cloud computing is a business oriented IT-technology, which is composed of multiple computing technologies accessed via internet. With the rapid increase in cloud usage, it becomes a challenge to deliver the cloud services effectively and efficiently to the cloud consumers on the pay-per usage basis. In this concern Balancing of load has become one of the essential components for the cloud computing environment to perform the effective operations. Scheduling of virtual machines or data centers has to be done properly by using an appropriate load balancing technique. Hence, several algorithms have been developed to process the client's request towards the cloud nodes.. In this present work, a hybridized swarm intelligence technique is proposed to evenly distribute the incoming task requests among the virtual machines or server. Additionally, the performance analysis has been performed using the CloudAnalyst simulator. This paper gives a comprehensive performance analysis of the proposed approach and compares its results with existing Round Robin (RR), Equally Spread Current Execution (ESCE) and ant colony optimization (ACO) techniques. Simulation results have demonstrated that the proposed technique shows a significant outcome in terms of response time, data center processing time and total cost in cloud computing.

**General Terms**

Algorithms Used: Priority based Bee Colony for Scheduling, For Load Balancing Ant Colony Optimization.

**Keywords**

Load Balancing, Cloud Computing, Priority based Bee Colony, ACO.

**1. INTRODUCTION**

Cloud Computing is an IT- Service, where infrastructure, software applications and platform is served over the internet on request, according to the requirement of the users at a particular time. This term is basically used in the case of the internet. The fundamental need of cloud computing is to share and provide the computational resources viz: virtual machines (VMs) as services on demand. Allocating a better VM on user’s request is being performed using the load balancing techniques in the cloud computing network. Since the load balancing algorithm plays a vital role while deciding which VM is to be alloted on demand to the client. It is necessary to facilitate and achieve the potentials of computational clouds using an effective scheduling system as to minimize

the job execution time. In cloud computing, the problem of scheduling of jobs is an NP-complete. As the number of users associated with the computing increases, the job requests to be scheduled are correspondingly increases, however the existing strategies of task scheduling cannot fulfil its requirements. For such reasons, there is a need of better algorithms for task scheduling for minimizing the computational time and the total processing cost related to that computing. A prominent task scheduling algorithm directly influences the entire cloud system. One of the best examples is a swarm intelligent technique, i.e. bee colony optimization which has been used in this paper for the scheduling purpose. Additionally, load balancing and resource management are the key areas of research in a Cloud Computing environment that is used to evenly dispense the workload on each cloud node, maximizing the resource utilization rate and minimizing the task execution time [1]. In cloud and grid computing, load balancing is a great challenge for dynamic, heterogeneous and complex grid and cloud computing environment. Ant Colony technique is an emerging methodology used for managing, job scheduling and load balancing. Hence, an effective use of this algorithm will lead to maximum usage of resources available, thus improve the throughput and the overall system performance [2] [3]. The main objective of load balancing technique is to uniformly distribute the load among the nodes to optimize the service time of the resources and the response time of the application. Thus ACO technique has been used for load balancing in [4].

The specific contributions of our work involve:

⮚ The novel hybrid approach of load balancing and scheduling of jobs in a cloud computing environment influenced by the foraging behavior of ant colony and honey bee.

⮚ A literature review about several existing load balancing techniques along with their pros and cons.

⮚ Systematic study of the proposed hybrid algorithm using a well defined flowchart depicting its behavioral control structures that how ants and bees foraging behavior inspire scheduling and load balancing for distributed cloud computing system.

⮚ Evaluation and performance analysis, of hybrid technique with respect to other existing load balancing algorithms.

The remaining paper is arranged as: The related works are discussed in Section 2; in Section 3 we discuss the problem formulation. Section 4 provides the brief description of the existing load balancing techniques, Section 5 provides proposed methodology, Section 6 deals with simulation

configuration, Section 7 depicts the performance analysis of the algorithms and their comparison. Section 8 will conclude and offers the future work. Section 9 provides the acknowledgement. Section 10 gives the references.

**2. RELATED WORKS**

In this section, we summarize the load balancing algorithms used in the cloud computing environment in a nutshell. The main focus is on the assignment of all incoming jobs among the available virtual machines with minimal response time. Load balancing is defined as a process of making effective utilization of computational resources by sharing the total workload among the individual nodes of the cloud setup and thereby minimizing the response time of the task.

It has been expressed in [5] that there are many examples of evolutionary algorithms, where genetic algorithm is one of them that is used for scheduling in a network. These algorithms may comprise a memory to retain the last status that helps in reducing the no of agents close to locations in optimal solutions that have been discovered earlier. But unlike swarm intelligence techniques, genetic algorithm don’t serve a large number of service consumers. Additionally, the authors in [6] depict that evolutionary algorithms are slower in nature for finding the optimal solutions since there is a need of handling the population movements. The authors in [7] designed algorithms for statically balancing the load on trees, considering that the total load is fixed. The paper [8] proposes an efficient algorithm for data migration in dynamic load balancing by calculating the Lagrange multiplier associated with the Euclidean form of transferred weight. This work can minimize the data movement in homogenous environments in an efficient manner, but it does not support the distributed heterogeneous environments. A novel distributed load balancing technique has been proposed in [9]. The major benefits of the proposed load balancing technique are the distributed structure, less complexity and optimal allocation of virtual machines for each user request. In the studies of [10], VMware based load balancing approach has been offered in order to generate the ants at hop level whenever required, and concurrently the mobile agents memorize every visited node and record their whole information for future reference. A Particle Swarm Optimization (PSO) technique for cloud computing has been developed in [11] which is being inspired by the movement of birds in flocks or fishes in school. In PSO search network, each solution is considered as “particle”. Each iteration updates the particle by considering the two “best” values i.e pbest and gbest and these two values are need not be evaluated in ACO algorithm. Unlike PSO technique, ACO reaches with guaranteed convergence to the optimal solution*.* Moreover ACO is more applicable for problems that require crisp results and PSO is applicable to problems that are fuzzy in nature. The authors in [12] proposed task load balancingbased on foraging nature of bees. In this work a bee nature based algorithm is designed i.e. (HBB-LB) to handle the priority of jobs and reducing the processing time. The proposed technique is more effective and serve less execution time and waiting time as compared to existing load balancing algorithm. However this approach can be further improved by considering the QoS factors.

From the related work it was observed that in most of the studies unlike ACO many algorithms like Genetic algorithm don’t serve a large number of clients , and don’t create a large number of VMs which the ant colony optimization technique can solve this issue. It has also been noticed that in most of the load balancing algorithms like throttled and round robin most of the time processor remains idle and performance degrades as no. of serve. Further, the two non-preemptive scheduling techniques like First-In-First-Out (FIFO) and RR Policy may lead to load variations between the VMs which results in reducing makespan and flowtime. Thus, such situations may

lead to the development of dynamic load balancing algorithms.

**3. PROBLEM FORMULATION**

The issue of load balancing occurs when the cloud clients try to access and send the request to the same cloud server while other cloud servers don’t receive the service request from the cloud clients which leads to the unbalanced workload on the cloud data centers. Therefore, it causes the development of numerous algorithms for scheduling and load balancing. But unlike swarm intelligent algorithm, one of the classes of evolutionary strategies like genetic algorithm does not support multiple users at an instant of time. Many authors have just focused on the availability of nodes and only few factors are taken into consideration like node’s memory, processing capacity, etc. Thus we added the some more factors like virtual machine bandwidth, virtual machine computing capacity which is being calculated in respect of millions of instructions per second, no of processors in a virtual machine and image size of a virtual machine therefore all these factors will easily provide the fittest resource for the job to be processed in a cloud. In some research papers of ABC, FCFS priority concept has been considered that could increase the response time. Hence Shortest Job Criteria has been considered to minimize the average response time of the cloudlets. In our approach the antnet concept of forward and backward movement is also taken into account for distributing the workload on the nodes, whereas ABC is applied for searching the optimal path towards the best suitable resource in the cloud network.

**4. EXSISTING LOAD BALANCING ALGORITHMS**

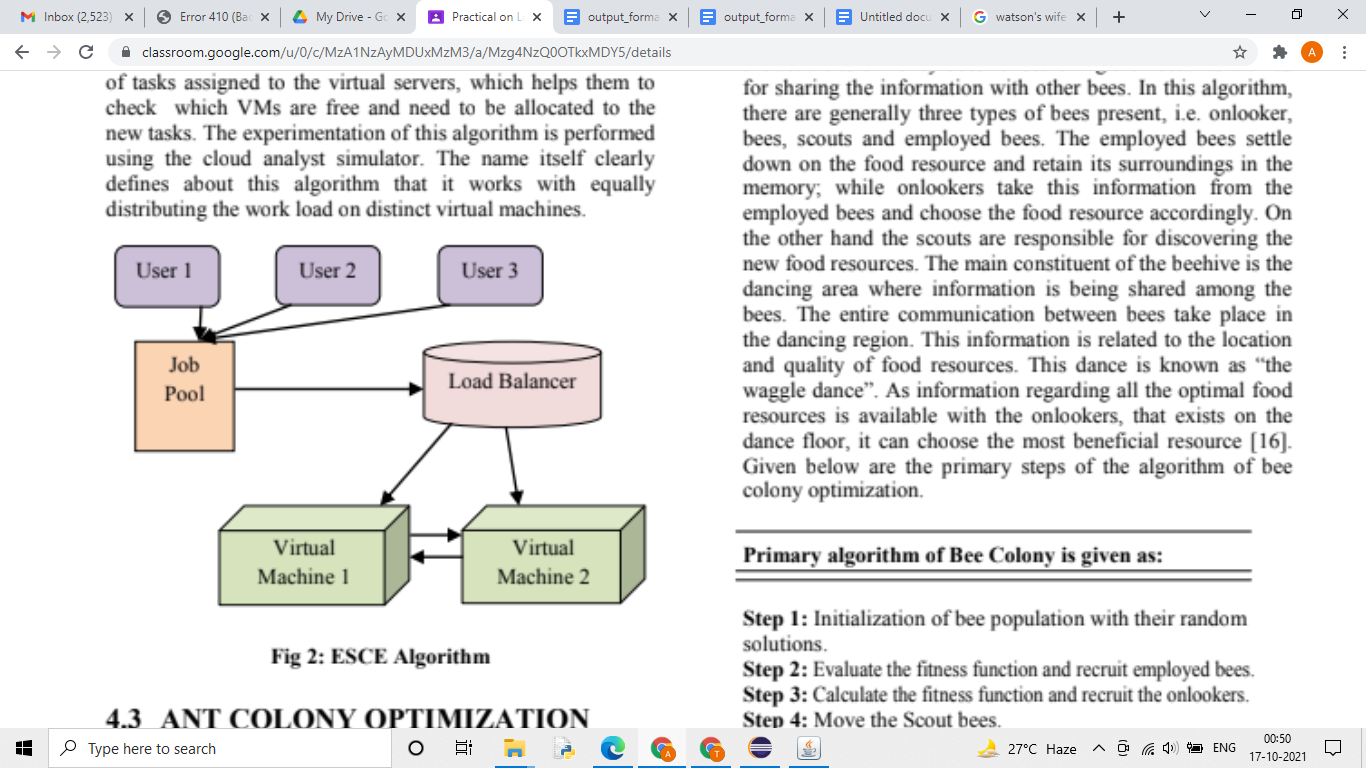
**4.1 Round Robin (RR) Algorithm**

The round robin algorithm in the cloud computing is quite similar as the round robin scheduling performed in the process scheduling. This algorithm performs on basis of random choice of the VMs. The datacenter controller (DCC) allots the service calls to a pool of VMs in a cyclic manner. The initial client request is assigned to a randomly selected VM from the group of VMs and then the DCC allots the requests in a round manner. Once the VM is allocated, it is moved to the bottom of the pool of VMs [13]. The major drawback of this type of allocation of tasks is this that it donot favor the advanced load balancing requisites viz: response time for each individual service request and processing time and if the VM is not free then incoming job should wait in the queue.

**Fig 1: Round Robin Policy** 

**4.2 Equally Spread Current Execution Load (ESCE)**

The algorithm of ESCE needs a load balancer that tracks the jobs which are asked for execution. The main requirement of load balancer is to put the tasks in the job pool and assign them to the distinct VMs [14]. The balancer keeps monitoring the job queue frequently for new tasks and then assign them to the pool of free VMs. The load balancer also handles the list of tasks assigned to the virtual servers, which helps them to check which VMs are free and need to be allocated to the new tasks. The experimentation of this algorithm is performed using the cloud analyst simulator. The name itself clearly defines about this algorithm that it works with equally distributing the work load on distinct virtual machines.



**Fig 2: ESCE Algorithm**

**4.3 THRESHOLD ALGORITHM**

According to this algorithm, the processes are assigned immediately upon creation to hosts. Hosts for new processes are selected locally without sending remote messages. Each processor keeps a private copy of the system’s load. The load of a processor can characterize by one of the three levels: underloaded, medium and overloaded. Two threshold parameters tunder and tupper can be used to describe these levels.

Under loaded - load < tunder Medium - tunder ≤ load ≤ tupper Overloaded - load > tupper

Initially, all the processors are considered to be under loaded. When the load state of a processor exceeds a load level limit, then it sends messages regarding the new load state to all remote processors, regularly updating them as to the actual load state of the entire system. If the local state is not overloaded then the process is allocated locally. Otherwise, a remote under loaded processor is selected, and if no such host exists, the process is also allocated locally. Thresholds algorithm have low inter process communication and a large number of local process allocations. The later decreases the overhead of remote process allocations and the overhead of remote memory accesses, which leads to improvement in performance. A disadvantage of the algorithm is that all processes are allocated locally when all remote processors are overloaded. A load on one overloaded processor can be much higher than on other overloaded processors, causing significant disturbance in load balancing, and increasing the execution time of an application.

# BEE COLONY OPTIMIZATION

The algorithm of artificial bee colony (ABC) is confined to the activities of honey bees for searching the nectar as well as for sharing the information with other bees. In this algorithm, there are generally three types of bees present, i.e. onlooker, bees, scouts and employed bees. The employed bees settle down on the food resource and retain its surroundings in the memory; while onlookers take this information from the employed bees and choose the food resource accordingly. On the other hand the scouts are responsible for discovering the new food resources. The main constituent of the beehive is the dancing area where information is being shared among the bees. The entire communication between bees take place in the dancing region. This information is related to the location and quality of food resources. This dance is known as “the waggle dance”. As information regarding all the optimal food resources is available with the onlookers, that exists on the dance floor, it can choose the most beneficial resource [16]. Given below are the primary steps of the algorithm of bee colony optimization.



### Primary algorithm of Bee Colony is given as:



**Step 1:** Initialization of bee population with their random solutions.

**Step 2:** Evaluate the fitness function and recruit employed bees.

**Step 3:** Calculate the fitness function and recruit the onlookers.

**Step 4:** Move the Scout bees.

**Step 5:** Evaluate the optimal solution.

**Step 6:** Check stopping condition, if met, then end the execution, otherwise repeat from 2nd step.

**6. SIMULATION CONFIGURATION**

The simulation and performance analysis has been performed using the cloud analyst toolkit by using given specified configuration. It models the cloud data centers, cloud service brokers and resource scheduling policies in [17] but it became necessary to own an easy to access tool with a user friendly GUI

that lead to Cloud Analyst tool. The load balancing technique is used by the VmLoadBalancer component of the cloud analyst toolkit. Cloud analyst simulator removes all the programming complexities by designing a user friendly GUI. It allows the user to do parameter sweep experiments. The cloud analyst framework allows setting the regions of cloud based data centers and usersbases. Several parameters can be configured viz: no of user bases, no of requests generated per user per hour, no of VMs, no of cpus, storage amount, bandwidth of the network and other significant parameters as shown in Table 1.

**Table 1: Paramter Settings in Simulation**

| **S.no** | **Parameter** | **Value** |
| --- | --- | --- |
| 1 | VM-Memory | 512 MB |
| 2 | Data Center OS | Linux |
| 3 | Data Center- VMM | Xen |
| 4 | Data Center Architecture | x86 |

On the basis of these parameters, cloud analyst evaluates the simulation and shows its results in a graphical format. Following are the statistical metrics derived as the output of the simulation in the initial version of the simulator:

* Overall Response Time of the system
* Total Data Center Processing Time
* Overall Processing Cost (Sum of Total virtual Machine Cost and Total Data Transfer Cost).

# 7.1 Response Time

It is defined as the time interval between the request sent and the response received by the cloud user/ consumer. Overall response time is calculated by the tool for each load balancing policy

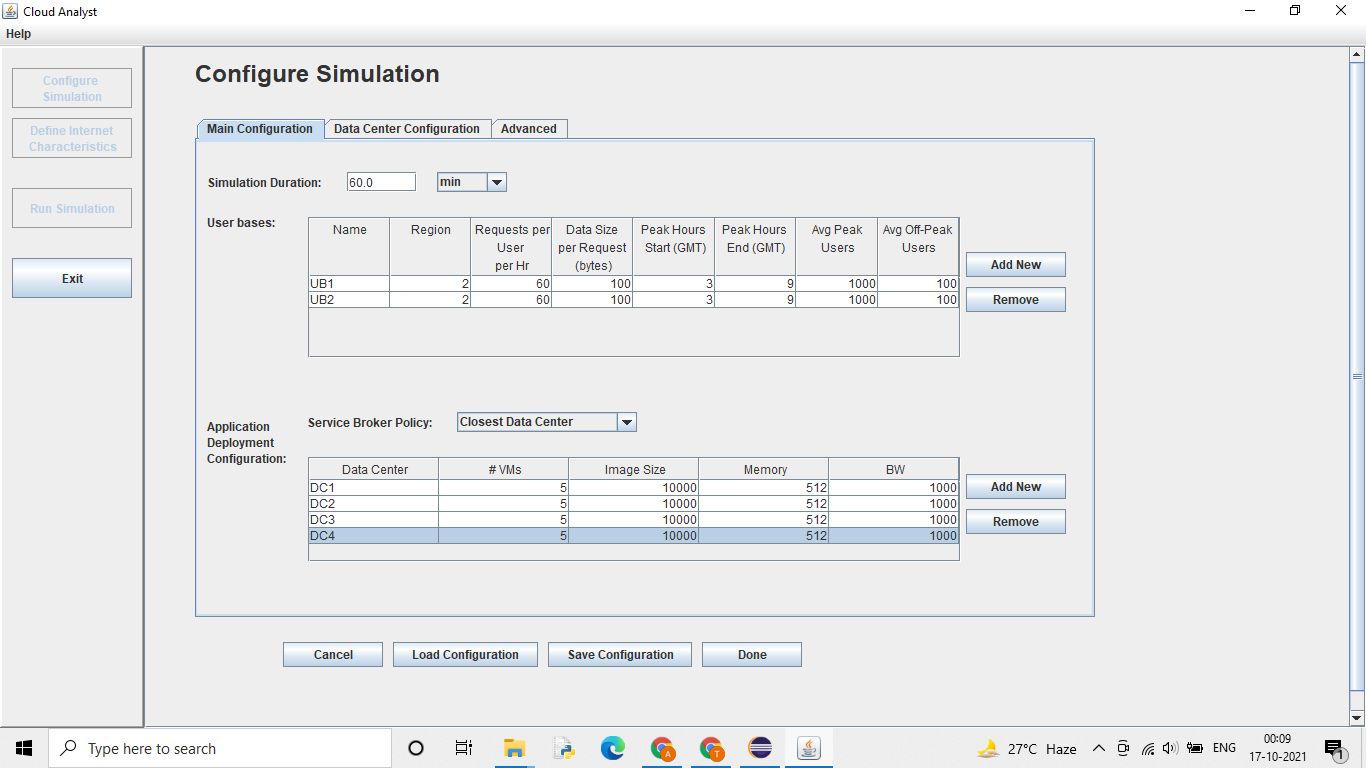
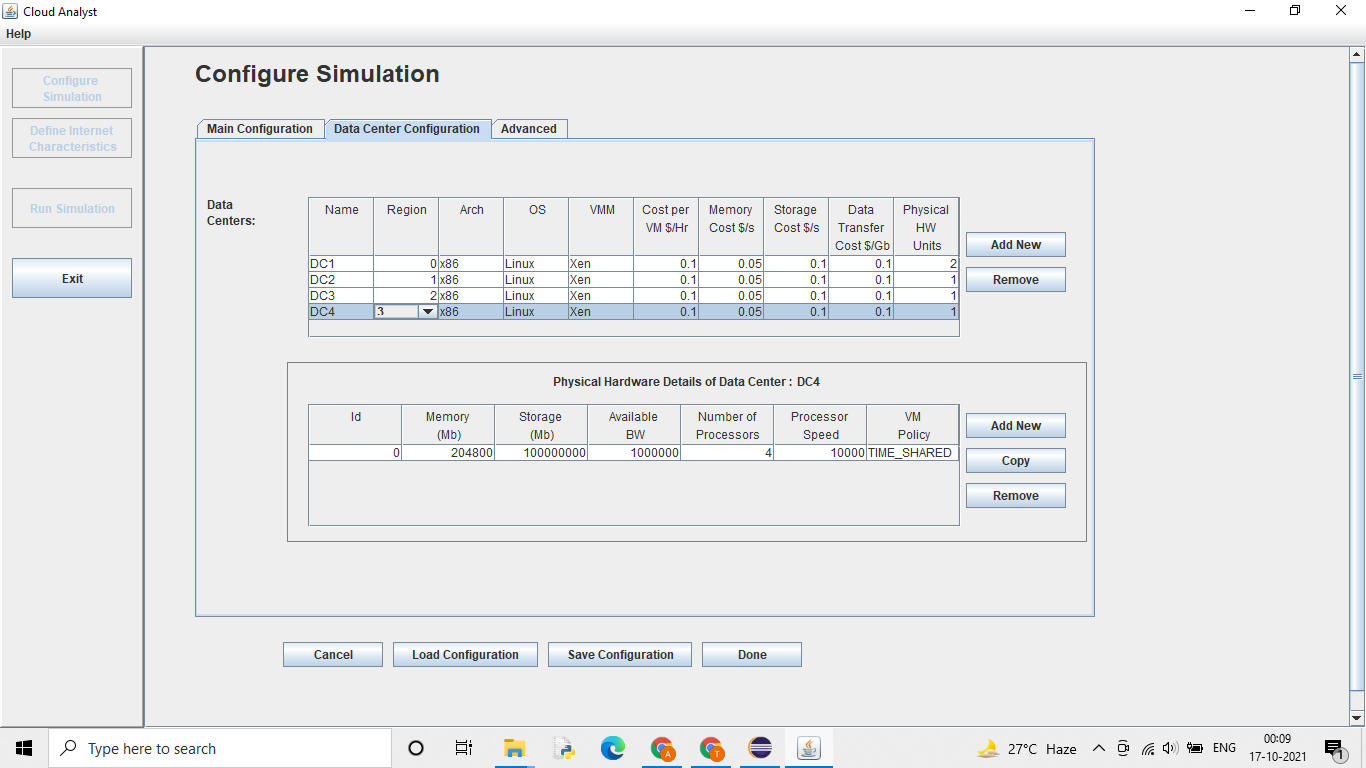
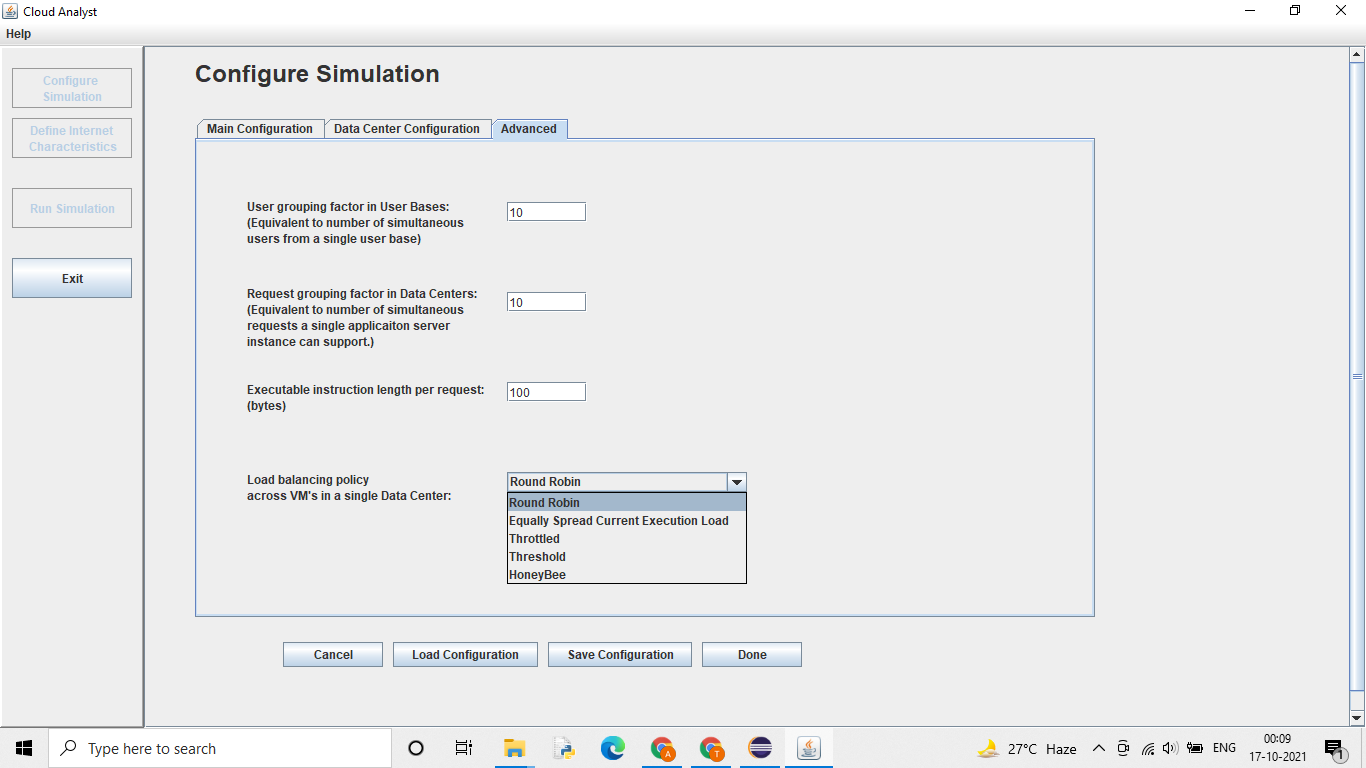
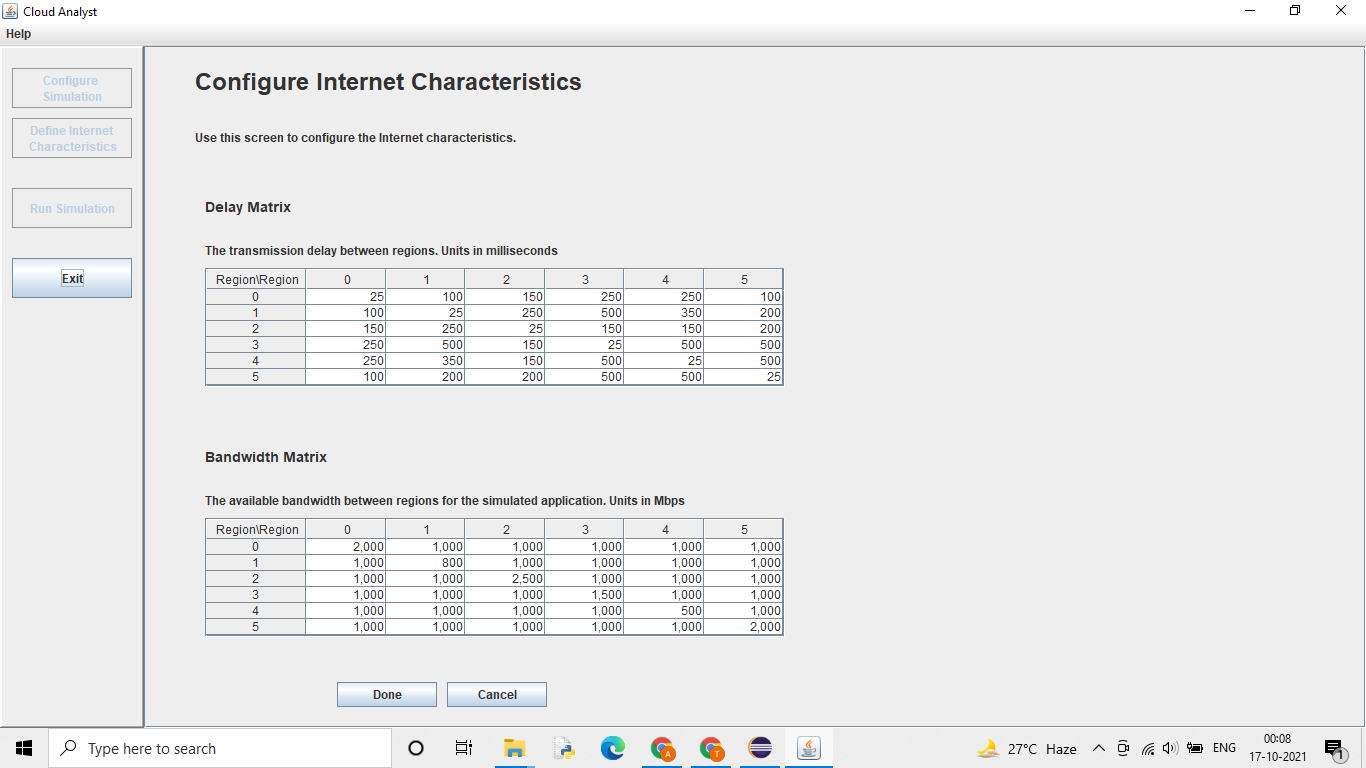
# 7.2Data Center Processing Time

For each load balancing algorithm Data Center Processing Time has been calculated by the tool

**7.3Total Cost**

The total processing cost for each load balancing technique is computed by the cloud analyst simulator

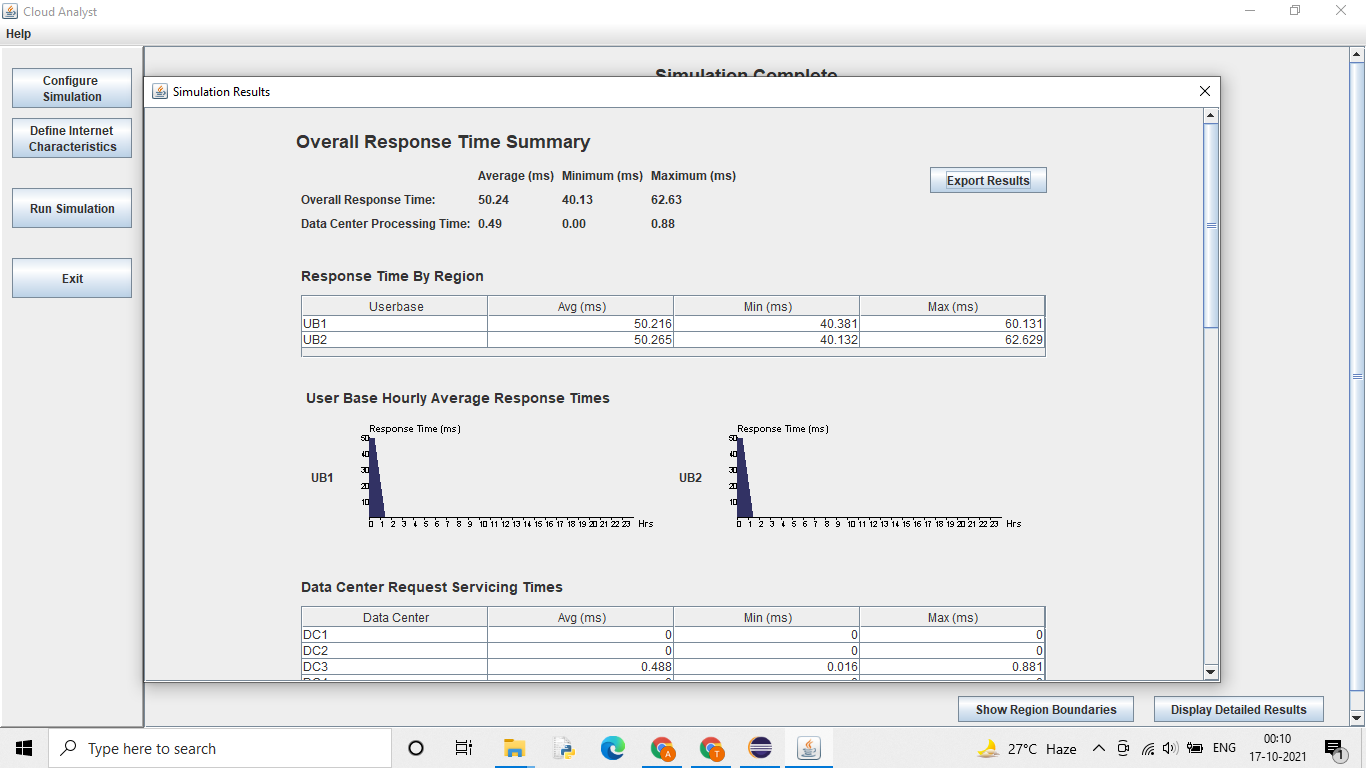
**Configuration Images**



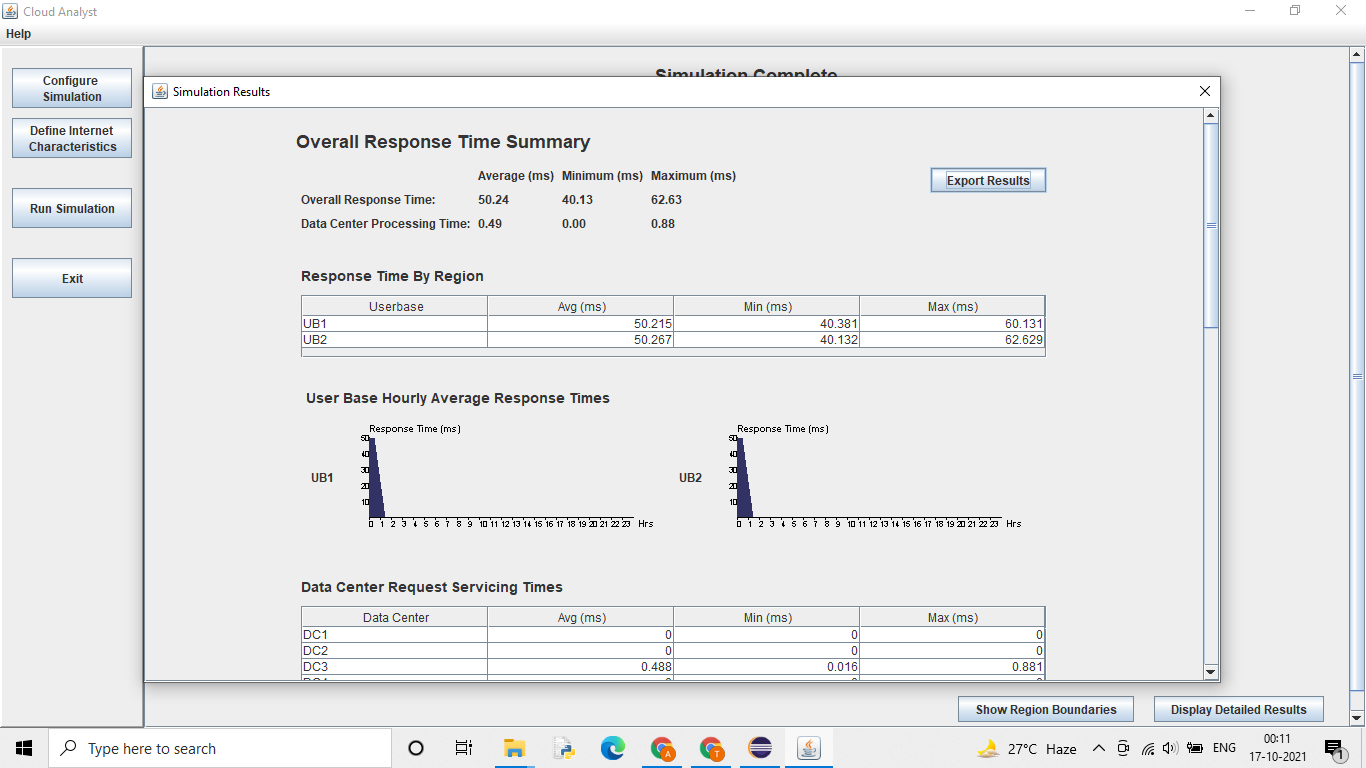
**9. ANALYSIS**

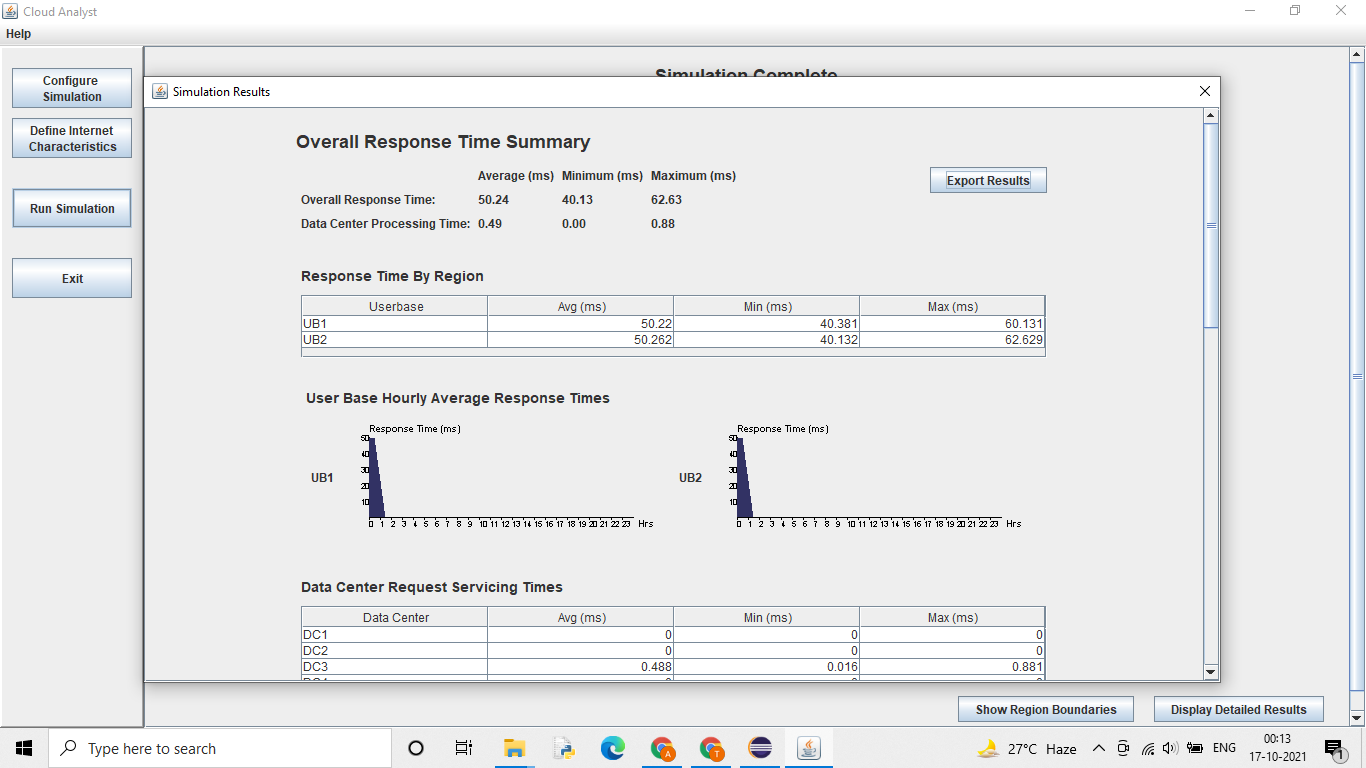
**[1]Datacenters with different location**

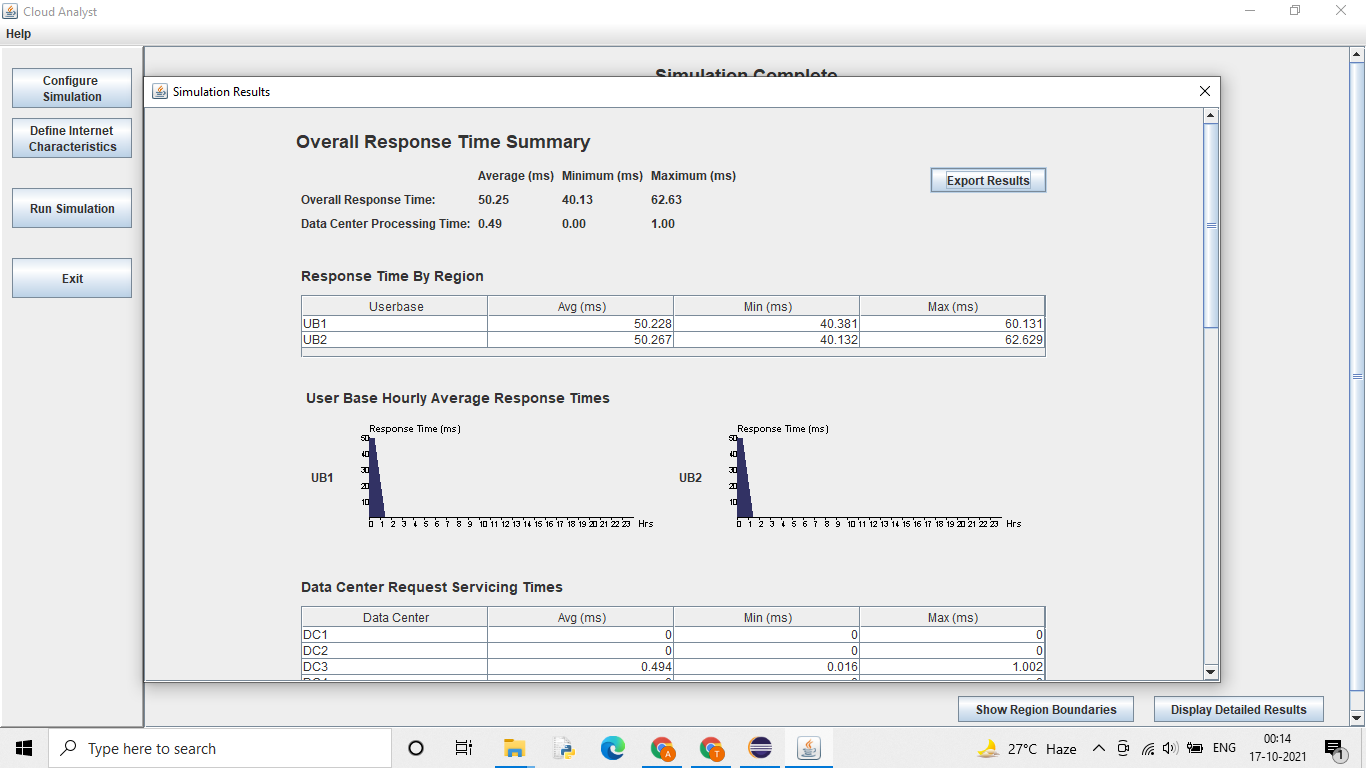
1. Round Robin



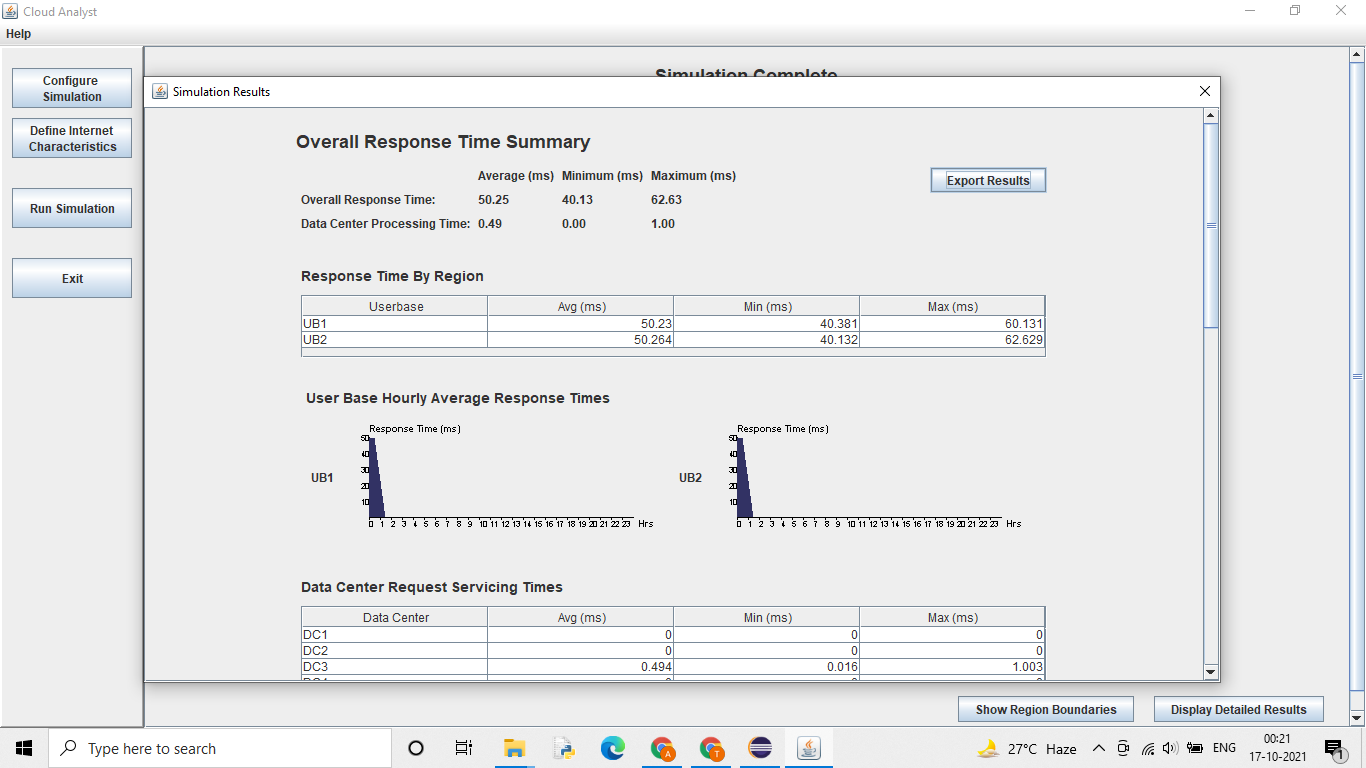
1. Equally Spread

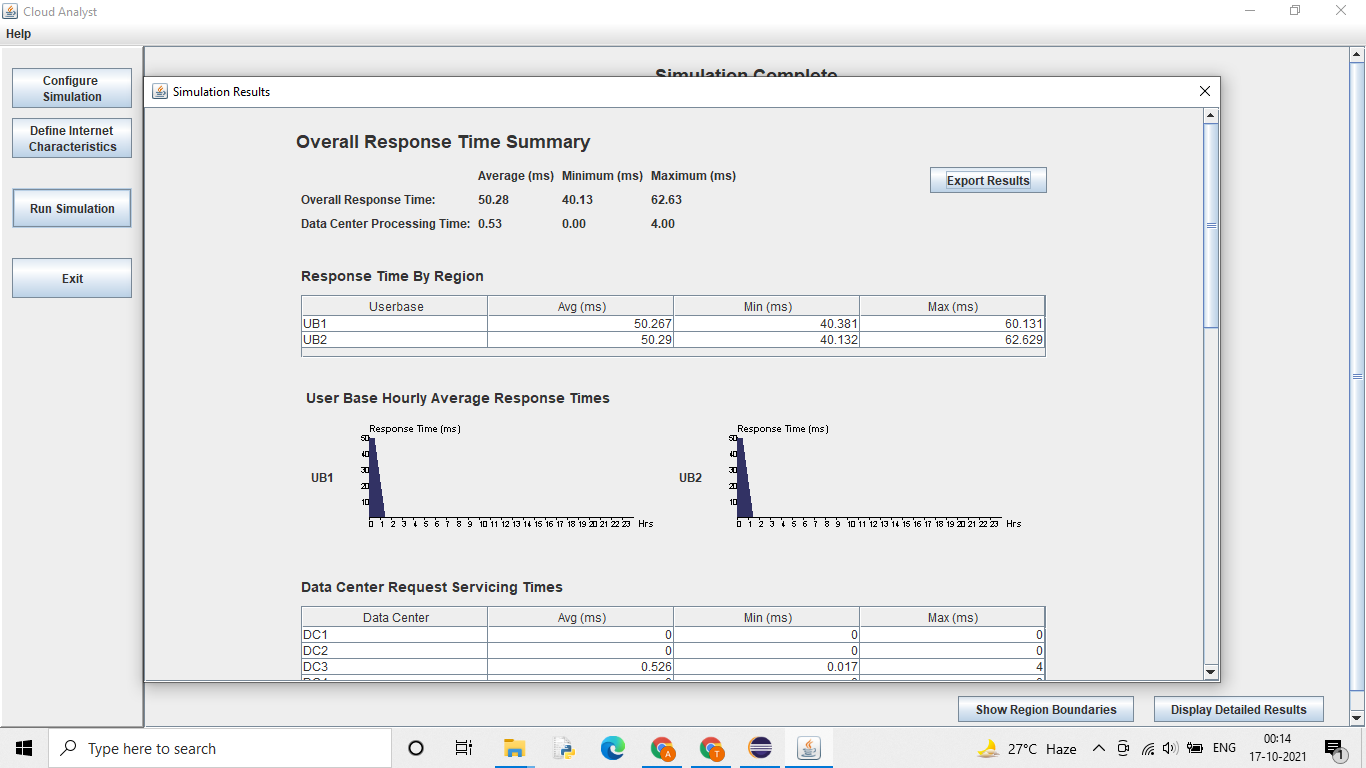




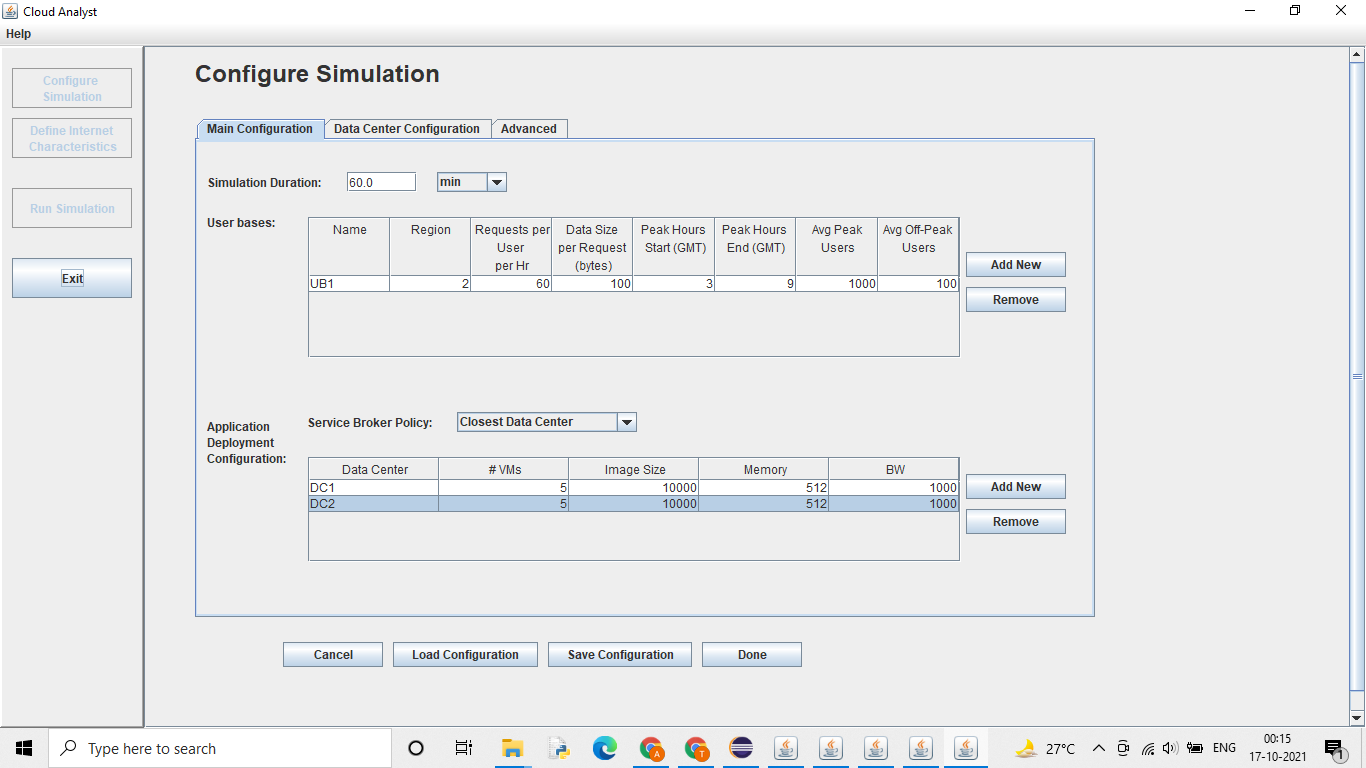
3.Throttled

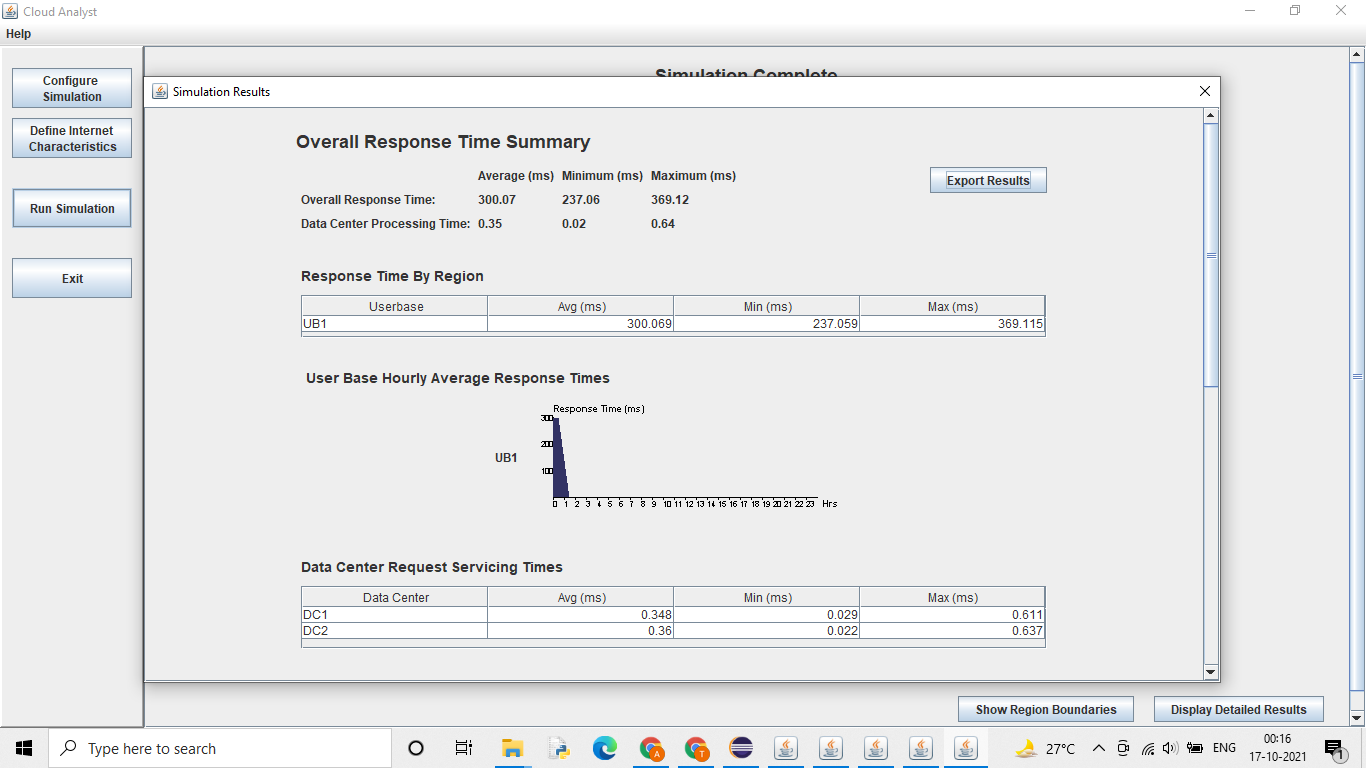
4.Threshold

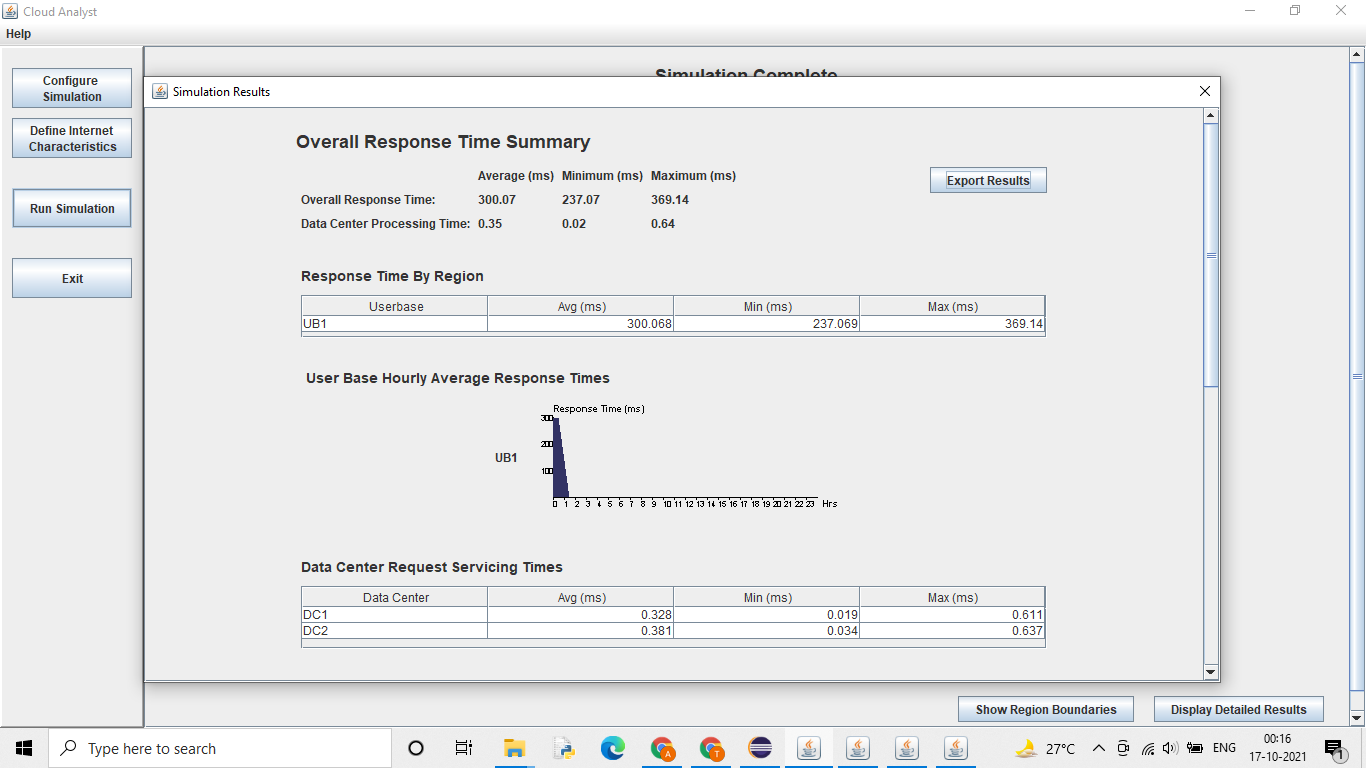


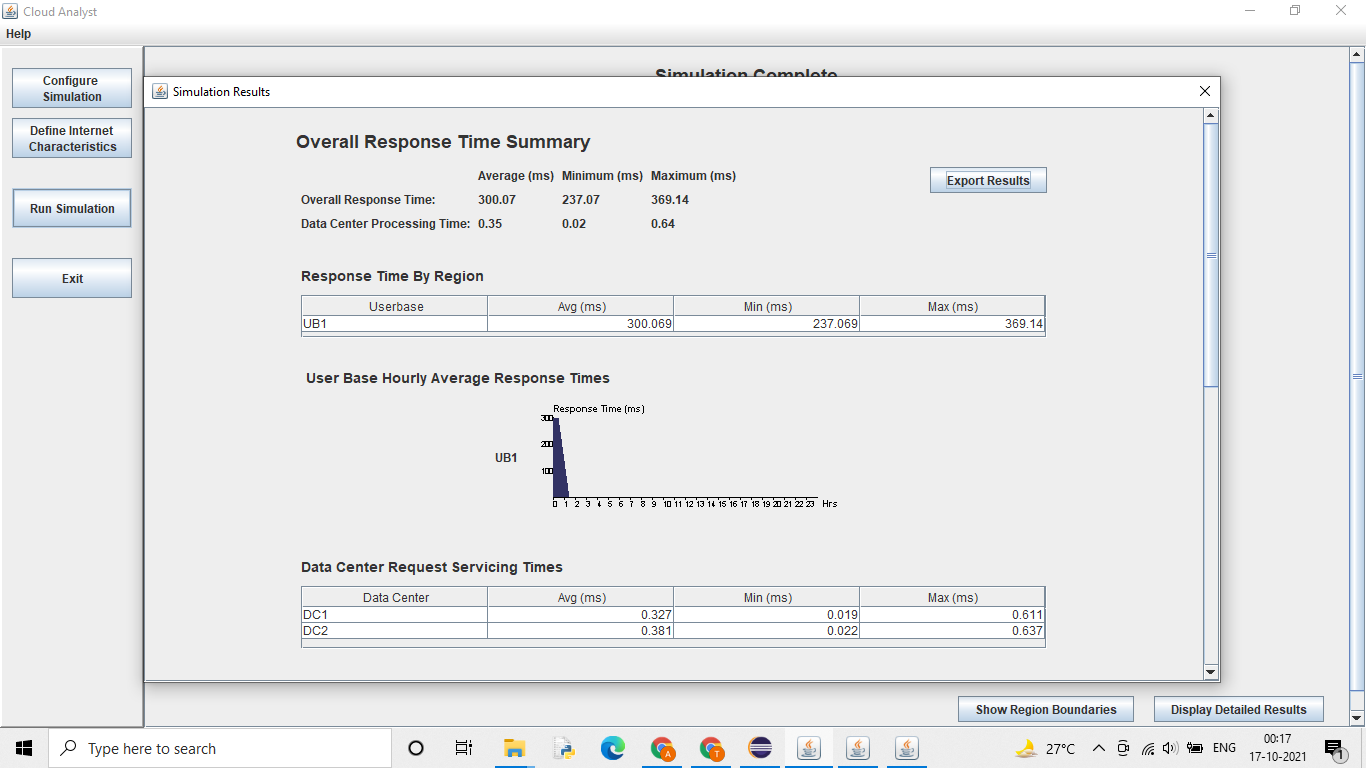
5.Honeybee

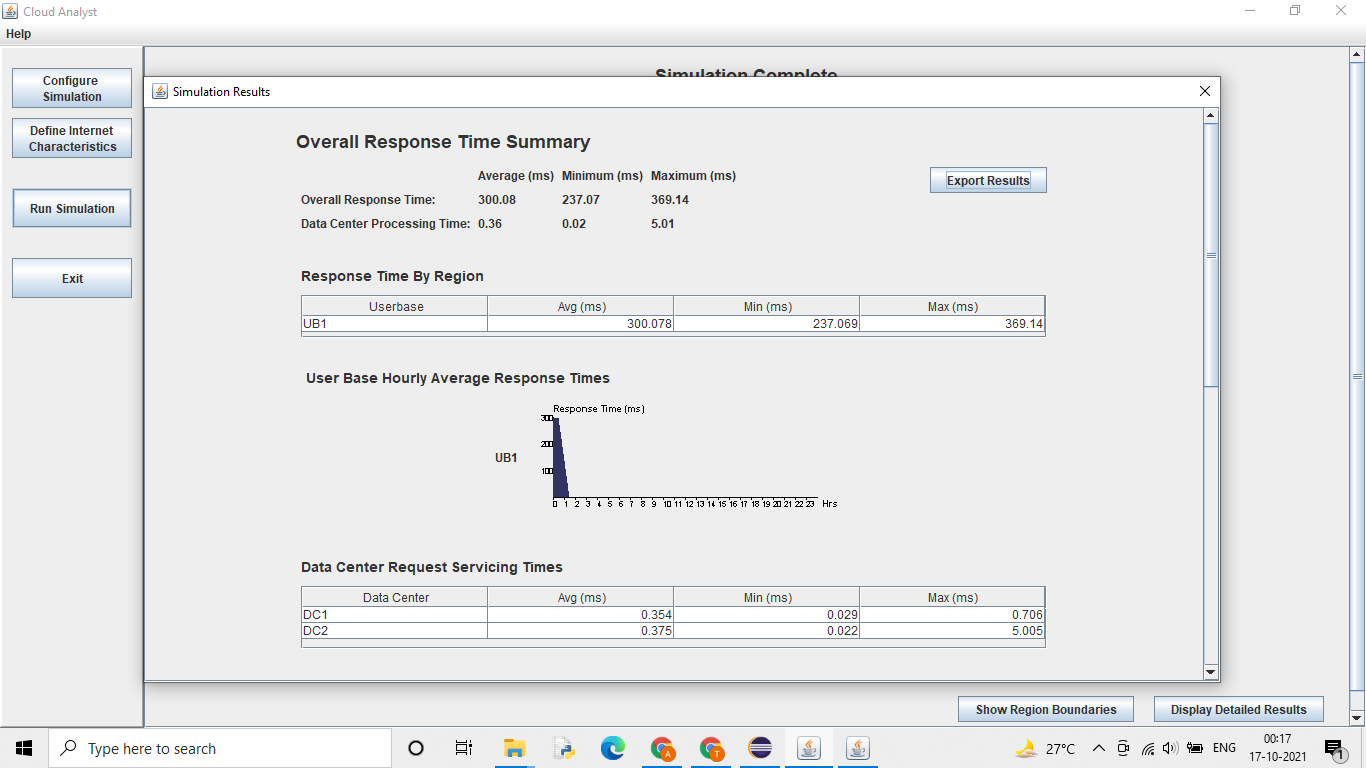
**[2]Data centers with same region**



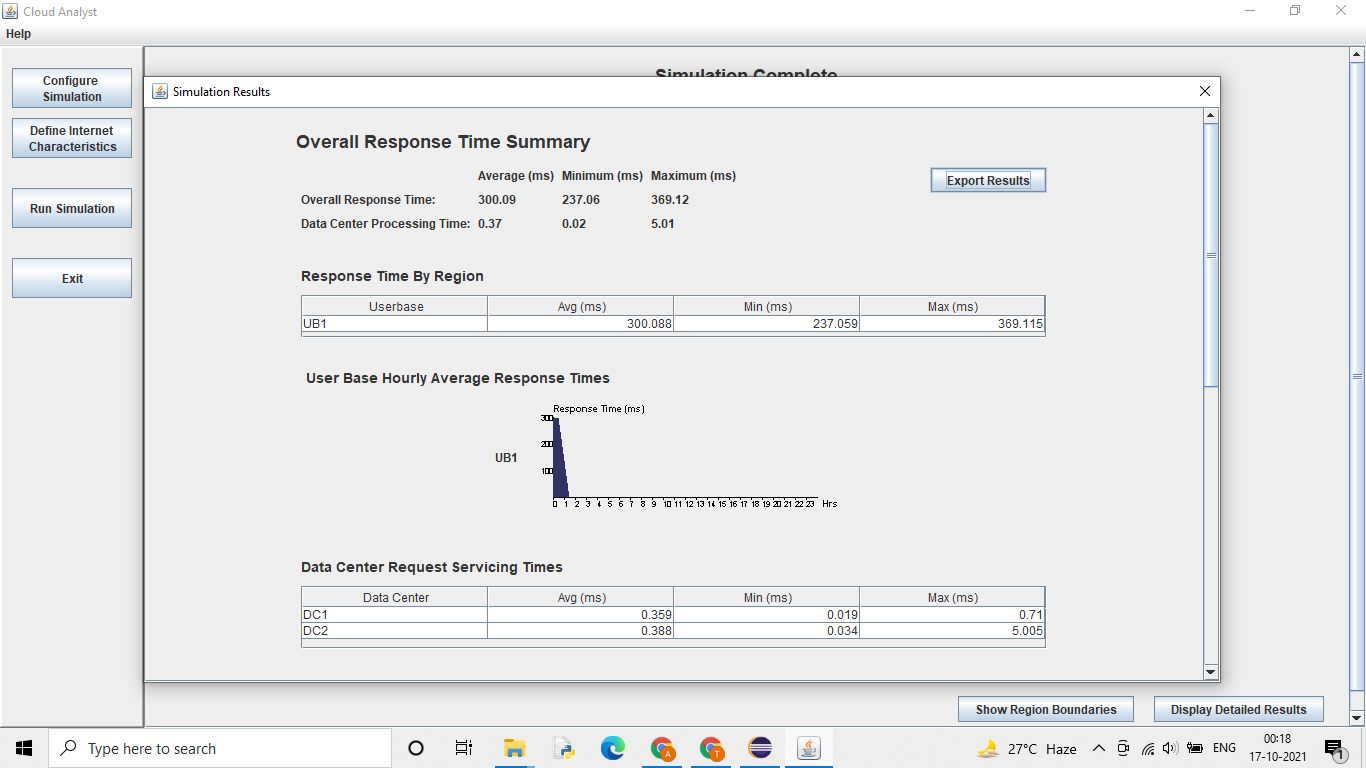
1. Round Robin
2. Equally spread



1. Throttled
2. Threshold



1. Honey bee



**Datacenters With different Regions**

| Name | Response Time | Processing Time | Cost |
| --- | --- | --- | --- |
| Round Robin | 50.24 | 0.49 | 1.85 |
| ESCE | 50.24 | 0.49 | 1.85 |
| Throttled | 50.24 | 0.49 | 1.85 |
| Threshold | 50.25 | 0.49 | 1.85 |
| Honey bee | 50.28 | 0.58 | 1.85 |

**Data Centers With same Regions**

| Name | Response Time | Processing Time | Cost |
| --- | --- | --- | --- |
| Round Robin | 300.07 | 0.35 | 1.08 |
| ESCE | 300.07 | 0.35 | 1.08 |
| Throttled | 300.07 | 0.35 | 1.08 |
| Threshold | 300.08 | 0.36 | 1.08 |
| Honey bee | 300.09 | 0.37 | 1.08 |

**Conclusion:**

1. Response time for datacenters with same region is more that with different regions
2. Processing time for datacenters with same region is less that with different regions
3. Round robin ,ESCE,Throttled have same avg Response time and avg processing Time
4. Honey bee has the most Response time and avg processing Time followed by threshold
5. Overall cost is the same for all algorithms.
6. Cost for datacenters with same region is less that with different regions

**10. REFERENCES**

[1] Sowmya Suryadevera, Jaishri Chourasia, Sonam Rathore, Abdul Jhummarwala. (2012). Load Balancing in Computational Grids Using Ant Colony Optimization Algorithm. International Journal of Computer & Communication Technology (IJCCT).

[2] Kumar Nishant, Pratik Sharma, Vishal Krishna, Chhavi Gupta and Kunwar Pratap Singh, Nitin and Ravi Rastogi. 2012 Load Balancing of Nodes in Cloud Using Ant Colony Optimization. 14th International Conference on Modelling and Simulation.

[3] Kuppani Sathish , A Rama Mohan Reddy, ( Octobe 2008) Enhanced ant algorithm based load balanced task scheduling in grid computing, IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.10, 21910.

[4] Ratan Mishra and Anant Jaiswal, April 2012 Ant colony Optimization: A Solution of Load balancing in Cloud, International Journal of Web & Semantic Technology (IJWesT) Vol.3, No.2.

[5] A. Y. Zomaya, & Y. H. Teh. (2001). Observations on using genetic algorithms for dynamic load-balancing. IEEE Transaction on Parallel and Distributed Systems, vol. 12, no. 9, pp. 899-911.

[6] Shanti Swaroop Moharana, Rajadeepan D. Ramesh & Digamber Powar. (May 2013) Analysis of Load Balancers in Cloud Computing, In International Journal of Computer Science and Engineering (IJCSE*),* ISSN 2278- 9960 Vol. 2, Issue 2, 101-108 ©IASE.

[7] M. Houle, A. Symnovis and D. Wood, (June 2002). Dimension-exchange algorithms for load balancing on trees, in: Proc. of 9th Int. Colloquium on Structural Information and Communication Complexity, Andros, Greece, pp. 181–196.

[9] Daniel Grosu, Anthony T. Chronopoulos et.al. (2005) ‘Noncooperative load balancing in distributed systems’, Journal of Parallel and Distributed Computing, Elsevier, Volume 65, Issue 9, pp 1022 – 1034.

[10] M. Salehi & H. Deldari. (2006) Grid Load Balancing using an Echo System of Intelligent Ants, Proceedings of the 24th IASTED International Conference on Parallel and Distributed Computing and Networks, pp. 47-52.

[11] Pandey S, Wu L, Guru S, Buyya R. (2010) A particle swarm optimization-based heuristic for scheduling workflow applications in Cloud Computing environments, In: International conference on advanced information networking and applications (AINA), IEEE Computer Society; p. 400–7.

[12] Dhinesh Babu L.D. and P. Venkata Krishna, 2013 Honey bee behaviour inspired load balancing of tasks in cloud computing environments, Applied Soft Computing, Vol. 13, No. 5, pp. 2292–2303.

[13] Mr.Manan D. Shah, February 2013 Allocation of Virtual Machines In Cloud Computing Using Load Balancing Algorithm, in International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555 Vol. 3, No.1.

[14] Ms.Nitika, Ms.Shaveta, Mr. Gaurav Raj, May 2012 Comparative Analysis of Load Balancing Algorithms in Cloud Computing, in International Journal of Advanced Research in Computer Engineering & Technology Volume 1, Issue 3,

[15] Dorigo, M., Gambardella, (1997). L.M.: Ant colony system: A cooperative learning approach to the traveling salesman problem. IEEE Transactions on Evolutionary Computation, 53–66.

[16] Salim Bitam, 2012 Bees Life Algorithm for Job Scheduling in Cloud Computing, Proceedings of The Third International Conference on Communications and Information Technology, pp. 186-191.

[17] Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg, and Ivona Brandic. (June 2009). Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility, Future Generation Computer Systems, Volume 25, And Number 6, ISSN: 0167-739X, Elsevier Science, Amstersdam, The Netherlands, Pages: 599-616.

27