An Research on the Effectiveness of the Different Algorithms and the Scheduling in Improving the Performance of Cloud Computing by using Cloud Analyst Simulator

Abdo H Guroob, Arpana Shetty, Manjaiah D. H.

Abstract— Cloud computing can be defined as a computing paradigm, where the various systems and large pool are connected to each other in private or public networks. The aim for that is to provide a dynamically scalable infrastructure, where it is used for applications, data and file storage. Cloud computing reduced the cost of computation and application hosting so that content storage and delivering services are handled faster and more flexibility. Load balancing is one of the challenges that affect the performance of cloud computing and the overcome it leads to better resource utilization and response time. The service broker policy plays an important role in accelerating the response time of customer requests by locating data centers or optimize the pattern of access to them. The contribution of this paper investigates the effectiveness of using the different algorithms and the approaches to improve the performance of cloud computing as it has been shown that there is a possibility to increase the performance of cloud computing by relying on certain criteria described in this paper. The results, which are presented in this paper were obtained using the cloud analyst simulator, where this simulator contains (Time duration, Load balancing algorithms, and Service Broker Algorithms, etc).

Keywords— Cloud Computing, Load balancing policy, Broker policy, Cloud analyst.

I. INTRODUCTION

Cloud computing is considered one of the modern technologies in the field of information technology, which provides the needs and requirements of end users through the Internet. There are other names for cloud computing such as On-Demand Computing, where services are provided to customers according to their needs so that all services of the company or enterprise are migrated to cloud computing (data centers). Of the services provided to customers in cloud computing whether it is on the level of hardware or software is Platform as a Service (PAAS), Infrastructure as a Service (IAAS), and Software as a Service (SAAS). The main goal of cloud computing lies in enabling customers

regardless of their number to access their data anywhere, anytime provided that the Internet service is available. Virtualization and grid computing is the base of cloud computing. The virtualization can be defined as the process of providing physical services over the Internet on a virtual platform. Where virtualization acts as an intermediate between a cloud service provider and a cloud user so that all the resources hosted on the virtual server can be scaled according to the request of the customers. The deployment models in cloud computing are different. There are four types of deployment models, in which each one of them works differently from the other in terms of method and architecture. The cloud deployment models are **Private cloud** is the main deployment model in cloud computing, there is provided the cloud service to a specific organization for a limited number of users (belonging to this organization) so that this network is blocked from other users or in other words, will not visible or available beyond that cloud users. This model is one of the safest types compared to other models and the leading providers of this type of services are Microsoft Azure, Amazon Web Services. A **community cloud** is an updated form of the previous model (Private cloud) where two or more organizations are connected in order to share their computational requirements within a single cloud framework. On the other hand, this model does not allow the user who does not belong to the community cloud to access to the services available. **Public Cloud** is one of the widest cloud services that is used in terms of the number of customers or spread. This service provides the giant companies, which are aiming to provide a cloud computing service and its features to a large group of audience. Although the service provided is almost free, it lacks security vulnerabilities, which increases the risk and threats associated with using this type of model. **Hybrid cloud** is one of the solutions used in cloud computing services especially in the security aspect where it is a combination of two or more deployment models. By using this type of models, the cloud service providers can

Abdo H Guroob, Department of Computer Science and Engineering Bearys Institute of Technology, Mangalore, India. (E-mail: abduohassan@yahoo.com)

Arpana Shetty, Department of Computer Science and Engineering Bearys Institute of Technology, Mangalore, India. (E-mail: arpana.shetty1@gmail.com)

Manjaiah D. H., Department of Post Graduate Studies and Research in Computer Science, Mangalore University, Mangalore, India. (E-mail: manju@mangaloreuniversity.ac.in)

present certain resources to a general audience and while providing other specific services to a particular group of customers. Highly secured services are made private and the rest are made as Public.

II. CLOUD ANALYST ARCHITECTURE

Cloud analyst considers one of an open source toolkit that used for analysis the cloud-computing environment in the simulation manner. Some features exist by using Cloud Analyst, where it provides for the researchers to develop new algorithms or methods to examine various scenarios in the cloud-computing environment. The simulator includes some configurations that must be performed before each simulation process which are four main settings:

- *Data Center Configuration*
- *User Bases Configuration*
- *Scheduling and Load Balancing Algorithms*
- *Internet Characteristics*

Data center configuration contains some variables such as data center location, the architecture used, physical components such as the number of CPUs, the operating system used, virtual machines, memory size, bandwidth consumption. User Bases configurations include the number of customers requesting cloud computing services, their location and the number of customers requesting service at the same time. The third section includes the algorithms of the scheduling and load balancing on the data centers where the simulator contains three algorithms for scheduling and three load balancing. Internet Characteristics includes the transmission delay between regions in ms and the available bandwidth between regions for the simulated application in Mbps.

Figure 1: Data Center configuration

Figure 2: User bases configuration

III. BUILD SCENARIOS FOR CLOUD COMPUTING ENVIRONMENT MORE COMPLEX

In this section, we suggest some more realistic scenarios to study the extent of their impact on throughput and cost ratio in cloud computing environments. In addition, sub-scenarios were proposed for each scenario in order to ascertain all the expected probability outcomes associated with each scenario, where data centers capabilities differ in terms of loads, available resources, and their location when the customers try to get cloud-computing services, which should have the lowest cost and in the fastest time during the execution. The proposed scenarios as shown in the figures (3-9) are as follows:

First Scenario: Increasing the number of data-centers as well as increasing the number of customers in the cloud-computing environment**.**

Second Scenario: Increasing the number of customers in the cloud-computing environment with a lack of numbers of data-centers**.**

Third Scenario: Increasing the number of available resources with a lack of data-centers.

Fourth Scenario: Increasing the number of customers in the cloud-computing environment who request the service at the same time with a lack of numbers of data-centers.

Fifth Scenario: Increasing the distance between the location of the data-centers and the users who request the cloud-computing service.

Sixth Scenario: Increasing the volume of data and requests sent by users with a lack of numbers of data-centers.

Seventh Scenario: Limited of the bandwidth consumption for data-centers regions with the increase in each of the following (the volume of data, the number of requests at the same time, and the number of users) with a lack of numbers of data-centers.
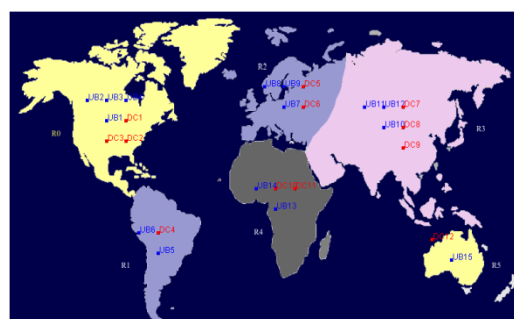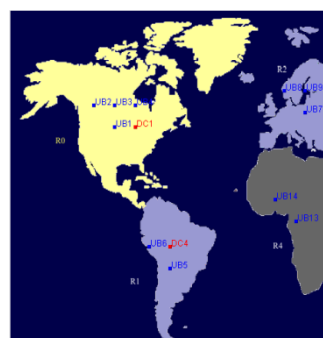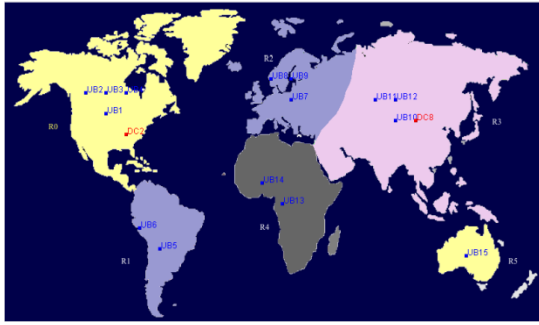

Figure 3: First Scenario
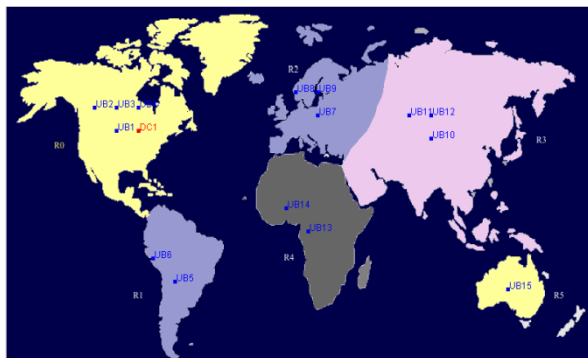
Figure 4: Second Scenario

Figure 5: Third Scenario

Table 1: The parameters of different scenarios

Scenarios	Number of UB's	Regions	Data Centers	Requests/hour	Req. in the same time	Request size (KB)	Memory (Mb)	Available BW	Number of Processors
Scenario 1	15	0,1,2,3,4,5	12	60	10	100	204800	1000	4
Scenario 2	15	0,1,2,3	4	60	10	100	204800	1000	4
Scenario 3	15	0,8	2	60	10	100	2048000	1000	20
Scenario 4	15	0,2,3	3	90	100	1000	2048000	1000	4
Scenario 5	15	0	1	100	100	1000	204800	1000	4
Scenario 6	15	0,2	5	200	100	2000	204800	1000	4
Scenario 7	15	0,2,3	3	100	100	1000	204800	100	4
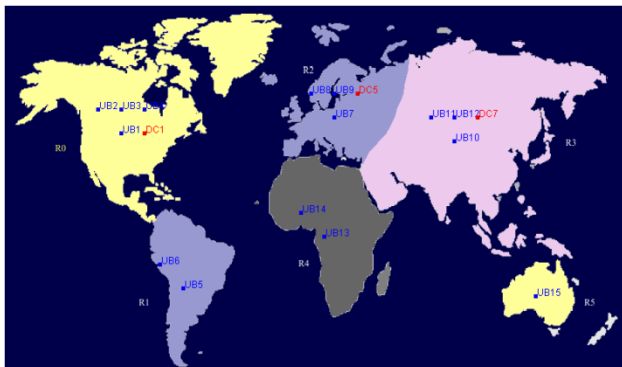
Figure 6: Fourth Scenario

Figure 8: Sixth Scenario

Figure 7: Fifth Scenario

Figure 9: Seventh Scenario

IV. SIMULATION EXPERIMENTS AND RESULTS

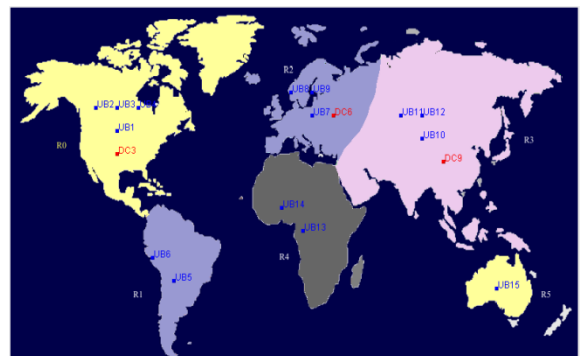Before simulations, the simulator configured to suit the requirements and scenarios proposed, where table 2 and table 3 show the configurations which are used in the current paper. Through what has been simulated for the cloud-computing environment using the Cloud Analyst for the proposed scenarios, which represent the most realistic scenarios for real environments, it turns out the following:

For the first scenario, an ideal cloud-computing environment was proposed by providing a number of data centers distributed in different geographical areas in a fair manner. In addition, the proportion of customers' presence in these regions whose seek to obtain the cloud computing services in the shortest possible time they have been distributed logically depending on the prevalence of modern technology in those regions.

Table 2: The matrix of delay, the delay of transmission between regions in milliseconds

Regions	0	1	2	3	4	5
Region 0	25	100	150	250	250	100
Region 1	100	25	250	500	350	200
Region 2	150	250	25	150	150	200
Region 3	250	500	150	25	500	500
Region 4	250	350	150	500	25	500
Region 5	100	200	200	500	500	25

Table 3: The matrix of bandwidth, the bandwidth that available between regions in Mbps

Regions	0	1	2	3	4	5
Region 0	2000	1000	1000	1000	1000	1000
Region 1	1000	800	1000	1000	1000	1000
Region 2	1000	1000	2500	1000	1000	1000
Region 3	1000	1000	1000	1500	1000	1000
Region 4	1000	1000	1000	1000	500	1000
Region 5	1000	1000	1000	1000	1000	2000

The simulation shows that the minimum time required to implement customer special requests with the capabilities available in the cloud computing environment equal to 50 milliseconds. However, other factors may affect the performance of executing the tasks in cloud computing environments such as the algorithms used to schedule the requests of customers or the load balancing policy algorithms on the data centers. We compared nine different possibilities, which are three algorithms for scheduling and three algorithms for load balancing. The scheduling algorithms are Reconfigure Dynamically with L, Optimize Response Time and Closest Data Center. The load balancing algorithms are Round Robin, Center, Equally Spread Current Execution Load, and Throttled. After conducting the simulations process was observed that both algorithms (Closest Data Center and Round Robin) for scheduling and load balancing respectively are considered more suitable than other algorithms as shown in the figure 10. Because that these algorithms have a fair load balancing and to the use of factor closest data center to the customers whose demand cloud computing services. This has made these two algorithms with each

other to provide the best performance compared to other algorithms. The most customers whose are benefited from these algorithms are the customers of the geographical region (R3) because to the fact that the number of data centers to the number of customers is fully commensurate with the requirements to be implemented. The worst algorithms in the performance which are available in the simulator for this scenario are (Reconfigure Dynamically with L, Round Robin) because reconfiguring the data centers dynamically to suit the current requirements lead to increases the time taken for implementation. The percentage difference between the performance of the best algorithms and worst algorithms, which are used here equal to 6.13%, where this percentage represents a large, value especially that these algorithms are applied in the ideal cloud computing environment. For the second scenario, the role of reducing the number of data centers in the cloud computing environment has had a significant impact on the fact that the customers whose presence is located in the two regions R4 and R5 are more affected regarding to execution time. Where the rate of implementation of requests between their regions and other regions is about 83% regardless of whether the algorithms used to schedule requests or the load balancing on data centers as shown in the figure 11. This leads us to the conclusion that data centers should be established in geographic regions R4 and R5 to meet customer requirements. Ironically, which we observed in the third scenario is that despite increased the resources, CPU capacity, and the limited number of distributed data centers in the different geographic regions, when we compare this scenario with the previous scenario (second scenario), the implementation time for the customers' requests was more than 40%. This leads us to the conclusion that even with increased data center resources and increased CPU capacity, the implementation time for customer requests remains constant. Nevertheless, what affects the execution time is the spatial dimension of data centers. In terms of scheduling algorithms and loads balancing algorithms, which have been applied to simulate the cloud-computing environment in this scenario, the ratio for the performance of the best-aggregated algorithms and worst aggregated algorithms that are used here are 75% and 89.8%, respectively based on the ideal scenario (first scenario) figure 12. In the fourth and fifth scenarios, we find the harmony between the implementation time and the location of the data centers. Where a match was observed in the implementation values of the various algorithms regarding the performance of customers' requirements regardless of remote or closest data centers even with an increase in the rate of requests that must be implemented at the same time. However, the time of implementation of customers' requirements in the fourth scenario is less than the time of implementation of customers' requirements in the fifth scenario by up to 38.6% due to the reason that the geographical dimension of the data centers and

customers is large as shown

in the figures 13 and 14. When applying the sixth scenario by increasing the size of the data and the sent requests, the performance of the various algorithms is almost equal except the Reconfigure Dynamically with L and Round Robin algorithm, especially for the customers of the geographical region (R2) despite the availability of data centers. Where the difference between the performance of the algorithms mentioned above and the rest of the other algorithms are about 33%. This is due to the nature of the algorithms used in this simulation, which depends on reconfiguring the data centers dynamically to suit the current requirements. Hence, it is clear to us here the role of the data volume and the requests sent that effect on the performance in the cloud-computing environment as well as the impact of algorithms used in scheduling operations even if the environment is ideal as shown in the figure 15. We have deliberately applied the fourth scenario and its implementation in the seventh scenario in terms of similarity in the number of customers and the number of data-centers and their presence in the same geographical regions. Nevertheless, the differences in the seventh scenario from the fourth scenario in limited of the bandwidth consumption for data centers regions with the increase in each of the following (the volume of data, the number of requests at the same time, and the number of users) with a lack of numbers of data-centers. However, we were surprised by matching the results obtained from simulations in the two scenarios. Besides, referring to the variables of the fourth scenario; it is clear that limited bandwidth consumption does not significantly affect the performance of cloud computing services due to the nature of data sent in the simulator as shown in figure 15.
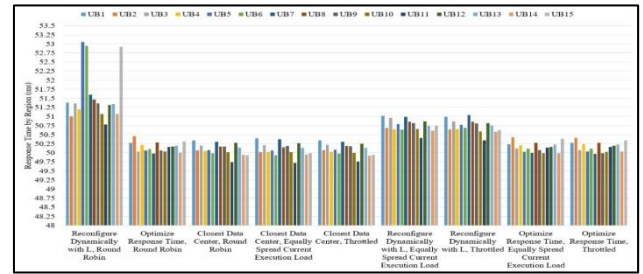

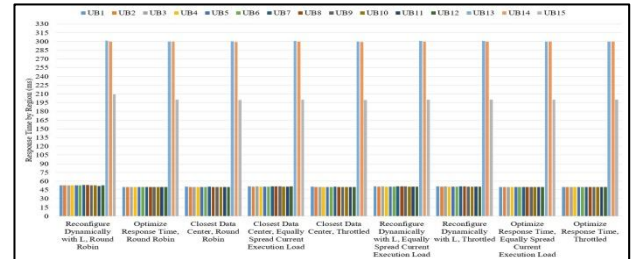**Figure10: First scenario_ Average response time.**

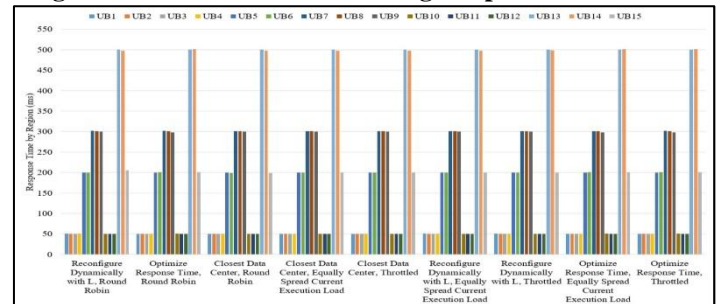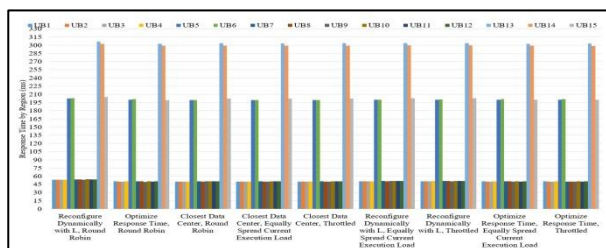**Figure11: Second scenario_ Average response time.**

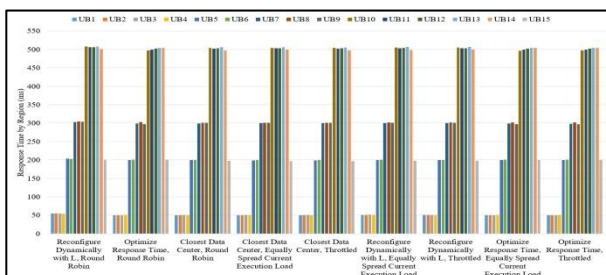**Figure12: Third scenario_ Average response time.**

Nevertheless, with changing data type it may possibly affect the performance of the implementation of the services in the cloud-computing environment. Of things that have also been observed is the bandwidth consumption applicable in the simulator is sufficient to conduct transmissions of large data in a complex manner that the performance of service delivery is not affected in the cloud-computing environment.

Table 3: Data Centre Configuration

| DC	Regions	Arch.	OS	VMM	Cost per VM $/Hr.	Cost of using Memory $/s	Cost of using Storage $/s	Cost of using Data Transfer $/Gb	Physical HW Units
D1	0	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D2	0	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D3	0	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D4	1	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D5	2	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D6	2	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D7	3	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D8	3	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D9	3	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D10	4	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D11	4	x86	Linux	Xen	0.5	0.05	0.1	0.25	1
D12	5	x86	Linux	Xen	0.5	0.05	0.1	0.25	1

Figure13: Fourth scenario_ Average response time

Figure14: Fifth scenario_ Average response time
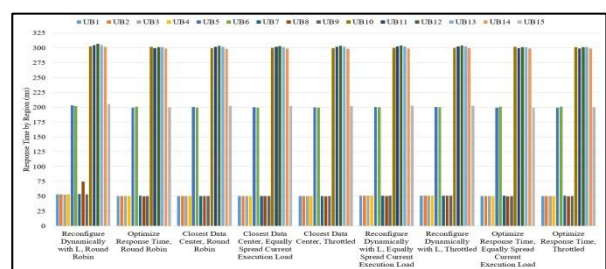
Figure15: Sixth scenario_ Average response time
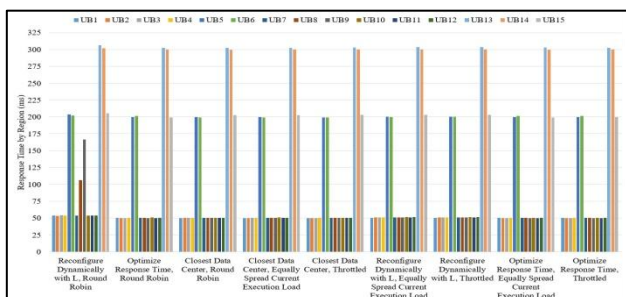
Figure16: Seventh scenario_ Average response time

V. CONCLUSION

Cloud computing at the present time is one of the important pillars in most modern techniques used in almost all fields, Therefore, researchers and those interested in developing technological environments trying to meet the needs of customers. However, there are some challenges that limit the fulfillment of all requirements such as the load balancing on data centers and scheduling the tasks to be implemented. From this point of view, there are many mechanisms and algorithms that have been developed to overcome these obstacles, which lead to accelerating the implementation of the requirements of customers in the shortest possible time. In this paper, seven different and more realistic scenarios have been proposed for the study of their effect on the ratio of throughput and performance. The obtained results showed a large difference in the implementation of the different algorithms available in the simulator according to the proposed scenarios. For future work, a research paper is prepared to examine the impact of cost and performance on data centers.

REFERENCES

1. Tushar Desai, Jignesh Prajapati, "A Survey of Various Load Balancing Techniques and Challenges in Cloud Computing," International Journal of Scientific & Technology Research, Volume 2, Issue 11, November 2013.
2. Dhaval Limbani, Bhavesh Oza, "A Proposed Service Broker Strategy in Cloud Analyst for Cost-Effective data center Selection," International Journal of Engineering Research and Applications. Vol. 2, Issue 1, Jan-Feb 2012, pp.793-797
3. Kishor Kunal, Thapar Vivek, "An Efficient Service Broker Policy for Cloud Computing Environment," International Journal of Computer Science Trends and Technology. Vol. 2, Issue 4, July-Aug 2014.
4. Bhathiya Wickremasinghe, "Cloud Analyst: A Cloud-Sim-Based Tool For Modeling And Analysis Of Large Scale Cloud Computing Environments. MEDC Project," Report 2010.
5. AmolJaikar, Seo-Young Noh, "Cost and performance effective data center selection system for scientific federated cloud", Springer Science+Business Media New York, Peer-to-Peer Network Appl. DOI 10.1007/s12083-014-0261-7, May 2014.
6. B. Wickremasinghe and R. N. Calheiros, "Cloud Analyst: a cloudSIm-based visual modeler for analyzing cloud computing environments and applications," 24th International Conference on Advanced Information Networking and Application, pp.446-452, 2010.
7. Rakesh Kumar Mishra, Sandeep Kumar, SreenuNaik B, "Priority Based Round-Robin Service Broker Algorithm for Cloud-Analyst", IEEE International Advance Computing Conference, pages 878-881, 2014.
8. P. Garbacki and V. K. Naik, "Efficient Resource virtualization and sharing strategies for heterogeneous Grid environments," inProc. IFIP/IEEE IMSymp, 2007, pp. 40–49.
9. Shyam Patidar Dheeraj Rane "A Survey Paper on Cloud Computing," 2012 Second International Conference on Advanced Computing & Communication Technologies.
10. Wei-Tek Tsai*. Xin Sun. lanaka Balasooriya "Service-Oriented Cloud Computing Architecture" Computer society 2010.
11. Tanveer Ahmed and Yogendra Singh, "Analytic Study Of Load Balancing Techniques Using Tool Cloud Analyst," International Journal of Engineering Research and Applications, April 2012.
12. Shipra Gupta and lndu Kashyap, "Cost and Time Evaluation of Load Balancing and Service Broker Strategies in Multiple Data Centers," International Journal of Computer Applications (0975 - 8887) Volume 103 - No 17, October 2014.
13. Rajwinder Kaur and Pawan Luthra, "Load Balancing in Cloud Computing," Association of Computer Electronics and Electrical Engineers, 2014, DOI: 02.ITC.2014.5.92
14. Ashalatha R; J. Agarkhed, "Dynamic load balancing methods for resource optimization in cloud computing environment," Annual IEEE India Conference (INDICON) 2015, Pages: 1 - 6, DOI: 10.1109/INDICON.2015.7443148.

15. K. Garala; N. Goswami; P. D. Maheta, "A performance analysis of load Balancing algorithms in Cloud environment," 2015 International conference on Computer Communication and Informatics (ICCCI), Pages: 1 - 6, DOI: 10.1109/ICCCI.2015.7218063.

16. G.Punetha Sarmila, N. Gnanambigai, P. Dinadayalan, "Survey on Fault Tolerant –Load Balancing Algorithms in Cloud Computing," IEEE Sponsored 2nd International Conference On Electronics And Communication System (ICECS 2015), Pages-1715-1720.

17. Ahmed, S., 2012. Enhanced Proximity-based Routing Policy for Service Brokering in Cloud Computing. Int. J. Eng. Res. Appl. (India) 2 (2), 1453–1455.

18. Dillon, T., Wu, C., Chang, E, "Cloud computing: issues and challenges," Paper presented at the Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on.

19. Mishra, R.K., Kumar, S., Sreenu Naik, B., "Priority based round robin service broker algorithm for cloud-analyst," Paper Presented at the Advance Computing Conference (IACC), 2014 IEEE International 21–22 Feb. 2014.

20. Rekha, P.M., Dakshayini, M., "Cost based data center selection policy for large scale networks," Paper presented at the Computation of Power, Energy, Information and Communication (ICCPEIC), 2014 International Conference on 16–17 April 2014.

21. Iosup, A., Ostermann, S., Yigitbasi, M.N., "Performance analysis of cloud computing services for many-tasks scientific computing," IEEE Trans. Parallel Distrib. Syst. 22 (6), 931–945.

22. Rajesh, G., Sreenivasulu, G., "The issues of cloud service delivery through virtualization of Dynamically Generated multiple virtual machine Services without missing deadline on the World Wide Web," Int. J. Curr. Eng. Tech. 4 (4), 2758–2762.

23. Beloglazov, A., Buyya, R., "Managing overloaded hosts for dynamic consolidation of virtual machines in cloud data centers under quality of service constraints," IEEE Trans. Parallel Distrib. Syst. 24 (7), 1366–1379.