ISSN 2277-3061

IMPROVED ANT COLONY LOAD BALANCING ALGORITHM IN CLOUD COMPUTING

Jaspreet Singh [1], Susmita Mishra [2]
Research Scholar, BMS College of Engineering, Sri Muktsar Sahib, Punjab [1]
jassp777@gmail.com
Assistant professor, BMS College of Engineering, Sri Muktsar Sahib, Punjab [2]
susmita.mshr@gmail.com

ABSTRACT

Cloud Computing is an emerging computing paradigm. It aims to share data, calculations, and service transparently over a scalable network of nodes. Since Cloud computing stores the data and disseminated resources in the open environment. Since, cloud has inherited characteristic of distributed computing and virtualization there is a possibility of machines getting unused. Hence, in this paper, a load balancing algorithm has been proposed to avoid deadlocks and to provide proper utilization of all the Virtual Machines (VMs) while processing the requests received from the users by VM migration. Further, this paper also provides the anticipated results with the implementation of the proposed algorithm. The main contributions of our work is to balance the entire system load while trying to minimize the make span of a given set of jobs. Compared with the other job scheduling algorithms, the improved Ant Colony Optimization algorithm can outperform them according to the experimental results.

Keywords: ANT COLONY; LOAD BALANCING ALGORITHM; CLOUD COMPUTING.

ISSN 2277-3061

INTRODUCTION

Cloud computing is Internet ("cloud") based development and use of computer technology ("computing"). It is a style of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet. Users need not have knowledge of, expertise in, or control over the technology infrastructure "in the cloud" that supports them. The concept incorporates infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS) as well as Web 2.0 and other recent technology trends which have the common theme of reliance on the Internet for satisfying the computing needs of the users. Examples of SaaS vendors include Salesforce.com and Google Apps which provide common business applications online that are accessed from a web browser, while the software and data are stored on the servers. The term cloud is used as a metaphor for the Internet, based on how the Internet is depicted in computer network diagrams, and is an abstraction for the complex infrastructure it conceals.

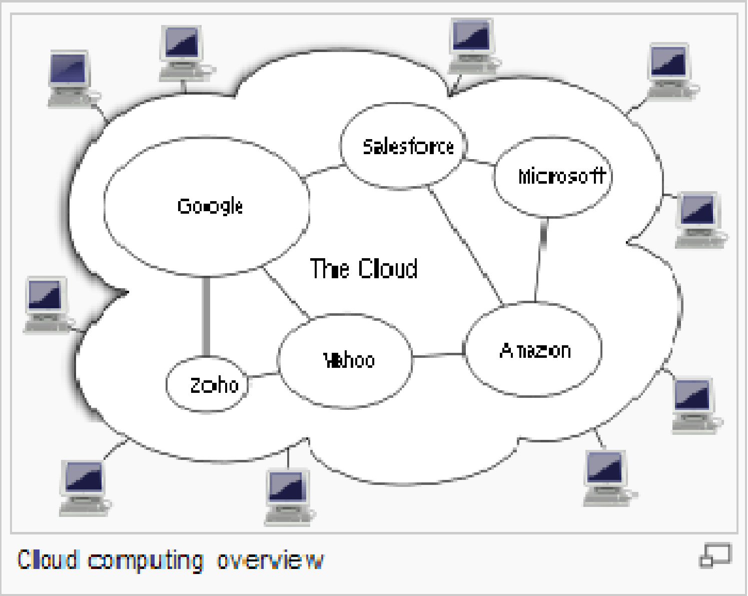


Fig 1: Cloud computing

A. Comparison

Cloud computing is often confused with grid computing ("a form of distributed computing whereby a 'super and virtual computer' is composed of a cluster of networked, loosely-coupled computers, acting in concert to perform very large tasks"), utility computing (the "packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility such as electricity") and autonomic computing ("computer systems capable of self-management")..

B. Implementation

The majority of cloud computing infrastructure as of 2009 consists of reliable services delivered through data centers and built on servers with different levels of virtualization technologies. The services are accessible anywhere that has access to networking infrastructure. The Cloud appears as a single point of access for all the computing needs of consumers. Commercial offerings need to meet the quality of service requirements of customers and typically offer service level agreements. Open standards are critical to the growth of cloud computing and open source software has provided the foundation for many cloud computing implementations

C. Characteristics

As customers generally do not own the infrastructure, they merely access or rent, they can avoid capital expenditure and consume resources as a service, paying instead for what they use. Many cloud-computing offerings have adopted the utility computing model, which is analogous to how traditional utilities like electricity are consumed, while others are billed on a subscription basis. Sharing "perishable and intangible" computing power among multiple tenants can improve utilization rates, as servers are not left idle, which can reduce costs significantly while increasing the speed of application development. A side effect of this approach is that "computer capacity rises dramatically" as customers do not have to engineer for peak loads. Adoption has been enabled by "increased high-speed bandwidth" which makes it possible to receive the same response times from centralized infrastructure at other sites.

D. Economics

Cloud computing users can avoid capital expenditure (CapEx) on hardware, software and services, rather paying a provider only for what they use. Consumption is billed on a utility (e.g. resources consumed, like electricity) or subscription (e.g. time based, like a newspaper) basis with little or no upfront cost. Other benefits of this time sharing style approach are low barriers to entry, shared infrastructure and costs, low management overhead and immediate access to a broad range of applications. Users can generally terminate the contract at any time (thereby avoiding return on investment risk and uncertainty) and the services are often covered by service level agreements with financial penalties. According to Nicholas Carr the strategic importance of information technology is diminishing as it becomes standardized and cheaper. He argues that the cloud computing paradigm shift is similar to the displacement of electricity generators by electricity grids early in the 20th century.

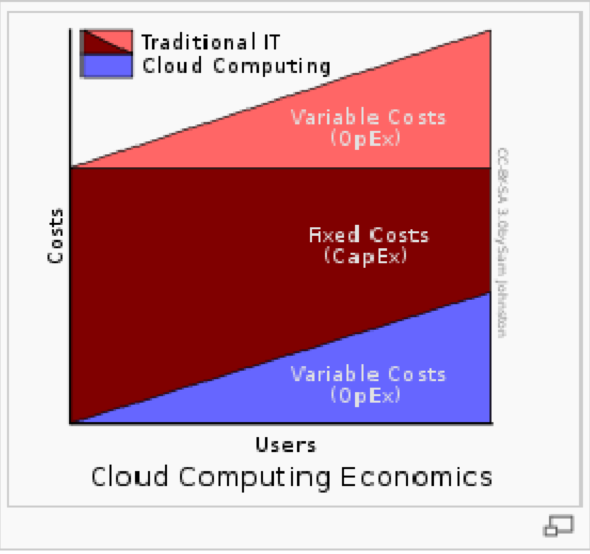


Fig 2: Cloud computing Economics

E. Companies

Providers including Amazon, Microsoft, Google, Sun and Yahoo exemplify the use of cloud computing. It is being adopted by individual users through large enterprises including General Electric, L'Oréal, and Procter & Gamble.

IMPROVED ACO ALGORITHM

Ant based control system was designed to solve the load balancing in cloud environment. Each node in the network was configured with

1) Capacity that accommodates a certain.

2) Probability of being a destination.

3) Pheromone (or probabilistic routing) table.

Each row in the pheromone table represents the routing preference for each destination, and each column represents the probability of choosing a neighbor as the next hop. Ants are launched from a node with a random destination. In this approach, incoming ants update the entries the pheromone table of a node. For instance, an ant traveling from (source) to (destination) will update the corresponding entry in the pheromone table in. Consequently, the updated routing information in can only influences the routing ants and calls that have as their destination. However, for asymmetric networks, the costs from to and from to may be different. Hence, In this approach for updating pheromone is only appropriate for routing in symmetric networks. If an ant is at a choice point when there is no pheromone, it makes a random decision However, When only pheromone from its own colony is present there is a higher probability that it will choose the path with the higher concentration of its own pheromone type. In addition, due to repulsion, an ant is less likely to prefer paths with (higher concentration of) pheromone from other colonies. Moreover, it is reminded that the degrees of attraction and repulsion are determined by two weighting parameters. **Pheromone tables** we replaced the routing tables in the network nodes by tables of probabilities, which we will call ‘pheromone tables', as the pheromone strengths are represented by these probabilities. Every node has a pheromone table for every possible destination in the network, and each table has an entry for every neighbor. For example, a node with four neighbors in a 30-node network has 29 pheromone tables with four entries each. One could say that an n-node network uses n different kinds of pheromones. The entries in the tables are the probabilities which influence the ants' selection of the next node on the way to their destination node. Figure 4 shows a possible network configuration and pheromone table. For example, ants travelling from node 1 to node 3 have a 0.49probability of choosing node 2 as their next node, and 0.51 of choosing node 4 ‘Pheromone laying' is represented by ‘updating probabilities

ISSN 2277-3061

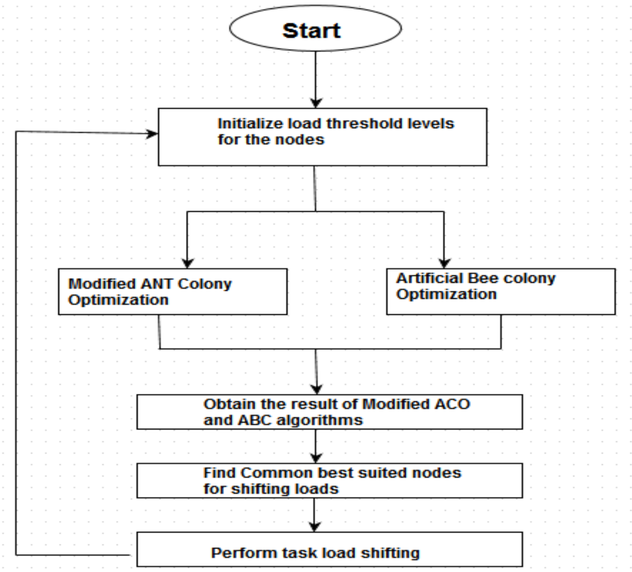


Fig 3: Ant Colony Algorithm

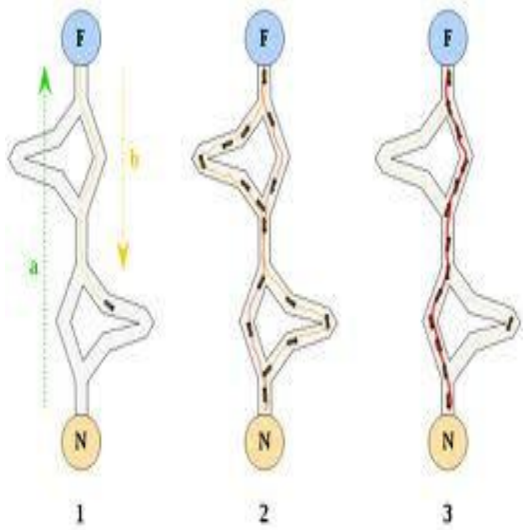


Fig 4: Ant Colony Optimization

SIMULATION AND RESULTS

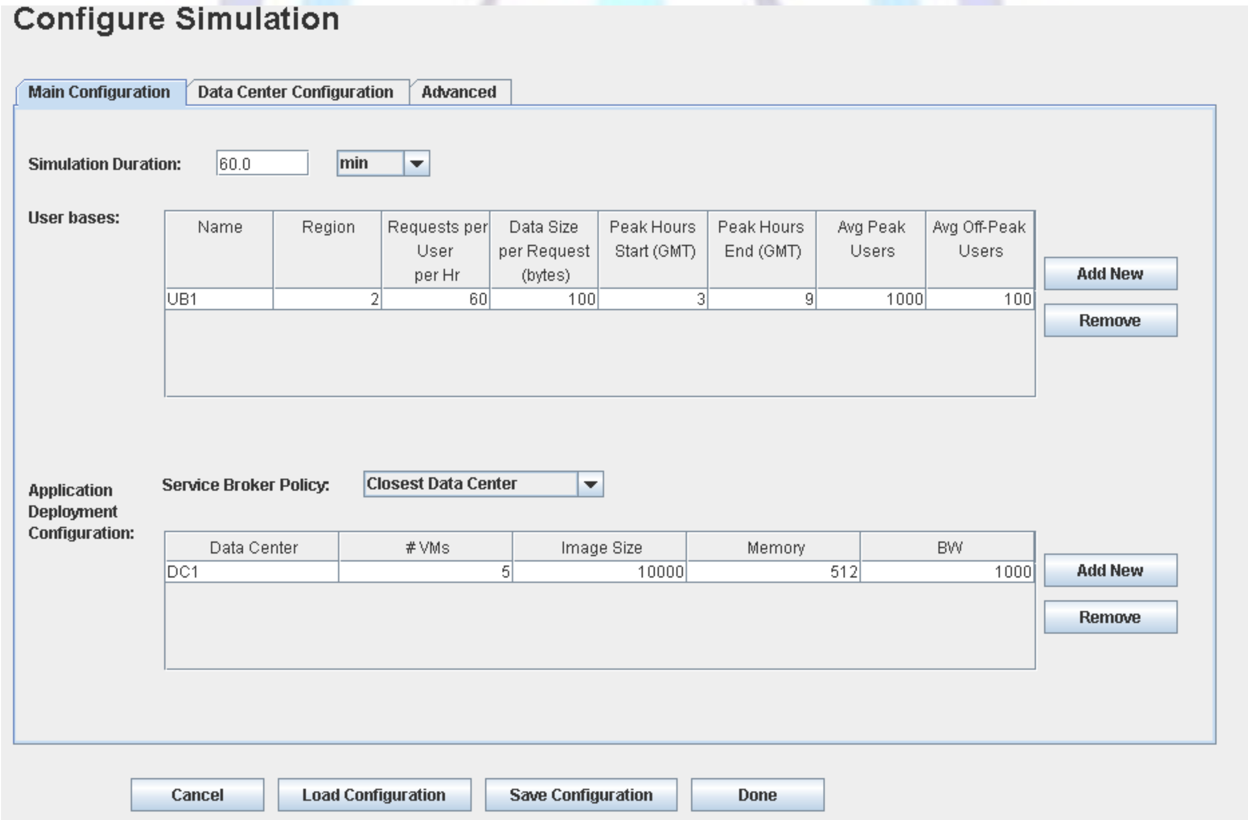
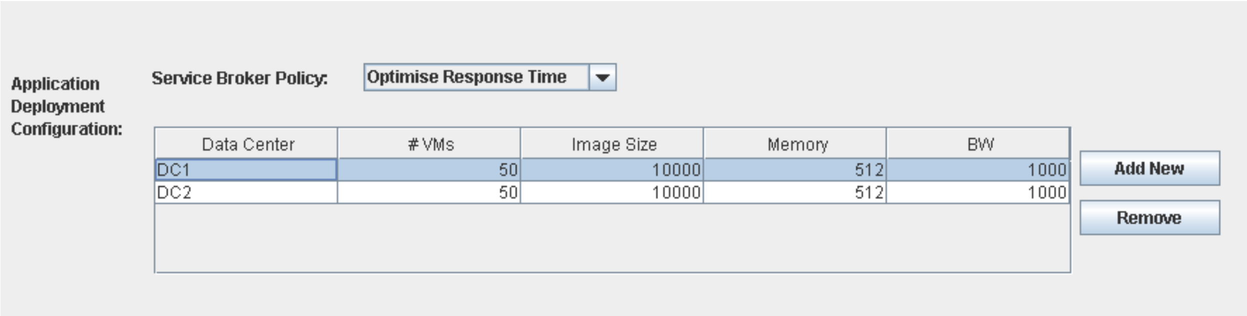


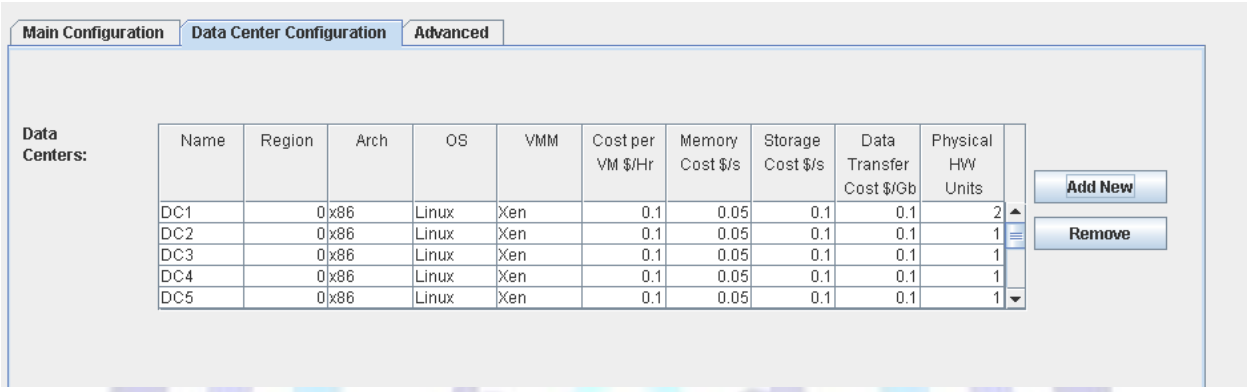
Fig 5: Configuration Screen in Cloud Analyst

ISSN 2277-3061



Data Center	# VMs	Image Size	Memory	BW
DC1	50	10000	512	1000
DC2	50	10000	512	1000

Fig 6: Broker Policy Configuration



Name	Region	Arch	OS	VMM	Cost per VM $/Hr	Memory Cost $/s	Storage Cost $/s	Data Transfer Cost $/Gb	Physical HW Units
DC1	0	x86	Linux	Xen	0.1	0.05	0.1	0.1	2
DC2	0	x86	Linux	Xen	0.1	0.05	0.1	0.1	1
DC3	0	x86	Linux	Xen	0.1	0.05	0.1	0.1	1
DC4	0	x86	Linux	Xen	0.1	0.05	0.1	0.1	1
DC5	0	x86	Linux	Xen	0.1	0.05	0.1	0.1	1

Fig 7: Datacenter Configuration

	Average (ms)	Minimum (ms)	Maximum (ms)
Overall Response Time:	300.06	237.06	369.12
Data Center Processing Time:	0.34	0.02	0.61

Fig 8: Response time of new Proposed Algorithm

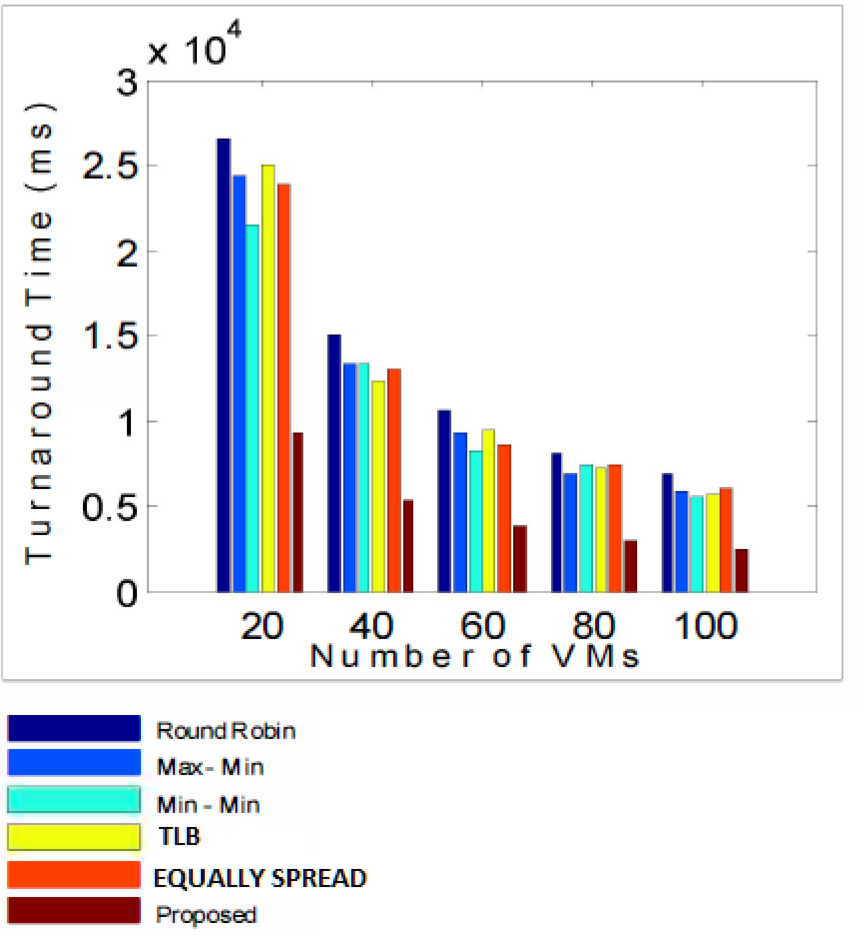


Fig 9. Turn Around time

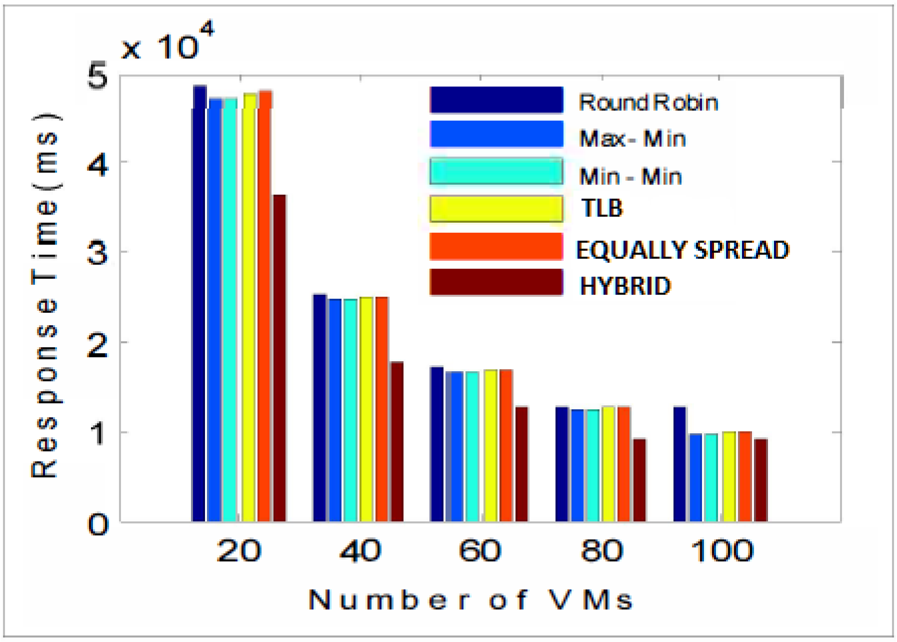


Fig 10. Response Time

ISSN 2277-3061

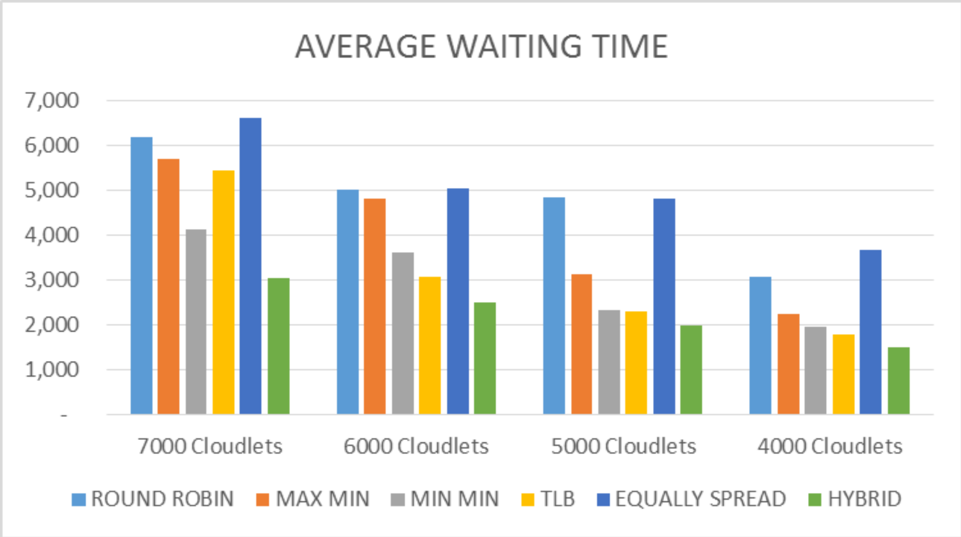


Fig 11. Average Waiting Time

	Average (ms)	Minimum(ms)	Maximum (ms)
Response Time	300.06	237.06	369.12
Data Center Processing Time	0.34	0.02	0.61

Table 1: Response and Data Processing Time

Parameters	Round Robin	Equally Spaced	Throttled	**Hybrid**
Data Centers	2	2	2	2
UB	5	5	5	5
H/W Unit	2	2	2	2
V.M	20	20	20	20
Avg (ms)	0.28	0.32	0.24	0.18
Max (ms)	0.64	0.69	0.83	0.56
Min (ms)	0.02	0.05	0.01	0.01
Total (Cost in $)	1.83	2.84	2.96	1.56

Table 2: Computed Parameters of different Algorithms

Conclusions

The performance of Improved Ant Colony optimization algorithm is studied in the paper. The request time for the policies applied (Round Robin, Equally spread current execution load, Throttled Load balancing) are same which means there is no effect on data centers request time after changing the algorithms. The cost analysis showed for each algorithm is calculated in the experimental work. The proposed load balancer (Improved Ant Colony Optimizer) wisely binds the cloudlets to virtual machines to minimize the turnaround time and response time so that the desired objective is fulfilled. The model proposed is based on centralized load balancing strategy. Before binding cloudlets to VM, the proposed load balancer first calculate the remaining capacities of all VMs and then dispatch the cloudlet to more powerful VM. Moreover, as load is first readjusted among VMs, the load redistribution process is fair. In order to evaluate performance of the proposed model simulation study has been put through various test conditions. It has been found that the model works well in ensuring an even distribution of the workload. In the done work, it has been assumed that all the incoming requests are independent to each other. This study is only concerned with the number of the resources to be odd or even and analyses the merits and drawbacks of two well-known traditional algorithms, Max-min and Min-min. In this paper, the deadline of each task, arriving rate of the tasks, cost of the task execution on each of the resource, cost of the

ISSN 2277-3061

communication and many other cases that can be a topic of research are not considered. Also, applying the proposed algorithm on actual grid environment for practical evaluation can be other open problem in this area.

References

[1] O.M. Elzeki . “Improved Max-Min Algorithm in Cloud Computing”. International Journal of Computer Applications(0975-8887) Volume 50-No.12,july 2012.

[2] “A technical support seminar on cloud computing technology” by Prashant Gupta.

[3] Amandeep Kaur Sidhu. “Analysis of load balancing techniques in cloud computing”. International Journal of computers & technology volume 4 No. 2, March-April, 2013, ISSN 2277-3061.

[4] Ektemal Al-Rayis. “Performance Analysis of load balancing Architectures in Cloud computing” 2013 European Modeling Symposium. 978-1-4799-2578-0/13$31.00@2013 IEEE.

[5] Haozheng Ren. “The load balancing Algorithm in cloud computing Environment” 2nd International Conference on computer science and network technology 2012.

[6] Tushar Desai. “A survey of various load balancing techniques and challenges in cloud computing” International Journals of scientific and technology research volume 2. Issue11,Nov2013.

[7] Upendra Bhoi. “Enhanced max-min Task scheduling Algorithm in cloud computing”. International Journal of Application or Innovation in Engineering & management(IJAIEM), April 2013.

[8] Klaithem Al Nuaimi, “A survey of load balancing in cloud computing challenges and algorithm”. 2012 IEEE second symposium on network cloud computing and applications.

[9] Gytis Vilutis, “Model of load balancing and scheduling in cloud computing”. Proceedings of the ITI 2012 34th Int.Conf. on Information Technology Interfaces, June 25-28,Cavat,Croatia.

[10] S. Banerjee, I. Mukherjee and P.K. Mahanti, Cloud Computing Initiative using Modified Ant Colony Framework, World Academy of Science and Technology, 56, pp. 221-224, 2009.

[11] Y. Li, A Bio-inspired Adaptive Job Scheduling Mechanism on a Computational Grid, International Journal of Computer Science and Network Security, 6(3B), pp. 1-7, 2006.

[12] S.C. Wang, K.Q. Yan, W.P. Liao and S.S. Wang, Towards a Load Balancing in a Three-level Cloud Computing Network, Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology, pp. 108-113, 2010.

[13] M. Salehi and H. Deldari, 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, 2006.

[14] M. Dorigo, V. Maniezzo and A. Colorni, Ant System: Optimization by a Colony of Cooperating Agents, IEEE Transactions on Systems, Man, and Cybernetics, PP. 29-41, 1996.

[15] C.W. Chiang, Y.C. Lee, C.N. Lee and T.Y. Chou, Ant Colony Optimization for Task Matching and Scheduling, IEE Proceedings on Computers and Digital Techniques, 153 (6), pp. 373- 380, 2006.

[16] M. Dorigo, M. Birattari and T. Stutzle, Ant Colony Optimization-Artificial Ants as a Computational Intelligence Technique, IEEE Computational Intelligence Magazine, pp. 1- 12. 2006.

[17] J. Sun, S. Xiong and F.M. Guo, A New Pheromone Updating Strategy In Ant Colony Optimization, Proceedings of the International Conference on Machine Learning and Cybernetics, pp. 620-625, 2004.

[18] K.u Ruhana, K. Mahamud, H. Jamal and A. Nasir, Ant Colony Algorithm for Job Scheduling in Grid Computing, Proceedings of the Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation, pp. 40-45, 2010.

[19] H. Jamal, A. Nasir, K. Ruhana, K. Mahamud and A.M. Din, Load Balancing Using Enhanced Ant Algorithm in Grid Computing, Proceedings of the Second International Conference on Computational Intelligence, Modelling and Simulation, pp. 160-165, 2010.

[20] C. Gong, J. Liu, Q. Zhang, H. Chen and Z. Gong, The Characteristics of Cloud Computing, Proceedings of the 39th International Conference on Parallel Processing Workshops, pp. 275-279, 2010.

[21] http://searchcloudcomputing.techtarget.com

[22] http://www.techsmith.com/morae/whitepaper/ux20.asp

[23] R.A. Arnous, H.A. Arafat and M.M. Salem, Improving the Load Balancing within the Data Network via Modified AntNet Algorithm, Proceedings of the 5th International Conference on Information and Communication Technology, pp. 189-195, 2007.

[24] R. Chang, J. Chang and P. Lin, Balanced Job Assignment Based on Ant Algorithm for Grid Computing, Proceedings of the IEEE Asia-Pacific Services Computing Conference, pp. 291-295, 2007.

[25] H. Yan, X. Shen, X. Li and M. Wu, An Improved Ant Algorithm for Job Scheduling in Grid Computing, Proceedings of the International Conference on Machine Learning and Cybernatics, 5, pp. 2957-2961, 2005.

[26] http://boinc.berkeley.edu/

[27] Singh A., Korupolu M. and Mohapatra D. (2008) ACM/IEEE conference on Supercomputing.

[28] Stanojevic R. and Shorten R. (2009) IEEE ICC, 1-6.

[29] Zhao Y. and Huang W. (2009) 5th International Joint Confer-ence on INC, IMS and IDC, 170-175.