

A Proposed Load Balancing Algorithm Over Cloud Computing (Balanced Throttled)



Shereen Yousef Mohamed, Mohamed Hamed N. Taha, Hesham N. Elmahdy, Hany Mohamed Mohieddin Harb

Abstract: Cloud computing refers to the services and applications that are accessible throughout the world from data centers. All services and applications are available online. Virtual machine migration is an important part of virtualization which is considered as essential part in cloud computing environment. Virtual Machine Migration means transferring a running Virtual Machine with all its applications and the operating system state as it is to target destination machine where it continues to run as if nothing happened. It makes balancing between servers. This improves the performance by redistributing the workload among available servers. There are many algorithms of load balancing classified into two types: static load balancing algorithms and dynamic load balancing algorithms. This paper presents the algorithm (Balanced Throttled Load Balancing Algorithm- BTLB). It compares the results of the BTLB with round robin algorithm, AMLB algorithm and throttled load balancing algorithm. The results of these four algorithms would be presented in this paper. The proposed algorithm shows the improvement in response time (75 μ s). Cloud analyst simulator is used to evaluate the results. BTLB was developed and tested using Java.

Keywords (Amlb Algorithm, Balanced Throttled Load Balancing Algorithm, Cloud Analyst Simulator, Cloud Computing, Load Balancing, Round Robin Algorithm, Throttled Load Balancing Algorithm.)

I. INTRODUCTION

Cloud computing is a technology that enables data exchange and provides users with a lot of resources. Users just pay for the services they have used. Archana showed that in an open environment: cloud computing stores data, software packages and distributed resources. Software-as-a-Service (SaaS), Infrastructure -as-a-Service (IaaS), and Platform-as-a-Service (PaaS) are the main categories of cloud computing services. There is no need to install software in any thin clients (PCs, Laptops, iPads and Mobile devices) [1]. It provides on-demand services for users with online applications. There is no worry about backup and restoring data. Sharing resources, information and internet applications are the key functions of cloud computing. The

goal of cloud computing is reducing operating costs, improving data processing time, decreasing computing power requirements for thin devices, enhancing security, and improving the efficiency. It also increases system stability for potential system changes, improving fault tolerance, increasing multiple data processing speed, reducing (hardware, software costs and their maintenance overheads), rationalizing energy and saving disk space.

Green information technology would be used as a default (3R: Recycle, Reuse and Reduce). As a result of increases on a website or business application, the full workload cannot eventually be supported by a single server [1].

Fig 1 (Elaborated from Elmahdy's Cloud Infrastructure Services class notes [2]) shows the main components of a cloud computing system.

The essential cloud characteristics are: On-demand self-service, Broad network access, Resource pooling, Rapid elasticity, and Measured service [2].

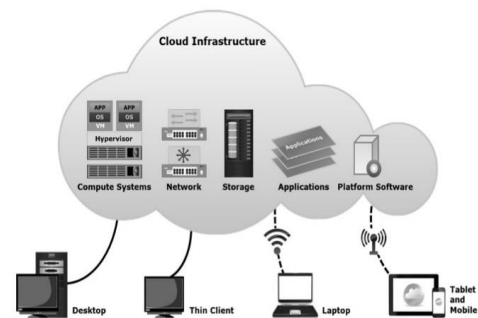


Fig. 1 The components of a cloud computing system.

Organizations are spreading the workload across multiple servers to meet demand. This called "load balancing" that prevents the overloading of a single server. Overloading could cause slowing down, dropping requests, and even crashing. Load Balancing is a method of spreading workload across network connections on multiple servers. It optimizes throughput and minimize overall response time [3]. The main purpose of load balancing is to prevent overloading and possibly crashing any server.

Load balancers are great benefit to cloud environments. Massive workloads can crash servers and get them out of service. High service availability and response time are critical to all processes. Load balancers can also identify disabled servers and redirect traffic to still running servers. Load balancers can determine whether any server is overloaded and redirect traffic to others.

Manuscript received on June 11, 2021.

Revised Manuscript received on June 16, 2021.

Manuscript published on July 30, 2021.

* Correspondence Author

Shereen Yousef Mohamed*, Department Information Technology, Cairo University, Cairo, Egypt. Email: sheren.yousef@must.edu.eg

Mohamed Hamed N. Taha, Faculty of Computers and Artificial Intelligence, Cairo University, Cairo, Egypt. Email: mnasrtaha@fci-cu.edu.eg

Hesham N. Elmahdy, Faculty of Computers and Artificial Intelligence, Cairo University, Cairo, Egypt. Email: ehesham@fci-cu.edu.eg

Hany Harb, Faculty of Information Technology, MUST University, Cairo, Egypt. Email: hany.harb@must.edu.eg

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The main purpose of load balancers is to guarantee healthier servers. A load balancer regulates users and servers. It handles incoming network and application traffic. Load balancer uses different approaches to manage traffic through several servers. Load balancing takes place in two levels: The first level, mapping between requested services or applications and virtual machines. The second level, mapping between virtual machines and hosts.[4]

LIVE VIRTUAL MACHINE MIGRATION

Is the process of moving a running VM from one physical host to another with minimal disruption to ongoing services. It is a powerful tool that facilitates hardware maintenance, load balancing, fault tolerance, and power saving in clusters and data centers.[5]

This paper is composed of six sections. The first section is Introduction. The Second section shows the literature review. The third section introduces related work. The fourth section demonstrates the proposed algorithm. The fifth section discusses.

II. LITERATURE REVIEW

In cloud computing there are many load balancing algorithms. The most used load balancer algorithms are: Round Robin Algorithm (RR), Active Monitoring Load Balancing algorithm (AMLB) and Throttled Algorithm (TLB) [4]. “Cloud Analyst Simulator” is used to evaluate the used algorithms.

A. Load balancers Algorithms:

A.1 Round Robin Algorithm (RR)

It is a simple way to distribute requests from clients across servers. A client request is forwarded to each server in turn. It is the most used algorithm, and it is easy to implement and conceptualize [4].

In this algorithm, requests from clients are routed cyclically to the available servers (Fig. 2). It works good when servers have roughly the same computing and storage capabilities.

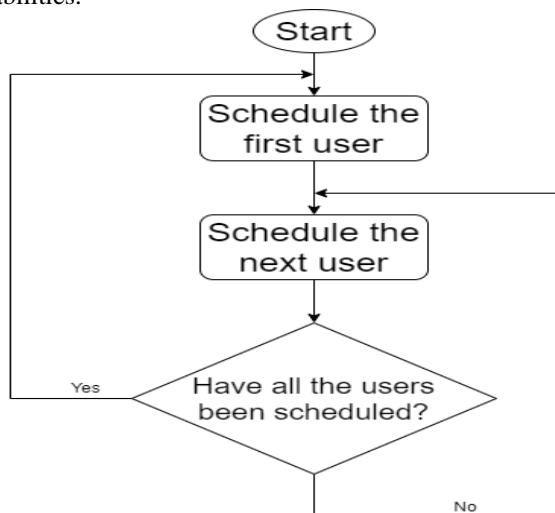


Fig. 2 Round Robin Algorithm (RR).

A.2 Active Monitoring Load Balancing Algorithm (AMLB)

It maintains a virtual machine index table and the number of allocations allocated to each virtual machine. Data Center Controller receives a new task from a client. When a

request from Data Center Controller arrives at AMLB, The AMLB scans the index table from top. When it finds the least loaded VM, AMLB returns VM id to datacenter controller. If there are more than one loaded VM, AMLB selects the first identified [14]. The datacenter controller sends the request to the selected VM (Fig. 3 Elaborated from [9]).

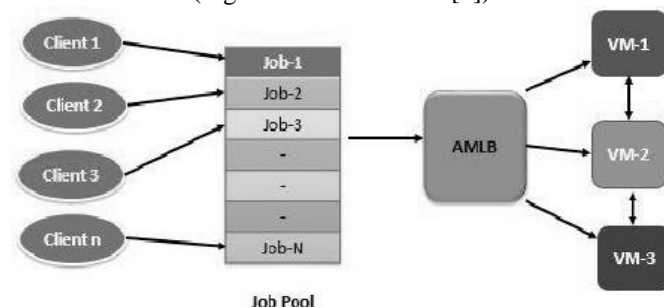


Fig. 3 AMLB Algorithm (AMLB).

A.3 Throttled Load Balancing Algorithm (TLB)

In this algorithm the load balancer responsible for preparing index table. It displays the information of the virtual machine state either Available or Busy. When the task is arrived, the load balancer assigns the task to suitable virtual machine. Which is available to execute the user task (Fig 4 Elaborated from [9]). But every time throttled checks the table from first index to determine the available virtual machine [7].

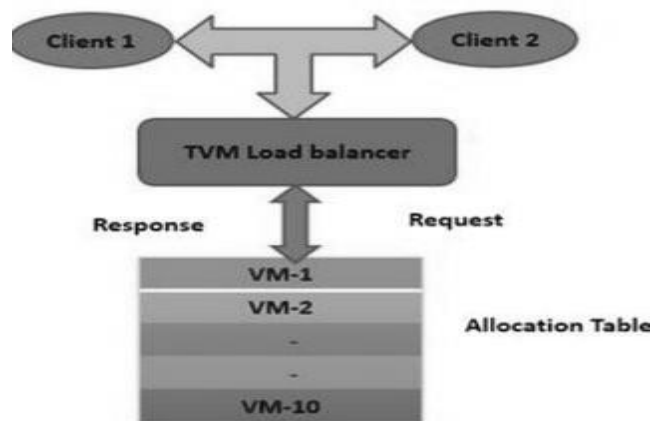


Fig. 4 Throttled Algorithm (TLB).

B. Related work

Different load balancing algorithms for cloud computing have been proposed. To provide an efficient distribution of load between available machines. The paper briefly summarizes some of the essential load balancing algorithms built in the cloud computing environment by various researchers.

- Vibhore Tyagi and Tarun Kumar [15] used throttled load balancing algorithm in multi data center across VM's, and optimized response time. They deduced that the throttled load balancing algorithm provides the best summary overall response time and data center processing time with low processing costs.

- S. A. Narale and P. K. Butey [8] presented an efficient algorithm for load balancing, depending on two different features: One is the response time to serve the requests and other is the load distribution among the existing VMs. they proposed modified throttled algorithm and compared it with Round Robin, Throttled algorithms. And observed that using Round Robin and Modified Throttled algorithm lead to efficient utilization of VMs compared to Throttled algorithm and Modified Throttled algorithm gives better average response time compared to other two algorithms.
- Violetta N. Volkova, Liudmila V. Chernenkaya, Elena N. Desyatirikova, Moussa Hajali, Almothana Khodar and Alkaadi Osama [4] displayed three load balancing algorithms used in cloud computing Round Robin Algorithm (RR), Active Load Balancing Monitoring (AMLB) and Throttled Load Balancing Algorithm (TLB), they compared the results of using the three load balancing algorithms then conclude that in the Throttled load balancing algorithm (TLB) the response time and the data center time is better than in other algorithms.
- Sambit Kumar Mishra, Bibhudatta Sahoo and Priti Paramita Parida [10] have described various load balancing techniques, in varying cloud computing environments, and suggested a classification for the algorithm for load balancing.
- Bhandari A. and Kaur K. [11] proposed a modified throttled balancing technique, used CloudAnalyst tool of CloudSim to implement it, comparing it with other load balancing techniques and validate that a modified throttled load balancing algorithm is more efficient and effective than other conventional algorithms for load balancing.
- Geeta and Shiva Prakash [12] discussed the major challenges facing the cloud computing environment and the development of many load balancing algorithms and various QoS management modeling techniques in cloud systems.
- Wenhong Tian, Minxian Xu, Aiguo Chen, Guozhong Li, Xinyang Wang and Yu Chen [13] compared four open-source simulators, CloudSim, GreenCloud, iCanCloud and CloudSched. These simulators can simulate different layers of Cloud data center instances, from their architectures, simulation process, elements modeling, performance metrics and outputs, then gave detailed comparisons about these simulators. They realized that none of them is completely perfect.

C. Cloud Analyst Simulator

The cloud analyst is a cloud simulator that offers an efficient simulation platform for the deployment of data centers in real time. It was developed to analyze the behavior of large-scale cloud applications under different implementation circumstances. It is a simulator that can be used to model and repeatedly perform simulations in easy way [13]. Cloud Analyst used to analyze number of load balancing algorithms.

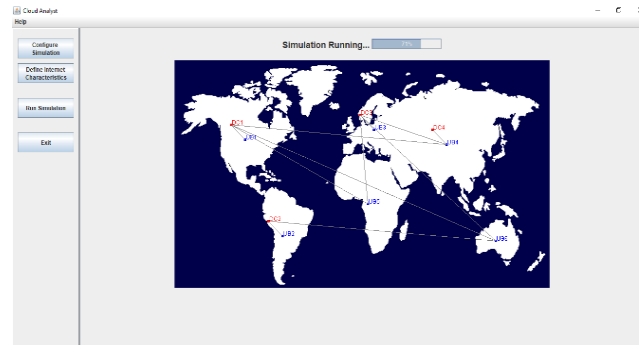


Fig. 5 Cloud Analyst Simulator.

The main features of Cloud Analyst are:

- Cloud analyst provides a graphical user interface that enable users to make experiments more easily.
- Cloud analyst allows easily making experiments with same and different parameters and allowing viewing graphical results.
- Cloud analyst has flexibility and configurability with high degree.

III. PROPOSED ALGORITHM

Balanced Throttled Load balancing algorithm (BTLB)

As a result of the drawback of using throttled load balancing algorithm we proposed the balanced throttled load balancing which work as follow:

1. The index table of all the virtual machines is maintained by Balanced Throttled Load Balancer. This also maintains the state of each virtual machine i.e. whether the virtual machine is busy or available. Initially, at the start of the algorithm, all the virtual machines have been present.
2. Then, Data Center Controller gets the fresh task.
3. Data Center Controller, then on receiving the call, contacts Balanced Throttled Load Balancer to do the next allocation of the virtual machine.
4. Then balanced throttled Load Balancer construct new map and start to add all available VM.
5. Then balanced throttled Load Balancer deconstructs virtual machine available VM map if this map length is greater than 0 then it gets the first available VM Id from the map, then:
 - a) VM id is returned by the Balanced Throttled Load Balancer to the Data Center Controller.
 - b) The Data Center Controller then transfers the call to a respective virtual machine which has been identified by that virtual machine id.
 - c) Then Data Center Controller gives the notification to the Balanced Throttle load balancer about the allotment about the new virtual machine id, then remove this VM Id from map of available VM.
 - d) On receiving the call from the Data Center Controller, Balanced Throttled Load Balancer then upgrades the virtual machines available VM map consequently.

When available virtual machine is not found in the available VM Map then:



A Proposed Load Balancing Algorithm Over Cloud Computing (Balanced Throttled)

- 1 is returned by the Balanced Throttled Load Balancer.
- Then the request is put into the queue by the Data Center Controller.
- When all the processing request is completed by the virtual machine and the response has also been received by the Data Center Controller, Balanced Throttled Load Balancer gets a notification from Data center controller to perform the de-allotment of the respective virtual machine and add this available VM Id to available VM Map.
- Now when the virtual machine is de-allocated, then the Data Center Controller examines the awaiting call queue. When some waiting calls in the pending queue exist, the processing of the call starts from 3rd step onwards.

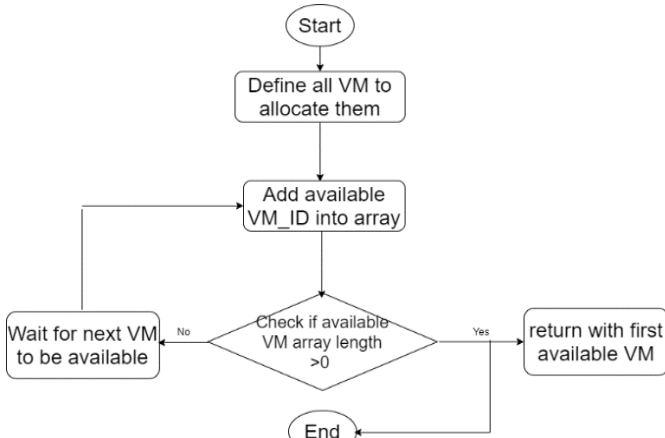


Fig. 6 Balanced Throttled Load balancing algorithm (BTLB)

The best way to test the algorithm is make simulation. We defined 6 user bases that represent the 6 regions of the world and 4 data centers. Each user base has several parameters showed in TABLE I, each data center has 800 VM's.

Table I, User Database Settings

Name	Region	Requests per User per Hr	Data Size per Request (bytes)	Peak Hours Start (GMT)	Peak Hours End (GMT)	Avg Peak Users	Avg Off-Peak Users
UB1	0	60	100	13	15	600000	60000
UB2	1	60	100	15	17	400000	40000
UB3	2	60	100	20	22	400000	40000
UB4	3	60	100	1	3	180000	18000
UB5	4	60	100	21	23	80000	8000
UB6	5	60	100	9	11	90000	9000

Figure 7 shows the parameters of main configure part Which contain User bases, Simulation Duration and Application Deployment Configuration.

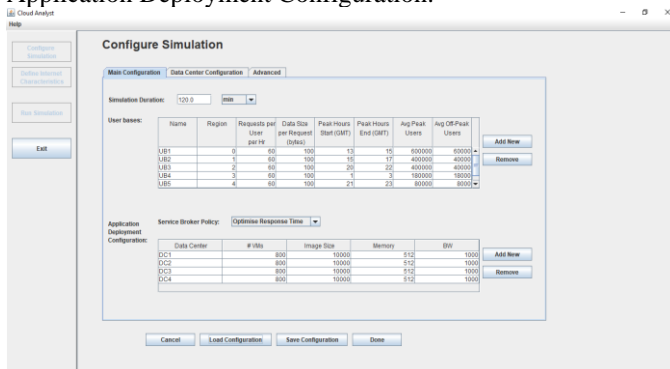


Fig. 7 User and VMs configuration setting

Figure 8 shows the parameters of Data Center Configuration.

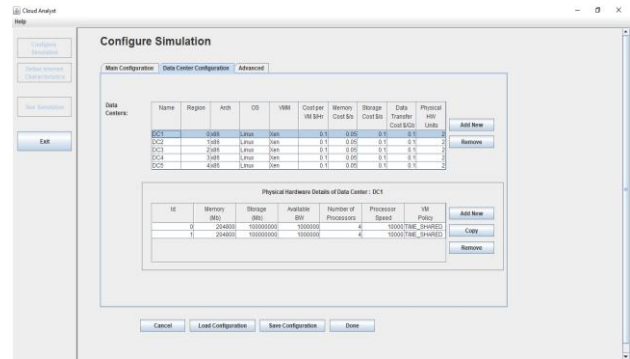


Fig. 8 User and VMs configuration setting

The simulation performed for every load balancing algorithm with the above parameters, then compared the result depending on Overall response time. Fig 7, Fig 8, Fig 9, Fig 10 are the results of Round Robin algorithm, Active Monitoring algorithm, Throttled Algorithms and The Proposed Algorithm, respectively.

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	173.79	48.55	642.74
Data Center processing time:	84.45	9.99	259.13

Fig. 9 Round Robin Algorithm

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	173.94	48.55	652.73
Data Center processing time:	84.78	9.99	259.13

Fig. 10 Active Monitoring algorithm

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	132.97	45.21	590.07
Data Center processing time:	45.22	9.99	242.02

Fig. 11 Throttled Algorithm

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	132.87	45.21	590.07
Data Center processing time:	45.34	9.99	242.00

Fig. 12 Balanced Throttled Algorithm

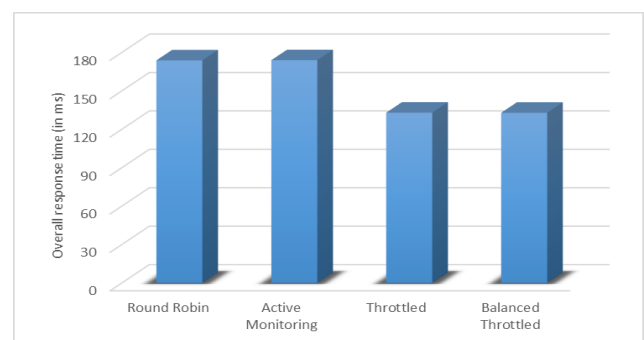


Fig. 13 Overall response time

IV. CONCLUSION

After using these four load balancing algorithms we conclude that creating new load balancing algorithm is important task in cloud computing where load balancing is complex task in cloud computing.



Comparing the above results of simulation using the four algorithms we can conclude that the overall response time of Balanced Throttled load balancing algorithm is better than the overall response time of throttled load balancing and the other algorithms.

REFERENCES

1. R.Archana, "A Survey of Fault Tolerance in Cloud Computing," *The Proceedings of the One Day International Level Seminar on Recent Innovations in Information Technology*, Vol 6, No 1, pp: 98–104, MORAPPUR KONGU COLLEGE OF ARTS AND SCIENCE, Morappur, India, Septemper 2018.
2. Hesham N. Elmahdy, "Cloud Infrastructure Services DELL-EMC2 course – Cloud Computing Master Academic Program," *Lecture Notes, Faculty of Computers and Artificial Intelligence Cairo University, Cairo, Egypt*, Feb. 2020.
3. Saumendu Roy, Dr. Md. Alam Hossain, Sujit Kumar Sen, Nazmul Hossain and Md. Rashid Al Asif, "Measuring the Performance on Load Balancing Algorithms," *Global Journal of Computer Science and Technology: B Cloud and Distributed*, Vol 19, No 1, Version 1.0, Year 2019.
4. Violetta. N. Volkova, L. V. Chemenkaya, E. N. Desyatirikova, M. Hajali, A. Khodar and A. Osama, "Load Balancing in Cloud Computing," *The Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, *IEEE Explorer*, pp: 387-390, doi: 10.1109/EIConRus.2018.8317113, Moscow and St. Petersburg, Russia, 29 Jan.-1 Feb. 2018.
5. Usman Bukar Usman, Faruku Umar Ambursa, Egwan Jessica Onyiyechi and Tanimu Jesse Jareemiah, "Virtual Machine Migration Techniques in Cloud Computing Environment: A Review of Load Balancing Algorithms," *International Journal of Artificial & Computational Intelligence*, Vol1, ISSUE4_03, September 2020.
6. A. Singh and R. Kumar, "Performance Evaluation of Load Balancing Algorithms Using Cloud Analyst," *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp: 156-162, doi: 10.1109/Confluence47617.2020.9058017, Noida, India, 2020.
7. Hota A., Mohapatra S., Mohanty S., "Survey of Different Load Balancing Approach-Based Algorithms in Cloud Computing: A Comprehensive Review," *Computational Intelligence in Data Mining. Advances in Intelligent Systems and Computing, Springer, Singapore*, ISBN 978-981-10-8055-5, Vol 711, pp:99-110, doi.org/10.1007/978-981-10-8055-5_10, India, 04 July 2018.
8. S. A. Narale and P. K. Butey, "Throttled Load Balancing Scheduling Policy Assist to Reduce Grand Total Cost and Data Center Processing Time in Cloud Environment Using Cloud Analyst," *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp: 1464-1467, doi: 10.1109/ICICCT.2018.8473062, Coimbatore, India, 2018,
9. Pradhan A., Bisoy S.K., Mallick P.K., "Load Balancing in Cloud Computing: Survey," *Innovation in Electrical Power Engineering, Communication, and Computing Technology. Lecture Notes in Electrical Engineering, Springer, Singapore*, ISBN978-981-15-2305-2, Vol 630, pp: 99-111, doi.org/10.1007/978-981-15-2305-2_8, India, 22 February 2020.
10. Sambit Kumar Mishra, Bibhudatta Sahoo and Priti Paramita Parida, "Load balancing in cloud computing: A big picture" *Journal of King Saud University – Computer and Information Science*, Vol 32, No 2, pp: 149-158, February 2020.
11. Bhandari A. and Kaur K. "An Enhanced Postmigration Algorithm for Dynamic Load Balancing in Cloud Computing Environment," *The Proceedings of the International Ethical Hacking Conference 2018. Advances in Intelligent Systems and Computing, Springer, Singapore*, ISBN 978-981-13-1543-5, vol 811, pp: 59-73, doi.org/10.1007/978-981-13-1544-2_6, Kolkata, India, 05 October 2018.
12. Geeta, Prakash S. "A Literature Review of QoS with Load Balancing in Cloud Computing Environment," *Advances in Intelligent Systems and Computing, Big Data Analytics*, ISBN 978-981-10-6619-1, Vol 654, pp: 667-675, Springer, doi.org/10.1007/978-981-10-6620-7_64, October 2017.
13. Wenhong Tian, Minxian Xu, Aiguo Chen, Guozhong Li, Xinyang Wang and Yu Chen, "Open-source simulators for Cloud computing: Comparative study and challenging issues" *Simulation Modelling Practice and Theory, ELSEVIER*, Vol 58, pp: 239-254, 2015.
14. Singh A.N. and Prakash S. "WAMLB: Weighted Active Monitoring Load Balancing in Cloud Computing," *Big Data Analytics, Advances in Intelligent Systems and Computing, Springer*, ISBN 978-981-10-6619-1, Vol 654, pp: 677-685, doi.org/10.1007/978-981-10-6620-7_65,2018.
15. Vibhore Tyagi, Tarun Kumar, "ORT Broker Policy: Reduce Cost and Response Time Using Throttled Load Balancing Algorithm", *Procedia Computer Science*, ISSN 1877-0509, Vol 48, pp: 217-221, doi.org/10.1016/j.procs.2015.04.173.2015.

AUTHORS PROFILE



Shereen Youssef, earned her B.Sc. from the Department of Information Technology, Helwan University, Egypt. Currently she is a teaching assistant in Faculty of Information Technology, Misr University for Science & Technology. She is pursuing her M.Sc. degree in Information Technology Department, Faculty of Computers and Artificial Intelligence Cairo University, Egypt.



Mohamed Hamed N. Taha, Assistant Professor of Information Technology, Faculty of Computers and Artificial Intelligence. Ph.D. Cairo University, Faculty of Computers and Information Dissertation: "Improving QoS of Data Transimission over Wireless Networks". (2013). MSc, Cairo University, Faculty of Computers and Information Specialization in Multimedia Security (2009) Thesis: "Subject QoS for Voice over IP". BSc, Cairo University, Faculty of Computers and Information, Department of Information Technology (May. 2006)



Hesham N. Elmahdy, received his B.Sc. in Automobile Engineering with honor degree in the Military Technical Collage, Cairo in 1981. He received his first M. Sc. in Computer Science from the Institute of Statistical Studies & Research, Cairo University, Cairo in 1992. He received his second M.Sc. in Computer Science from the University of Mississippi in August 1996. He received his Ph D in Computer Science from the University of Mississippi in December 1997. He is a Professor at the Information Technology Department in Faculty of Computers and Information, Cairo University. Dr. Eng. Elmahdy was nominated to get the "King Faisal International" prize in Islamic Studies in 1993. He was elected to join the Upsilon Pi Epsilon Chapter of the University of Mississippi in April 1995. He has been joined to The Order of The Engineer of USA since May 1997. He has been included in the 2006-2007 (9th.) Edition of Who's Who in Science and Engineering, the Outstanding Scientists of The 21st Century Cambridge UK, 2007, and in the 2009 (26th.) Edition of Who's Who in The World. He got many prizes: The Best Invotive Idea to Develop Cairo University in 2008, Supervising the Best M.Sc. In FCI Cairo University in 2010, and Ideal Faculty Member Medal 2011. His current research interests are Networks, Distance Learning, and Multimedia. Hesham was honored "The Best Professor of Information Technology in Africa" award from The African Education Leadership Awards in Dec, 2012. His name was included many times for his contributions in: the 2006-2007 (9th.) Edition of Who's Who in Science and Engineering, the Outstanding Scientists of The 21st. Century, Cambridge, UK, 2007, and the 2009 (26th.) Edition of Who's Who in The World. He was selected to be the Professor of the Year in 2011 and in 2012. He was awarded and nominated for many prizes (locally and internationally). Hesham was selected as THE BEST INFORMATION TECHNOLOGY PROFESSOR OF AFRICA by THE AFRICA EDUCATION LEADERSHIP AWARDS | 12th DECEMBER 2012 | MAURITIUS.





Hany Harb, Dean of Information Technology Faculty, MUST University (Misr University for Science and Technology). Doctor of philosophy (Ph.D.), Computer Science, Illinois Institute of Technology (IIT), Chicago, Illinois, USA, 1986. Master of Science in Operations Research (MSOR), IIT, 1987. Master of Science in Computers and Systems Engineering, Faculty of

Engineering, Azhar University, Cairo, Egypt, 1981. Bachelor of Science in Computers and Control Engineering, Faculty of Engineering, Ain Shams University, Cairo, Egypt, 1978. He got many Awards: The best computer science faculty member in faculty of science emirates university 2004–2005, Oracle Data warehousing certificate-Project Manager award from azhar university (2009, 2010).