UNIVERSITI PUTRA MALAYSIA

AN OPTIMAL TASKS SCHEDULING ALGORITHM BASED ON QOS IN CLOUD COMPUTING NETWORK

MOHAMMED AMEEN MOHAMMED ABDO ALHAKIMI

FSKTM 2018 36

AN OPTIMAL TASKS SCHEDULING ALGORITHM BASED ON QOS IN CLOUD COMPUTING NETWORK

By

MOHAMMED AMEEN MOHAMMED ABDO ALHAKIMI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Master of Computer Science
2017

DEDICATIONS

This thesis is dedicated to my parents and my wife who have supported me all the way since the beginning of my studies.

In addition, this thesis is dedicated to all those who believe in the richness of learning.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia, in fulfillment of the requirement for the degree of Master of Computer Science

AN OPTIMAL TASKS SCHEDULING ALGORITHM BASED ON QOS IN CLOUD COMPUTING NETWORK

By

MOHAMMED AMEEN MOHAMMED ABDO ALHAKIMI

Chair: Rohaya Binti Latip

Faculty: Computer Science and Information Technology

Cloud Computing has emerged as a service model that offers online accessible resources to the clients. These resources contain storage, servers, and other applications and it provides security, flexibility, and scalability. In Max-Min algorithm where the large tasks have their priority to be scheduled first, this leads small tasks to stay longer in the queue until all large length tasks finished their execution.

This study presents an optimal task scheduling algorithm by enhancing Max-Min and TS algorithm. Our proposed algorithm isolates the resources into two different groups where the first group contains the resources with maximum execution time while the second group contains the resources with minimum execution time. The main idea here is to choose the resource that takes less time to execute the selected job/task. Therefore, if the resource is from the first group then map the average length task to it and if the choosing resource is from the second group, then map the largest length task to it.

The simulation tool used for testing the algorithm is WorkflowSim. We tested averages of execution time span of the proposed algorithm for 10 running times with 200-1000 tasks in 50 or 100 VMs. Test results show that the proposed algorithm represents enhanced resource utilization with better execution time.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk Ijazah Komputer Sains

AN OPTIMAL TASKS SCHEDULING ALGORITHM BASED ON QOS IN CLOUD COMPUTING NETWORK

Oleh

MOHAMMED AMEEN MOHAMMED ABDO ALHAKIMI

Pengerusi: Rohaya Binti Latip

Fakulti: Computer Science and Information Technology

Pengkomputeran Awan telah muncul sebagai model perkhidmatan yang menawarkan sumber yang boleh diakses secara online kepada pelanggan. Sumber-sumber ini mengandungi storan, pelayan, dan aplikasi lain dan menyediakan keselamatan, fleksibiliti dan skala. Dalam algoritma Max-Min di mana tugas-tugas besar mempunyai keutamaan mereka untuk dijadualkan terlebih dahulu ini membawa tugas-tugas kecil untuk berada lebih lama dalam barisan sehingga semua tugas besar selesai pelaksanaan mereka.

Kajian ini membentangkan algoritma penjadualan tugas yang optimum dengan meningkatkan algoritma Max-Min. Dalam algoritma yang dicadangkan ini, sumber akan dibahagikan kepada dua kumpulan yang berlainan di mana kumpulan pertama mengandungi sumber dengan masa pelaksanaan maksimum manakala kumpulan kedua mengandungi sumber dengan masa pelaksanaan minimum. Idea utama di sini adalah untuk memilih sumber yang memerlukan sedikit masa untuk melaksanakan tugas yang dipilih. Oleh itu, jika sumber adalah dari kumpulan pertama, kemudian peta tugas panjang rata-rata kepadanya dan jika sumber pilihan adalah dari kumpulan kedua, kemudian peta tugas terbesar kepadanya.

Alat simulasi yang digunakan untuk menguji algoritma adalah WorkflowSim. Kami menguji purata Makespan algoritma yang dicadangkan untuk 10 kali dengan 200-1000 tugas dalam 50 atau 100 VMs. Hasil ujian menunjukkan bahawa algoritma yang dicadangkan mewakili penggunaan sumber yang lebih baik dengan makespan yang lebih baik.

ACKNOWLEDGMENTS

I managed to finish this thesis due to the effort of a number of people who have always given their valuable advice or lent a helping hand. I would like to take this opportunity to extend my sincere gratitude to those who have helped me directly or indirectly in preparing my final thesis.

Firstly, I would like to express my appreciation to my supervisor, **Assoc. Prof. Dr. Rohaya Binti Latip** for her support and guidance in completing this thesis through various stages. She has given me many sources of ideas and references for this thesis. Her brilliant ideas and suggestions to improve my master are much appreciated.

I would also like to thank my Assessor, **Dr. Masnida Hussin** for her suggestions and comments throughout the whole process of developing my thesis. She has given me helpful suggestions in order to enhance my thesis. I appreciate the concern and suggestions from her.

Moreover, I am glad to have such supportive **friends** who would help me when I have difficulties developing my implementation. Their time and effort are much appreciated. Not to forget my beloved **wife and family**, who always give their moral support and advises which motivates me to complete my studies.

TABLE OF CONTENTS

LIST OF FIGURE

FIGURE **PAGE**

LIST OF TABLES

LIST OF ABBREVIATIONS

QoS	Quality of Service
SooS	Software as a Service
PaaS	Platform as a Service
IaaS	Infrastructure as a Service
OTS	Optimal tasks scheduling
TS	Task scheduling
GTS	Grouped tasks scheduling
DLT	Divisible Load Theory
PCSO	Parallel Cat Swarm Optimization
MET	Minimum execution time
MCT	Minimum Completion Time
OLB	Opportunistic Load Balancing
RASA	Resource Aware Scheduling Algorithm
O.S	Operating System
VM	Visual Machine
MIPS	Million instructions per second

CHAPTER 1

INTRODUCTION

1.1. BACKGROUND

The concept of Cloud Computing leads many business and organizations to turn toward using this technique. There is nothing essentially new in any of the technologies that make up cloud computing as most of these technologies have been used before. Numerous experts in the academic field have attempted to define precisely what "cloud computing" is and what unique attributes that cloud computing presents. Cloud computing is an addition of different techniques such as parallel, distributed, and grid computing, it is a recent model for enabling access to a shared computing, storage resources and services which can be accessed through the internet using a set of applications. Virtualization, dispersal and dynamic extensibility are the basic features of this environment (Buyya et al.; 2009).

"Software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS)" are three dissimilar services where it can be provided by cloud computing. Clients can submit their works into the cloud for computational transmutation or rather leave the data in the cloud for future use. Therefore, different users may have different requirements. The requirement in cloud computing environment is to schedule the current jobs/tasks with the given constraints. The meaning of constrains here is applying Quality of Service that users need and balancing between these QoS and fairness among the tasks (Hend et al. 2016).

In addition, to satisfy the requirement of this dynamic environment, many researches have worked within this area and many algorithms have proposed which meet this requirement. An optimal tasks scheduling algorithm is one of the algorithms that execute the tasks based on its

attribute, average length and long length tasks then schedule the tasks with the resource that will takes less execution time.

The focus of this research is to implement an algorithm for tasks scheduling in cloud computing environment that can solve the problem of time execution for all tasks to satisfy users need.

1.2. DESCRIPTION OF THE RESEARCH PROBLEM

In cloud computing where service and storage can be access by the client, Clients can submit their tasks into the cloud for computation processing purpose or rather for future storage (Monir et al. 2013). Those clients who is using this service have different requirements, this requirement knows as Quality of Service, which is a cooperative effort of service performance that defines the degree of gratification of a user for the service (Hend et al. 2016). The dynamic nature of this environment can endure a big challenge as some assigned resources might overloaded during the execution that may cause the failure of the whole execution process (Khadija et al. 2016). Hence, Scheduler in cloud must have the ability to schedule clients' tasks/jobs in such a way that cloud provider can earn maximum advantage for his service (Monir et al. 2013). In Max-Min algorithm where the large tasks have their priority to schedule first this leads small tasks to stay longer in the queue until all large length tasks finished their execution. Thus, mapping this task/job to the resource should take it in the consideration in order to increase the performance of this environment.

1.3. THE OBJECTIVES OF THE STUDY

With the increasing number of users over the cloud, scheduling a huge tasks still a critical issue. In the world of web tremendously large number of tasks from various aspects are processed which

may affect the performance of the cloud, and as a result choosing the optimal scheduling algorithm may play an important role to avid this problem. Therefore, to solve the scheduling issue of tasks, several researches have work where they focused in how to implement an algorithm for tasks scheduling in cloud computing that can solve the problem of time execution for all tasks to satisfy users need. The main aims of this study as the following:

1.3.1 To propose a tasks scheduling algorithm called optimal tasks scheduling OTS in cloud computing network that gets minimum execution time to all tasks.

1.3.2 To analyze the algorithm in different number of jobs and different number of services which may lead to have different perspectives.

1.4. RESEARCH QUESTIONS

In this thesis, there are many questions that need to be answered, as the following:

1.4.1. How does the system determine the task that has priority when the user used the same QoS?

1.4.2. How does the system determine the task that has a higher priority compared with another task?

1.4.3. How does the system determine the task that is important, when the QoS is equal?

1.5. RESEARCH SCOPE

In this thesis, the study was limited into the following:

1.5.1. Allocate tasks into services with driven QoS.

1.5.2. Two attributes will be applying in the algorithm (long and average length tasks).

1.5.3. The algorithm can work with independent tasks.

1.6. THESIS ORGANIZATION

This thesis comprises five chapters. This **chapter** provides a basic introduction of cloud computing and briefly presents the motivation for the study, its goals and objectives, and research focus. **Chapter 2** explains the main concepts of cloud computing through the background that includes tasks scheduling technique. In addition, it contains Related Work that presents a study of various existing tasks scheduling algorithms and associated work. **Chapter 3** contains Methodology that describes proposed tasks scheduling algorithm, discrete event simulation, and design specification. **Chapter 4** presents the Implementation that presents all the information and steps regarding proposed scheduling algorithm implementation. Moreover, this chapter presents the result and discussion of all experiments with analyzing the differences between them. **Chapter 5** concludes this thesis. It summarizes and further elaborates the implications of the findings from Chapter 4. In addition, the chapter identifies future work and possible next steps in research.

REFERENCES

R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems,25:599-616, 2009

Selvarani S, Sudha Sadhasivam G. Improved cost-based algorithm for task scheduling in cloud computing. In: International conference. IEEE; 2010.

Gamal El Din Hassan Ali Hend, Saroit Imane Aly, Kotb Amira Mohammed. Grouped tasks scheduling algorithm based on QoS in cloud computing network. In: Egyptian Informatics Journal. eij; 2016.

Abdullah Monir, Othman Mohamed. Cost-based multi-QOS job scheduling using divisible load theory in cloud computing. In: International conference on computational science. ICCS; 2013.

Bousselmi Khadija, Brahmi Zaki, Gammoudi Mohamed Mohsen. QoS-aware scheduling of Workflows in CloudComputing environment. In International Conference on Advanced Information Networking and Applications. IEEE; 2016

Wu Xiaonian, Deng Mengqing, Zhang Runlian, Zeng Bing, Zhou Shengyuan. A task scheduling algorithm based on QoS-driven in cloud computing. In: International conference on information technology and quantitative management. ITQM 2013.

Liu Gang, Li Jing, Xu Jianchao. An Improved Min-Min Algorithm in Cloud Computing. In: Proceedings of the 2012 international conference of modern computer science and applications, Zhenyu Du; 2013.

Selvarani S, Sudha Sadhasivam G. Improved cost-based algorithm for task scheduling in cloud computing. In: International conference. IEEE; 2010.

Mandal Tripti, Acharyya Sriyankar. Optimal Task Scheduling in Cloud Computing Environment: Meta Heuristic Approaches. In Proceedings of International Conference on Electrical Information and Communication Technology (EICT 2015).

Razaque Abdul, Vennapusa Nikhileshwara Reddy, Soni Nisargkumar. Task Scheduling in Cloud Computing. In Long Island Systems, Applications and Technology Conference (LISAT), 2016 IEEE.

Hemamalini, M. and Srinath, M. Memory constrained load shared minimum execution time grid task scheduling algorithm in a heterogeneous environment, Indian Journal of Science and Technology 8(15), 2015.

Etminani, K., Naghibzadeh, M. and Yanehsari, N. R. A hybrid min-min max-min algorithm with improved performance, Department of Computer Engineering, University of Mashad, 2009.

Moharana, S. S., Ramesh, R. D. and Powar, D. Analysis of load balancers in cloud computing, International Journal of Computer Science and Engineering 2(2): 101-108, 2013

Maheswaran, M., Ali, S., Siegel, H. J., Hensgen, D. and Freund, R. F. Dynamic mapping of a class of independent tasks onto heterogeneous computing systems, Journal of parallel and distributed computing 59(2): 107-131, 1991.

Salot, P. A survey of various scheduling algorithm in cloud computing environment, International Journal of research and engineering Technology (IJRET), ISSNpp. 2319-1163, 2013.

Sindhu, S. Task scheduling in cloud computing, International Journal of Advanced Research in Computer Engineering Technology 4, 2015.

Parsa, S. and Entezari-Maleki, R. Rasa: A new task scheduling algorithm in grid environment, World Applied sciences journal 7: 152-160, 2009.

Topcuoglu, H., Hariri, S. andWu, M.-y. Performance-effective and low-complexity task scheduling for heterogeneous computing, IEEE transactions on parallel and distributed systems 13(3): 260-274, 2002.

Bittencourt, L. F., Sakellariou, R. and Madeira, E. R. Dag scheduling using a lookahead variant of the heterogeneous earliest finish time algorithm, 2010 18thEuromicro Conference on Parallel, Distributed and Network-based Processing, IEEE,pp. 27-34, 2010.

Rajwinder Kaur, P. L. Load balancing in cloud system using max-min and min-min algorithm, National Conference on Emerging Trends in Computer Technology(NCETCT-2014) pp. 31-34, 2014.

Sviji, K. M. A. Resource management system in cloud environment: An overview, International Journal of Advanced Research in Biology, Engineering, Science and Technology 2: 357-362, 2016.

Warneke, D. and Kao, O. Exploiting dynamic resource allocation for efficient parallel data processing in the cloud, IEEE transactions on parallel and distributed systems 22(6): 985-997, 2011.

Cao, Q., Wei, Z.-B. and Gong, W.-M. An optimized algorithm for task scheduling based on activity-based costing in cloud computing, 2009 3rd International Conference on Bioinformatics and Biomedical Engineering, IEEE, pp. 1-3, 2009.

Sagar, M. S., Singh, B. and Ahmad, W. Study on cloud computing resource allocation strategies, International Journal 1(3): 107-114, 2013.

Priyadarsini, R. J. and Arockiam, L. Performance evaluation of min-min and max-min algorithms for job scheduling in federated cloud, International Journal of Computer Applications (0975-8887) Volume, 2014.

Tilak, S. and Patil, D. A survey of various scheduling algorithms in cloud environment, International Journal of Engineering Inventions 1(2): 36-39, 2012.

Calheiros, R. N., Ranjan, R., Beloglazov, A., De Rose, C. A. and Buyya, R .Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms, Software: Practice and Experience41(1): 23-50, 2011.

Chen, W. and Deelman, E. Workflowsim: A toolkit for simulating scientific workflows in distributed environments, E-Science (e-Science), 2012 IEEE 8th InternationalConference on, IEEE, pp. 1-8, 2012.