



***PERFORMANCE EVALUATION OF TASK SCHEDULING USING HYBRID
META-HEURISTIC IN THE HETEROGENEOUS CLOUD ENVIRONMENT***

FATEN AMEEN SAIF MOHAMMED

FSKTM 2019 32



**Performance evaluation of task scheduling using Hybrid
meta-heuristic in the heterogeneous cloud environment**

By

Faten Ameen Saif Mohammed

Thesis Submitted to the School of Graduate Studies, University Putra
Malaysia, in Fulfillment of the Requirement for the Degree of Master of
Computer science

January 2018

COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of University Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of University Putra Malaysia.

Copyright ©University Putra Malaysia

Abstract of thesis presented to the Senate of University Putra Malaysia in
Fulfilment of the Requirement for the Degree of Master of Computer
Science

**Performance evaluation of task scheduling using Hybrid
meta-heuristic in the heterogeneous cloud environment**

By

FATEN AMEEN SAIF MOHAMMED

JANUARY 2019

Supervisor : Mohd. Noor Derahman

Faculty: Computer Science and Information Technology

Cloud computing is a platform in which it provides services, information and software over the Internet. The essential role of cloud computing is enabling sharing of resources on-demand over the network (e.g. servers, applications, storage, services and database) to the end-users that are distributed geographically. Task scheduling is a significant function in the cloud computing that plays a vital role to raise the rate of efficiency and the performance of the system. Task scheduling is considered as an NP-complete problem. However, the heterogeneity of resources in the cloud environment put the scheduling in a critical issue. Furthermore, heuristic algorithms do not have the required level of efficiency to optimize the scheduling and the performance in this environment. Thus, this study focuses on optimizing the

hybrid meta-heuristic (genetic algorithm along with DE algorithm that minimizes the completion time and enhances the performance of the task scheduling. The results will be compared with a three heuristic algorithms. The performance evaluation in this work is a statically analysis that used in an experimental comparison. The expected result of this study is optimizing the overall of completion time and enhancing resource efficiency.



APPROVAL FORM

This thesis was submitted to the Senate of University Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Computer science. The members of the Supervisory Committee were as follows:

Mohd. Noor Derahman, Mr

Faculty of computer science and IT technology

University Putra Malaysia

(Supervisor)

Nor Asila Wati Abdul Hamid, PhD

Faculty of computer science and IT technology

University Putra Malaysia

(Assessor)

Abu Bakar Md. Sultan, PhD

Dean

School of Graduate Studies

University Putra Malaysia

Date:

DECLARATION

Declaration by a graduate student

I hereby confirm that:

- This thesis is my original work;
- Quotations, illustrations and citations have been duly referenced;
- This thesis has not been submitted previously or concurrently for any other Degree at any other institutions;
- Intellectual property from the thesis and copyright of thesis are fully-owned by University Putra Malaysia, as according to the University Putra Malaysia. (Research) Rules 2012;
- Written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the University Putra Malaysia (Research) Rules 2012;
- There is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the University Putra Malaysia (Graduate Studies) Rules 2003(Revision 2012-2013) and the University Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: Faten Ameen Saif Mohammed (GS48308)

ACKNOWLEDGEMENT

Firstly, all praises and gratitude go to the almighty Allah, whose favors are never ending for me, He is the one to give me all the abundances in my life to be become the person that I am today. After Allah, and what He has given me during my lifetime, I should not forget to mention those whom Allah put in my way to help me and support me. This thesis would not be completed without the help and supervision of my beloved supervisor Mr. Mohammed Noor Derahman, who provided me with support, time, encouragement, unending guidance, and productive discussion.

I would like to thank eternally my father who was the backbone for me and thank him for his encouragement, support, help, and his spiritual and substantial support to continue and complete my master journey. My step mother who deserves to be my mother my best friend for what she did to support me. My sisters and my brothers deeply thanks from my heart for sharing my happiness and sadness moments. Khaled AL-Madani, I cloud not find the suitable words to express my grateful for what you did and the situations that prove I have gained real brother.

Even more, MR. Yousef AL-Omran my teacher and who believes on me and has predicted my future. Thanks for your continuously supporting, and encouraging me specially when I break down in the weakness. My success is your success. Thanks from the bottom of my heart.

No doubt that they deserve all my thanks because they are the reason for me to be here today. In the end, May Allah bless my mother's spirit, wellbeing and in the end the highest level of Jannah.

DEDICATION

I am dedicating this thesis to my family and all my friends. Starting with the utmost gratitude to my father who never failed to give me all types of support. My sisters and my brothers who never forget to support me all along.

I likewise dedicate this thesis to many of my friends who always had been by my side whenever I needed them. I will not forget what was given to me by them either morally or physically, especially my brother Awadh Bajaher for his cooperation since I have studied in the faculty.

Even more, Dr.Ali Aljuboori for empowering me to get a hand on my paper and improve my work on CloudSim for helping me understand the algorithm details in spite of his busyness with his work.

Thanks to all of you

Thank you for being a huge part of this work

List of Abbreviations

VM	Virtual Machine
DCB	Data Center Broker
MIP	Million Instruction per Second
Bw	Bandwidth
QoS	Quality of Service
VMs	Virtual Machines
MI	Million Instruction
CIS	Cloud Information Service
DC	Data Center
GA-DE	Genetic algorithm & Differential Evolution
DAG	Directed acyclic graph
HEFT	Heterogeneous Earliest Finish Time
EST	Earliest Start Time
EFT	Earliest Finish Time

List of figures

Figure	page
3.1 Cloudsim structure and entities	12
3.2: The Cybershake workflow	14
3.3: The Cybershake workflow	14
3.4: The montage workflow	15
4.1: The proposed simple directed acyclic graph	21
4.2: Flowchart of the proposed algorithm	22
4.3: the crossover operator	24
4.4: the mutation operator	25
5.1: Bar graph representing the make-span for various task	29
5.2: Bar graph representing the make-span on various scientific workflows on GA-DE	30

List of Tables

Table	Page
Table 2.1: scheduling objectives	7
Table 7.1: Simulation Parameters	17
Table 4.1: Task priorities	23
Table 5.1: The Result of Total Executed Jobs	27
Table 5.2: The Result of Total Executed scientific workflows on GA-DE	29

Table of Contents

COPYRIGHT	iii
Abstract.....	iv
APPROVAL FORM	vi
DECLARATION	vii
ACKNOWLEDGEMENT	viii
List of Abbreviations.....	x
List of Figures	xi
List of Tables.....	xi
1 INTRODUCTION	1
1.1 Background.....	1
1.2 Problem Statement	3
1.3 Objective of Research	3
1.4 Scope of Research	3
1.5 Thesis Organization	4
2 LITERATURE REVIEW	5
2.1 Introduction.....	5
2.2 Task Scheduling.....	5
2.3 Task scheduling algorithm.....	5

2.3.1	Heuristic algorithm.....	7
2.3.	Meta-Heuristic algorithm.....	8
2.3.3	Hybrid strategy	9
3	METHODOLOGY.....	11
3.1	Introduction	11
3.2	The CloudSim Model	11
3.2.1	Cloud Information Service (CIS)	11
3.2.2	Data center (DC)	11
3.2.3	The Datacenter Broker (DCB).....	12
3.2.4	Host	12
3.2.5	Virtual machine (VM)	12
3.2.6	Cloudlet	12
3.2.8	CloudSim model:.....	12
3.3	The Workflow.....	13
3.3.1	Business Workflow.....	13
3.3.2	Scientific Workflows	13
3.3.2.1	Cybershake Workflow.....	13
3.3.2.2	Montage Workflow	14
3.3.2.3	Epigenomics Workflow.....	15
3.4	Discrete Event Simulation (DES).....	15
3.5	Datasets	16

3.5.1 The parameters used.....	17
3.5.2 Workload.....	17
4 IMPLEMENTATION	18
4.1 Introduction	18
4.2 implementation of the GA-DE algorithm	18
4.2.1 HEFT-Upward Rank	18
4.2.2 HEFT-Downward Rank.....	19
4.2.3 HEFT-Level Rank	19
4.2.4 Genetic Algorithm-Differential Evolution Method	20
4.2.4.1 Production of initial population	22
4.2.4.2 Make-span.....	23
4.2.4.3 Fitness Function.....	23
4.2.4.4 Selection Operator	23
4.2.4.5 Crossover Operator	24
4.2.4.6 Mutation Operator	25
4.2.4.7 Termination Conditions	25
4.3 The Performance Evaluation	26
5 RESULTS AND DISCUSSION.....	27
5.1 Introduction	27
5.2 Success Rate	27
5.3 Evaluate the performance of GA-DE.....	29

CONCLUSION	31
References	32
6 Appendix A	34



© COPYRIGHT UPM

CHAPTER 1

INTRODUCTION

1.1 Background

Cloud computing is a new technology that facilitates sharing resources to the end user over the Internet. The core function of the cloud computing is the scheduling which it mapping tasks to proper resources in polynomial time with maintaining the desired quality of service (QoS) that is the total achievement of service performance that defines the level satisfaction of the service-user . Scheduling is considered as an NP-complete problem especially in the large or complex task. This type of problem needs an approximate solution with the existing of the constraints to optimize the objectives of the scheduling (Bhowmik,2017) (e.g. Reduce the completion time, minimize the communication cost, maximize the throughput, decrease the energy consumption, fault tolerance , resources utilization, deadline , Laxity , load balancing and tardiness). Generally, there are two common classification algorithms for scheduling the task in the heterogeneous resources in the cloud system (Gupta & Garg, 2017): Heuristic algorithms that provides solution almost optimal solution (e.g. Minimum completion time MCP, HEFT, Graham algorithm, POP etc.) And Meta-Heuristic that gain the popularity to solve NP-hard problem and provides good solution with less effort of computational that make it proper for large or complex task (e.g. Ant colony optimization, Genetic algorithm, Particle Swarm optimization, Honeybee etc.) Unfortunately, the heterogeneous resources in the cloud environment have made a challenge for scheduling the tasks and it has become more

complicated. The common heuristic and meta-heuristic don't have the required efficiency to schedule the task in this environment. The critical issue attracts the attention of the researchers to propose several hybrid Meta-Heuristic to find an optimal solution that enhances the performance of the system. This method has become the recent trend that gathers two or more of the meta-heuristic algorithm to utilize the strength of them that leads to provide a better optimal solution.

The work in this study refers to discuss the issue of task scheduling in heterogeneous resources in cloud computing. The proposed we evaluate the hybrid meta-heuristic method (genetic algorithm along with DE algorithm) by using the scientific workflow as a data set. After implementing phase the proposed algorithm compared with the Genetic algorithm to investigate the effectiveness rate.

The rest of the work consists of subsections, problem statement, objectives and scope of the work. While the main sections, the literature review, methodology, implementation, results and analysis and conclusion.

1.2 Problem Statement

The heterogeneous cloud environment does not obtain the satisfying solution for task scheduling by utilizing the traditional heuristic and meta-heuristic algorithms specifically with the large-scale or complex task.

The researcher employs the methods to improve and achieve better results by proposing a hybrid meta-heuristic to provide a maximum rate of efficiency and approximate optimal solution. However, the existing study has not used the standard dataset to verify the effectiveness of the algorithm performance.

In this study, we evaluate the performance of the proposed algorithm (hybrid meta-heuristic) by using the scientific workflow as a data set.

1.3 Objective of Research

Evaluate the performance effectiveness of the proposed algorithm (GA-DE) to obtain the minimum make-span for task scheduling in the heterogeneous cloud environment by using the scientific workflows (Cybershake, Montage, and Epigenomics).

1.4 Scope of Research

The current focus of the study is the case of task scheduling in heterogeneous cloud computing. Even more, evaluate the effectiveness of hybrid meta-heuristic (genetic algorithm along with DE algorithm) by comparing with other heuristics algorithms. Furthermore, the evaluation will be implemented using a discrete event cloud simulator based on the CloudSim framework for modeling and cloud computing simulation.

1.5 Thesis Organization

The thesis is organized as the following:

Chapter 1, Introduction that elaborates the cloud computing environment, task scheduling and the summary of work.

Chapter 2, the literature review focuses on explaining more about the cloud and task scheduling algorithm. Also, it includes a critically reviewed related work to this study.

Chapter 3, The methodology illustrated more about the definition of the problem, the algorithm, the method used, the parameters and workloads included as well as the simulator used.

Chapter 4, Implementation which shows how algorithms works, how this study simulation is performed and the experimentation executed.

Chapter 5, Results and Discussion illustrates the results of the simulation done in Chapter 4 as well as analyzing it.

In chapter 6, the conclusion includes a summary of the whole work.

References

- Akbari, M., Rashidi, H., & Alizadeh, S. H. (2017). An enhanced genetic algorithm with new operators for task scheduling in heterogeneous computing systems. *Engineering Applications of Artificial Intelligence*, 61(March 2016), 35–46. <https://doi.org/10.1016/j.engappai.2017.02.013>
- Alla, H. Ben, Alla, S. Ben, & Alla, H. Ben. (2018). A novel task scheduling approach based on dynamic queues and hybrid meta-heuristic algorithms for cloud computing environment. *Cluster Computing*, 3456789. <https://doi.org/10.1007/s10586-018-2811-x>
- Alkhanak, E. N., & Lee, S. P. (2018). A hyper-heuristic cost optimisation approach for Scientific Workflow Scheduling in cloud computing. *Future Generation Computer Systems*, 86, 480–506. <https://doi.org/10.1016/j.future.2018.03.055>
- Amipara, H. (2015). A Survey on CloudSim Toolkit for Implementing Cloud Infrastructure, 1(12), 239–243.
- Calheiros, R. N., Ranjan, R., Beloglazov, A., De Rose, C. A. F., & Buyya, R. (2011). CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience*, 41(1), 23–50. <https://doi.org/10.1002/spe.995>
- Cho, K. M., Tsai, P. W., Tsai, C. W., & Yang, C. S. (2015). A hybrid meta-heuristic algorithm for VM scheduling with load balancing in cloud computing. *Neural Computing and Applications*, 26(6), 1297–1309. <https://doi.org/10.1007/s00521-014-1804-9>
- Choi, B. K., & Kang, D. (2013). *Modeling and Simulation of Discrete Event Systems* (1 edition). Hoboken, NJ: Wiley.
- Choudhary, A., Gupta, I., Singh, V., & Jana, P. K. (2018). A GSA based hybrid algorithm for bi-objective workflow scheduling in cloud computing. *Future Generation Computer Systems*, 83, 14–26. <https://doi.org/10.1016/j.future.2018.01.005>
- Eawna, M. H., Mohammed, S. H., & El-Horbaty, E. S. M. (2015). Hybrid Algorithm for Resource Provisioning of Multi-tier Cloud Computing. *Procedia Computer Science*, 65(Iccmit), 682–690. <https://doi.org/10.1016/j.procs.2015.09.012>

Gupta, A., & Garg, R. (2017). Workflow scheduling in heterogeneous computing systems : A survey. 2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN), 319–326. <https://doi.org/10.1109/IC3TSN.2017.8284499>

Jaybhaye, S.M. and Attar, V.Z., 2017, October. A review on scientific workflow scheduling in cloud computing. In Communication and Electronics Systems (ICCES), 2017 2nd International Conference on (pp. 218-223). IEEE.

Juve, G., Chervenak, A., Deelman, E., Bharathi, S., Mehta, G., & Vahi, K. (2013). Characterizing and profiling scientific workflows. *Future Generation Computer Systems*, 29(3), 682–692. <https://doi.org/10.1016/j.future.2012.08.015>

Kamalinia, A., & Ghaffari, A. (2017). Hybrid Task Scheduling Method for Cloud Computing by Genetic and DE Algorithms. *Wireless Personal Communications*. <https://doi.org/10.1007/s11277-017-4839-2>

Keshanchi, B., Souri, A., & Navimipour, N. J. (2017). An improved genetic algorithm for task scheduling in the cloud environments using the priority queues: Formal verification, simulation, and statistical testing. *Journal of Systems and Software*, 124, 1–21. <https://doi.org/10.1016/j.jss.2016.07.006>

Kum, M., & Sharma, S. C. (2018). ScienceDirect for Task Scheduling in Clouds, *Environment Kalk Kalk*, 00. <https://doi.org/10.1016/j.procs.2017.12.093>

Mangla, P. (2017). Heuristic vs Meta-Heuristic Approaches for Load Balancing in Cloud Environment, 6(8), 297–301.

Margaret Rouse. (n.d.). What is cloud computing? - Definition from WhatIs.com. Retrieved July 5, 2018, from <https://searchcloudcomputing.techtarget.com/definition/cloud-computing>

Md Enzai, N. I., & Tang, M. (2016). A heuristic algorithm for multi-site computation offloading in mobile cloud computing. *Procedia Computer Science*, 80, 1232–1241. <https://doi.org/10.1016/j.procs.2016.05.490>

Medium. (2018). *Getting Started with CloudSim – Ingkwan's Knowledge Hub – Medium*. [online] Available at: <https://medium.com/ingkwan/getting-started-with-cloudsim-631e7f6b85d6> [Accessed 25 Dec. 2018].

Cloud Computing. (2017), (Icces), 218–223.

Moschakis, I. A., & Karatza, H. D. (2015). A meta-heuristic optimization approach to the scheduling of bag-of-tasks applications on heterogeneous clouds with multi-level arrivals and critical jobs. *Simulation Modelling Practice and Theory*, 57, 1–25. <https://doi.org/10.1016/j.simpat.2015.04.009>

Tyagi, R., & Gupta, S. K. (n.d.). A Survey on Scheduling Algorithms for Parallel and Distributed Systems, 51–64.

Vashistha, A., Porwal, R., & Soni, A. K. (2015). A Taxonomy of Scheduling Algorithms for Cloud Computing, 12(1), 67–71.