TASK SCHEDULING IN CLOUD COMPUTING USING HYBRID GENETIC ALGORITHM AND BALD EAGLE SEARCH (GA-BES)

KAMAL KHAIRI BIN SUPPARHMAN

FACULTY OF COMPUTING AND INFORMATICS

UNIVERSITY MALAYSIA SABAH 2022

TASK SCHEDULING IN CLOUD COMPUTING USING HYBRID GENETIC ALGORITHM AND BALD EAGLE SEARCH (GA-BES)

KAMAL KHAIRI BIN SUPPARHMAN

THESIS SUBMITTED IN PARTIAL FULFILLMENT FOR THE DEGREE OF BACHELOR OF COMPUTER SCIENCE WITH HONOURS

(NETWORK ENGINEERING)

FACULTY OF COMPUTING AND INFORMATICS

UNIVERSITY MALAYSIA SABAH 2022

NAME	**:** KAMAL KHAIRI BIN SUPPARHMAN
MATRIC NUMBER	**:** BI18110026
TITLE	: TASK SCHEDULING IN CLOUD COMPUTING USING HYBRID GENETIC ALGORITHM AND BALD EAGLE SEARCH (GA-BES)
DEGREE	**:** BACHELOR OF COMPUTER WITH HONOURS (NETWORK ENGINEERING)
VIVA'S DATE	**:** 17 01 2022

CERTIFIED BY;

1. **SUPERVISOR** — Signature

Mr. Nordin Saad

DECLARATION

I, **Kamal Khairi bin Supparhman** hereby declare that the Project Research titled "**TASK SCHEDULING IN CLOUD COMPUTING USING HYBIRD GENETIC ALGORITHM AND BALD EAGLE SEARCH (GA-BES)**" is the original work done by me and submitted to Universiti Malaysia Sabah in partial to the fulfilment of requirements for this Semester 2 2020/2021 session in Faculty of Computing and informatics. This project is done by me under the Supervision of **Mr. Nordin Saad**, Senior Lecturer of Universiti Malaysia Sabah.

17 01 2022

KAMAL KHAIRI
BIN
SUPPARHMAN

BI18110026

ACKNOWLEDGEMENT

First and forever, I would like to express my deepest gratitude to Allah S.W.T for his grace and love I was able to complete my research work successfully.

I would also like to thank my supervisors, Mr. Nordin Saad, Assoc. Prof. Dr. Azali Saudi as a co-supervisor for this project, all senior lecturers, panel, and examiner, at Universiti Malaysia Sabah, Faculty Computer for giving me the opportunity and teaching me to make this research a success. With his guidance, words of passion, and guidance I was able to understand deeply and Successfully inspired me to complete this research. He taught me a lot about how to prepare literature reviews and also gave a lot of insightful explanations about the methodology to further facilitate my research success.

My sincere thanks also to my parents Supparhman bin Abd Hapid and my mother Faridah binti Jafri who has not given up on praying for my success. Thank you also to my friends Muhammad Syahril, Sharon June, Mohd Izuddin, and Mohd Ozanizie for not having the guidance and discussion with them I would not be able to make this research a success.

Lastly, I would like to thank all involved directly or indirectly for the success of this research.

KAMAL KHAIRI BIN SUPAPRHMAN

17 01 2022

ABSTRACT

Cloud computing is on-demand resources available for the computer system, especially in the data storage without management from the human. Cloud computing one of the fastest growing technologies and now managing file and services on the local storage devices is no longer use because of the cloud computing that can do the same things over the internet where anyone and everyone can access to it. Task scheduling and resource allocation are essential aspects of cloud computing. This Study proposes task scheduling in cloud computing using a hybrid genetic algorithm, and bald eagle search proposed to solve the task scheduling problem. The genetic algorithm was widely used because of its accuracy and simplicity. However, it will become slower in some instances that include a more significant problem size. Hence, Bald Eagle Search (BES) can increase efficiency and performance because it provides an efficient scheduling mechanism. The natural evolution optimization algorithm which is genetic algorithm can be improve by combining the nature meta-heuristic algorithms which is bald eagle search to improve the makespan of genetic algorithm using cloudsim that need to be implement on the eclipse platform. As for the running method the compilation of the code will be run by using cmd and Ant Apache and the total average result of 30 simulation will be view on the web base application will be run using xampp.

ABSTRAK

PENJADUALAN TUGAS DALAM PENGKOMPUTERAN AWAN MENGGUNAKAN HIBRID GENETIK ALGORITHMA DAN CARIAN HELANG BOTAK (GA-BES)

Pengkomputeran awan ialah sumber atas permintaan yang tersedia untuk sistem komputer, terutamanya dalam storan data tanpa pengurusan daripada manusia. Pengkomputeran awan salah satu teknologi yang paling pesat berkembang dan kini mengurus fail serta perkhidmatan pada peranti storan tempatan tidak lagi digunakan kerana pengkomputeran awan yang boleh melakukan perkara yang sama melalui internet di mana sesiapa sahaja dan semua orang boleh mengaksesnya. Penjadualan tugasdan peruntukan sumber adalah aspek penting dalam pengkomputeran awan. Kajian ini mencadangkan penjadualan tugas dalam pengkomputeran awan menggunakan algoritma genetik hibrid, dan carian helang botak dicadangkan untuk menyelesaikan masalah penjadualan tugas. Algoritma genetik digunakan secara meluas kerana ketepatan dan kesederhanaannya. Walau bagaimanapun, ia akan menjadi lebih perlahan dalam beberapa keadaan yang merangkumi saiz masalah yang lebih ketara. Oleh itu, Bald Eagle Search (BES) boleh meningkatkan kecekapan dan prestasi kerana ia menyediakan mekanisme penjadualan yang cekap. Algoritma pengoptimuman evolusi semula jadi iaitu algoritma genetik boleh ditambah baik dengan menggabungkan algoritma meta-heuristik alam iaitu pencarian helang botak untuk menambah baik makespan algoritma genetik menggunakan cloudsim yang perlu dilaksanakan pada platform eclipse. Bagi kaedah larian, kompilasi kod akan dijalankan dengan menggunakan cmd dan Ant Apache dan jumlah purata hasil 30 simulasi akan dilihat pada aplikasi pangkalan web yang dijalankan menggunakan xampp.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Chapter Overview

Cloud Computing is one of the popular things nowadays. Rather than own computing infrastructure or data centres. Nowadays, big or small companies can rent a cloud service provider to access anything stored in the cloud itself. The benefit of using cloud computing is that the company can avoid the upfront cost and complexity of maintaining the IT infrastructure of the cloud. Cloud computing service is available from the primary storage, networking, and processing of natural language and artificial intelligence (AI) and, last but not least, office applications. The service does not need physical close to the computer hardware and can be delivered via the cloud.

Task scheduling is an essential matter in cloud computing itself. While task scheduling acts as one of the techniques that will improve cloud computing performance, workload division for all resources will be provided fairly and effectively to reduce waiting time, execution time, maximum throughput, and the effectively undesirable use of resources. The main problem with task scheduling is the waiting time and the execution time. For example, in First Come First Serve (FCFS), the problem for this algorithm is that tasks have high waiting time, and for Shortest Job First (SJF), the waiting time is shorter, but the Total Finish Time (TFT) is longer than FCFS.

In this project, we will do research base on the combination of two algorithms which is Genetic Algorithm (GA) and Bald Eagle Search (BES) and this algorithm will become a Hybrid Genetic Algorithm and Bald Eagle Search (GA-BES). Hopefully, this algorithm will counter the highest waiting time and longer total finish time for the task scheduling.

1.2 Project Background

Normally, the task scheduling system in cloud computing goes through three levels. The first task level is a set of tasks (Cloudlets) sent by cloud users to perform tasks. The second scheduling level is responsible for mapping tasks to appropriate resources to obtain the highest resource utilization with the smallest manufacturing time. The generation time is the total completion time of all tasks from beginning to end.

Now days there are to many algorithm that involve in task scheduling, our purpose to see can the hybrid genetic algorithm bald eagle search can work better compare to the genetic algorithm and bald eagle search algorithm optimization itself. The final result is to see either combining Genetic Algorithm (GA) and Bald Eagle Search (BES) can produce the less best duration compare to the two base algorithm itself.

1.3 Problem Statement

Task scheduling in Cloud Computing using Genetic Algorithm and Bald Eagle Search could be challenging for several reasons:

1.3.1 The task scheduling algorithm is still developed from time to time to provide a better solution for the virtual machine to schedule its task.

1.3.2 The output for the result is not organized very well.

1.3.3 The research on hybrid genetic algorithm for Genetic Algorithm bald eagle search still can't be found.

1.4 Project Objectives

The aims for task scheduling in cloud computing is:-

1.4.1 To implement the GA, BES and hybrid GA-BES Optimization algorithm to

optimize task scheduling using the cloudsim in java platform

1.4.2 To develop the web base application for a better output for the algorithm for admin to view the total average time in table for 50 times simulation

1.4.3 To test and evaluate the performance of propose research algorithms in term of the best duration time on the Cloudsim simulator.

1.5 Project scope

This project will lead to see the advantages of hybrid algorithm which is explore from a previous research on Genetic Algorithm for task scheduling and will improvise the list of the objective itself. the target user for this project is network administrator, where network administrator can manage the task scheduling algorithm for a better use of the cloud computing its self. For example like JTMK.

Table 1.1: Properties and Quantity to be used in Task Scheduling In Cloud Computing Using Hybrid Genetic Algorithm and Bald Eagle Search (GA-BES)

Properties	**Quality**
Output	**Show the output that already has been organize to make the network administrator easy to do a compression between other Task Scheduling optimization algorithms**

CHAPTER 2

LITERATURE REVIEW

2.1 Chapter Overview

Task scheduling is the crucial part in cloud computing which involved a massive number of user and data, the most efficient task scheduling mechanism need to meets the user's requirement and improve the utilization for the resource and enhance the whole performance for the cloud computing, waiting time and the total finish time is the most critical things in the task scheduling. Although some research creates a set of task schedules and evaluates the quality of each task schedule for the user satisfaction by experimenting on genetic algorithm in term of base task scheduling model to make a comparison with the existing task scheduling, which are the round-robin task scheduling, load index-based task scheduling and the ABC base task scheduling model (Sung Ho Jang, Tae Young Kim, Jae Kwon Kim & Jong Sik Lee, 2012).

For Example, some researchers also create an algorithm to countermeasure this problem by making a compression job spanning time and load balancing genetic algorithm (JLGA) and adaptive genetic algorithm (AGA) result bring proves the validity of the task scheduling algorithm and the effectiveness of the optimization method (Tingting Wang, Zhaobin Liu, Yi Chen & Et al, 2014). Some researchers also said that the energy-efficient techniques also prime importance in cloud computing because of the increase of the high- performance computing resources and to counter this by creating a harmony- inspired genetic algorithm (HIGA) to provides better energy savings which make the application performance with 39% less execution overhead (Mohan Sharma & Ritu Garg,2019).

Some researchers also Aggrey that the Bald Eagle Search (BES) algorithm competes well with the advanced meta-heuristic algorithm and conventional methods for the optimization by evaluates the algorithm based on mean, standard deviation,

best point and Wilcoxon signed-rank test statistic of the function values (H. A. Alsattar, A. A. Zaidan & B. B.Zaidan, 2019). These recent studies prove task schedule is needed to be developed in term of performance. The availability for the task scheduling to bring satisfaction for the user that using cloud computing.

2.2 Cloud Computing

Cloud computing is a on-demand availability of computer system that involve in data storage (cloud storage) and computing without management by user. As we know now days cloud computing is involve in large company to provide a better solution for resources and for storage. Large cloud, predominant today, this often have function to distributed over multiple location from servers. The availability of high-capacity networks, low-cost computers and storage devices, and the widespread adoption of hardware virtualization, service-oriented architecture and autonomic and utility computing has led to growth in cloud computing. Rather than owning own computing infrastructure or data centres, companies also can rent an access to any form of application or storage from a cloud service provider. This can avoid the firms from upfront cos and complexity of owning and maintaining the infrastructure of IT. The benchmarking of the cloud computing solutions exist for a different layer for example like Infrastructure as a Service (Iaas), Platform as a Service (Paas), and Software as a Service (David Bermbach, 2015).

In an IaaS service model, a cloud provider hosts the infrastructure components traditionally present in an on-premises data centre. This includes servers, storage and networking hardware, as well as the virtualization or hypervisor layer. IaaS providers also supply a range of services to accompany those infrastructure components, such as detailed billing, monitoring, log access, security, load balancing, clustering, and storage resiliency (backup, replication, and recovery). IaaS customers access resources and services through a wide area network (WAN), such as the internet, and can use the cloud provider's services to install the remaining elements of an application stack. For example, the user can log in to the IaaS platform to create virtual machines (VMs); install operating systems in each VM; deploy middleware, such as databases; create storage buckets for workloads and backups; and install the

enterprise workload into that VM. Customers can then use the provider's services to track costs, monitor performance, balance network traffic, troubleshoot application issues and manage disaster recovery. The advantages of IaaS is Organizations choose IaaS because it is often easier, faster and more cost-efficient to operate a workload without having to buy, manage and support the underlying infrastructure, with IaaS, a business can simply rent or lease that infrastructure from another business, some IaaS providers also charge customers based on the amount of virtual machine space they use.

Conversely, the business could commit that piece of software to a long- term IaaS deployment if the costs of a long-term commitment are less, in general, IaaS customers pay on a per-user basis, typically by the hour, week or month, and When a business cannot use.

third-party providers, a private cloud built on premises can still offer the control and scalability of IaaS -- though the cost benefits no longer apply. The significant issues commonly associated with IaaS in cloud systems are virtualization and multi-tenancy, resource management, network infrastructure management, data management, APIs, interoperability etc. (Sunilkumar S. Manvi, Gopal Krishna Shyam, 2013).

In the Platform-as-a-Service (PaaS) model, developers essentially rent everything they need to build an application, relying on a cloud provider for development tools, infrastructure, and operating systems. This is one of the three service models of cloud computing. PaaS vastly simplifies web application development; from the developer's perspective, all back-end management occurs behind the scenes. Although PaaS has some similarities with serverless computing, there are many critical differences between them.

PaaS can be accessed over any internet connection, making it possible to build an entire application in a web browser. Because the development environment is not hosted locally, developers can work on the application from anywhere in the world. This enables teams that are spread out across geographic locations to collaborate. It also means developers have less control over the development environment, though this comes with far lessoverhead.

The main things that PaaS offering is development tools, middleware, operating system, database management, and infrastructure.

2.2.1 Development Tools

PaaS vendors offer various tools necessary for software development, including a source code editor, a debugger, a compiler, and other essential tools. These tools may be provided together as a framework. The specific tools provided will depend on the vendor, but PaaS offerings should include everything developers need to build their application.

2.2.2 Middleware

Platforms offered as a service usually include middleware so that developers don't have to build it themselves. Middleware is software that sits between user-facing applications and the machine's operating system; for example, middleware allows the software to access input from the keyboard and mouse. Middleware is necessary for running an application, but end users don't interact with it.

2.2.3 Operating System

A PaaS vendor will provide and maintain the operating system developers work on and the application runs on.

2.2.4 Databases

PaaS providers administer and maintain databases. They will usually provide developers with a database management system as well.

2.2.5 Infrastructure

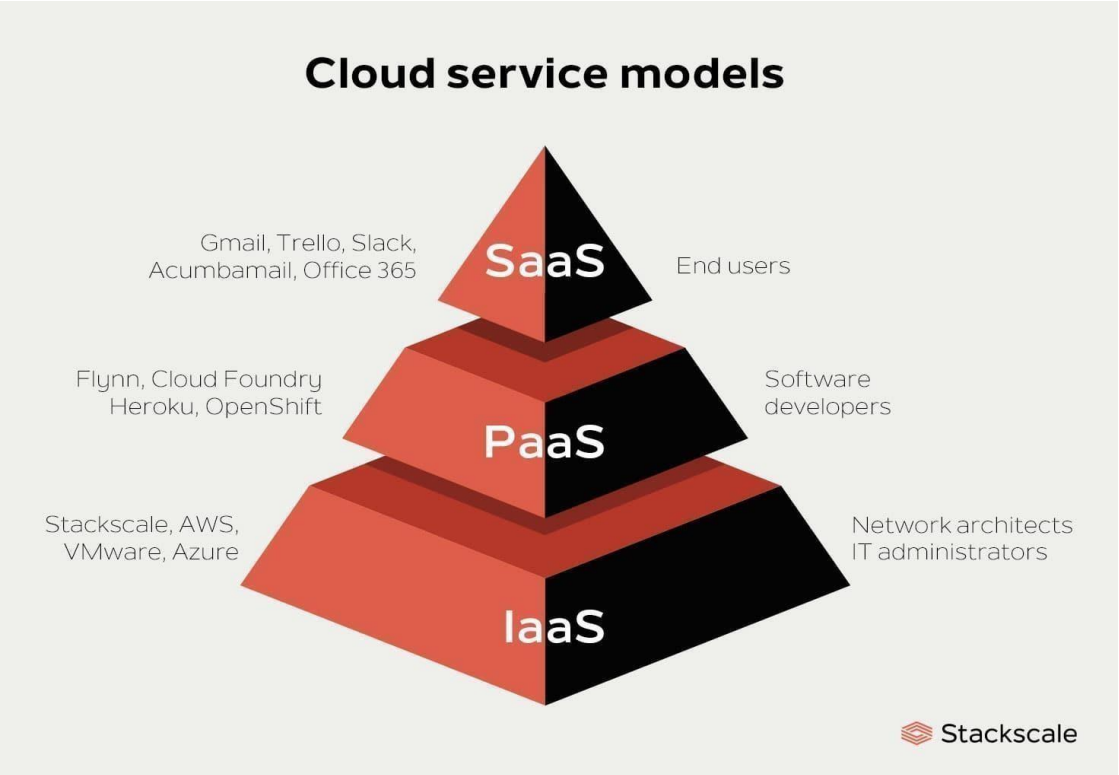
PaaS is the next layer up from IaaS in the cloud computing service model, and everything included in IaaS is also included in PaaS. A PaaS provider either manages servers, storage, and physical data centres, or purchases them from an IaaS provider.

Software-as-a-Service (SaaS) is a software licensing model in which access to the software is provided on a subscription basis, with the software being located on external servers rather than on servers located in-house.

Software-as-a-Service is typically accessed through a web browser, with users logging into the system using a username and password. Instead of each user having to install the software on their computer, the user is able to access the program via the Internet. The rise of Software-as-a-Service (SaaS) coincides with the rise of cloud-based computing. Cloud computing is the process of offering technology services through the Internet, which often includes data storage, networking, and servers.

Before SaaS was available, companies looking to update the software on their computers had to purchase compact disks containing the updates and download them onto their systems. For large organizations, updating software was a time-consuming endeavor. Over time, software updates became available for download through the Internet, with companies purchasing additional licenses rather than additional disks. However, a copy of the software still needed to be installed on all devices that needed access.

Figure 2.1: Cloud Service models



2.3 Task Scheduling in Cloud Computing

Task Scheduling are defined as the mechanism used to select resources and execute the task to get less waiting and execution time. There a two level of task scheduling the first level is set of policies to distribute virtual machine in host and the second level is set of policies to distributed task to virtual machine (Tahani Aladwani, 2020). The biggest challenge in cloud is to control the performance such as execution time, response time, waiting time, network, bandwidth, and service cost. The task scheduling has been developed in the 2013 to 2020 witch is a 3% in year 2013 and in the year of 2014 is 8%, in 2015 is 5% and 8% contribution that has been provide by the researcher (Prashant B. Jawade, Sai Kumar D, Ramachandram S, 2021).

Tasks scheduling algorithms are defined as the mechanism used to select there sources to execute tasks to get less waiting and execution time.

In the cloud computing environment, there are two levels of scheduling algorithms.

2.3.1 First level: in host level where a set of policies to distribute VMs in the host.

2.3.2 Second level: in VM level where a set of policies to distribute tasks to VM.

2.4 Genetic Algorithm

Genetic Algorithm is inspired by evolution that are a family of computational models, and bring a potential of solution to a specific problem for a simple chromosome. This Algorithm also viewed as function optimizer, even though the range of problems to which this Algorithm applied are large. The implementation of this algorithm begins with a random population of chromosomes and it will evaluates the structures and allocated and to find out the better solutions to the target problem so that more chances to reproduce than to a poorer solution.

The process of natural selection will start with the selection of fit test individual from a population and will produce offspring which inherit the characteristics of parents and will be added to next generation. There a three component that involve in Genetic Algorithm which is gene, chromosome and population. Chromosome is a complicated long thread of DNA (deoxyribonucleic acid) and each trait is get from combination of the DNA its self. There a five phase that are considered in a genetic algorithm which is initial population, fitness function, selection, crossover and mutation.

Selection phase will select the fittest individual and pass their genes to the next generation, the parents are selected base on their fitness score for a reproduction. The crossover is the most important phase, crossover point is chosen at random from the genes which their parents mated from. Crossover is need to combine two string to get a better string. This recombination create a different individuals in the next generation by combine the material from two individual that are from previous generation. For example before crossover the string 1 is 1110000 and the string 2 is 0000111 but after crossover happened string 1 will become 00001111 and string 2 will become 11110000 (Tom V. Mathew).

2.5 Bald Eagle Search

Bald eagle search was a mimics to the behavior of bald eagle during hunting. This algorithm divided into three part which is selecting stage, searching stage and swooping stage.

In the select stage bald eagle will identify and select the best area. Where the A is parameter for controlling the changes will take the value between 1.5 and 2 and r is a random number that take value between 0 and 1 (H. A. Alsattar, A. A. Zaidan, B. B. Zaidan).

The 2nd stage is search stage. In this stage the bald eagle will search for the best prey within the selected area. The points move around the centre point during the search stage. When parameters a and R are changed, the algorithm increases diversification to escape from the local optimum and to continuously obtain an efficient solution (H. A. Alsattar, A. A. Zaidan, B. B. Zaidan).

The last stage is the swooping stage where the bald eagle will swing from the best position in the search space to get the target pray.

2.6 Cloud Computing Optimization algorithm

Zhong Zong et al. (Zong, 2020) propose combining dynamic fusion mission planning methodology, genetic approach, and the ant-colony system in 2020. Cloud computing data and storage facilities use less energy as a result of this. The test findings show that the suggested task programming approach would significantly reduce the amount of time and energy spent on cloud computing devices.

Vijayalakshmi A. Lepaksh et al. (Vijayalakshmi A. Lepaksh, 2020) proposed an Efficient Resource Allocation with Score (ERAS) for task scheduling in cloud environments in 2020, which takes into account Virtual Machines (VM) temporary operational availability by suggesting different types of delays and using EFT to set the processor for task scheduling toa standardized score. The results show that the improved dependability of the ERAS algorithm delivers superior efficiency than current systems that just consider EFT for allocations.