



UNIVERSITY OF SCIENCE AND TECHNOLOGY
COLLEGE OF GRADUATE STUDIES AND ACADEMIC
ADVANCEMENT

Faculty of Computer Science and Information Technology

Scheduling Jobs on Cloud Computing Using Firefly Algorithm

by

Demyana Izzat Eisa Azir

A Thesis

Submitted to the College of Graduate Studies and Academic Advancements
in Partial Fulfillment of the Requirement for the Degree of Master of Science in
Computer Science

Supervisor

Dr. Adil yousif

November 2015

Abstract

Cloud computing is a new technology, instead of all the computer hardware and the software that used on desktop, or somewhere within company's network, it's presented as a service by cloud service providers and accessed via the Internet. Exactly where hardware and software are located and how everything works does not matter.

In cloud computing there are many jobs that requires to be executed by the available resources to achieve best performance, minimal execution time for completion, shortest response time, utilization of resources etc. Several optimization methods an available for cloud job scheduling. However, the job scheduling process is still need to be optimized.

This research proposes a job scheduling mechanism using Firefly Algorithm to minimize the execution time of jobs. The proposed mechanism based on information of jobs (cloudlets) and resources (virtual machines) such as length of job, speed of resource and identifiers. The scheduling function in the proposed job scheduling mechanism firstly creates a set of jobs and resources to generate the population by assigning the jobs to resources randomly and evaluates the population by using the fitness value which represent the execution time of jobs. Secondly the function used iterations to regenerate populations based on fireflies behavior to produce the best job schedule that gives the minimum execution time of jobs.

Several scenarios are implemented using Java Language and CloudSim simulator, different scenarios have been considered in the evaluation and experimentation phase to examine the proposed mechanism in different work load. The first phase of the evaluation process describes how the proposed mechanism can be used to minimize the execution time of jobs. The second phase of the evaluation process compares the proposed mechanism with First Come First Serves (FCFS) algorithm. The results showed that the proposed mechanism minimizes the execution time significantly. Furthermore, the proposed mechanism outperformed the FCFS algorithm.

المستخلص

الحوسبة السحابية هي تكنولوجيا جديدة، بدلاً من كل أجهزة الحاسوب والبرمجيات التي تستخدم على سطح المكتب، أو في مكان ما داخل شبكة الشركة، هو تقديمها كخدمة من قبل مقدمي الخدمات السحابية والوصول إليها عبر شبكة الإنترنت . بالضبط أين تقع الأجهزة والبرمجيات وكيف يعمل كل شيء لا يهم. في الحوسبة السحابية هناك العديد من الوظائف التي يحتاج إليها ليعتمدها من قبل الموارد المتاحة لتحقيق أفضل أداء، تقليل وقت التنفيذ، أقصر زمن إستجابة، وإستفادة من الموارد الخ. العديد من الطرق الأمثل المتاحة لجدولة الوظائف السحابية. ومع ذلك، فإن عملية جدولة الوظائف لا تزال في حاجة إلى أن يتم تحسينها. إقتراح هذا البحث آلية جدولة الوظائف بإستخدام خوارزمية اليراع لتقليل وقت تنفيذ الوظائف . الآلية المقترحة تعتمد على معلومات من الوظائف (cloudlets) والموارد (الأجهزة الظاهرية) مثل طول الوظيفة، وسرعة المورد ومعرف لكلاهما. دالة الجدولة في آلية جدولة الوظائف المقترحة يخلق أولاً مجموعة من الوظائف والموارد لتوليد السكان عن طريق تعيين الوظائف إلى الموارد بشكل عشوائي وتقييم السكان بإستخدام قيمة اللياقة البدنية التي تمثل وقت تنفيذ الوظائف.

ثانياً تكرر إعادة توليد السكان على أساس سلوك اليراعات لإنتاج أفضل جدول الوظيفة الذي يعطي وقت تنفيذ أقل للوظائف. وتستند المنهجية المستخدمة في هذا العمل البحثي على عدة سيناريوهات بإستخدام لغة جافا والمحاكاة بإستخدام جهاز محاكاة CloudSim. تم النظر في سيناريوهات مختلفة في مرحلة التقييم والتجريب لدراسة الآلية المقترحة في عبء العمل المختلفة. يصف المرحلة الأولى من عملية التقييم كيف يمكن إستخدام الآلية المقترحة لتقليل وقت تنفيذ الوظائف. المرحلة الثانية من عملية التقييم يقارن الآلية المقترحة مع خوارزمية First Come First Serves (FCFS). وأظهرت النتائج أن الآلية المقترحة يقلل من وقت التنفيذ إلى حد كبير. وعلاوة على ذلك، تفوقت الآلية المقترحة على خوارزمية FCFS.

1.1 Overview

This chapter introduces the research work, state the problem, defines research objectives, significant and describe the thesis structure.

1.2 Problem Background

As the IT technologies are growing day by day, the need of computing and storage are rapidly increasing. Cloud Computing has become a widely accepted paradigm for high performance computing, because in Cloud Computing all type of IT facilities are provided to the users as a service. The services of the cloud are provided through the Internet. In Cloud Computing the term Cloud is used for the service provider, which holds all types of resources for storage, computing etc. Mainly three types of services are provided by the cloud. First is Infrastructure as a Service (IaaS), which provides cloud users the infrastructure for various purposes like the storage system and computation resources. Second is Platform as a Service (PaaS), which provides the platform to the clients so that they can make their applications on this platform. Third is Software as a Service (SaaS), which provides the software to the users and hence the user don't need to install the software on their own machines and they can use the software directly from the cloud. Cloud Computing provides many benefits: it results in cost savings because there is no need of initial installation of much resource; it provides scalability and flexibility, the users can increase or decrease the number of services as per requirement; maintenance cost is very less because all the resources are managed by the Cloud providers[1].

Job scheduling is one of the major activities performed in all the computing environments. Cloud computing is one the upcoming latest technology which is developing drastically. To efficiently increase the working of cloud computing environments, job scheduling is one the jobs performed in order to gain maximum profit.

The goal of scheduling algorithms in distributed systems is spreading the load on processors and maximizing their utilization while minimizing the total job execution time Job scheduling, one of the most famous optimization problems [2]. Job scheduling has been considered as one of crucial problems in cloud computing. An optimized scheduler would improve many factors in scheduling of jobs in a cloud system such as throughput and performance. Different Approaches have tried to solve this problem like Genetic algorithm, Ant colony optimization, Particle swarm optimization and etc.

1.3 Problem Statement

There are several resources on cloud and huge number of jobs, suppose that $R = \{r_1, r_2, r_3 \dots r_m\}$ are m cloud resources and $J = \{j_1, j_2, j_3 \dots j_n\}$ are n independent jobs. The speed of each resource is expressed in form of MIPS (Million Instructions Per Second), and the length of each job is expressed in the form of number of instructions. Now the problem is how to allocate the submitted jobs to the available resources in order to complete the jobs efficiently. The current scheduling mechanism are still need to be optimized, such that an optimum execution time is achieved.

1.4 Research Objectives

The objective of this research is to propose a new job scheduling mechanism based on Firefly Algorithm with modified distance to minimize the execution time and to evaluate the proposed mechanism using CloudSim simulator.

1.5 Thesis Structure

This thesis contains six chapters, Chapter 2 Gives an overall idea of cloud computing, and scheduling algorithm on cloud computing. Chapter 3 Describes the Research Methodology. Chapter 4 Describes the Proposed Scheduling Mechanism. Chapter 5 Presents the Results and Discussion. Chapter 6 Provide the Conclusion and Recommendation of the study.