Kingdom of Saudi Arabia
Majmaah University
Ministry of Higher Education
College of Science Al Zulfi

المملكة العربية السعودية
جامعة المجمعة
وزارة التعاليم العالي
كلية العلوم بالزلفي

Cloud computing scheduling using modified PSO algorithm

Student Affairs System
For College of science Al Zulfi
Department of Computer Science and Information

Graduation Project

Submitted in partial fulfillment of the requirements for the award of
Bachelor's degree of the Majmaah University
(Semester 2, 2019-2020)

Submitted by:

Student Name: Adel Teheeni

Student ID: 352106226

Under the supervision of:
Dr. Yousef Qawqzeh

Contents

Abstract

Task scheduling problem is one of the most important steps in using cloud computing environment capabilities. Different experiments showed that PSO algorithm is reliable for scheduling in cloud computing environments. However, having an optimum solution is almost impossible but having a sub-optimal solution using heuristic algorithms seems possible. This approach seeks the modification of PSO algorithm for efficient task scheduling. This algorithm aims to generate an optimal schedule in order to minimize completion time.

Keywords: Scheduling; Cloud computing; PSO; Optimization; Execution Time.

CHAPTER ONE

1.1 Introduction

- Scheduling

Scheduling is a method that is used to distribute valuable computing resources, usually processor time, bandwidth and memory, to the various processes, threads, data flows and applications that need them. Scheduling is done to balance the load on the system and ensure equal distribution of resources and give some prioritization according to set rules. This ensures that a computer system is able to serve all requests and achieve a certain quality of service. Scheduling is also known as process scheduling.

- cloud computing

Cloud computing is the on-demand availability of computer system resources, especially data storage and computing power, without direct active management by the user. The term is generally used to describe data centers available to many users over the Internet. Large clouds, predominant today, often have functions distributed over multiple locations from central servers. If the connection to the user is relatively close, it may be designated an edge server.

Clouds may be limited to a single organization (enterprise clouds) or be available to many organizations (public cloud).Cloud computing relies on sharing of resources to achieve coherence and economies of scale. Advocates of public and hybrid clouds note that cloud computing allows companies to avoid or minimize up-front IT infrastructure costs. Proponents also claim that cloud computing allows enterprises to get their applications up and running faster, with improved manageability and less maintenance, and that it enables IT teams to more rapidly adjust resources to meet fluctuating and unpredictable demand, providing the burst computing capability: high computing power at certain periods of peak demand.

- swarm intelligence

Swarm intelligence (SI) is the collective behavior of decentralized, self-organized systems, natural or artificial. The concept is employed in work on artificial intelligence. The expression was introduced by Gerardo Beni and Jing Wang in 1989, in the context of cellular robotic systems. SI systems consist typically of a population of simple agents or boids interacting locally with one another and with their environment. The inspiration often comes from nature, especially biological systems. The agents follow very simple rules, and although there is no centralized control structure dictating how individual agents should behave, local, and to a certain degree random, interactions between such agents lead to the emergence of "intelligent" global behavior, unknown to the individual agents. Examples of swarm intelligence in natural systems include ant colonies, bird flocking, hawks hunting, animal herding, bacterial growth, fish schooling

and microbial intelligence. The application of swarm principles to robots is called swarm robotics, while 'swarm intelligence' refers to the more general set of algorithms. 'Swarm prediction' has been used in the context of forecasting problems. Similar approaches to those proposed for swarm robotics are considered for genetically modified organisms in synthetic collective intelligence.

- PSO (particle swarm optimization)

In computational science, particle swarm optimization (PSO) is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. It solves a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Each particle's movement is influenced by its local best-known position but is also guided toward the best-known positions in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions.

PSO is originally attributed to Kennedy, Eberhart and Shi and was first intended for simulating social behaviour, as a stylized representation of the movement of organisms in a bird flock or fish school. The algorithm was simplified, and it was observed to be performing optimization. The book by Kennedy and Eberhart describes many philosophical aspects of PSO and swarm intelligence. An extensive survey of PSO applications is made by Poli. Recently, a comprehensive review on theoretical and experimental works on PSO has been published by Bonyadi and Michalewicz. PSO is a metaheuristic as it makes few or no assumptions about the problem being optimized and can search very large spaces of candidate solutions. However, metaheuristics such as PSO do not guarantee an optimal solution is ever found. Also, PSO does not use the gradient of the problem being optimized, which means PSO does not require that the optimization problem be differentiable as is required by classic optimization methods such as gradient descent and quasi-newton methods.

1.2 Problem statement

In this vast and complex computational task, along with new technologies and scheduling needs, there is a big need to have scheduling schemes in which they are fast and fully utilizing the available resources in cloud computing environments. The growth of cloud computational complexities makes it necessary to optimize scheduling methods to better solve scheduling problems and to fully benefit from the available resources. This work tries to utilize the particle swarm intelligence (PSO) to improve and optimize scheduling problems in cloud computing environments.

1.3 Objectives

The main goal of this work is to better utilize the PSO algorithm in task scheduling inside cloud computing environment. To achieve this task, the following objectives have to be met:

- Explore the recent literature reviews concerning the scheduling in cloud environment based on PSO algorithm.
- Evaluate the inertia weight factor that alters PSO performance.
- Propose an optimized PSO algorithm to better solve scheduling issues.
- Evaluate the proposed algorithm in cloudsim environment.

1.4 scope of study

This study focuses on the use of PSO algorithm for cloud computing environment scheduling.

1.5 Research hypothesis

The modified PSO algorithm could improve the performance of scheduling techniques in cloud computing environment.

CHAPTER TWO

Literature Reviews

Task scheduling problem is one of the most important steps in using cloud computing environment capabilities. Different experiments show that although having an optimum solution is almost impossible but having a sub-optimal solution using heuristic algorithms seems possible. [1]compared three heuristic approaches for task scheduling on cloud environment with each other. These approaches are PSO algorithm, genetic algorithm and modified PSO algorithm for efficient task scheduling. In all these three algorithms the goal is to generate an optimal schedule in order to minimize completion time of task execution.

In cloud computing environment, there is a large quantity of submitted tasks by users. How to schedule these massive tasks efficiently and reasonably becomes a serious challenge? [2]they are proposed a Chaotic Particle Swarm Optimization algorithm (CPSO) to overcome the problems of Standard Particle Swarm algorithm such as premature convergence and low accuracy. Firstly, in initial process, chaotic sequence is introduced to enhance the diversity of particles. Then, an effective diagnosis mechanism of premature is adopted to determine local convergence and algorithm correction is performed by chaotic mutation, which could activate the particles in stagnation and make them escape from local optimum. Simulation experiments show that the proposed approach is feasible and effective.

The task scheduling policy is the important factors for achieving efficient calculation in a cloud computing environment[3]. This article put forwards a task scheduling method based on improved particle swarm algorithm against the present inefficiency. Particle Swarm Optimization (PSO) algorithm is used to solve task scheduling optimization by introducing the iterative selection operator. Improved particle swarm optimization algorithm (IPSO) can improve the ability of the optimization, as much as possible avoiding falling into a local optimum. The convergence effect is so better that the task scheduling time costs can be reduced. By simulation on a Cloud Sim simulation platform, the experimental results show that the algorithm has the advantages of improving optimization and taking less time. So, it can be used to research and practice about cloud computing problem for complex scheduling optimization.

In today's IT industry, cloud computing has become trending and a driving force. Many services are offered by IT vendors and supported by service level agreements and uptime assurance such as storage services, computations and hosting services covering multiple continents. One of the challenging issues is to have efficient task scheduling algorithm for cloud computing. Particle Swarm Optimization (PSO) is among of the best scheduling algorithm in cloud computing. However, PSO has generated the first population, which is produced randomly. As randomness decreases the probability of the algorithm to converge to the best solution. This paper surveys the existing of modified PSO scheduling algorithms to solve randomly issue in cloud computing environment. The comparison shows that the modified algorithm successfully makes some improvement compared with standard PSO [4].

Task scheduling is a very important part of the cloud computing environment. Aiming at the characteristics of task scheduling and considering both users and cloud service providers [5]. They proposed an improved particle swarm optimization algorithm based on adaptive weights. The algorithm uses adaptive weights to make the weight change with the increase of the number of iterations, and introduces random weights in the later stage, which avoids the situation that the particle swarm algorithm may be trapped in the local optimum when it comes to late stage. Applying the algorithm to task scheduling in cloud computing can achieve a better scheduling plan. The experiment results show that under the same conditions, the improved particle swarm optimization algorithm is better than the standard particle swarm optimization algorithm, which improves the using efficiency of resource while ensuring the task completion time.

The Cloud computing has become the fast spread in the field of computing, research and industry in the last few years. As part of the service offered, there are new possibilities to build applications and provide various services to the end user by virtualization through the internet. Task scheduling is the most significant matter in the cloud computing because the user has to pay for resource using on the basis of time, which acts to distribute the load evenly among the system resources by maximizing utilization and reducing task execution Time. Many heuristic algorithms have been existed to resolve the task scheduling problem such as a Particle Swarm Optimization algorithm (PSO), Genetic Algorithm (GA), Ant Colony Optimization (ACO) and Cuckoo search (CS) algorithms, etc. In this paper, a Dynamic Adaptive Particle Swarm Optimization algorithm (DAPSO) has been implemented to enhance the performance

of the basic PSO algorithm to optimize the task runtime by minimizing the make span of a particular task set, and in the same time, maximizing resource utilization. Also, .a task scheduling algorithm has been proposed to schedule the independent task over the Cloud Computing. The proposed algorithm is considered an amalgamation of the Dynamic PSO (DAPSO) algorithm and the Cuckoo search (CS) algorithm; called MDAPSO. According to the experimental results, it is found that MDAPSO and DAPSO algorithms outperform the original PSO algorithm. Also, a comparative study has been done to evaluate the performance of the proposed MDAPSO with respect to the original PSO [6].

CHAPTER THREE

Research Methodology

Agile Software

What is it & how is it work:

Agile methodology is a process by which a team can manage a project by breaking it up into several stages and involving constant collaboration with stakeholders and continuous improvement and iteration at every stage, The Agile methodology begins with clients describing how the product will be used and what problem it will solve. This clarifies the customer's expectations to the project team. Once the work begins, teams' cycle through a process of planning, executing, and evaluating — which might just change the final deliverable to fit the customer's needs better. Continuous collaboration is key, both among team members and with project stakeholders, to make fully informed decisions.

The stages:

- Requirements: is a singular documented physical or functional need that a particular design, product or process aims to satisfy.
- Design: is a plan or specification for the construction of an object or system or for the implementation of an activity or process, or the result of that plan or specification in the form of a prototype, product or process.
- Develop: that means to build up or improve gradually over time.
- Test: is an investigation conducted to provide owners with information about the quality of the software product or service under test.
- Deploy: in this step, all of the activities that make a software system available for use.
- Review: is an evaluation of an application or the program before submitted.

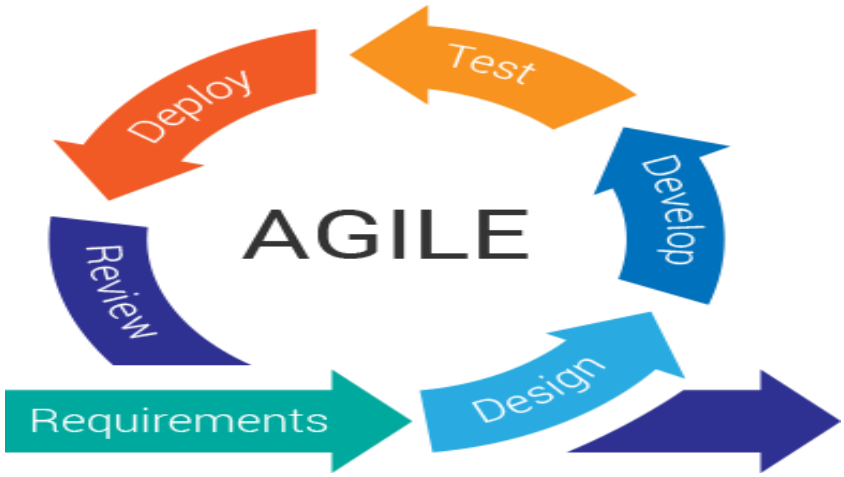

Figure 1. Agile Methodology

Agile works best when:

- we can't estimate the time we'll need and don't know the full scope of requirements.
- we don't know whether there's a need on the market for us software.
- we can't map out the business needs, so the design needs to emerge through trial and error.
- we have unlimited access to us customer who's ready for extensive involvement.
- we can afford to iterate and don't need to deliver a fully functional software at once.
- Neither we nor us client has a complex bureaucracy that delays decision-making.
- Clients don't have a fixed budget/schedule.
- we need to capture the market before there's any competition.
- we customers don't have trouble updating their software (or don't even notice it, e.g., they use a web app).

Advantages:

- we can deploy software quicker, so us customer can get value sooner rather than later.
- we waste fewer resources because we always work on up-to-date tasks.
- we can better adapt to change and respond faster.
- Faster turnaround times.
- we can detect and fix issues and defects faster.
- we spend less time on bureaucracy and busywork.
- There's a big community of Agile practitioners with whom we can share knowledge.
- we can get immediate feedback (which also improves team morale).
- Developers can improve their skills based on QA feedback.
- we don't have to worry about premature optimization.
- we can experiment and test ideas because its costs are low.

Disadvantages:

- Documentation tends to get sidetracked, which makes it harder for new members to get up to speed.
- It's more difficult to measure progress than it is in Waterfall because progress happens across several cycles.

- Agile demands more time and energy from everyone because developers and customers must constantly interact with each other.
- When developers run out of work, they can't work on a different project since they'll be needed soon.
- Projects can become everlasting because there's no clear end.
- Clients who work on a specified budget or schedule can't know how much the project will actually cost, which makes for a very complicated sales cycle ("Until iteration ends," is not something clients like to hear).
- Product lacks overall design, both from a UX and architecture point of view, which leads to problems the more you work on the product.
- Teams can get sidetracked into delivering new functionalities, which increases the amount of unplanned work.
- Features that are too big to fit into one or even several cycles are avoided because they don't fit in nicely into the philosophy.
- we need a long-term vision for the product and actively work on communicating it.
- Products lack cohesion, and the user journey is fragmented because the design is fragmented. The more time passes, the more disjointed the software ends up becoming.
- Short cycles don't leave enough time for the design thinking process, so designers have to redevelop the experience over and over due to negative feedback.

References

[1] Solmaz Abdi, Seyyed Ahmad Motamedi, and Saeed Sharifian ″Task Scheduling using Modified PSO Algorithm in Cloud Computing Environment″(2014).

[2] Li Yingqiu, Li Shuhua, Gao Shoubo ″Cloud Task Scheduling Based on Chaotic Particle Swarm Optimization Algorithm″(2016).

[3] Wu.DaQin ″Cloud Computing Task Scheduling Policy Based on Improved Particle Swarm Optimization″(2018).

[4] Mohd Farhan Ngatman, Dr Johan Mohd Sharif, and Dr Md Asri Ngadi ″A Study on Modified PSO Algorithm in Cloud Computing″(2017).

[5] Fei Luo, Ye Yuan, Weichao Ding ″An Improved Particle Swarm Optimization Algorithm Based on Adaptive Weight for Task Scheduling in Cloud Computing″(2018).

[6] Ali Al-maamari, and Fatma A.Omara ″Task Scheduling Using PSO Algorithm in Cloud Computing Environments″(2015).

MAJMAAH UNIVERSITY,
COLLEGE OF SCIENCE AL ZULFI,
DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION

(CERTIFICATE BY STUDENT)

This is to certify that the project titled **"Cloud computing scheduling using modified PSO algorithm"** submitted by me(**Adel Teheeni, 352106226**) under the supervision of **Dr. Yousef Qawqzeh** for award of Bachelor degree of the Majmaah University carried out during the Semester 2, 2019-2020 embodies my original work.

Signature in full: ----------

Name in block letters:

Student ID:

Date: