



University of Science and Technology
Faculty of Post Graduate Studies & Academic Advancement

High Exploration Genetic Algorithm For Job Scheduling on Grid Computing

**Thesis submitted in Partial Fulfillment of the
Requirements for The Degree of Master of
((Computer Science))**

Prepared By:

Hanaa Abdlerahman Fadl Alseed

Supervised By:

Dr. Adil Yousif

December 2014

Abstract

Grid computing presents a new trend to distribute and Internet computing to coordinate large scale heterogeneous resources providing sharing and problem solving in dynamic, multi- institutional virtual organizations. Scheduling is one of the most important problems in computational grid. To increase the performance, of the grid system an efficient grid schedule is needed . However, the dynamic nature of the grid and the various demands of users, due to the complexity of the grid schedule.

Genetic Algorithms are adaptive methods that can be used to solve optimization problems, based on the genetic process of biological organisms. Job scheduling on grid computing represents a great challenge. Genetic algorithms is one of the widely used algorithms for scheduling on grid computing.

The objectives of this research is to develop a job scheduling algorithm using genetic algorithm with high exploration processes. To evaluate the proposed scheduling algorithm we conducted a simulation using gridsim Simulator and a number of different work load.

The research found that genetic algorithm get best results in the best time when increasing the mutation and these result directly proportional with the increase in the number of job. Meaning that the mutation and exploration process has a good effect on the final execution time when we have large number of jobs. However, in small number of job mutation has no effects.

المستخلص

تعتبر الحوسبة الشبكية من الشبكات الصاعدة في شبكات الحواسيب مما تقدمه من مميزات عن الشبكات الأخرى من امكانية اتاحة الموارد المختلفه في الشبكة مجانا ، لطالما كانت معظم الحوسبات الشبكية تتيح لنا تبادل المعلومات وهذا ماجعلها تاخذ طابعا اخر من حيث الاستفادة من الموارد التي قد لا تكون متاحة بل مستحيله لامكانيات الجهاز الواحد.

هناك عدة خوارزميات تعمل علي بيئة الشبكة الحاسوبية لتقوم بجدوله المهام وتنظيم عمل الشبكة ومن الخوارزميات المستخدمة في ذلك الخوارزمية الجينية وهي واحدة من خوارزميات الامثلية حيث انها تقوم بالبحث عن افضل حلول للجدوله المهام في الشبكة الحاسوبية مستخدمة في ذلك مفهوم الوراثة في الكائنات الحية حيث يتم عمل عينة من الحلول ومن ثم دمجها لايجاد حلول اخري و عمل تغير في هذه الحلول للحصول علي حلول جديدة وكل ذلك للبحث عن افضل جدوله للمهام.

الهدف من هذا البحث تطوير وتعديل الخوارزمية الجينية باستكشاف عالي وزيادة نسبة الطفرة في الخوارزمية بدلا من النسبة القليلة التي اعتمدت عليها الخوارزمية الجينية، ولتقييم اداء الخوارزمية تم استخدام المحاكى Gridsim وتحميل عدد مختلف من المهام .

توصلت الى انه عند زيادة الاستكشاف (الطفرة) وهو البحث عن حل بعيد عن الحلول المتاحة كان اداء

الخوارزمية الجينية افضل وزمن التنفيذ الكلي اقصر وهذه الاداء كان يتناسب طرديا مع زيادة عدد المهام وهذا

يعني انه كلما كان عدد المهام كبير من الافضل زيادة الطفرة للوصول الي حلول جيدة، اما اذا كان عدد المهام

صغير فزيادة الاستكشاف غير مؤثرة

1.1 Introduction

This chapter introduces the research work; describe the research problem and objectives.

1.2 Problem Statement

Job scheduling on grid computing represents a great challenge. Genetic algorithms is one of the widely used algorithms for scheduling on grid computing. The process of exploring new solution in the genetic search space is an important issue. Because sometimes the initial population so far from the best solution because initial population is random.

1.3 Research Objectives

To develop a job scheduling algorithm for genetic algorithm with high exploration processes.

1.4 Thesis Structure:

Chapter 2: provide the literature view.

Chapter 3: describe research Methodology.

Chapter 4: explain the proposed method.

Chapter 5: provide the evaluation and results.

Chapter 6: conclude the research presents the recommendations