UNIVERSITI PUTRA MALAYSIA

LOAD-BALANCING MODELS FOR SCHEDULING DIVISIBLE LOAD ON LARGE SCALE DATA GRIDS

MONIR ABDULLAH ABDUH KAID

FSKTM 2009 8

LOAD-BALANCING MODELS FOR SCHEDULING DIVISIBLE LOAD ON LARGE SCALE DATA GRIDS

MONIR ABDULLAH ABDUH KAID

DOCTOR OF PHILOSOPHY
UNIVERSITI PUTRA MALAYSIA

2009

LOAD-BALANCING MODELS FOR SCHEDULING DIVISIBLE LOAD ON LARGE SCALE DATA GRIDS

By

MONIR ABDULLAH ABDUH KAID

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for Degree of Doctor of Philosophy

June 2009

TO MY PARENTS, MY WIFE, MY KIDS, AND TO ALL MY BROTHERS AND SISTERS

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

LOAD-BALANCING MODELS FOR SCHEDULING DIVISIBLE LOAD ON LARGE SCALE DATA GRIDS

By

MONIR ABDULLAH ABDUH KAID

June 2009

Chairman: Mohamed Othman, PhD

Faculty: Computer Science and Information Technology

In many data grid applications, data can be decomposed into multiple independent sub datasets and distributed for parallel execution. This property has been successfully employed using Divisible Load Theory (DLT), which has been proven to be a powerful tool for modeling divisible load problems in large scale data grid. Load balancing in such environment plays a critical role in achieving high utilization of resources to schedule the applications efficiently through join consideration of communication and computation time. There are some scheduling models, which have been studied, such as Constraint DLT (CDLT), Task Data Present (TDP) and Genetic Algorithm (GA). However, there has been no optimal solution reached. At the same time, effective schedulers are not only required to minimize the maximum completion time (*makespan*) of the jobs, but also the execution time of the schedulers.

This thesis proposes several load balancing models for scheduling divisible load on large scale data grids, when both processor and communication link speed are heterogeneous. The proposed models can be decomposed into three stages. The first stage is to develop new DLT based models for multiple sources scheduling. Closed form solutions for the load allocation are derived. The new models are called Adaptive DLT (ADLT) and A^{2}DLT models. In the second stage, an Iterative DLT (IDLT) model is proposed. Recursive numerical equations are derived to find the optimal workload assigned to the grid node. The closed form solutions are derived for the optimal load allocation. Although the IDLT model is proposed for single source, it has been applied in the case of multiple sources. The third stage integrates the proposed DLT based models with GA algorithm to solve the time consuming problem. In addition, the integration of the proposed DLT model with Simulated Annealing (SA) algorithm has been also developed.

The experimental results have proven that the proposed models yield better performance than previous models in terms of *makespan* and scheduler execution time. The ADLT and A^{2}DLT models have reduced the *makespan* by 21% and 37% respectively compared to CDLT model. The IDLT model is capable of producing almost optimal solution for single source scheduling with low time complexity. In addition, the integration of the proposed DLT model with GA and SA algorithms has also significantly improved the performance. The SA is 64.70% better than GA in terms of *makespan*. Thus, the proposed models can balance the processing loads efficiently so that they can be integrated in the existing data grid schedulers to improve the performance.

Abstrak thesis yang diserahkan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

MODEL KESEIMBANGAN MUATAN UNTUK MENJADUAL MUATAN YANG BOLEH DIBAHAGI PADA GRID DATA BERSKALA BASAR

Oleh

MONIR ABDULLAH ABDUH KAID

June 2009

Pengerusi: Mohamed Othman, PhD

Fakulti: Sains Komputer dan Teknologi Maklumat

Dalam kebanyakan aplikasi grid data, data boleh diuraikan kepada pelbagai sub dataset bebas dan disebarkan untuk dilaksanakan secara selari. Ciri-ciri ini berjaya diguna-pakai melalui penggunaan Teori Muatan Mampu-Dibahagi (Divisible Load Theory - DLT) yang telah terbukti sebagai alat yang berkuasa untuk memodelkan masalah muatan yang boleh dibahagi pada grid data berskala besar. Keseimbangan muatan dalam persekitaran ini memainkan peranan yang kritikal bagi mencapai kekerapan penggunaan yang tinggi bagi sesuatu sumber untuk penjadualan aplikasi yang berkesan dengan mengambil kira komunikasi dan tempoh penaksiran. Terdapat beberapa model penjadualan seperti DLT Desakan (Constraint DLT - CDLT), Kehadiran Data Bertugas (Task Data Present - TDP), dan Algoritma Genetik (Genetic Algorithma - GA). Walaubagaimanapun, tiada penyelesaian berkesan dapat dicapai. Pada masa yang sama, penjadualan yang berkesan tidak hanya memerlukan pengu-

rangan masa penyelesaian maksima bagi sesuatu tugas, tetapi juga meminimakan masa perlaksanaan bagi penjadualan tersebut.

Kajian ini mencadangkan beberapa model muatan seimbang untuk menjadualkan muatan yang boleh dibahagi pada grid data berskala besar apabila kelajuan pemproses dan talian komunikasi adalah berlainan. Model yang dicadangkan dapat diuraikan kepada tiga tahap. Tahap pertama ialah mengembangkan model DLR baru untuk menjadualkan pelbagai sumber. Kaedah berdasarkan pendekatan tertutup bagi pembahagian beban telah dihasilkan. Pada tahap ini, dua model baru dikenali sebagai DLT Adaptasi (Adaptive DLT - ADLT) dan A^2DLT dicadangkan. Pada tahap kedua, model DLT Iteratif (Iterative DLT - IDLT) telah dicadangkan. Persamaan ber-angka berulang-semula dihasilkan bagi mencari muatan tugas yang optima berhubung dengan nod grid. Kaedah berdasarkan pendekatan tertutup ini dihasilkan untuk mencari penempatan muatan tugas yang optima. Walaupun model IDLT ini dicadangkan untuk sumber tunggal, ia turut digunakan bagi kes-kes sumber yang pelbagai. Tahap ketiga pula menyatupadukan model DLT dan GA untuk menyelesaikan masalah pengiraan masa. Sehubungan itu, integrasi antara model DLT yang dicadangkan dengan algoritma Penguatan Di-simulasi (Simulated Annealing - SA) juga telah dibangunkan.

Keputusan eksperimen menunjukkan model yang dicadangkan telah memperoleh pencapaian yang lebih baik berbanding dengan model terdahulu dalam konteks masa penyelesaian maksima dan masa pelaksanaan oleh penjadual. Model ADLT dan

A^2DLT telah mengurangkan masa penyelesaian maksima masing-masing antara 21% dan 37% berbanding dengan model CDLT. Model IDLT pula berkeupayaan untuk menghasilkan penyelesaian paling baik bagi penjadualan sumber tunggal dengan kerumitan masa yang rendah. Selain itu, integrasi antara model DLT yang dicadangkan dengan GA dan algoritma SA telah meningkatkan prestasi secara signifikasi. Algoritma SA adalah 64.70% lebih baik berbanding GA dari segi masa penyelesaian maksima. Oleh yang demikian, model yang dicadangkan mampu mengimbangi muatan pemprosesan dengan berkesan seterusnya dapat disatupadukan dalam penjadualan grid data yang sedia ada bagi memperbaiki prestasinya.

ACKNOWLEDGMENTS

All Praise is due to Almighty ALLAH as He is all Mercifull, Most Gracious and Most Compassionate who gathered all knowledge for eternity. May ALLAH praise our Prophet Muhammad, and render him and his household safe and secure from all evil. All Grace and Thanks belongs to Almighty Allah.

Many special thanks and gratitude go to my supervisor Associate Professor Dr. Mohamed Othman, for his extensive help and valuable comments and advice at every stage of my research.

I would like to take this opportunity to express my sincere appreciation and thanks to the supervisory committee members, Associate Professor Dr. Hamidah Ibrahim and Dr. Shamala Subramaniam for their advices and comments during the completion of this thesis.

Thanks to every person who has supported me to produce my thesis. I am very grateful to the Faculty of Computer Science and Information Technology and the staff of Postgraduate office, Library and Universiti Putra Malaysia, for providing a studying and research environment.

Finally, many thanks to my parents, my wife, my kids, my brothers, my sister, all the family members and friends for their love, constant support and encouragement in all my endeavors.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADLT	Adaptive Divisible Load Theory
ATDP	Adaptive Task Data Present
CDLT	Constraint Divisible Load Theory
CERN	Conseil Europen pour la Recherche Nuclaire
CMS	Compact Muon Solenoid
DIANE	Distributed Analysis Environment
DLT	Divisible Load Theory
DS	Dataset Scheduler
EGEE	Enabling Grids for E-sciencE
ES	External Scheduler
GA	Genetic Algorithm
HEP	High Energy Physics
IDLT	Iterative Divisible Load Theory
LAN	Local Area Network
LHC	Large Hadron Collider
LS	Local Scheduler
Mbps	Mega bye per second
NAREGI	National Research Grid Initiative
NP	None Polynomial
RADIS	Resource-Aware Dynamic Incremental Scheduling
SA	Simulated Annealing
TDP	Task Data Present
TSP	Traveling Salesperson Problem
VO	Virtual Organization
WAN	Wide Area Network

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 Background

Grid technology promises geographically distributed scientists to the access and sharing of physically distributed resources such as computing, communication and storage, and most importantly, data collections for large scale data intensive problems [1]. During the early stages, grid computing had research focused mainly on coordination of geographically distributed computing resources for high performance. However, in many areas of science and engineering such as the High-Energy Physics (HEP) and aerospace, requirements have emerged for collaborating and sharing huge amount of geographically distributed data as well as sharing high performance computing resources. For example, in Large Hadron Collider (LHC) experiment at Conseil Europen pour la Recherche Nuclaire (CERN), huge amounts of data in terabyte or petabyte scale are collected by observing particle collisions, and analyzed through different levels of data processing operations.

Because of the massive size and distributed nature of data in data grid, scheduling data grid applications must simultaneously consider communication and computation time to achieve high performance. One issue of the data grid system is the development of effective load balancing and scheduling techniques for the distribution of multiple processes on multiple processors. At the same time, these schedulers

should maintain the expected performance goals such as minimising execution time and communication delays, which will lead to maximization of resource utilisation. As the maximum completion time (*makespan*) minimization problem is known as NP-complete when the node and communication link speeds are considered [53, 65], this problem can only heuristically dealt with [66].

This thesis focuses on load scheduling and balancing on large scale data grid. Specifically, it focuses on issues associated with scheduling and the distribution of arbitrarily divisible loads among processors in such environment. This work is an extension of the existing DLT based model for scheduling divisible load in large scale data grid. The main objective in the research of the DLT based model is to determine the optimal fractions of the entire load to be assigned to each processor so that the maximum completion time of the entire load is at the lowest point possible.

1.1.1 Data Grid Computing

A classical grid is a collection of connected resources that are available for an application to perform a task. Nodes are often geographically distributed and heterogeneous in nature. Administrative domains are well connected throuth links of various performance and reliability.

A data grid is a system that deals with the control of sharing and managing large amounts of distributed data. Many scientific and engineering applications require