

การจัดลำดับกระแสนระดับสูงด้วยหลักการเลือกผลลัพธ์ระหว่างกลางที่ดีที่สุด  
สำหรับโปรแกรมประยุกต์ที่เป็นกระแสนแบบกวาดพารามิเตอร์



นาย คุณาพร ศรีมโนธรรม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2549

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

HIGH-LEVEL WORKFLOW SCHEDULING WITH BEST-INTERMEDIATE-RESULT-FIRST  
POLICY FOR WORKFLOW-BASED PARAMETER-SWEEP APPLICATIONS

Mr. Kunaporn Srimanotham

A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Engineering Program in Computer Engineering  
Department of Computer Engineering  
Faculty of Engineering  
Chulalongkorn University  
Academic Year 2006

**490717**

Thesis Title High-level Workflow Scheduling with Best-Intermediate-Result-First Policy for Workflow-Based Parameter-Sweep Applications

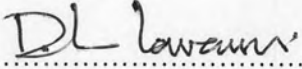
By Mr. Kunaporn Srimanotham

Field of Study Computer Engineering

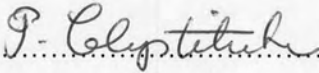
Thesis Advisor Veera Muangsin, Ph.D.

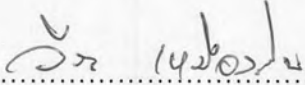
---

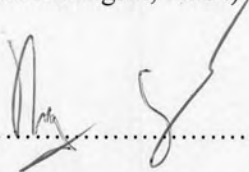
Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

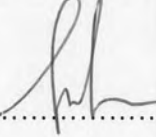
  
..... Dean of the Faculty of Engineering  
(Professor Direk Lavansiri, Ph.D.)

THESIS COMMITTEE

  
..... Chairman  
(Associate Professor Prabhas Chongstitvatana, Ph.D.)

  
..... Thesis Advisor  
(Veera Muangsin, Ph.D.)

  
..... Member  
(Natawut Nupairoj, Ph.D.)

  
..... Member  
(Assistant Professor Putchong Uthayopas, Ph.D.)

คุณาพร ศรีมโนธรรม : การจัดลำดับกระแสงงานระดับสูงด้วยหลักการเลือกผลลัพธ์ระหว่างกลางที่ดีที่สุดสำหรับโปรแกรมประยุกต์ที่เป็นกระแสงงานแบบกวาดพารามิเตอร์  
(HIGH-LEVEL WORKFLOW SCHEDULING WITH BEST-INTERMEDIATE-RESULT-FIRST POLICY FOR WORKFLOW-BASED PARAMETER-SWEEP APPLICATIONS)  
อาจารย์ที่ปรึกษา: ดร.วีระ เหมืองสิน, 97 หน้า.

โปรแกรมประยุกต์ที่เป็นกระแสงงานแบบกวาดพารามิเตอร์เป็นหนึ่งในเรื่องที่สำคัญในงานประเภทขนานบนระบบคลัสเตอร์และกริดในปัจจุบัน ระบบกระจายงานโดยทั่วไปและเครื่องมือของงานประเภทกวาดพารามิเตอร์ยังไม่มีประสิทธิภาพพอที่จะจัดการกับประเภทโปรแกรมประยุกต์ที่งานวิจัยกำลังสนใจอยู่ในขณะนี้ และยิ่งไปกว่านั้นนโยบายการกระจายงานโดยทั่วไปนั้นได้ถูกออกแบบเพื่อลดเวลาการทำงานโดยรวมของการกวาดพารามิเตอร์ทั้งหมด อย่างไรก็ตามโปรแกรมประยุกต์ที่เป็นแบบกวาดพารามิเตอร์ทั้งหลายนั้นมีวัตถุประสงค์ที่เหมือนกันคือต้องการได้ผลลัพธ์ที่ดีที่สุดหรือผลลัพธ์ที่อยู่ในอันดับต้นๆ งานวิจัยนี้จึงได้นำเสนอวิธีการกระจายงานโปรแกรมประยุกต์ที่เป็นกระแสงงานแบบกวาดพารามิเตอร์ เพื่อที่จะลดเวลาการประมวลผลกระแสงงานที่ให้ผลลัพธ์ที่ดีที่สุด กระบวนการทำงานนั้นมีหลักการทำงานพื้นฐานคือ การปรับระดับความสำคัญอยู่ตลอดเวลา ความสัมพันธ์ที่ได้รับระหว่างกลางซึ่งจะได้รับจากช่วงใดช่วงหนึ่งของกระแสงงาน เทคนิคนี้ได้ถูกนำไปประยุกต์ใช้สำหรับโปรแกรมประยุกต์ทางการค้นหา ผลลัพธ์จากการทดลองแสดงให้เห็นอย่างชัดเจนว่า เทคนิคในงานวิจัยนี้สามารถพัฒนาความสัมพันธ์ระหว่างอันดับคุณภาพของผลลัพธ์สุดท้ายและลำดับการทำงานเสร็จสิ้นของกระแสงงานได้อย่างมีประสิทธิภาพ

ภาควิชา.... วิศวกรรมคอมพิวเตอร์.....ลายมือชื่อนิสิต.....*Munpon Sirimanotham*.....  
สาขาวิชา....วิศวกรรมคอมพิวเตอร์.....ลายมือชื่ออาจารย์ที่ปรึกษา.....*Dr. Ueak*.....  
ปีการศึกษา .....2549.....ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

## 4770235421 : MAJOR COMPUTER ENGINEERING

KEY WORD: HIGH-LEVEL WORKFLOW SCHEDULING, BEST-INTERMEDIATE-RESULT-FIRST-POLICY, PARAMETER-SWEEP APPLICATION

KUNAPORN SRIMANOTHAM: HIGH-LEVEL WORKFLOW SCHEDULING WITH BEST-INTERMEDIATE-RESULT-FIRST POLICY FOR WORKFLOW-BASED PARAMETER-SWEEP APPLICATIONS, THESIS ADVISOR: VEERA MUANGSIN, Ph.D., 97 pp.

Workflow-based parameter-sweep applications are an important class of parallel jobs on clusters and grid today. Conventional batch schedulers and parameter study tools are not effective for this type of application. Especially, their scheduling policies are usually designed to minimize the makespan of the whole parameter study. However, many parameter-sweep applications also have a primary objective to obtain the best or a few top-ranked results from a large parameter space. This research describes a new heuristic for scheduling parameter-sweep workflows in order to minimize the turnaround time of the workflows that give the best results. The algorithm is based on dynamically adjusted priority according to intermediate data obtained at some stage in the workflow. The technique is applied on a high-throughput drug screening application. The experimental results show that our technique can significantly improve the correlation between the ranking of the final results and the order of completion of the workflows.

Department.... Computer Engineering.... Student's signature.... *Kunaporn Srivanotham*  
 Field of study.... Computer Engineering.... Advisor's signature.... *Dr. Veera Muangsin*  
 Academic year ...2006.....Co-advisor's signature.....

## ACKNOWLEDGMENTS

I gratefully wish to acknowledge my thesis advisor, Veera Muangsin, Ph.D., for his beneficial advice throughout this research.. I also appreciate many intellectual suggestions from all of my thesis committee, Associate Professor Prabhas Chongstitwatana, Ph.D, Natawut Nupairoj, Ph.D., and Assistance Professor Putchong Uthayopas, Ph.D.

There are a lot of thanks I would like to extend for all my friends especially the member of Scientific Parallel Computer Engineering Laboratory (SPACE Lab) for their abundant helps, encouragement and truly relationships.

Most of all, I deeply wish to express my thanks to my parents for their love, understanding and invaluable supports throughout my graduate study.

## TABLE OF CONTENTS

	Page
ABSTRACT (THAI) .....	iv
ABSTRACT (ENGLISH) .....	v
ACKNOWLEDGMENTS .....	vi
TABLE OF CONTENTS.....	vii
LIST OF FIGURES .....	x
LIST OF TABLES .....	xii
CHAPTER 1 INTRODUCTION .....	1
1.1 Overview .....	1
1.2 Motivation .....	1
1.3 Problem Statement.....	2
1.4 Objectives .....	2
1.5 Scopes .....	2
1.6 Research Plan .....	3
1.7 Expected Benefits .....	3
1.8 Thesis Structure .....	3
1.9 Publications .....	3
CHAPTER 2 BACKGROUND AND RELATED WORKS.....	5
2.1 Background.....	5
2.1.1 Job Scheduling on a Cluster Computer.....	5
2.1.2 Parameter-sweep Applications and Scheduling.....	6
2.1.3 Workflow and Directed Acyclic Graph .....	6
2.1.4 Workflow Scheduling .....	7
2.2 Related Works .....	7
2.2.1 Batch Schedulers.....	7
2.2.2 Workflow Schedulers.....	8
2.2.3 Parameter-sweep Schedulers .....	11
CHAPTER 3 THE BEST-INTERMEDIATE-RESULT-FIRST ALGORITHM.....	13
3.1 Assumptions .....	13
3.2 The Algorithm.....	13

CHAPTER 4 GENERIC APPLICATION EVALUATION.....	19
4.1 Simulation Tool .....	19
4.1.1 Programmation environments layer.....	20
4.1.2 Simulation kernel layer .....	21
4.1.3 Base layer.....	22
4.2 Workflow Scheduling Algorithms Simulation .....	22
4.2.1 Configuration process .....	23
4.2.2 Simulation process .....	30
4.2.3 Iteration process .....	35
4.2.4 Analytical process .....	36
4.3 Experiment.....	38
4.3.1 Environment.....	38
4.3.2 Experiment setting .....	38
4.3.3 Experiment design .....	39
4.4 Results and Discussion .....	39
4.4.1 Application class.....	39
4.4.2 The effect of the evaluation point.....	43
4.4.3 The effect of workflow computation size .....	47
CHAPTER 5 APPLICATION-SPECIFIC EVALUATION.....	51
5.1 High-level Workflow Scheduling Implementation.....	51
5.1.1 A design of the High-level workflow scheduling.....	51
5.1.2 Management of the High-level workflow scheduling .....	55
5.2 Drug Discovery Application.....	56
5.3 Experiment Settings .....	57
5.4 Results and Discussion .....	58
CHAPTER 6 CONCLUSION.....	62
REFERENCES .....	63
APPENDICES .....	66
APPENDIX A: The effect of correlation intermediate results and final results.....	67
APPENDIX B: The effect of evaluation point .....	83
APPENDIX C: The effect of workflow computation size .....	94



BIOGRAPHY .....97

## LIST OF FIGURES

Figure	Page
2-1 The scheduling system .....	6
2-2 A simple directed acyclic graph.....	7
2-3 Classical batch scheduler architecture .....	8
2-4 (a) DAGMan script (b) Job dependency represented by DAG.....	9
2-5 TrellisDAG architecture .....	10
2-6 Overall structure of Nimrod .....	12
3-1 Workflow and intermediate evaluation .....	14
3-2 Workflow scheduling system .....	16
4-1 Relationships between the SimGrid components .....	19
4-2 Simulation system.....	22
4-3 (a) Task dependency description (b) Dependency graph .....	25
4-4 Procedure of assigning workflow computation size .....	26
4-5 Simulation process.....	31
4-6 Input and output of before-evaluation-queue and after-evaluation-queue .....	33
4-7 Flowchart of iteration process.....	35
4-8 Relationship between average final ranking and completion order for class A application (correlation of results = 0.875).....	41
4-9 Relationship between average final ranking and completion order (a) class D application (correlation of results = 0.5), (b) class G application (correlation of results = 0.125).....	42
4-10 Relationship between average workflow completion time and final ranking for application class A (correlation of results = 0.875).....	42

4-11	Relationship between average workflow completion time and final ranking (a) application class D (correlation of results = 0.5) (b) application class G (correlation of results = 0.125).....	43
4-12	Correlation between order of workflow completion and final ranking.....	44
4-13	Relationship between final ranking and average completion time (hours) when evaluation point 30%.....	46
4-14	Relationship between final ranking and average completion time (hours) when evaluation point 50%.....	46
4-15	Relationship between final ranking and average completion time (hours) when evaluation point 70%.....	47
4-16	Correlation between completion order and final ranking when using difference workflow computation sizes.....	48
4-17	Relationship between final ranking and average completion time (hours) when using different workflow computation size.....	49
5-1	The framework of high-level workflow scheduling and management .....	52
5-2	Workflow description .....	54
5-3	The event of workflow scheduling and management.....	56
5-4	Drug screening workflow .....	57
5-5	Intermediate and final ranking .....	58
5-6	Average workflow ranking of completed workflows .....	59
5-7	Improved average workflow ranking .....	59
5-8	Average start and finish times.....	60

## LIST OF TABLES

Table		Page
3-1	Summarizes the algorithms and their probability value .....	17
4-1	Probability of queue selection for each scheduling type.....	34
4-2	Simulation results .....	35
4-3	Results from iteration process .....	36
4-4	Analytical results .....	37