

MARA University of Technology

HEAT EQUATION ANALYSIS

USING

PARALLEL PROGRAMMING TOOLS

Emilia Binti Mohamad Shukree

2004617904

Thesis submitted in fulfillment of the requirements for Bachelor of Science (Hons) Data Communications and Networking

> Faculty of Information Technology And Quantitative Science

> > November 2007

DECLARATION

I certify this thesis and the research to which it refers are the product of my own work and that any ideas or quotation from work of the other people, published or otherwise are fully acknowledge in accordance with the standard referring practices of the discipline.

.....

DECEMBER 3, 2007

EMILIA BINTI MOHAMAD SHUKREE

2004617904

APPROVAL

HEAT EQUATION ANALYSIS USING PARALLEL PROGRAMMING TOOLS

BY

EMILIA BINTI MOHAMAD SHUKREE

This thesis was prepared under the supervision of thesis coordinator, En. Adzhar Abd. Kadir, Department of Computer Technology and Networking and has been approved by the thesis supervisor, Pn. Siti Arpah Ahmad. It was submitted to the Faculty of Information and Quantitative Science and was accepted in partial fulfillment of the requirement for the degree of Bachelor of Science.

Approved by:

fet

Siti Arpah Ahmad Thesis Supervisor

Date: DECEMBER 3, 2007

ACKNOWLEDGEMENT

بالسمشالرحمان الرحيم

(In the name of Allah, the most Gracious and the most Merciful)

Alhamdulilah, praise and blessing of Allah Al-Mighty for giving me good health of body and mind, strength, patience and ability to complete this thesis project. The project delivered in this paper could not have been accomplished without the help from Him.

First and foremost, I would like to express my appreciations and million of thanks to my supervisor, Puan Siti Arpah Ahmad for her guidance, ideas, advices and patience in assisting me in my thesis from the beginning until the completion of this project. I feel very lucky to be her supervisee because the completion of this project run smoothly with her continuous courage and any problem arose could be solved straight away. Her relentless effort in guiding me could not be forgotten and will be remembered for the rest of my life.

Highest form of gratitude is dedicated to Encik Adzhar Abd Kadir for coordinating this thesis project, besides giving advices and guidance along the producing of this thesis report. His numerous invaluable advice, positive comments, continuous guidance and persistence encouragement throughout the course of this project give me a lot of information and could assist me in producing better report writing. His critics could be taken as an advice and helped brightens me up the unclear part of my project.

Special thanks also dedicated to my family, especially to my mum, Puan Azizah Mohamed Nor for her great support towards me even though this is the second chance for me to complete this thesis project. Even though she was quite disappointed before, she still encourages me to accomplish this thesis and achieve great results. After she prays day and night for me, finally I could finish completing this project and hopefully it will be great news for her. I love u mum. Not forgetting also to my lovely sister, Puan Edleen Mohamad Shukree for giving me a nonstop of moral support and who had always helped me in difficult moments. She is also the person who always had the patience to listen to my problems during accomplishing my thesis.

To my dearly loved friends of B. Sc. (Hons.) Data Communications and Networking intake of June 2004 and July 2005, thank you for all your support and brilliant ideas throughout the duration of the thesis project starting from the proposal stage until the completion of my thesis. Special gratitude also dedicated to my best friend, Hazrin Mohammad for kindly helping me in technically and mentally support. Without his helped, I do not think that I will carry out with this thesis successfully. Your kindheartedness will be remembered and appreciated.

Lastly, my gratitude extends to all individuals who have contributed either directly or indirectly in preparing this research successfully. I thank you to all of you for your support.

Wassalam.

Emilia Binti Mohamad Shukree November 2007

ABSTRACT

Many new technologies, in terms of software and hardware had been introduced that had made computing power doubles every year. Programmers and software developers had introduced more high quality software's, vivid and clear animations, more real time simulations, and any other programs that need more powerful processing. Sequential processing and computing is a thing of the past and cannot afford to today's software's and programs that has the quality as mentioned earlier. That's what had made parallel computing and processing is the choice in the world of today. However, due to cost constrains, some organizations could not afford to have today's parallel supercomputers. Hence, urging developers and programmers to come up with a solution which is to virtualize the use of parallel processing which is why the Parallel Virtual Machine had come out as of today, to benefit organizations by connecting a few personal computers (PC's) via today's standard networking architectures.

This research project is conducted to find the time difference between the parallel heat equation program that was built by the author derived from the former parallel coding by using C language and the modified sequential heat equation program in peer-to-peer (P2P) network and star topology network. SSH communications standard via RSA key will be used between those PC's in both network architecture to ensure that the PVM daemons on those PC's can be contacted and making sure that the PVM run smoothly. It is expected that the parallel heat equation programming will execute faster than the latter sequential heat programming code.

TABLE OF CONTENTS

CONTENTS

DECLARATION		
APPROVAL		
ACKNOWLEDGEMENT		
ABSTRACT.		vi
TABLE OF C	CONTENTS	vii
LIST OF FIG	URES	xii
LIST OF TAI	BLES	xiii
LIST OF GRA	APHS	xiv
LIST OF ABI	BREVIATION	XV
1. CHAP	PTER 1: INTRODUCTION	1
1.1	Background	1
1.2	Problem Statement	2
1.3	Objective of the project	3
1.4	Scope of the project	4
1.5	Significance of the project	4
1.6	Chapter overview	5
1.7	Conclusion	6
2. CHAP	PTER 2: LITERATURE REVIEW	7
2.1	Introduction	7
2.2	Serial Computing	7
2.3	Parallel Computing	8
2.4	Parallel Computing Architectures	10

	2.4.1	Single Instruction, Single Data (SISD)	11
	2.4.2	Multiple Instruction, Single Data (MISD)	12
	2.4.3	Single Instruction, Multiple Data (SIMD)	13
	2.4.4	Multiple Instruction, Multiple Data (MIMD)	14
2.5	Memor	ry Architectures	15
	2.5.1	Shared Memory	16
	2.5.2	Distributed Memory	17
2.6	Virtual	Machine	18
2.7	Paralle	l Virtual Machine (PVM) Overview	19
	2.7.1	PVM System	20
		2.7.1.1 Daemon	20
		2.7.1.2 Library	22
2.8	Master	/Slave Computational Model	23
2.9	Secure	Shell (SSH)	24
2.10	RSA Algorithm		
2.11	Heat Equation 2		
2.12	Overview of Heat Equation Application 2		26
	2.12.11	FEMLAB® Heat Transfer Module Simulates	
		Complex Temperature Fields in Parallel with	
		Other Physics	26
2.13	Related	d Projects	27
	2.13.1	Parallelization Using Parallel Virtual Machine by	
		Norikhti Faqriah Azime, Mara University of	
		Technology, Malaysia (2006)	27
	2.13.2	Non-stationary Iterative Solvers on a PC Cluster	
		by B.V. Rathish Kumar, Shalini and Mani	
		Mehra, Indian Institute of Technology, Kanpur,	
		India (2004)	28
	2.13.3	3-D Heat Conduction on a Linux Cluster using	
		PVM by Timothy Salter, Alabama A&M	
		University, U.S.A (2003)	28

- 2.13.4 Crystal Growth Simulation on Clusters by Dana
 Petchu, Daniel Vizman, Western
 University, Timisoara, Jochen Friedrich, Institute
 for Intergrated Circuits Erlangen, Germany, and
 Monica Popescu, Aurel Vlaicu University,
 Romania (2002).....
- 2.13.5 The AGEB Algorithm for Solving the Heat Equation in Three Space Dimensions and Its Parallelization Using PVM by Mohd Salleh Sahimi, Universiti Tenaga Nasional, Malaysia, Norma Alias, Universiti Kebangsaan Malaysia, Malaysia and Elankovan Sundararajan (2001)..... 29

ix

29

	3.2	Metho	odology and Research Approach	32
	3.3	Phase	1: Project Planning	34
		3.3.1	Problem Identification	34
	3.4	Phase	2: Data Gathering	35
	3.5	Phase	3: Implementation	39
	3.6	Phase	4: Testing	39
	3.7	Phase	5: Analysis and Evaluation	40
	3.8	Summ	nary	40
4.	CHAP	PTER 4	: IMPLEMENTATION AND TESTING	41
	4.1	Backg	ground	41
	4.2	Hardy	vare Design	41
	4.3	Softw	are Installation and Configuration	43
		4.3.1	Operating System	43
		4.3.2	Ubuntu Software Package	46
		4.3.3	Communication Protocol	47
		4.3.4	Parallel Programming Tools	56
		4.3.5	PVM GUI	61
		4.3.6	Graph Plotting Software	62
	4.4	Pilot I	Run	63
	4.5	Parall	el and Sequential Heat Equation Program	
		Expla	nation	64
		4.5.1	Parallel Heat Equation Program	64
		4.5.2	Modified Sequential Heat Equation Program	65
	4.6	Testin	ıg	66
		4.6.1	Peer-to-peer Testing	67
			4.6.1.1 Parallel Heat Equation Program	67
			4.6.1.2 Sequential Heat Equation Program	71
			4.6.1.3 Comparison between the Time Taken for	
			Parallel (P2P) and Sequential Program	76

		4.6.2	Star Topology Testing	76
			4.6.2.1 Parallel Heat Equation Program	77
			4.6.2.2 Comparison between the Time Taken for	
			Parallel (Star Topology) and Sequential Program	78
	4.7	Summ	nary	78
5	. CHAI	PTER 5	: FINDINGS AND DISCUSSION	79
	5.1	Backg	ground	79
	5.2	Time	Difference / Comparison	79
	5.3	Summ	nary	88
6	. CHAI	PTER 6	CONCLUSION AND RECOMMENDATION	89
	6.1	Backg	ground	89
	6.2	Concl	usion	89
	6.3	Recor	nmendation	90
7	. REFE	RENCE	ES	91
8	. APPE	NDICE	S	95

LIST OF FIGURES

FIGURES

Figure 2.1 :	Process of serial computing	8
Figure 2.2 :	Process of parallel computing	9
Figure 2.3 :	Flynn's Taxonomy	10
Figure 2.4 :	Single Instruction, Single Data (SISD)	11
Figure 2.5 :	Multiple Instruction, Single Data (MISD)	12
Figure 2.6 :	Single Instruction, Multiple Data (SIMD)	13
Figure 2.7 :	Multiple Instruction, Multiple Data (MIMD)	15
Figure 2.8 :	Distributed Memory Architecture	16
Figure 2.9 :	Shared Memory Architecture	18
Figure 2.10 :	Master/Slave Computational Model	23
Figure 3.1 :	Project Phase Diagram	33
Figure 4.1 :	Peer-to-peer	42
Figure 4.2 :	Star topology	43
Figure 4.3 :	Disk Partition on Master PC	45
Figure 4.4 :	OpenSSH in Synaptic Package Manager	48
Figure 4.5 :	Process Creating RSA Key on Master	54
Figure 4.6 :	Establish Connection with Slave1 (P2P)	60
Figure 4.7 :	Establish Connection with Slave1 and Slave2 (Star	
	Topology)	61
Figure 4.8 :	XPVM Interface	62
Figure 4.9 :	Pilot Run on Hello World Program (P2P)	64
Figure 4.10 :	Spawn Parallel Heat Equation Program (Peer-to-peer)	68
Figure 4.11 :	Graph1 ($\Delta t = 5.0 \text{xe-1}$)	69
Figure 4.12 :	Graph2 ($\Delta t = 5.0 \text{xe}-3$)	70
Figure 4.13 :	Graph3 (Δt =5.0xe-6)	70
Figure 4.14 :	Graph4 ($\Delta t = 5.0 \text{xe-9}$)	71
Figure 4.15 :	Compilation of Modified Sequential Heat Equation	
	Program	73
Figure 4.16 :	GraphA ($\Delta t = 5.0 \text{xe-1}$)	74
Figure 4.17 :	GraphB ($\Delta t = 5.0 \text{xe-}3$)	74
Figure 4.18 :	GraphC ($\Delta t = 5.0 \text{ xe-} 6$)	75
Figure 4.19 :	GraphD ($\Delta t = 5.0 \text{ xe-9}$)	75
Figure 5.1 :	Time taken for parallel heat equation code	80
Figure 5.2 :	XPVM with heat equation program being run (P2P)	81
Figure 5.3 :	XPVM with heat equation program being run (Star	
	Topology)	81

LIST OF TABLES

TABLES

Table 3.1 :	Master (mimi-laptop) Hardware Specifications	36
Table 3.2 :	Slave1 (hazrinmd-laptop) Hardware Specifications	37
Table 3.3 :	Slave2 (ubuntu-pc) Hardware Specification	37
Table 3.4 :	Software Specifications	38
Table 4.1 :	Time Runs for Problem Size (Parallel in P2P)	68
Table 4.2 :	Time Runs for Δt (Parallel P2P)	69
Table 4.3 :	Time Runs for Problem Size (Sequential)	72
Table 4.4 :	Time Runs for Δt (Sequential)	72
Table 4.5 :	Time Runs for Problem Size (Parallel Star Topology)	77
Table 5.1 :	Results from 10 Problem Sizes.	83

LIST OF GRAPHS

TABLES

Parallel and Sequential Heat Equation (Peer-
Parallel and Sequential Heat Equation (Star
on between Parallel (Peer-to-Peer and Star
equential

LIST OF ABBREVIATION

Abbreviation Full Text

2D	2 Dimensions
3D	3 Dimensions
AMD	Advanced Micro Devices
СРИ	Central Processing Unit
DAEMON	Disk And Execution Monitor
DDR	Double Data Rate
DDR2	Double Data Rate 2
FORTRAN	FORmula TRANslation
FTMSK	Fakulti Teknologi Maklumat dan Sains Kuantitatif
G++	GNU C++ Compiler
GCC	GNU C Compiler
GNU	GNU's Not Unix
GTK	GIMP Tool Kit for C++
IP	Internet Protocol
IT	Information Technology
LCD	Liquid Cristal Display
MIMD	Multiple Instruction, Multiple Data
MISD	Multiple Instruction, Single Data
MPI	Message Passing Interface
NIC	Network Interface Card
OS	Operating System
P2P	Peer-to-Peer

PC	Personal Computer
PU	Processing Unit
PVM	Parallel Virtual Machine
RAM	Random Access Memory
RSA	Rivest, Shamir, & Adleman
RSH	Remote Shell
SIMD	Single Instruction, Multiple Data
SISD	Single Instruction, Single Data
SMP	Symmetric MultiProcessor
SSH	Secure SHell
UNIX	UNIpleX information and computer services
UTP	Unshielded Twisted Pair
XPVM	Executable Parallel Virtual Machine (GUI)

CHAPTER 1

INTRODUCTION

1.1 Background

Back to these past several years, there are many new technologies has been launched to the computer technology field. This is due to the evolution of technological advances of information technology and people are getting excited to have a better performance in terms of speediness as time passed by. The eagerness of these peoples have encourage the IT (Information Technology) experts to produce much better equipment with low price but boost up the performance and generates faster speed time rather than higher price equipment with the same performance. As a result, we can see the significant growth in multiprocessing computing and it has been the introduction to the parallel computing.

Parallel computing is the simultaneous execution of the same task (split up and specially adapted) on multiple processors in order to obtain results faster. The idea is based on the fact that the process of solving a problem usually can be divided into smaller tasks, which may be carried out simultaneously with some coordination.¹ Today, many large companies or labs need to solve complicated tasks which require large amount of processing. Parallel computing could truly help them to solve their problems with processing and many of developed parallel programming tools such as Parallel Virtual Machine (PVM) could be used in order to the implementation of parallel computing system.

¹ http://en.wikipedia.org/wiki/Parallel_computing

PVM is a software tool for parallel networking of computers. It is designed to allow a network of heterogeneous machines to be used as a single distributed parallel processor.² PVM permits a heterogeneous collection of Unix and/or Windows computers hooked together by a network to be used as a single large parallel computer. Thus, large computational problems can be solved more cost effectively by using the aggregate power and memory of many computers rather than we use a supercomputer which needs a higher cost. PVM is free software and could be downloaded from the internet and easy to use. Furthermore, the source code freely available and people can use the existing source code and modify it so that it can be used for their similar project or work area.

In this research project, two types of network architecture; P2P and star topology will be used. In P2P network, two personal computers (PC's) will be networked together and parallel programming tool; PVM will be installed and configured for parallelization of both PC's while three (3) PC's will be used for star topology network and also will be parallelized using PVM. By using networked PC rather than supercomputer or new PC with high performance, this could provide large benefits to companies or labs. The sample of parallel heat equation program will be compared between the developed sequential heat equation program in terms of speed and performance on both types of network architecture. Large or complex computational problem could be solved by utilize the use of old computers and put it together and parallelized them.

1.2 Problem Statement

The idea of the come out with this research project is derived from the problems encountered in sequential computing. Sequential computing may lead to low performance for large computational problems such as the complex calculation of heat equation. For sequential computing, complex calculations needs extremely high of CPU (Central

² http://en.wikipedia.org/wiki/Parallel_Virtual_Machine

Processing Unit) usage and if just a single processing is used, it will burden and may lead to the damage of the machine. These kinds of problems will increase the time taken and slower the production of the results especially when the machine were crashed or hanged.

If we networked together the old machine and parallelized them, we may utilize the used of many processors and share the tasks where it could increase the performance and no problems will occur if new tasks were added. The cost also could be decreased if we networked together the old or used machines rather than we use a supercomputer with single high performance processor which is extremely expensive. By gathering many workstations like the one that we use for personal computing today, we can save the cost but have almost the same performance as a supercomputer.

1.3 Objective of the project

This project is done to achieve three main objectives:

- To install and configure parallel programming tools in Linux environment for three hosts which will be used in two types of network architectures which are peer-to-peer and star topology networks in order to run the parallel heat equation program.
- 2. To modify the sample of parallel heat equation program in order to build a sequential heat equation program.
- 3. To analyze the performance and time consumed between both of the parallel and sequential program in those two types of network architecture.

1.4 Scope of the project

The scope that has be considered in this project:

- The entire project will be limited to Linux environment because those PC's were installed with Linux-based operating system.
- 2. The modification of sequential heat equation program is limited to C language because the sample of parallel heat equation program is written in C language.
- 3. This project also limited to simple heat equation program.

1.5 Significance of the project

This project could benefits mostly to the communities:

- 1. This project will contribute to scientific labs or non-profit organizations which do complex calculations due to the effectiveness of the project in terms of cost and performance.
- 2. This project will also contribute to the FTMSK (Fakulti Teknologi Maklumat & Sains Kuantitatif) students which could be exposed with parallel computing and they could expand the scope of this research.

1.6 Chapter overview

Following is the introductory of the chapters:

Chapter 1: INTRODUCTON

Introduce and overview the background of the project and main objectives that need to be accomplish and achieve in finishing the project.

Chapter 2: LITERATURE REVIEW

Covers literature review and general information of the technology and terminology that will be used in the development of the project. The related works also will be discussed in this chapter.

Chapter 3: METHODOLOGY

Includes the methodology and design involved in this project. The techniques used to complete the project efficiently within time given will be explained further in this chapter.

Chapter 4: IMPLEMENTATION

Explains the installation, configuration and testing of the project as a whole.

Chapter 5: FINDINGS AND DISCUSSION

Contains the discussion of the results from project's output. The outputs are figured up based on implementation in the project, either it achieves the objective or not.

Chapter 6: CONCLUSION AND RECOMMENDATION

Comprises the conclusion of the project as a whole and it will be evaluated based on the achievement of the goals and objective of the project.

References: Lists of further information sources.

Appendices: Collections of data and source code. This part also includes the project outline.

1.7 Conclusion

With the understanding the concept of this project, the exploration of parallel computing could be implemented. It follows up by the problems aroused that might be solved by the existing of this research. The objectives of this research were stated to give a clear overview of what is need to be achieved within the area of the limited scope. The successful of this project will give benefits to the community especially to scientific labs and FTMSK students in parallelizing computers and gain benefits from it.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter will be a review on some literature that is related to the research area which includes the supporting technologies including its definition and explanations. This will follow by the review on the similar works and research from the other researchers will be included in this chapter. From the particular to the area of research, literature review is the documentation of a comprehensive review of the published and unpublished work from secondary source of information. Most of these articles or literatures are gathered on browsing through the internet and reference books.

2.2 Serial Computing

What is serial? Serial means one by one. Serial data transfer refers to transmitting data one bit at a time.³ Serial computing which is also known as sequential computing means a single computer having a single Central Processing Unit (CPU) to be run on (*Barney, 2007*).⁴ Traditionally, software has been written for serial computation. A problem is broken into a discrete series of instructions and it will be executed one after another. Only one instruction may execute at any moment in time. The example of serial computing is show in Figure 2.1 below:

³ http://www.webopedia.com/TERM/s/serial.html

⁴ http://www.llnl.gov/computing/tutorials/parallel_comp/



Figure 2.1: Process of serial computing

2.3 Parallel Computing

Parallel computing is the simultaneous execution of some combination of multiple instances of programmed instructions and data on multiple processors in order to obtain results faster. The idea is based on the fact that the process of solving a problem usually can be divided into smaller tasks, which may be carried out simultaneously with some coordination.⁵ In the simplest sense, parallel computing is the simultaneous use of multiple compute resources to solve a computational problem. A problem is broken into discrete parts that can be solved concurrently and each part is further broken down to a series of instructions. Those instructions from each part will execute simultaneously on different CPUs. Figure 2.2 below will be the detailed explanation of the process:

⁵ http://en.wikipedia.org/wiki/Parallel_computing