PARALLEL MATRIX MULTIPLICATION ANALYSIS

HAZRIN BIN MOHAMMAD

2004617831

Thesis report submitted in fulfillment for the requirements for

Bachelor of Science (Hons) Data Communications and Networking

Faculty of Information Technology and Quantitative Sciences

Mara University of Technology

NOVEMBER 2007

DECLARATION

In the name of Allah the Almighty, I hereby certify that this thesis and the research to which it refers are the produce of my own and any parts, ideas, or quotes from the research or work which belongs to any other people that will be cited here in this research hereafter will be acknowledged in full accordance of the discipline and the standard of referring practices.

29th OCTOBER 2007 ……..

HAZRIN BIN MOHAMMAD

2004617831

APPROVAL

PARALLEL MATRIX MULTIPLICATION ANALYSIS

BY

HAZRIN BIN MOHAMMAD

The thesis has been prepared by the author under the supervision of both of En. Adzhar Bin Abd Kadir and Pn. Siti Arpah Binti Ahmad, the thesis coordinator and the thesis supervisor, respectively, and the title of the thesis had been approved by the thesis supervisor. It had been submitted to the Faculty of Information Technology and Quantitative Sciences as a partial requirement for the Bachelors Degree (with honours) in Data Communication and Networking.

Approved By: .....

Pn. Siti Arpah Binti Ahmad

Thesis Supervisor

Date: 3rd DECEMBER 2007

ACKNOWLEDGEMENT

بالسمﷲ الرحمان الرحيم

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

السلام علايكم ورحمة ﷲ وبركاتة

(May peace be upon all of you, the reader)

Alhamdulillah, soon after a period of four months (in around a semester) and after a year (two semesters) doing both research proposal and completing the research report, I finally can breath in relief as the hardest part in completing this bachelor's degree program had come to the end. But I always do bear in my mind that the success did not come through my works alone, but instead, through other people's contribution in ideas and their hard work in making sure that this research is a success. As the completion of this research and its report, I would like to take this opportunity to express my form of gratitude to those whom, without their help, would compromise the success of this research.

As one of His humble creation, I would like to express my deepest gratitude to Allah S.W.T. for bestowing me with all the strengths and motivation for me to successfully complete this research. Obviously, without His blessings, this research, from the proposal to the completion of the report, wouldn't have gone this far. Alhamdulillah, Thank You Allah.

Deepest form of gratitude also goes for my research project supervisor, Puan Siti Arpah Binti Ahmad for her direct and continuous support in terms of ideas and helps just to make sure that this research project goes on its track as scheduled. Her relentless effort and support gave me major motivation for me to complete this project as scheduled. Her ideas and advices gave me moral support and brighten and enlightened me through dark times especially whenever this research came to a halt.

She is one good lecturer and advisor and tackles all problems professionally through all times without hassle.

Not forgetting also to all my friends and classmates of B. Sc. (Hons.) Data Communications and Networking intake of June 2004 and July 2005 for all their advices and ideas and mental support, not forgetting all their critics and skepticism that made me go further in this research project towards the completion of this research report. Here I would like to express highest gratitude especially to Emilia Binti Mohamad Shukree as her research project also in similar field as mine, only to few different scope and characteristics. She had given me so many advices and had helped me a lot in this research project in terms of technical aspects. Besides, her contribution in terms of hardware and software meant a lot to me and without her help I am sure that this research project would have never been this far. As one saying goes, "A friend indeed is a friend in need" and "Two heads are better than one". I thank you to all of you for your support.

Last but not to leave the least, my sincere gratitude also goes to my family which consists of my father, Mohammad Bin Sabli and my mother, Ramlah Binti Wasly for their support financially and morally for me that keeps me fighting in this research project towards the completion. All their patience will be paid off when seeing me completing this research project and report and finally became a university graduate. I am proud of them and I hope that they will be proud of me and I hope that they will be living happily together ever after. Mom and dad, thank you.

Lastly but not forgotten also, to all those where names weren't stated here but had contributed to me either directly or indirectly whether how small the contribution were, it doesn't mean that their contribution was not important but instead, were vital to me in completing this research project as one saying that I invented goes, "Without 1 cent that one anonymous person gave you, you'll never get to reach RM1000.00 when you already and only have RM999.99 with you."

Thank you…

Hazrin Bin Mohammad

Faculty of Information Technology and Quantitative Sciences

Mara University of Technology

October 2007

ABSTRACT

The first computer was meant to solve equations sequentially and was programmed to do sequential equations. However, as computations are getting more complex and the demand of having more powerful processors doubles, it seems like the move to solve problems and equations through parallelization is the best solution we have in the world of computing today. Therefore, various parallel computing solutions had been introduced in terms of hardware and software. Since then, computers are getting more advanced than they were before thus enabling software developers and programmers to produce with vivid quality software's, real time simulations, and more advanced modeling's. However, the cost of having a parallel supercomputer is way beyond the limits or budget restrains of some small and medium organizations. Therefore, a low cost parallel computing solution had been introduced by developers which involve a few personal computers or PC's networked together to imitate the producing power of a supercomputer. This research project was conducted to compare the time difference between the sample parallel matrix calculation programming and the sequential matrix calculation programming that was developed by the author in C language. The parallel matrix multiplication program will be run using 3 Ubuntu Linux PC's connected together with communications protocol, SSH, been configured using RSA key via 100Mbps Ethernet either through the peer-to-peer connections and star topology connections and it is expected that the parallel matrix multiplication programming will run faster than the sequential matrix multiplication programming.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF GRAPHS

LIST OF ABBREVIATIONS

Abbrev.	Full Text	Page
3D	3 Dimensional	14
AMD	Advanced Micro Devices	13
CD	Compact Disk	33
CGI	Common Gateway Interface (web scripting facility)	67
CLI	Command Line Interface	45
CPU	Central Processing Unit	3
DAEMON	Disk And Execution Monitor	19
DDR	Double Data Rate	29
DDR2	Double Data Rate 2	29
Eth0	Ethernet 0 address or interface	42
Ext2	Second Extended file system (*nix)	37
Ext3	Third Extended file system (journaling)	37
G++	GNU C++ Compiler	30
GB	Gigabyte (1024 megabytes)	31
GCC	GNU C Compiler	30
GHz	Gigahertz (thousands of hertz)	31
GNU	GNU's Not Unix	17
GRUB	GRand Unified Bootloader	36
GTK	GIMP Tool Kit for C++	30
GUI	Graphical User Interface	30
IO	Input Output	7
IP	Internet Protocol	41
KB	Kilobyte (thousands of bytes)	31
L1	Level One (Cache on or near processor die)	31
L2	Level Two (Cache on or near processor die)	31
LAN	Local Area Network	7
LARPBS	Linear Array Reconfigurable Pipelined Bus System	8
LCD	Liquid Crystal Display	29
lo	Loopback address or interface	42
MB	Megabyte (millions of bytes)	31

Mbps	Megabits per second	31
MIMD	Multiple Instruction, Multiple Data Stream	11
MISD	Multiple Instruction, Single Data Stream	11
MMX	Multimedia Extension	13
MPI	Message Parsing Interface	4
MPP	Massively Parallel Processor	8
OS	Operating System	17
P2P	Peer-to-Peer Networks	4
PC	Personal Computer	4
PCI	Peripheral Component Interconnect	30
PCMCIA	Personal Computer Memory Card International Association	29
PERL	Practical Extraction and Report Language	67
PLAPACK	Parallel Linear Algebra Package	22
PRISM	Parallel Research on Invariant Subspace Methods	9
PVM	Parallel Virtual Machine	3
Ra0	Ralink chipset wireless interface or address	42
RAM	Random Access Memory	29
RSA	Rivest, Shamir, and Adelman	21
RSH	Remote Shell	21
SIMD	Single Instruction, Multiple Data Stream	11
SISD	Single Instruction, Single Data Stream	11
SPMD	Single Program Multiple Data	19
SSH	Secure Shell	21
UNIX	UNIpleX information and computer services	4
UTP	Unshielded Twisted Pair	30
WXGA	Wide Extended Graphics Array (1366 by 768 pixels; 1.78:1 aspect ratio)	31
XPVM	Executable Parallel Virtual Machine (GUI)	30

All the acronyms for the abbreviated words were sourced from The Free Dictionary at http://acronyms.thefreedictionary.com/

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The first computer made on earth was meant to solve equations sequentially. However, as we saw rapid growth in computers and computational methods and apparatus, equations are getting more complex and need more time and resources and the need for faster processors is increasing and the cost of having a state of the art computational technology had also risen beyond the thoughts of human brains ever since. In 1965, Gordon Moore had created a law in which he predicts that the numbers of transistors in a single processor will double every 18 months and the law is correct until today (Moore's Law, 1965). Imagine that today's state of the art Intel Core 2 Duo processor already contains nearly up to 300 millions of transistors (Anand Lal, 2006)[1] as computations and processing had become more and more complex as equations need to be as accurate as a fraction of millions and even billions. This is because software's are getting more complicated, simulations are getting near reality, and color displays are getting more vivid than a couple of years ago.

From then we can see that by having a single processor in a computer is not enough for humans to do calculations and computation. There are still many equations and calculations that are so enormous even beyond the capabilities of human brain. Since then we can see that number of processors in a single computer doubles and even triples or quadruples just so that processing speed can be faster and computers will be more reliable.

However, having double or even quadruple processors in a single chip would even be costly for us, not to mention quoting the cost of a supercomputer or mainframe alone.

[1] http://www.anandtech.com/cpuchipsets/showdoc.aspx?i=2795

Therefore, an ingenious solution had been produced by engineers at the Oak Ridge National Laboratory during summer in 1989 which to gather or group a few computers and networked them together to produce computational powers as the same as of a supercomputer at a competitively low cost (Manchek, 1995)[2]. Further explanations of this parallel computing are detailed in chapter two in this report.

By the introduction of parallel computing it is easier for laboratories and small and medium organizations to commence research without having costly apparatus such as a mainframe or a supercomputer to perform complex calculations such the matrix calculation.

1.2 PROBLEM STATEMENT

Sequential computations as we all know of today, consists of instructions built by the compiler and queued to be processed by a single processor in a single computer and those instructions need to be hold by the memory until the previous instructions were done processed by the processor and this may lead to a lead time in which a user have to wait for some moment of time before all instructions were completed by the processor and results granted to the user.

As we all know, calculations are getting more complex and the need for parallel computing is growing fast while sequential processing for today's calculations and programs would lead to timing constraint which means the user would have to wait for some period of time before the user could obtain the results. Imagine if the calculation is so complex or big such as having a 100 x 100 or even 1000 x 1000 matrix multiplications, the calculation would take to a sum of 2 to 3 seconds before obtaining the results and in the world where time is moving fast, the amount of 2 to 3 seconds would be vital for organizations or laboratories. Besides, complex

[2] http://www.netlib.org/pvm3/faq_html/node4.html

calculations that is processed sequentially would cause stress to the processor or the CPU and may result to the period of inactivity for the system or in other words, the systems may hang during the processing period and in worse situations, could lead the whole system to crash and it would be costly for organizations to repair the system to its healthy working condition by means of time, effort, and money.

Mentioning about the cost, We are sure that we all know that the cost of having a supercomputer or a mainframe would be beyond the budget figure of any small and medium organizations and laboratories for them to perform complex calculations and these calculations such as matrix multiplication calculation is vital in our everyday work since we use mathematical equations in our everyday work. That's why a low cost parallel computing solution is needed by researchers and research organizations which is the introduction of PVM (Parallel Virtual Machine) which could cut the costs of these organizations to 4 or 5 figures of their budget. This is said so because most of workstations or PC's could be networked together and by the help with an appropriate architecture or framework, these workstations could imitate the use of a single supercomputer or mainframe.

In terms of performance, sequential processing may lead to low performance in obtaining results. For example, a calculation or equation that would normally took 4 seconds using a single processor single pc would take less time by using two processors, 4 seconds divide by 2 would result in only 2 seconds per calculation. Therefore, performance would be increased and productivity also would be increased. This is because; by having multiple processors in a single CPU, or by having multiple computers paralleled together combined to be a supercomputer would enhance the stability and reliability of the system.

1.3 PROJECT OBJECTIVES

The research project objectives are as follows:

1) To install and configure parallel programming software named Parallel Virtual Machine (PVM) in Ubuntu in Linux environment for three hosts that will be arranged in two types of networks, Peer-to-Peer (P2P) and Star Topology), in order to run the parallel matrix multiplication program.

2) To modify the parallel matrix multiplication coding to develop a sequential matrix multiplication program derived from the parallel matrix multiplication program.

3) To analyze the time difference marked between the parallel matrix multiplication program that was run in both type of networks and the sequential matrix multiplication program and to determine which program type and in which type of network produce the fastest time and better performance (parallel programming) for the matrix multiplication programming.

1.4 SCOPE OF THE PROJECT

The scopes of the research project are as follows:

1) The entire project will only be limited to the Linux / UNIX environment where the hosts that were networked together were using Ubuntu Linux and the parallel PC software were developed and fully supported by the Linux environment.

2) The project will also be limited to the parallel programming software named Parallel Virtual Machine (PVM) although there are other parallel programming software's such as the Message Parsing Interface (MPI).

3) Because the sample parallel matrix multiplication programming code was written in C language, the developed sequential matrix multiplication programming code will also be written in C language as to make sure that the instructions made to the CPU's were the same in order to perform comparisons between these two architectures.

4) This project will also be limited to matrix multiplication programming since both the sequential and parallel matrix multiplication programs can be compared in terms of performance and time consumption.

1.5 PROJECT SIGNIFICANCE

1) This project will contribute to other researchers in this field that do researches about complex calculations and the effectiveness of the project in terms of cost, performance, and reliability.

2) This project would attract more students from this faculty to further the scope of this project or do research from a different angle of view to enhance the application of parallel programming to everyday usage.

3) Besides students from this faculty, this project would also benefit engineering students to utilize the usage of parallel programming concept since engineering students are more familiar with complex equations and calculations.

1.6 CHAPTER SUMMARY

In this chapter, we had explained the introduction of this research which is the characteristics of sequential and parallel programming and the difference between those types of programming. It also explains the problems which are CPU optimization, performance, stability and reliability of the system using both means of computing, and other factors that lead to this research. This chapter also explained the objective of this research and its scope and limitations and also the benefits of this research to the community.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

In this chapter, all the overviews from the supporting technologies or all the definitions will be explained thoroughly in order to ensure that all the definitions and technologies are understandable. Besides the supporting technologies, related previous projects will also be stated in this chapter. Explanations regarding parallel computing and its architectures and frameworks will be written in this chapter besides most, if not all, of the supporting technologies such as the encryption keys and the Local Area Networks (LAN) technologies.

Before going into the definitions of sequential computing, parallel computing, and the terms relating to the project, we would like to give an overview of the applications of the parallel matrix multiplication.

2.2 OVERVIEW OF PARALLEL MATRIX MULTIPLICATION APPLICATION

According to Charles Crowell (1989), matrix multiplication is useful in applications such as graphics, numerical analysis, or high speed control[3]. It is said that in order to multiply large matrices, high speed input output (IO) and high speed processing was needed. By then, the only solution available was from a supercomputer.

[3] http://focus.ti.com/lit/an/spra008/spra008.pdf

Since the virtualization of parallel computing had been developed, there were various projects that had commenced among laboratories, utilizing the performance and heterogeneity of the parallel computing built by these Linux clusters. Shortly after, we will describe projects that utilize parallel computing in the form of matrix multiplication.

2.2.1 MEMORY EFFICIENT PARALLEL MATRIX MULTIPLICATION OPERATION FOR IRREGULAR PROBLEMS

The project report, authored by Manojkumar Krishnan and Jarek Nieplocha in Pacific Northwest National Laboratory in the United States explains the project of designing an algorithm for better performance in irregular parallel matrix multiplication program using Massively Parallel Processor (MPP)[4]. The algorithm designed was meant to faster the message passing procedure so that the performance of the application would be increased, thus, producing faster output for large scale matrix multiplications and reliability for irregular distributions of matrix.

2.2.2 PARALLEL MATRIX MULTIPLICATION ON A LINEAR ARRAY WITH A RECONFIGURABLE PIPELINED BUS SYSTEM

The project was conducted to develop an efficient parallelization of the fastest sequential matrix multiplication program on a linear array with a reconfigurable pipelined optical bus system[5]. The system developed was using the Linear Array Reconfigurable Pipelined Bus System (LARPBS) where the researcher researches both of the Strassen's algorithm and Winograd's algorithm to produce the parallel

[4] Krishnan M, Nieplocha J, **Memory efficient parallel matrix multiplication operation for irregular problems,** Proceddings of the 3rd conference on Computing Frontiers, P229 - P240, 2006
[5] Li K, Pan V Y, **Parallel Matrix Multiplication on a Linear Array With a Reconfigurable Pipelined Bus System,** IEEE Transactions on Computers, Vol. 50 No. 5, P519 – P525, May 2001