

**UNIVERSITI TEKNOLOGI MARA**

**DEVELOPMENT OF A LOW-COST PARALLEL  
COMPUTING SYSTEM FOR DIGITAL  
MAMMOGRAMS VISUALIZATION**

**HANIFAH SULAIMAN**

Thesis submitted in fulfillment of the requirements

for the degree of

**Master of Science**

**Faculty of Computer and Mathematical Sciences**

**March 2010**

## ABSTRACT

Parallel Computing System (PCS) is currently being used widely in many applications of complex problems in image processing. It is now a trend towards the utilization of PCS to solve intermediate and large scale mathematical problems because this system can generate the output in shorter periods as compared with sequential computing systems. Two ARS clusters which are both low-cost parallel computing systems have been developed at Pusat Pengajian Matematik using the Beowulf model. The clusters are composed of used processors and implement the open source software as the operating system which incurs no cost on the system. Both clusters are internally connected using LINUX and the standard Ethernet network.

The first ARS cluster is a heterogeneous PCS which consists of five (5) Vectra Pentium III and one (1) NEC Pentium III with 450MHz processors each of which has an Intel 486 microprocessor. The cluster is programmed using PVM (Parallel Virtual Machine) for the cluster to communicate. This cluster is successfully benched with the standard skyvase.pov image. This system can be used to run programs in C, C++ and FORTRAN language.

The second ARS cluster is a homogeneous PCS which consists of eight (8) IBM Pentium 4 with 2.80 GHz processors each of which has an Intel 486 microprocessor too. This cluster is incorporated with Matlab Distributed Computing Server (MDCS) in which the MPI (Message Passing Interface) is embedded for the cluster to communicate. This system is used to process digital images programmed in Matlab.

The Matlab Distributed Computing Server (MDCS) enables multiple programs to run in parallel on different processors to generate output from different simulation runs. The ARS cluster with the MDCS version is used to process digital mammograms to detect abnormalities via suitable mathematical algorithms. The 40 digital mammograms on Malaysian women breast cancer patients between 20 to 40 years

## ACKNOWLEDGEMENTS

First, I am grateful to Allah s.w.t that I have finally succeeded to complete my research on time. I owe a huge debt to Prof. Madya Dr Arsmah Ibrahim, who has supervised me during the two years of research period. I am grateful for her guidance and support and for all the trust that has been given to me in doing other tasks related to this research. I have learned and have been given the chance to explore new knowledge especially in grid computing and Open Source application.

I am also indebted to Dr Norma Alias who is my second supervisor who has given me a lot of knowledge in Cluster Computing.

Thank you also to my friends at Universiti Teknologi Malaysia (UTM) Skudai Johor, Nor Hafizah Binti Hamzah, Rosnaida Binti Sharil and Noriza Binti Satam for their advice, guidance and support.

Last but not least, I thank my beloved parents, - and Sulaiman Bin Urif for the endless support and inspiration given to me.

# TABLE OF CONTENTS

	<b>Page</b>
<b>ABSTRACT</b>	iii
<b>ACKNOWLEDGEMENTS</b>	v
<b>TABLE OF CONTENTS</b>	vi
<b>LIST OF TABLES</b>	x
<b>LIST OF FIGURES</b>	xi
<b>CHAPTER I: INTRODUCTION</b>	
1.1 Overview	1
1.2 Background	3
1.3 Problem Description	5
1.4 Objectives of the Research	6
1.5 Research Scope	6
1.6 Research Methodology	7
1.7 Significance of the Research	10
1.8 Limitation of the Research	11
1.9 Organization of the Thesis	11
1.10 Summary	11
<b>CHAPTER II: LITERATURE REVIEW</b>	
2.1 Overview	12
2.2 Parallel Computing Evolution	13
2.2.1 Beowulf Cluster Computing	15
2.2.2 Performance of PCS	17
2.2.3 Parallel Image Processing	21

# CHAPTER I

## INTRODUCTION

### 1.1 Overview

Today many mathematical models are used to solve many problems computationally using High Performance Computing (HPC) systems. A HPC is simply a computer system with superb computing power and speed that is used to solve complex mathematical computations efficiently. This system adopts the use of parallel processing in order to run several application programs efficiently and reliably at great speed.

A Parallel Computing System (PCS) is a HPC system in which different parts of a program run simultaneously on more than one computers that communicate with each other over a network (Wilkinson and Allen, 2005). Conceptually, it is a system that uses the technique of deploying concurrent tasks on either multi-processor hardware or on a multi-processing operating system. The idea is based on the assumption that the more processors are used to perform a task, the faster it will complete. In PCS, the process of solving a problem can be divided into smaller tasks, which may be carried out simultaneously with some coordination. It is anticipated that if a computing process can be decomposed into smaller, concurrent tasks then theoretically a performance increase can be achieved.

There are many real problems still unsolved in today's world. Breast cancer is one of the major problems faced by mankind today. In general, breast cancer is the spread of malignant or cancer cells within the breast area. This cancer type is the