Beowulf Cluster:
An Implementation and Testing

by

Abdul Haddis bin Abdul Aziz

BSc Engineering (Mechanical)
University of Warwick

In partial fulfillment of the requirements
for the degree of
Master of Science in Information Technology

Faculty of Information Technology and Quantitative Sciences

UNIVERSITY TEKNOLOGI MARA

October 2003

ABSTRACT

The use of computers within our society has developed from the very first usage in 1945 when the modern computer era began, until about 1985 when computers were large and expensive. Two modern age technologies, the development of high-speed networking and the personal computer have allowed us to break-down these price barriers and construct cost effective clusters of PCs which provide comparable performance to super-computers at a fraction of the cost. As PC's and networks are in common use, this allows most commercial organizations, governments, and educational institutions access to high performance super-computers.

The major difference between a network of PC's and a super-computer is the software which is loaded on each machine, and the way in which an application is processed, namely in parallel. Parallel processing is the method of breaking down problems or work into smaller components to be processed in parallel thus taking only a fraction of the time it would take to run on a stand-alone PC.

The only drawback to this cost-effective way of computing is the lack of all in one package solution that can be implemented at once. The hardware keeps changing, the software and its documentation keeps changing and lack of integrity between its software environments [24] pose a great challenge for any Beowulf implementer.

This paper describes the technologies and methodologies employed to achieve this breakthrough. Both opportunities afforded by this approach and the challenges confronting its application to real-world problems are discussed in the framework of hardware and software systems as well as the results from benchmarking experiments.

Tabic of Contents

CHAPTER 3: METHODOLOGY

CHAPTER 4: TESTING AND RESULTS

Abbreviations

The following words and acronyms have been used throughout this document:

Word List	Definition
COTS	Commercial Of The Shelf
GPL	General Public Licence
HPCC	High Performance Cluster Computing
HPF	High Performance FORTRAN
I/O	Input Output
LAN	Local Area Network
MPI	Message Passing Interface
NCSA	National Computer and Science Alliance
OS	Operating System
PC	Personal Computer
PVM	Parallel Virtual Machine
SMP	Symmetric Multiprocessing Processor
TFCC	Task Force on Cluster Computing