TABLE OF CONTENTS

COVER PAGE i

STATEMENT OF THESIS ORIGINALITY ii

APPROVAL BY THESIS SUPERVISORS iii

EXAMINATION BOARD APPROVAL iv

ABSTRACT v

PREFACE vi

TABLE OF CONTENTS viii

LIST OF FIGURES xiii

LIST OF TABLES xv

CHAPTER I INTRODUCTION 1

1.1 Background 1

1.2 Statement of the Problem 2

1.3 Purpose of the Research 3

1.4 Scope of the Research 4

1.5 Research Methodology 5

1.6 Structure of the Report 6

CHAPTER II THEORY REVIEW 8

2.1 Distributed Computing 8

2.2 Peer-to-Peer 9

2.3 Code Migration 10

2.4 User Datagram Protocol 10

2.5 Multicast 11

2.6 Socket 13

2.7 Load Balancing 14

2.8 Algorithm Analysis 15

2.9 Speedup and System Efficiency 15

2.10 Shared Library 17

2.11 Java™ Platform 17

2.11.1 Java™ Platform, Standard Edition 6 (Java SE) 18

2.11.2 Java Fundamental Classes (`java.lang`) 19

2.11.3 Java Input and Output (`java.io`) 21

2.11.4 Java Networking (`java.net`) 23

2.11.5 Java Utility Classes (`java.util`) 25

2.11.6 Java Concurrent Programming 25

CHAPTER III THE DESIGN OF GARCINIA FRAMEWORK 26

3.1 Garcinia Peer-to-Peer Overlay Network Architecture 26

3.2 Task and Assignment 27

3.3 Garcinia Framework Components Design 27

3.3.1 Garcinia Shared Library 28

3.3.2 Garcinia Agent... 29

3.3.3 Consumer Application 32

3.4 Assignment Scheduling Strategy... 33

CHAPTER IV THE IMPLEMENTATION OF GARCINIA FRAMEWORK 35

4.1 The Network Overview 35

4.2 Overlay Network Datagram 36

4.3 Garcinia Worker-Dispatcher Protocol.. 37

4.4 Garcinia Shared Library 38

4.4.1 `garcinia.Assignment` Class 39

4.4.2 `garcinia.AssignmentResult` Class............................ 40

4.4.3 `garcinia.Sender` Class.. 42

4.4.4 `garcinia.Task` *Abstract* Class 47

4.4.5 `garcinia.TaskMetadata` Class 49

4.4.6 `garcinia.WorkerInfo` Class 49

4.4.7 `garcinia.util.ByteClassLoader` Class 49

4.4.8 `garcinia.util.Helper` Class....................................... 50

UNIVERSITAS PELITA HARAPAN

4.5 Garcinia Agent 50

4.5.1 `garcinia.Main` Class 52

4.5.2 `garcinia.overlay.DatagramProcessor` Interface. 52

4.5.3 `garcinia.overlay.OverlayManager` Class 53

4.5.4 `NodeUpdateTimerTask` Inner Private Class 53

4.5.5 `garcinia.netloader.ConnectionHandler` Class . 54

4.5.6 `garcinia.netloader.Manager` Class 57

4.6 Garcinia Framework Test Application 57

4.6.1 Reverse String Test Application 58

4.6.2 Distributed Sort Test Application 58

4.6.3 LCS “Genetic Matching” Test Application 59

CHAPTER V FUNCTIONAL ANALYSIS AND EVALUATION 62

5.1 Test Environment 62

5.2 Overlay Network Connectivity Test 64

5.3 Overlay Network Data Traffic 69

5.3.1 User Datagram Protocol Traffic 70

5.3.2 Internet Group Management Protocol Traffic 72

5.3.3 Overlay Network Data Traffic Summary 74

5.4 Local Agent Functional Test 76

5.5 Peer-to-Peer Functional Test 79

UNIVERSITAS PELITA HARAPAN

5.6 Speedup and System Efficiency ... 81

5.6.1 Test Parameters ... 82

5.6.2 Execution Time Result ... 87

5.6.3 Speedup ... 90

5.6.4 System Efficiency ... 94

5.7 Multiplatform Functional Test ... 99

5.8 Evaluation of the Garcinia Framework ... 105

5.8.1 The Benefits ... 105

5.8.2 The Detriments ... 107

CHAPTER VI CONCLUSIONS AND RECOMMENDATIONS ... 108

6.1 Conclusions ... 108

6.2 Recommendations ... 109

BIBLIOGRAPHY ... 110

LIST OF FIGURES

Figure 2.4.1 The format of fields in a UDP datagram. 11

Figure 2.5.1 A simple internet with three networks connected by a router. 13

Figure 2.11.1 Java™ Platform, Standard Edition 6 JDK component diagram. 19

Figure 2.11.2 The Java collections framework. 25

Figure 3.3.1 Garcinia Framework Components. 28

Figure 3.3.2 The worker service operation diagram. 31

Figure 3.3.3 The interaction of Consumer Application and Shared Library. 33

Figure 4.4.1 The Garcinia Shared Library Class Structure 39

Figure 4.4.2 Flow chart of the `getWorkerInfoList()` method. 43

Figure 4.4.3 Flow chart of the sender's #1 `doAssignment()`. 45

Figure 4.4.4 Flow chart of the sender's #3 doAssignment(). 47

Figure 4.5.1 The Garcinia Agent Class Structure 51

Figure 4.5.2 Flowchart of the connection handler `run()` method execution process. 55

Figure 5.1.1 The physical network topology. 63

Figure 5.2.1 The agent activity display in a console session. 65

Figure 5.2.2 The "list" command to show the node list in a console session. 67

Figure 5.2.3 The detailed logging output showing the neighbouring nodes activities. 68

Figure 5.2.4 The detailed logging output showing HELLO-HI messages............ 69

Figure 5.3.1 The UDP packet capture. 70

Figure 5.3.2 A packet capture showing UDP datagram containing a HI message. 72

Figure 5.3.3 The IGMP packet capture. 73

Figure 5.3.4 Overlay packets capture statistics summary. 75

Figure 5.4.1 The output of the agent runs on a single machine. 76

Figure 5.4.2 The output of reverse string test application.................................... 77

Figure 5.4.3 The output of the agent on the local agent functional test. 78

Figure 5.6.1 Empirical number of operations per second. 83

Figure 5.6.2 The number of assignments and operations relative to the number of strands. 86

Figure 5.6.3 Distributed LCS "Genetic Matching" execution times. 88

Figure 5.6.4 The execution time relative to the number of processing elements.. 89

Figure 5.6.5 Speedup relative to the number of processing elements.................. 91

Figure 5.6.6 Speedup relative to the number of strands. 93

Figure 5.6.7 System efficiency relative to the number of processing elements.... 96

Figure 5.6.8 System efficiency relative to the number of strands. 98

Figure 5.7.1 The physical network topology of the multiplatform test. 101

Figure 5.7.2 The node list output on the multiplatform test................................ 102

Figure 5.7.3 The distributed sort consumer test application output. 103

Figure 5.7.4 The log output of the Garcinia Agent task execution. 105

LIST OF TABLES

Table 4.2.1 Overlay network datagram message format...................................... 37

Table 4.3.1 Garcinia Worker-Dispatcher Protocol 38

Table 5.1.1 Node name, IP address, packet loss, and TCP latency. 64

Table 5.3.1 Summary of UDP traffic on port 9011.. 71

Table 5.3.2 Summary of IGMP traffic of multicast address `239.7.12.90`...... 73

Table 5.5.1 Peer-to-peer functional test results.. 79

Table 5.5.2 Worker service load using random assignment distribution............. 81

Table 5.6.1 The node state in each test configuration.. 84

Table 5.6.2 Test cases for the speedup and system efficiency analysis............... 85

Table 5.6.3 Distributed LCS "Genetic Matching" execution times..................... 87

Table 5.6.4 Speedup in different treatments. 90

Table 5.6.5 Correlation between the numbers of processing elements with the speedup grouped according to the number of strands.. 92

Table 5.6.6 Correlation between the numbers of strands with the speedup grouped according to the number of processing elements... 94

Table 5.6.7 System efficiency in different treatments. 95

Table 5.7.1 Node name, IP address, and specifications of the multiplatform test. 101