Parallel Computing in Local Area Networks

PhD Thesis
Computer Science Department
Computer Architecture and Operating Systems Group

Fernando G. Tinetti

Barcelona, Spain, Dec. 2003

Acknowledgements

To my advisor, Emilio Luque, by his support during the development of the present work.

To the involved laboratories and their members, to whom -in a way or another- I have affected with this work:

- Computer Architecture and Operating System Unit, Faculty of Sciences, Autonomous University of Barcelona, Barcelona, Spain.
- Analog-Digital Techniques Center, Faculty of Engineering, National University of La Plata, La Plata, Argentina.
- Laboratory of Research and Development on Computer Sciences, Faculty of Computer Sciences, National University of La Plata, La Plata, Argentina.
- Laboratory of Theoretical Chemistry, Faculty of Exact Sciences, National University of La Plata, La Plata, Argentina.

To professors Alicia Jubert and Reinaldo Pis Diez, for their disinterested contribution to this work.

To professors Antonio Quijano and Armando De Giusti, for their myriad contributions to several areas of this work.

Table of Contents

Chapter 1: Introduction.....1

1.1 Parallel Applications and Architectures.....2

1.2 Parallel Computing Costs in Installed Local Area Networks.....9

1.3 Summary of the Thesis Objectives and Contributions and Organization 12

Chapter 2: Matrix Multiplication.....16

2.1 Matrix Multiplication Definition.....17

2.2 Linear Algebra Operations.....18

2.2.1 BLAS: Basic Linear Algebra Subprograms and Performance.....18
2.2.2 L3 BLAS and Matrix Multiplication.....20

2.3 Matrix Multiplication as Benchmark.....22

2.3.1 Level 3 BLAS “Benchmark”.....23
2.3.2 As a “General” Benchmark.....23

2.4 Matrix Multiplication Parallelization.....24

2.4.1 Parallel Algorithms *for* Multiprocessors.....25
2.4.2 Parallel Algorithms *for* Multicomputers.....30

Chapter 3: Heterogeneous Clusters.....39

3.1 Clusters Characteristics.....40

3.1.1 Characteristics of the Processors Interconnection Network.....41
3.1.2 Homogeneous Cluster as a Parallel Machine.....44
3.1.3 Heterogeneous Cluster as a Parallel Machine.....47

3.2 Parallel Computing in Heterogeneous Clusters.....48

3.2.1 Loosely Coupled Multicomputer.....49
3.2.2 Ethernet Network for Processors Interconnection.....49
3.2.3 Interconnection Network Low Performance.....51
3.2.4 Processing Heterogeneity.....54

3.3 Parallelization Guidelines for Clusters.....56

3.4 Parallel Matrix Multiplication.....57

3.4.1 Data Distribution.....57
3.4.2 Computation.....61
3.4.3 Overlapping Computation and Communication.....64
3.4.4 Reduction of Memory Requirement for Messages.....66
3.4.5 General Characteristics.....67

3.4.6 Other Ways of Distributing Data....68
3.5 Chapter Summary....70

Chapter 4: Experimentation....72

4.1 Characteristics of the Local Area Networks....73
4.1.1 CeTAD Local Area Network....73
4.1.2 LQT Local Area Network....75
4.1.3 LIDI Local Area Network....76
4.2 Computers' Sequential Performance....77
4.2.1 Matrix Sizes....77
4.2.2 CeTAD Local Area Network....79
4.2.3 LQT Local Area Network....80
4.2.4 LIDI Local Area Network....82
4.3 Parallel Performance Analysis of Local Area Networks....82
4.3.1 *Actual* Speedup Calculation....83
4.3.2 *Optimal* Speedup Calculation....84
4.4 Analysis of Algorithms Performance....90
4.4.1 SeqMsg: Sequential Computing and Communication....91
4.4.2 OverMsg: Overlapped Computing and Communications....92
4.5 Local Area Networks and Algorithms....92
4.5.1 CeTAD Local Area Network....93
4.5.2 LQT Local Area Network....95
4.5.3 LIDI Local Area Network....97
4.6 Actual Performance of Local Networks Using PVM....99
4.6.1 CeTAD Local Area Network....100
4.6.2 LQT Local Area Network....102
4.6.3 LIDI Local Area Network....103
4.7 Execution Profiles in Local Aera Networks with PVM....105
4.7.1 CeTAD Local Area Network....106
4.7.2 LQT Local Area Network....112
4.7.3 LIDI Local Area Network....116
4.8 Real Performance of Local Networks using “UDP”....121
4.8.1 CeTAD Local Area Network....124
4.8.1.1 Matrices of 2000×2000 Elements....124
4.8.1.2 Matrices de 3200×3200 Elements....132
4.8.1.3 General Conclusions of the Experimentation in the CeTAD....135
4.8.2 LQT Local Area Network....135
4.8.3 LIDI Local Area Network....138

4.9 Conclusions - Experimentation Summary.....140

Chapter 5: Comparison with ScaLAPACK.....141

5.1 5.1 ScaLAPACK General Characteristics.....142

5.2 LU Factorization Parallelization.....143

5.2.1 Block LU Factorization Sequential Algorithm.....143
5.2.2 LU Factorization Parallel Algorithm *for* Multicomputers.....145
5.2.3 LU Factorization Parallel Algorithm for Clusters 146

5.3 Experimentation.....149

5.3.1 Set of Experiments 150
5.3.2 Results: ScaLAPACK-PVM.....153
5.3.3 Results: ScaLAPACK-MPICH.....156
5.3.4 ScaLAPACK-MPICH and Scalability.....159

5.4 Summary of the Comparison with ScaLAPACK.....161

Chapter 6: Conclusions and Further Work.....163

6.1 Conclusions.....164

6.2 Further Work.....171

References.....178

Appendix A: Characteristics of Local Area Networks.....191

A.1 Introduction: Hardware and Software Characteristics.....192

A.2 Local Area Networks.....193

A.2.1 CeTAD Local Area Network.....194
A.2.2 LQT Local Area Network.....195
A.2.3 LIDI Local Area Network.....195

A.3 Computers' Details.....196

A.3.1 CeTAD Local Area Network's Computers.....196
A.3.2 LQT Local Area Network's Computers.....198
A.3.3 LIDI Local Area Network's Computers.....200

Appendix B: Sequential Processing Performance of Computers.....201

B.1 Introduction.....202

B.2 Computers used.....203

B.3 Description of Experimentats....204
B.3.1 Code without Optimization....204
B.3.2 Compiler Optimization....205
B.3.3 Source Code Optimization....206
B.3.4 Sizes of Matrices to be Multiplied....207
B.4 Matrix Multiplication Performance....209
B.4.1 Performance without Optimization....209
B.4.2 Performance with Compiler Optimization....211
B.4.3 Performance with Source Code Optimization....214
B.4.4 Bigger Matrices in Computers with the Highest Computing Performance....217
B.5 LIDI Computers Performance....219
B.6 Conclusions....220
References....221

Appendix C: Communications in the CeTAD Local Area Network.....223

C.1 Introduction....224
C.2 Ethernet Networks....225
C.3 Performance Evaluation....226
C.4 Performance Evaluation with the ping-pong Method....228
C.5 Different Ways of Message Transmission with PVM....229
C.5.1 Message Data Coding in PVM....230
C.5.2 Direct Translation of Data Representation....232
C.5.3 Data Routing of a Message in PVM....234
C.6 Different Ways of Message Transmission with MPI....235
C.7 Initial Experimentation with PVM....235
C.7.1 Performance with Routing Between pvmds and Codification....236
C.7.2 Performance with Routing Between pvmds and Representation Translation. 238
C.7.3 Performance with Routing Between Tasks and Codification....239
C.7.4 Performance with Routing Between Tasks and Representation Translation. .241
C.7.5 PVM Experimentation Conclusions....242
C.8 Experimentation with the Linux ping Command....243
C.9 Conclusions of PVM and Linux ping Experimentation....245
C.10 UDP-based Broadcasts....246
C.10.1 A Single Receiver (Point-to-Point Messages)....249
C.10.2 Broadcast Messages....251
C.11 Broadcasts in the LQT Local Area Network....253
C.12 Broadcasts in the LIDI Local Area Network....253

References.....254