Patterns for Parallel Programming

Timothy G. Mattson

Beverly A. Sanders

Berna L. Massingill

Addison-Wesley

Boston • San Francisco • New York • Toronto • Montreal
London • Munich • Paris • Madrid
Capetown • Sydney • Tokyo • Singapore • Mexico City

Contents

Preface x

1 A Pattern Language for Parallel Programming 1
- 1.1 Introduction . . . 1
- 1.2 Parallel Programming . . . 3
- 1.3 Design Patterns and Pattern Languages . . . 4
- 1.4 A Pattern Language for Parallel Programming . . . 5

2 Background and Jargon of Parallel Computing 7
- 2.1 Concurrency in Parallel Programs Versus Operating Systems . . . 7
- 2.2 Parallel Architectures: A Brief Introduction . . . 8
  - 2.2.1 Flynn's Taxonomy . . . 8
  - 2.2.2 A Further Breakdown of MIMD . . . 9
  - 2.2.3 Summary . . . 12
- 2.3 Parallel Programming Environments . . . 12
- 2.4 The Jargon of Parallel Computing . . . 16
- 2.5 A Quantitative Look at Parallel Computation . . . 18
- 2.6 Communication . . . 21
  - 2.6.1 Latency and Bandwidth . . . 21
  - 2.6.2 Overlapping Communication and Computation and Latency Hiding . . . 22
- 2.7 Summary . . . 23

3 The *Finding Concurrency* Design Space 24
- 3.1 About the Design Space . . . 24
  - 3.1.1 Overview . . . 25
  - 3.1.2 Using the Decomposition Patterns . . . 26
  - 3.1.3 Background for Examples . . . 26
- 3.2 The *Task Decomposition* Pattern . . . 29
- 3.3 The *Data Decomposition* Pattern . . . 34
- 3.4 The *Group Tasks* Pattern . . . 39
- 3.5 The *Order Tasks* Pattern . . . 42
- 3.6 The *Data Sharing* Pattern . . . 44
- 3.7 The *Design Evaluation* Pattern . . . 49
- 3.8 Summary . . . 55

4 The *Algorithm Structure* Design Space 57
- 4.1 Introduction . . . 57
- 4.2 Choosing an *Algorithm Structure* Pattern . . . 59
  - 4.2.1 Target Platform . . . 59

4.2.2 Major Organizing Principle 60
4.2.3 The *Algorithm Structure* Decision Tree 60
4.2.4 Re-evaluation 62
4.3 Examples 62
4.3.1 Medical Imaging 62
4.3.2 Molecular Dynamics 63
4.4 The *Task Parallelism* Pattern 64
4.5 The *Divide and Conquer* Pattern 73
4.6 The *Geometric Decomposition* Pattern 79
4.7 The *Recursive Data* Pattern 97
4.8 The *Pipeline* Pattern 103
4.9 The *Event-Based Coordination* Pattern 114

5 The *Supporting Structures* Design Space 121
5.1 Introduction 121
5.1.1 Program Structuring Patterns 122
5.1.2 Patterns Representing Data Structures 123
5.2 Forces 123
5.3 Choosing the Patterns 125
5.4 The *SPMD* Pattern 126
5.5 The *Master/Worker* Pattern 143
5.6 The *Loop Parallelism* Pattern 152
5.7 The *Fork/Join* Pattern 167
5.8 The *Shared Data* Pattern 173
5.9 The *Shared Queue* Pattern 183
5.10 The *Distributed Array* Pattern 198
5.11 Other Supporting Structures 211
5.11.1 SIMD 211
5.11.2 MPMD 212
5.11.3 Client-Server Computing 214
5.11.4 Concurrent Programming with Declarative Languages 214
5.11.5 Problem-Solving Environments 215

6 The *Implementation Mechanisms* Design Space 216
6.1 Overview 217
6.2 UE Management 217
6.2.1 Thread Creation/Destruction 218
6.2.2 Process Creation/Destruction 220
6.3 Synchronization 221
6.3.1 Memory Synchronization and Fences 221
6.3.2 Barriers 226
6.3.3 Mutual Exclusion 229
6.4 Communication 237
6.4.1 Message Passing 238
6.4.2 Collective Communication 245
6.4.3 Other Communication Constructs 251

Appendix A A Brief Introduction to OpenMP **253**
A.1 Core Concepts 254
A.2 Structured Blocks and Directive Formats 257
A.3 Worksharing 259
A.4 Data Environment Clauses 262
A.5 The OpenMP Runtime Library 265
A.6 Synchronization 266
A.7 The Schedule Clause 270
A.8 The Rest of the Language 272

Appendix B A Brief Introduction to MPI **273**
B.1 Concepts 273
B.2 Getting Started 275
B.3 Basic Point-to-Point Message Passing 277
B.4 Collective Operations 279
B.5 Advanced Point-to-Point Message Passing 283
B.6 MPI and Fortran 288
B.7 Conclusion 290

Appendix C A Brief Introduction to Concurrent Programming in Java **291**
C.1 Creating Threads 293
C.2 Atomicity, Memory Synchronization, and the `volatile` Keyword . . 297
C.3 Synchronized Blocks 297
C.4 Wait and Notify 299
C.5 Locks 301
C.6 Other Synchronization Mechanisms and Shared Data Structures . . 303
C.7 Interrupts 304

Glossary **307**

Bibliography **317**

About the Authors **333**

Index **335**