<table>
<thead>
<tr>
<th>List of Figures</th>
<th>xvii</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>xxv</td>
</tr>
</tbody>
</table>

**PART I: Background**

1. **INTRODUCTION**
   - 1.1 Why Parallel Processing? 3
   - 1.2 Parallel Architectures
     - 1.2.1 SIMD Systems 5
     - 1.2.2 MIMD Systems 6
   - 1.3 Job Scheduling 8
   - 1.4 Software Architectures 10
   - 1.5 Overview of the Monograph 11

2. **PARALLEL AND CLUSTER SYSTEMS**
   - 2.1 Introduction 13
   - 2.2 Parallel Architectures
     - 2.2.1 UMA Systems 15
     - 2.2.2 NUMA Systems 16
     - 2.2.3 Distributed-Memory Systems 16
     - 2.2.4 Distributed Shared Memory 18
   - 2.3 Example Parallel Systems
     - 2.3.1 IBM SP2 System 19
     - 2.3.2 Stanford DASH System 21
     - 2.3.3 ASCI Systems 22
   - 2.4 Interconnection Networks
     - 2.4.1 Dynamic Interconnection Networks 26
     - 2.4.2 Static Interconnection Networks 29
2.5 Interprocess Communication 36
  2.5.1 PVM 36
  2.5.2 MPI 40
  2.5.3 TreadMarks 43
2.6 Cluster Systems 45
  2.6.1 Beowulf 46
2.7 Summary 48

3. PARALLEL JOB SCHEDULING 49
  3.1 Introduction 49
  3.2 Parallel Program Structures 51
    3.2.1 Fork-and-Join Programs 51
    3.2.2 Divide-and-Conquer Programs 52
    3.2.3 Matrix Factorization Programs 53
  3.3 Task Queue Organizations 55
    3.3.1 Basic Task Queue Organizations 55
    3.3.1.1 Improving Centralized Organization 57
    3.3.1.2 Improving Distributed Organization 59
  3.4 Scheduling Policies 63
    3.4.1 Space-Sharing Policies 63
      3.4.1.1 Static Policies 63
      3.4.1.2 Dynamic Policies 64
    3.4.2 An Example Space-Sharing Policy 65
      3.4.2.1 Adaptive Space-Sharing Policy 66
      3.4.2.2 A Modification 67
      3.4.2.3 An Improvement 67
    3.4.3 Time-Sharing Policies 78
      3.4.4 Hybrid Policies 80
  3.5 Example Policies 81
    3.5.1 IBM SP2 81
    3.5.2 ASCI Blue-Pacific 82
    3.5.3 Portable Batch System 83
  3.6 Summary 84

PART II: Hierarchical Task Queue Organization 85
Contents

4. HIERARCHICAL TASK QUEUE ORGANIZATION 87
   4.1 Motivation 87
   4.2 Hierarchical Organization 89
   4.3 Workload and System Models 93
   4.4 Performance Analysis 96
      4.4.1 Queue Access Overhead 96
      4.4.2 Utilization Analysis 97
      4.4.2.1 Centralized Organization 98
      4.4.2.2 Distributed Organization 98
      4.4.2.3 Hierarchical Organization 99
      4.4.3 Contention Analysis 99
      4.4.3.1 Centralized Organization 99
      4.4.3.2 Distributed Organization 100
      4.4.3.3 Hierarchical Organization 100
   4.5 Performance Comparison 101
      4.5.1 Impact of Access Contention 102
      4.5.2 Effect of Number of Tasks 104
      4.5.3 Sensitivity to Service Time Variance 107
      4.5.4 Impact of System Size 109
      4.5.5 Influence of Branching and Transfer Factors 111
   4.6 Performance of Dynamic Task Removal Policies 114
   4.7 Summary 117

5. PERFORMANCE OF SCHEDULING POLICIES 121
   5.1 Introduction 121
   5.2 Performance of Job Scheduling Policies 122
      5.2.1 Policies 122
      5.2.2 Results 123
      5.2.2.1 Performance Sensitivity to System Load 123
      5.2.2.2 Sensitivity to Task Service Time Variance 124
      5.2.2.3 Sensitivity to Variance in Task Distribution 125
   5.3 Performance of Task Scheduling Policies 126
      5.3.1 Task Scheduling Policies 126
      5.3.2 Results and Discussion 131
      5.3.2.1 Principal Comparison 132
      5.3.2.2 Impact of Variance in Task Service Time 133
      5.3.2.3 Impact of Variance in Task Distribution 134
      5.3.2.4 Effect of Window Size 135
5.3.2.5 Sensitivity to Other Parameters 137
5.4 Conclusions 138

6. PERFORMANCE WITH SYNCHRONIZATION WORKLOADS 141
6.1 Introduction 141
6.2 Related Work 142
6.3 System and Workload Models 145
6.4 Spinning and Blocking Policies 147
   6.4.1 Spinning Policy 147
   6.4.2 Blocking Policies 148
6.5 Lock Accessing Workload Results 148
   6.5.1 Workload Model 149
   6.5.2 Simulation Results 149
   6.5.2.1 Principal Comparison 150
   6.5.2.2 Sensitivity to Service Time Variance 153
   6.5.2.3 Impact of Granularity 154
   6.5.2.4 Impact of Queue Access Time 155
6.6 Barrier Synchronization Workload Results 156
   6.6.1 Workload Model 156
   6.6.2 Simulation Results 157
   6.6.2.1 Impact of System Load 157
   6.6.2.2 Sensitivity to Service Time Variance 160
   6.6.2.3 Impact of Granularity 160
   6.6.2.4 Impact of Queue Access Time 161
6.7 Cache Effects 162
6.8 Summary 163

PART III: Hierarchical Scheduling Policies 165

7. SCHEDULING IN SHARED-MEMORY MULTIPROCESSORS 167
7.1 Introduction 167
7.2 Space-Sharing and Time-Sharing Policies 168
   7.2.1 Equipartitioning 168
   7.2.2 Modified RRJob 170
7.3 Hierarchical Scheduling Policy 170
7.4 Performance Evaluation 174
   7.4.1 System and Workload Models 174
### Contents

- **7.4.1.1 System Model** 174
- **7.4.1.2 Workload Model** 174
- **7.4.2 Performance Analysis** 176
  - **7.4.2.1 Effect of Scheduling Overhead** 178
  - **7.4.2.2 Impact of Variance in Service Demand** 181
  - **7.4.2.3 Effect of Task Granularity** 183
  - **7.4.2.4 Effect of the ERF Factor** 184
  - **7.4.2.5 Effect of Quantum Size** 186
  - **7.4.2.6 Sensitivity to Other Parameters** 186
- **7.5 Performance with Lock Accessing Workload** 187
  - **7.5.1 Lock Accessing Workload** 187
  - **7.5.2 Results** 188
- **7.6 Conclusions** 190

- **8. SCHEDULING IN DISTRIBUTED-MEMORY MULTICOMPUTERS** 193
  - **8.1 Introduction** 193
  - **8.2 Hierarchical Scheduling Policy** 195
  - **8.3 Scheduling Policies for Performance Comparison** 200
    - **8.3.1 Space Partitioning** 200
    - **8.3.2 Time-Sharing Policy** 200
  - **8.4 Workload Model** 201
  - **8.5 Performance Comparison** 203
    - **8.5.1 Performance with Ideal Workload** 203
    - **8.5.2 Performance with Non-Uniform Workload** 204
    - **8.5.2.1 Performance with 50–50 distribution** 205
    - **8.5.2.2 Sensitivity to variance in job service demand** 206
    - **8.5.2.3 Performance under 50–25 distribution** 208
    - **8.5.2.4 Performance under 50–75 distribution** 209
    - **8.5.3 Discussion** 210
  - **8.6 Conclusions** 211

- **9. SCHEDULING IN CLUSTER SYSTEMS** 213
  - **9.1 Introduction** 213
  - **9.2 Hierarchical Scheduling Policy** 215
    - **9.2.1 Job Placement Policy** 216
    - **9.2.2 Dynamic Load Balancing Algorithm** 218
  - **9.3 Space-Sharing and Time-Sharing Policies** 220
    - **9.3.1 Space-Sharing Policy** 221