Contents

Part I  High Performance Computing (HPC) with the Message Passing Interface (MPI)

1  A Glance at High Performance Computing (HPC) .................. 3
   1.1 What is High Performance Computing (HPC)? ................. 3
   1.2 Why Do We Need HPC? ...................................... 4
   1.3 Big Data: The Four Vs (Volume, Variety, Velocity, Value) ........................................ 5
   1.4 Parallel Programming Paradigms: MPI and MapReduce ...... 6
   1.5 Granularity: Fined-Grained Versus Coarse-Grained Parallelism ........................................ 6
   1.6 Architectures for Supercomputing: Memory and Network .... 7
   1.7 Speedup ......................................................... 11
      1.7.1 Scalability and Iso-efficiency Analysis .................. 11
      1.7.2 Amdahl’s Law: Characterizing the Asymptotic Speedup for Fixed Data-Size Problems .......... 12
      1.7.3 Gustafson’s Law: Scaled Speedup, Increasing Data Size with Resources .......................... 14
      1.7.4 Simulating Parallel Machines on Sequential Computers ........................................... 15
      1.7.5 Big Data and Parallel Inputs/Outputs (I/O) ............. 16
   1.8 Eight Common Fallacies on Distributed Systems ............... 17
   1.9 Notes and References .......................................... 19
   1.10 Summary ....................................................... 19
   1.11 Exercises ....................................................... 20

2  Introduction to MPI: The Message Passing Interface ............ 21
   2.1 MPI for Parallel Programming: Communicating with Messages ....................................... 21
   2.2 Parallel Programming Models, Threads and Processes ...... 22
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3</td>
<td>Global Communications Between Processes</td>
<td>23</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Four Basic MPI Primitives: Broadcast, Gather, Reduce, and Total Exchange</td>
<td>24</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Blocking Versus Non-blocking and Synchronous Versus Asynchronous Communications</td>
<td>26</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Deadlocks from Blocking Communications</td>
<td>28</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Concurrency: Local Computations Can Overlap with Communications</td>
<td>32</td>
</tr>
<tr>
<td>2.3.5</td>
<td>Unidirectional Versus Bidirectional Communications</td>
<td>33</td>
</tr>
<tr>
<td>2.3.6</td>
<td>Global Computations in MPI: Reduce and Parallel Prefix (Scan)</td>
<td>33</td>
</tr>
<tr>
<td>2.3.7</td>
<td>Defining Communication Groups with Communicators</td>
<td>35</td>
</tr>
<tr>
<td>2.4</td>
<td>Synchronization Barriers: Meeting Points of Processes</td>
<td>37</td>
</tr>
<tr>
<td>2.4.1</td>
<td>A Synchronization Example in MPI: Measuring the Execution Time</td>
<td>37</td>
</tr>
<tr>
<td>2.4.2</td>
<td>The Bulk Synchronous Parallel (BSP) Model</td>
<td>38</td>
</tr>
<tr>
<td>2.5</td>
<td>Getting Started with the MPI: Using OpenMPI</td>
<td>39</td>
</tr>
<tr>
<td>2.5.1</td>
<td>The “Hello World” Program with MPI C++</td>
<td>39</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Programming MPI with the C Binding</td>
<td>41</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Using MPI with C++ Boost</td>
<td>41</td>
</tr>
<tr>
<td>2.6</td>
<td>Using MPI with OpenMP</td>
<td>42</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Programming MPI with the Python Binding</td>
<td>43</td>
</tr>
<tr>
<td>2.7</td>
<td>Main Primitives in MPI</td>
<td>44</td>
</tr>
<tr>
<td>2.7.1</td>
<td>MPI Syntax for Broadcast, Scatter, Gather, Reduce and Allreduce</td>
<td>44</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Other Miscellaneous MPI Primitives</td>
<td>45</td>
</tr>
<tr>
<td>2.8</td>
<td>Communications on the Ring Topology with MPI</td>
<td>47</td>
</tr>
<tr>
<td>2.9</td>
<td>Examples of MPI Programs with Their Speed-Up Analysis</td>
<td>48</td>
</tr>
<tr>
<td>2.9.1</td>
<td>The Matrix–Vector Product in MPI</td>
<td>49</td>
</tr>
<tr>
<td>2.9.2</td>
<td>Example of MPI Reduce Operations: Computing the Factorial and Minimum Value of an Array</td>
<td>50</td>
</tr>
<tr>
<td>2.9.3</td>
<td>Approximating $\pi$ with Monte-Carlo Stochastic Integration</td>
<td>52</td>
</tr>
<tr>
<td>2.9.4</td>
<td>Monte-Carlo Stochastic Integration for Approximating the Volume of a Molecule</td>
<td>54</td>
</tr>
<tr>
<td>2.10</td>
<td>References and Notes</td>
<td>59</td>
</tr>
<tr>
<td>2.11</td>
<td>Summary</td>
<td>60</td>
</tr>
<tr>
<td>2.12</td>
<td>Exercises.</td>
<td>60</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>62</td>
</tr>
</tbody>
</table>
3 Topology of Interconnection Networks ........................................ 63

3.1 Two Important Concepts: Static Versus Dynamic Networks, and Logic Versus Physical Networks ........................................ 63

3.2 Interconnection Network: Graph Modeling ................................ 64

3.3 Some Attributes Characterizing Topologies .............................. 65

3.3.1 Degree and Diameter .............................................. 65

3.3.2 Connectivity and Bisection ........................................ 65

3.3.3 Criteria for a Good Network Topology ............................ 66

3.4 Common Topologies: Simple Static Networks ......................... 66

3.4.1 Complete Graphs: The Cliques .................................... 66

3.4.2 Star-Shaped Graph .................................................. 67

3.4.3 Rings and Chordal Rings .......................................... 67

3.4.4 Meshes (Grids) and Torii (Plural of Torus) .......................... 68

3.4.5 The 3D Cube and the Cycle Connected Cube (CCC) .................. 69

3.4.6 Trees and Fat Trees ................................................ 69

3.5 Topology of Hypercubes and Node Labeling Using the Gray Code ................................................................. 71

3.5.1 Recursive Construction of the Hypercube ........................... 71

3.5.2 Numbering Hypercube Nodes Using the Gray Code .................. 72

3.5.3 Generating Gray Codes in C++ ..................................... 73

3.5.4 Converting Gray Codes to and from Binary Codes ................. 75

3.5.5 Cartesian Product of Graphs ....................................... 76

3.6 Some Communication Algorithms on Topologies ..................... 78

3.6.1 Communication Primitives on the Oriented Ring .................... 79

3.6.2 Broadcasting on the Hypercube: Tree-Like Communications ......... 86

3.7 Embedding (Logical) Topologies in Other (Physical) Topologies ...................... 91

3.8 Complex Regular Topologies ............................................ 92

3.9 Interconnection Network on a Chip ..................................... 93

3.10 Notes and References ................................................. 96

3.11 Summary .......................................................... 96

References ............................................................. 97

4 Parallel Sorting .................................................................... 99

4.1 Quick Review of Sequential Sorting ..................................... 99

4.1.1 Main Sequential Sorting Algorithms .............................. 99

4.1.2 Complexity of Sorting: A Lower Bound ........................... 101

4.2 Parallel Sorting by Merging Lists ...................................... 102

4.3 Parallel Sorting Using Ranks ............................................ 102

4.4 Parallel Quicksort ........................................................ 104

4.5 HyperQuickSort ....................................................... 108

4.6 Parallel Sort Regular Sampling (PSRS) ................................. 109
4.7 Sorting on Grids: **ShearSort** ........................................ 111
4.8 Sorting Using Comparison Network: Odd–Even Sorting .......... 113
4.9 Merging Sorted Lists Using a Comparison Network ............ 115
4.10 The Bitonic Merge Sort ........................................... 116
4.11 Notes and References ............................................ 118
4.12 Summary ...................................................... 120
4.13 Exercises .................................................... 120

References .................................................. 120

5 **Parallel Linear Algebra** ........................................ 121

5.1 Distributed Linear Algebra ....................................... 121
   5.1.1 Linear Algebra for Data Science ............................ 121
   5.1.2 Classic Linear Algebra .................................. 125
   5.1.3 The Matrix–Vector Product: $y = Ax$ ................. 126
   5.1.4 Data Patterns for Parallelism ............................ 127

5.2 Matrix–Vector Product on the Topology of the Oriented Ring .... 128

5.3 Matrix Product on the Grid: The Outer Product Algorithm ....... 134

5.4 Matrix Products on the Topology of the 2D Torus .............. 134
   5.4.1 Cannon’s Algorithm .................................... 136
   5.4.2 Fox’s Algorithm: The Broadcast-Multiply-Roll Matrix Product .... 138
   5.4.3 Snyder’s Algorithm: Accumulating Local Products on the Diagonal .... 141
   5.4.4 Comparisons of Cannon’s, Fox’s and Snyder’s Algorithms .......... 143

5.5 Notes and References ............................................ 143
5.6 Summary ...................................................... 144
5.7 Exercises .................................................... 144

References .................................................. 144

6 **The MapReduce Paradigm** ..................................... 147

6.1 The Challenge of Processing Big Data Fast! ..................... 147
6.2 The Basic Principle of MapReduce ................................ 148
   6.2.1 Map and Reduce Processes .............................. 148
   6.2.2 A Historical Perspective: Map and Reduce in Functional Programming Languages .... 149

6.3 Data Types and the MapReduce Mechanism ..................... 150
6.4 A Complete Example of MapReduce in C++ ...................... 152
6.5 Launching MapReduce Jobs and Overview of the MapReduce Architecture .......... 154
6.6 Using MapReduce in MPI with the MR-MPI Library ............ 156
6.7 Notes and References ............................................ 158
6.8 Summary ...................................................... 159

References .................................................. 160
8.2 Strategies to Define a Good Linkage Distance ............. 198
  8.2.1 A Generic Algorithm for Agglomerative
          Hierarchical Clustering .............................. 199
  8.2.2 Choosing the Appropriate Elementary Distance
          Between Elements ...................................... 200
8.3 Ward Merging Criterion and Centroids ...................... 203
8.4 Retrieving Flat Partitions from Dendrograms ............... 204
8.5 Ultra-metric Distances and Phylogenetic Trees .............. 204
8.6 Notes and References ................................... 206
8.7 Summary ................................................ 207
8.8 Exercises ............................................... 208
References .................................................. 208
9 Supervised Learning: Practice and Theory of Classification
         with the $k$-NN Rule .................................. 213
  9.1 Supervised Learning ...................................... 213
  9.2 Nearest Neighbor Classification: NN-Rule ................... 213
     9.2.1 Optimizing Euclidean Distance Computation
              for Nearest Neighbor Queries ...................... 214
     9.2.2 Nearest Neighbor (NN) Rules and Voronoi
              Diagrams ........................................... 215
     9.2.3 Enhancing the NN-Rule with the $k$-NN Rule
              by Voting! ........................................ 216
  9.3 Evaluating the Performance of Classifiers .................. 217
     9.3.1 Misclassification Error Rate ......................... 218
     9.3.2 Confusion Matrices and True/False Positive/
              Negative ........................................ 218
  9.4 Precision, Recall and $F$-Score ............................ 219
  9.5 Statistical Machine Learning and Bayes' Minimal
         Error Bound ........................................... 220
     9.5.1 Non-parametric Probability Density Estimation ........ 220
     9.5.2 Probability of Error and Bayes' Error ................ 222
     9.5.3 Probability of Error for the $k$-NN Rule .............. 223
  9.6 Implementing Nearest Neighbor Queries on a Computer
         Cluster ............................................... 224
  9.7 Notes and References ..................................... 224
  9.8 Summary ............................................... 227
  9.9 Exercises ............................................... 227
References .................................................. 229
10 Fast Approximate Optimization in High Dimensions
with Core-Sets and Fast Dimension Reduction

10.1 Approximate Optimization for Large-Scale Data-Sets

10.1.1 An Example that Illustrates the Needs for High Dimensions

10.1.2 Some Distance Phenomena in High Dimensions

10.1.3 Core-Sets: From Big Data-Sets to Tiny Data-Sets!

10.2 Core-Sets (Coresets, Core sets): Definition

10.3 Core-Sets for the Smallest Enclosing Balls

10.4 A Simple Iterative Heuristic for Approximating the Smallest Enclosing Ball

10.4.1 Convergence Proof

10.4.2 Small Enclosing Balls and Good Margin Linear Separators for SVMs

10.5 Core-Sets for k-Means

10.6 Fast Dimension Reduction via Random Projection Matrices

10.6.1 The Curse of Dimensionality

10.6.2 Two Illustrating Examples of High-Dimensional Tasks

10.6.3 Linear Dimension Reduction

10.6.4 The Johnson–Lindenstrauss’ Theorem

10.6.5 Random Projection Matrices

10.7 Notes and References

10.8 Summary

10.9 Exercises

References

11 Parallel Algorithms for Graphs

11.1 Finding Dense(st) Sub-graphs in Large Graphs

11.1.1 Problem Statement

11.1.2 Densest Sub-graph Complexity and a Simple Greedy Heuristic

11.1.3 A Parallel Heuristic for the Densest Sub-graph

11.2 Testing (Sub)graph Isomorphisms

11.2.1 General Principles of Enumerating Algorithms

11.2.2 Ullman’s Algorithm to Test the Sub-graph Isomorphism

11.2.3 Parallelizing Enumerating Algorithms

11.3 Notes and Discussions

11.4 Summary

11.5 Exercises

References
Appendix A: Written Exam (3 h) ........................................ 261

Appendix B: SLURM: A Resource Manager and Job Scheduler on Clusters of Machines ........................................ 273

Index .................................................................................. 277
Introduction to HPC with MPI for Data Science
Nielsen, F.
2016, XXXIII, 282 p. 101 illus. in color., Softcover
ISBN: 978-3-319-21902-8