

# Contents

- Introduction** **1**
  
- 1 Vector and Parallel Processing** **3**

  - 1.1 Traditional Computers and Their Limitations . . . . . 3
  - 1.2 Parallelism within a Single Processor . . . . . 4
    - 1.2.1 Multiple Functional Units . . . . . 4
    - 1.2.2 Pipelining . . . . . 4
    - 1.2.3 Overlapping . . . . . 6
    - 1.2.4 RISC . . . . . 7
    - 1.2.5 VLIW . . . . . 8
    - 1.2.6 Vector Instructions . . . . . 8
    - 1.2.7 Chaining . . . . . 9
    - 1.2.8 Memory-to-Memory and Register-to-Register Organizations . . . . . 10
    - 1.2.9 Register Set . . . . . 10
    - 1.2.10 Stripmining . . . . . 11
    - 1.2.11 Reconfigurable Vector Registers . . . . . 11
    - 1.2.12 Memory Organization . . . . . 12
  - 1.3 Data Organization . . . . . 14
    - 1.3.1 Main Memory . . . . . 14
    - 1.3.2 Cache . . . . . 16

1.3.3	Local Memory . . . . .	18
1.4	Memory Management . . . . .	18
1.5	Parallelism through Multiple Pipes or Multiple Processors . . . . .	21
1.6	Interconnection Topology . . . . .	22
1.6.1	Crossbar Switch . . . . .	23
1.6.2	Timeshared Bus . . . . .	24
1.6.3	Ring Connection . . . . .	25
1.6.4	Mesh Connection . . . . .	25
1.6.5	Hypercube . . . . .	26
1.6.6	Multistaged Network . . . . .	27
1.7	Programming Techniques . . . . .	29
<b>2</b>	<b>Overview of Current High-Performance Computers</b>	<b>33</b>
2.1	Supercomputers . . . . .	33
2.2	Mini-Supercomputers . . . . .	36
2.3	Vector Mainframes . . . . .	37
2.4	Novel Parallel Processors . . . . .	37
<b>3</b>	<b>Implementation Details and Overhead</b>	<b>43</b>
3.1	Parallel Decomposition and Data Dependency Graphs . . . . .	43
3.2	Synchronization . . . . .	46
3.3	Load Balancing . . . . .	48
3.4	Recurrence . . . . .	49
3.5	Indirect Addressing . . . . .	51
<b>4</b>	<b>Performance: Analysis, Modeling, and Measurements</b>	<b>53</b>
4.1	Amdahl's Law . . . . .	54
4.1.1	Simple Case of Amdahl's Law . . . . .	54
4.1.2	General Form of Amdahl's Law . . . . .	55

4.2	Vector Speed and Vector Length . . . . .	56
4.3	Amdahl's Law—Parallel Processing . . . . .	57
4.3.1	A Simple Model . . . . .	57
4.3.2	Gustafson's Model . . . . .	60
4.4	Examples of $(r_\infty, n_{1/2})$ -values for Various Computers . . . . .	60
4.4.1	CRAY-1 and CRAY-2 (one processor) . . . . .	62
4.4.2	CRAY X-MP (one processor; clock cycle time 8.5 nsec) . . . . .	63
4.4.3	CYBER 205 (2-pipe) and ETA-10P (single processor) . . . . .	63
4.4.4	IBM 3090/VF (1 processor; clock cycle time 18.5 nsec) . . . . .	64
4.4.5	NEC SX/2 . . . . .	65
4.4.6	Convex C-1 and Convex C-210 . . . . .	66
4.4.7	Alliant FX/80 . . . . .	66
4.4.8	General Observations . . . . .	68
4.5	LINPACK Benchmark . . . . .	68
4.5.1	Description of the Benchmark . . . . .	68
4.5.2	Calls to the BLAS . . . . .	69
4.5.3	Asymptotic Performance . . . . .	69
<b>5</b>	<b>Building Blocks in Linear Algebra</b> . . . . .	<b>75</b>
5.1	Basic Linear Algebra Subprograms . . . . .	75
5.1.1	Level 1 BLAS . . . . .	76
5.1.2	Level 2 BLAS . . . . .	77
5.1.3	Level 3 BLAS . . . . .	78
5.2	Levels of Parallelism . . . . .	81
5.2.1	Vector Computers . . . . .	81
5.2.2	Parallel Processors with Shared Memory . . . . .	82
5.2.3	Parallel-Vector Computers . . . . .	83
5.2.4	Clusters of Parallel Vector Processors . . . . .	84

5.3	Basic Factorizations of Linear Algebra . . . . .	84
5.3.1	Point Algorithm: Gaussian Elimination with Partial Pivoting . . . . .	84
5.3.2	Special Matrices . . . . .	86
5.4	Blocked Algorithms: Matrix-Vector and Matrix-Matrix Versions . . . . .	89
5.4.1	Right-Looking Algorithm . . . . .	90
5.4.2	Left-Looking Algorithm . . . . .	92
5.4.3	Crout Algorithm . . . . .	93
5.4.4	Typical Performance of Blocked LU Decomposition . . . . .	94
5.4.5	Blocked Symmetric Indefinite Factorization . . . . .	95
5.4.6	Typical Performance of Blocked Symmetric Indefinite Factorization . . . . .	98
5.5	Linear Least Squares . . . . .	98
5.5.1	Householder Method . . . . .	99
5.5.2	Blocked Householder Method . . . . .	100
5.5.3	Typical Performance of the Blocked Householder Factorization . . . . .	101
5.6	Organization of the Modules . . . . .	102
5.6.1	Matrix-Vector Product . . . . .	102
5.6.2	Matrix-Matrix Product . . . . .	103
5.6.3	Typical Performance for Parallel Processing . . . . .	104
5.6.4	Benefits . . . . .	105
5.7	LAPACK . . . . .	106
<b>6</b>	<b>Direct Solution of Sparse Linear Systems</b>	<b>109</b>
6.1	Introduction to Direct Methods for Sparse Linear Systems . . . . .	111
6.1.1	Three Approaches . . . . .	111
6.1.2	Description of Sparse Data Structure . . . . .	112
6.1.3	Manipulation of Sparse Data Structure . . . . .	114
6.2	General Sparse Matrix Methods . . . . .	116
6.3	Methods for Symmetric Matrices and Band Systems . . . . .	124

6.3.1	The Clique Concept in Gaussian Elimination . . . . .	125
6.3.2	Code Performance and Symmetry . . . . .	128
6.4	Frontal Methods . . . . .	130
6.4.1	Organization . . . . .	130
6.4.2	Vector Performance . . . . .	132
6.5	Multifrontal Methods . . . . .	135
6.5.1	Performance on Vector Machines . . . . .	139
6.5.2	Performance on Parallel Machines . . . . .	139
6.6	Other Approaches for Exploitation of Parallelism . . . . .	141
6.7	Software . . . . .	141
6.8	Brief Summary . . . . .	142
<b>7</b>	<b>Iterative Solution of Sparse Linear Systems</b>	<b>143</b>
7.1	Iterative Methods . . . . .	145
7.1.1	Conjugate Gradient . . . . .	145
7.1.2	Least Squares Conjugate Gradients . . . . .	148
7.1.3	Biconjugate Gradients . . . . .	150
7.1.4	Conjugate Gradient Squared . . . . .	152
7.1.5	GMRES and GMRES(m) . . . . .	154
7.1.6	Adaptive Chebychev . . . . .	156
7.2	Vector and Parallel Aspects . . . . .	158
7.2.1	General Remarks . . . . .	158
7.2.2	Sparse Matrix-Vector Multiplication . . . . .	160
7.2.3	Performance of the Unpreconditioned Methods . . . . .	164
7.3	Preconditioning . . . . .	165
7.3.1	General Aspects . . . . .	165
7.3.2	Efficient Implementations . . . . .	168
7.3.3	Partial Vectorization . . . . .	170

7.3.4	Reordering the Unknowns . . . . .	172
7.3.5	Changing the Order of Computation . . . . .	174
7.3.6	Some Other Vectorizable Preconditioners . . . . .	180
7.3.7	Parallel Aspects . . . . .	183
7.4	Experiences with Parallelism . . . . .	186
7.4.1	General Remarks . . . . .	186
7.4.2	Overlapping Local Preconditioners . . . . .	186
7.4.3	Repeated Twisted Factorization . . . . .	188
7.4.4	Twisted and Nested Twisted Factorization . . . . .	189
7.4.5	Hyperplane Ordering . . . . .	189
<b>A</b>	<b>Acquiring Mathematical Software</b>	<b>191</b>
<b>B</b>	<b>Glossary</b>	<b>197</b>
<b>C</b>	<b>Information on Various High-Performance Computers</b>	<b>213</b>
<b>D</b>	<b>Level 1, 2, and 3 BLAS Quick Reference</b>	<b>221</b>
<b>E</b>	<b>Operation Counts for Various BLAS and Decompositions</b>	<b>227</b>
	<b>Index</b>	<b>247</b>