

Contents

Foreword	xix
Preface	xxi
Acknowledgments.....	xxvii
CHAPTER 1 Introduction	1
1.1 High Performance Computing Disciplines.....	3
1.1.1 Definition	3
1.1.2 Application Programs.....	4
1.1.3 Performance and Metrics	4
1.1.4 High Performance Computing Systems.....	5
1.1.5 Supercomputing Problems	7
1.1.6 Application Programming.....	8
1.2 Impact of Supercomputing on Science, Society, and Security	10
1.2.1 Catalyzing Fraud Detection and Market Data Analytics	10
1.2.2 Discovering, Managing, and Distributing Oil and Gas	10
1.2.3 Accelerating Innovation in Manufacturing.....	10
1.2.4 Personalized Medicine and Drug Discovery	11
1.2.5 Predicting Natural Disasters and Understanding Climate Change.....	12
1.3 Anatomy of a Supercomputer.....	14
1.4 Computer Performance	16
1.4.1 Performance.....	16
1.4.2 Peak Performance.....	17
1.4.3 Sustained Performance	18
1.4.4 Scaling	18
1.4.5 Performance Degradation.....	19
1.4.6 Performance Improvement.....	20
1.5 A Brief History of Supercomputing	21
1.5.1 Epoch I—Automated Calculators Through Mechanical Technologies.....	22
1.5.2 Epoch II—von Neumann Architecture in Vacuum Tubes.....	24
1.5.3 Epoch III—Instruction-Level Parallelism.....	29
1.5.4 Epoch IV—Vector Processing and Integration.....	30
1.5.5 Epoch V—Single-Instruction Multiple Data Array.....	33
1.5.6 Epoch VI—Communicating Sequential Processors and Very Large Scale Integration.....	34
1.5.7 Epoch VII—Multicore Petaflops.....	37
1.5.8 Neodigital Age and Beyond Moore’s Law	37
1.6 This Textbook as a Guide and Tool for the Student.....	38

1.7	Summary and Outcomes of Chapter 1	39
1.8	Questions and Problems	40
	References.....	41
CHAPTER 2	HPC Architecture 1: Systems and Technologies	43
2.1	Introduction.....	44
2.2	Key Properties of HPC Architecture.....	44
2.2.1	Speed	45
2.2.2	Parallelism	45
2.2.3	Efficiency.....	46
2.2.4	Power	46
2.2.5	Reliability	47
2.2.6	Programmability	48
2.3	Parallel Architecture Families—Flynn’s Taxonomy.....	48
2.4	Enabling Technology	51
2.4.1	Technology Epochs	51
2.4.2	Roles of Technologies	55
2.4.3	Digital Logic.....	55
2.4.4	Memory Technologies.....	58
2.5	von Neumann Sequential Processors	62
2.6	Vector and Pipelining	64
2.6.1	Pipeline Parallelism.....	65
2.6.2	Vector Processing	68
2.7	Single-Instruction, Multiple Data Array	69
2.7.1	Single-Instruction, Multiple Data Architecture	69
2.7.2	Amdahl’s Law	70
2.8	Multiprocessors.....	73
2.8.1	Shared-Memory Multiprocessors	74
2.8.2	Massively Parallel Processors	76
2.8.3	Commodity Clusters.....	77
2.9	Heterogeneous Computer Structures.....	78
2.10	Summary and Outcomes of Chapter 2.....	78
2.11	Questions and Problems	80
	References.....	82
CHAPTER 3	Commodity Clusters	83
3.1	Introduction.....	84
3.1.1	Definition of “Commodity Cluster”	84
3.1.2	Motivation and Justification for Clusters.....	84
3.1.3	Cluster Elements.....	85

3.1.4	Impact on Top 500 List.....	86
3.1.5	Brief History.....	88
3.1.6	Chapter Guide.....	90
3.2	Beowulf Cluster Project.....	91
3.3	Hardware Architecture.....	93
3.3.1	The Node.....	93
3.3.2	System Area Networks.....	94
3.3.3	Secondary Storage.....	95
3.3.4	Commercial Systems Summary.....	95
3.4	Programming Interfaces.....	97
3.4.1	High Performance Computing Programming Languages.....	97
3.4.2	Parallel Programming Modalities.....	97
3.5	Software Environment.....	98
3.5.1	Operating Systems.....	98
3.5.2	Resource Management.....	99
3.5.3	Debugger.....	101
3.5.4	Performance Profiling.....	101
3.5.5	Visualization.....	101
3.6	Basic Methods of Use.....	104
3.6.1	Logging On.....	104
3.6.2	User Space and Directory System.....	105
3.6.3	Package Configuration and Building.....	110
3.6.4	Compilers and Compiling.....	112
3.6.5	Running Applications.....	113
3.7	Summary and Outcomes of Chapter 3.....	113
3.8	Questions and Exercises.....	114
	References.....	114
CHAPTER 4	Benchmarking.....	115
4.1	Introduction.....	115
4.2	Key Properties of an HPC Benchmark.....	117
4.3	Standard HPC Community Benchmarks.....	120
4.4	Highly Parallel Computing Linpack.....	120
4.5	HPC Challenge Benchmark Suite.....	123
4.6	High Performance Conjugate Gradients.....	126
4.7	NAS Parallel Benchmarks.....	130
4.8	Graph500.....	132
4.9	Miniapplications as Benchmarks.....	135
4.10	Summary and Outcomes of Chapter 4.....	138
4.11	Exercises.....	139
	References.....	139

CHAPTER 5	The Essential Resource Management	141
5.1	Managing Resources	142
5.2	The Essential SLURM	146
5.2.1	Architecture Overview	147
5.2.2	Workload Organization	148
5.2.3	SLURM Scheduling	149
5.2.4	Summary of Commands	151
5.2.5	SLURM Job Scripting	166
5.2.6	SLURM Cheat Sheet	171
5.3	The Essential Portable Batch System	172
5.3.1	Portable Batch System Overview	172
5.3.2	Portable Batch System Architecture	173
5.3.3	Summary of PBS Commands	174
5.3.4	PBS Job Scripting	184
5.3.5	PBS Cheat Sheet	186
5.4	Summary and Outcomes of Chapter 5	187
5.5	Questions and Problems	189
	References	190
CHAPTER 6	Symmetric Multiprocessor Architecture	191
6.1	Introduction	191
6.2	Architecture Overview	192
6.3	Amdahl's Law Plus	196
6.4	Processor Core Architecture	199
6.4.1	Execution Pipeline	200
6.4.2	Instruction-Level Parallelism	201
6.4.3	Branch Prediction	201
6.4.4	Forwarding	202
6.4.5	Reservation Stations	202
6.4.6	Multithreading	203
6.5	Memory Hierarchy	204
6.5.1	Data Reuse and Locality	204
6.5.2	Memory Hierarchy	205
6.5.3	Memory System Performance	207
6.6	PCI Bus	209
6.7	External I/O Interfaces	213
6.7.1	Network Interface Controllers	213
6.7.2	Serial Advanced Technology Attachment	215
6.7.3	JTAG	218
6.7.4	Universal Serial Bus	220
6.8	Summary and Outcomes of Chapter 6	222
6.9	Questions and Exercises	223
	References	224

CHAPTER 7 The Essential OpenMP	225
7.1 Introduction	225
7.2 Overview of OpenMP Programming Model	226
7.2.1 Thread Parallelism	226
7.2.2 Thread Variables	228
7.2.3 Runtime Library and Environment Variables	228
7.3 Parallel Threads and Loops	231
7.3.1 Parallel Threads	231
7.3.2 Private	232
7.3.3 Parallel “For”	233
7.3.4 Sections	239
7.4 Synchronization	241
7.4.1 Critical Synchronization Directive	242
7.4.2 The Master Directive	242
7.4.3 The Barrier Directive	243
7.4.4 The Single Directive	243
7.5 Reduction	244
7.6 Summary and Outcomes of Chapter 7	245
7.7 Questions and Problems	246
Reference	247
CHAPTER 8 The Essential MPI	249
8.1 Introduction	250
8.2 Message-Passing Interface Standards	251
8.3 Message-Passing Interface Basics	253
8.3.1 mpi.h	253
8.3.2 MPI_Init	253
8.3.3 MPI_Finalize	254
8.3.4 Message-Passing Interface Example—Hello World	254
8.4 Communicators	255
8.4.1 Size	256
8.4.2 Rank	256
8.4.3 Example	257
8.5 Point-to-Point Messages	258
8.5.1 MPI Send	259
8.5.2 Message-Passing Interface Data Types	259
8.5.3 MPI Recv	259
8.5.4 Example	260
8.6 Synchronization Collectives	262
8.6.1 Overview of Collective Calls	262
8.6.2 Barrier Synchronization	263
8.6.3 Example	264

8.7	Communication Collectives	265
8.7.1	Collective Data Movement.....	265
8.7.2	Broadcast	268
8.7.3	Scatter	269
8.7.4	Gather	271
8.7.5	Allgather	272
8.7.6	Reduction Operations.....	274
8.7.7	Alltoall	277
8.8	Nonblocking Point-to-Point Communication.....	279
8.9	User-Defined Data Types.....	281
8.10	Summary and Outcomes of Chapter 8.....	283
8.11	Exercises	283
	References.....	284
CHAPTER 9	Parallel Algorithms	285
9.1	Introduction.....	285
9.2	Fork—Join	286
9.3	Divide and Conquer.....	287
9.4	Manager—Worker	291
9.5	Embarrassingly Parallel.....	292
9.6	Halo Exchange.....	294
9.6.1	The Advection Equation Using Finite Difference.....	295
9.6.2	Sparse Matrix Vector Multiplication.....	297
9.7	Permutation: Cannon’s Algorithm.....	301
9.8	Task Dataflow: Breadth First Search.....	306
9.9	Summary and Outcomes of Chapter 9.....	310
9.10	Exercises	311
	References.....	311
CHAPTER 10	Libraries.....	313
10.1	Introduction	313
10.2	Linear Algebra.....	315
10.2.1	Basic Linear Algebra Subprograms	317
10.2.2	Linear Algebra Package	324
10.2.3	Scalable Linear Algebra Package	326
10.2.4	GNU Scientific Library	326
10.2.5	Supernodal LU	326
10.2.6	Portable Extensible Toolkit for Scientific Computation.....	327
10.2.7	Scalable Library for Eigenvalue Problem Computations.....	328
10.2.8	Eigenvalue SoLvers for Petaflop-Applications.....	328
10.2.9	Hypre: Scalable Linear Solvers and Multigrid Methods.....	328
10.2.10	Domain-Specific Languages for Linear Algebra.....	329

10.3	Partial Differential Equations.....	329
10.4	Graph Algorithms.....	329
10.5	Parallel Input/Output.....	330
10.6	Mesh Decomposition.....	333
10.7	Visualization.....	334
10.8	Parallelization.....	334
10.9	Signal Processing.....	334
10.10	Performance Monitoring.....	341
10.11	Summary and Outcomes of Chapter 10.....	342
10.12	Exercises.....	343
	References.....	344
CHAPTER 11	Operating Systems.....	347
11.1	Introduction.....	347
11.2	Operating System Structures and Services.....	349
11.2.1	System Components.....	349
11.2.2	Process Management.....	349
11.2.3	Memory Management.....	350
11.2.4	File Management.....	350
11.2.5	I/O System Management.....	351
11.2.6	Secondary Storage Management.....	351
11.3	Process Management.....	351
11.3.1	Process States.....	352
11.3.2	Process Control Block.....	353
11.3.3	Process Management Activities.....	354
11.3.4	Scheduling.....	355
11.4	Threads.....	357
11.5	Memory Management.....	358
11.5.1	Virtual Memory.....	359
11.5.2	Virtual Page Addresses.....	359
11.5.3	Virtual Address Translation.....	359
11.6	Summary and Outcomes of Chapter 11.....	361
11.7	Exercises.....	362
CHAPTER 12	Visualization.....	363
12.1	Introduction.....	363
12.2	Foundational Visualization Concepts.....	364
12.3	Gnuplot.....	365
12.4	Matplotlib.....	369
12.5	The Visualization Toolkit.....	372
12.6	ParaView.....	379
12.7	VisIt.....	380

12.8	Summary and Outcomes of Chapter 12.....	381
12.9	Exercises	381
	References.....	382
CHAPTER 13	Performance Monitoring	383
13.1	Introduction.....	383
13.2	Time Measurement	385
13.3	Performance Profiling.....	390
	13.3.1 Significance of Application Profiling.....	390
	13.3.2 Essential <i>gperftools</i>	391
13.4	Monitoring Hardware Events	398
	13.4.1 Perf.....	398
	13.4.2 Performance Application Programming Interface	404
13.5	Integrated Performance Monitoring Toolkits.....	407
13.6	Profiling in Distributed Environments	411
13.7	Summary and Outcomes of Chapter 13.....	417
13.8	Questions and Problems	418
	References.....	419
CHAPTER 14	Debugging.....	421
14.1	Introduction.....	421
14.2	Tools.....	423
	14.2.1 The GNU Debugger	423
	14.2.2 Valgrind	430
	14.2.3 Commercial Parallel Debuggers	431
14.3	Debugging OpenMP Example: Accessing an Unprotected Shared Variable	433
14.4	Debugging MPI Example: Deadlock	434
14.5	Compiler Flags for Debugging.....	439
14.6	System Monitors to Aid Debugging	441
14.7	Summary and Outcomes of Chapter 14.....	445
14.8	Exercises	446
	References.....	450
CHAPTER 15	Accelerator Architecture.....	451
15.1	Introduction.....	451
15.2	A Historic Perspective	454
	15.2.1 Coprocessors.....	456
	15.2.2 Accelerators in Processor I/O Space.....	461
	15.2.3 Accelerators with Industry-Standard Interfaces.....	462
15.3	Introduction to Graphics Processing Units	464
15.4	Evolution of Graphics Processing Unit Functionality	466
15.5	Modern Graphics Processing Unit Architecture	471

15.5.1	Compute Architecture	471
15.5.2	Memory Implementation.....	474
15.5.3	Interconnects.....	475
15.5.4	Programming Environment	476
15.6	Heterogeneous System Architecture	477
15.7	Summary and Outcomes of Chapter 15.....	480
15.8	Problems and Questions	480
	References.....	481
CHAPTER 16	The Essential OpenACC	483
16.1	Introduction.....	483
16.1.1	CUDA	485
16.1.2	OpenCL	485
16.1.3	C++ AMP.....	486
16.1.4	OpenACC.....	486
16.2	OpenACC Programming Concepts	487
16.3	OpenACC Library Calls	489
16.4	OpenACC Environment Variables.....	491
16.5	OpenACC Directives	492
16.5.1	Parallel Construct	493
16.5.2	Kernels Construct.....	495
16.5.3	Data Management.....	497
16.5.4	Loop Scheduling.....	501
16.5.5	Variable Scope.....	504
16.5.6	Atomics.....	504
16.6	Summary and Outcomes of Chapter 16.....	506
16.7	Questions and Problems	506
	References.....	508
CHAPTER 17	Mass Storage.....	509
17.1	Introduction.....	509
17.2	Brief History of Storage	512
17.3	Storage Device Technology.....	514
17.3.1	Hard Disk Drives.....	514
17.3.2	Solid-State Drive Storage.....	520
17.3.3	Magnetic Tape	524
17.3.4	Optical Storage	529
17.4	Aggregated Storage	534
17.4.1	Redundant Array of Independent Disks.....	534
17.4.2	Storage Area Networks	541
17.4.3	Network Attached Storage	543
17.4.4	Tertiary Storage	544

17.5	Summary and Outcomes of Chapter 17.....	545
17.6	Questions and Problems	546
	References.....	547
CHAPTER 18	File Systems	549
18.1	Role and Function of File Systems	549
18.2	The Essential POSIX File Interface	554
18.2.1	System Calls for File Access	554
18.2.2	Buffered File I/O	559
18.3	Network File System	562
18.4	General Parallel File System	565
18.5	Lustre File System.....	569
18.6	Summary and Outcomes of Chapter 18.....	575
18.7	Questions and Problems	576
	References.....	577
CHAPTER 19	MapReduce.....	579
19.1	Introduction.....	579
19.2	Map and Reduce.....	579
19.2.1	Word Count	581
19.2.2	Shared Neighbors	581
19.2.3	K-Means Clustering	582
19.3	Distributed Computation	584
19.4	Hadoop.....	585
19.5	Summary and Outcomes of Chapter 19.....	588
19.6	Exercises	589
	References.....	589
CHAPTER 20	Checkpointing.....	591
20.1	Introduction.....	591
20.2	System-Level Checkpointing.....	592
20.3	Application-Level Checkpointing	598
20.4	Summary and Outcomes of Chapter 20.....	602
20.5	Exercises	602
	References.....	603
CHAPTER 21	Next Steps and Beyond	605
21.1	Introduction.....	606
21.2	Expanded Parallel Programming Models.....	606
21.2.1	Advance in Message-Passing Interface.....	606
21.2.2	Advances in OpenMP.....	607
21.2.3	MPI+X	607

21.3	Extended High Performance Computing Architecture	608
21.3.1	The World's Fastest Machine	608
21.3.2	Lightweight Architectures	608
21.3.3	Field Programmable Gate Arrays	609
21.4	Exascale Computing	610
21.4.1	Challenges to Exascale Computing	611
21.4.2	Doing the Math, How Big Is Exascale?	611
21.4.3	The Accelerated Approach	612
21.4.4	Lightweight Cores	612
21.5	Asynchronous Multitasking	613
21.5.1	Multithreaded	613
21.5.2	Message-Driven Computation	613
21.5.3	Global Address Space	614
21.5.4	Actor Synchronization	615
21.5.5	Runtime System Software	615
21.6	The Neodigital Age	616
21.6.1	Dataflow	617
21.6.2	Cellular Automata	618
21.6.3	Neuromorphic	619
21.6.4	Quantum Computing	619
21.7	Exercises	620
	References	621
	Appendix A: Essential C	623
	Appendix B: Essential Linux	639
	Glossary	667
	Index	677