Distributed and Cloud Computing

From Parallel Processing to the Internet of Things

Kai Hwang

Geoffrey C. Fox

Jack J. Dongarra

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Morgan Kaufmann is an imprint of Elsevier

Contents

Preface xv
About the Authors xix
Foreword xxi

PART 1 SYSTEMS MODELING, CLUSTERING, AND VIRTUALIZATION 1

CHAPTER 1 Distributed System Models and Enabling Technologies 3
Summary 4
1.1 Scalable Computing over the Internet 4
1.1.1 The Age of Internet Computing 4
1.1.2 Scalable Computing Trends and New Paradigms 8
1.1.3 The Internet of Things and Cyber-Physical Systems 11
1.2 Technologies for Network-Based Systems 13
1.2.1 Multicore CPUs and Multithreading Technologies 14
1.2.2 GPU Computing to Exascale and Beyond 17
1.2.3 Memory, Storage, and Wide-Area Networking 20
1.2.4 Virtual Machines and Virtualization Middleware 22
1.2.5 Data Center Virtualization for Cloud Computing 25
1.3 System Models for Distributed and Cloud Computing 27
1.3.1 Clusters of Cooperative Computers 28
1.3.2 Grid Computing Infrastructures 29
1.3.3 Peer-to-Peer Network Families 32
1.3.4 Cloud Computing over the Internet 34
1.4 Software Environments for Distributed Systems and Clouds 36
1.4.1 Service-Oriented Architecture (SOA) 37
1.4.2 Trends toward Distributed Operating Systems 40
1.4.3 Parallel and Distributed Programming Models 42
1.5 Performance, Security, and Energy Efficiency 44
1.5.1 Performance Metrics and Scalability Analysis 45
1.5.2 Fault Tolerance and System Availability 48
1.5.3 Network Threats and Data Integrity 49
1.5.4 Energy Efficiency in Distributed Computing 51
1.6 Bibliographic Notes and Homework Problems 55
Acknowledgments 56
References 56
Homework Problems 58

CHAPTER 2 Computer Clusters for Scalable Parallel Computing 65
Summary 66
2.1 Clustering for Massive Parallelism 66
2.1.1 Cluster Development Trends 66
2.1.2 Design Objectives of Computer Clusters 68
2.1.3 Fundamental Cluster Design Issues 69
2.1.4 Analysis of the Top 500 Supercomputers 71
2.2 Computer Clusters and MPP Architectures 75
2.2.1 Cluster Organization and Resource Sharing 76
2.2.2 Node Architectures and MPP Packaging 77
2.2.3 Cluster System Interconnects 80
2.2.4 Hardware, Software, and Middleware Support 83
2.2.5 GPU Clusters for Massive Parallelism 83
2.3 Design Principles of Computer Clusters 87
2.3.1 Single-System Image Features 87
2.3.2 High Availability through Redundancy 95
2.3.3 Fault-Tolerant Cluster Configurations 99
2.3.4 Checkpointing and Recovery Techniques 101
2.4 Cluster Job and Resource Management 104
2.4.1 Cluster Job Scheduling Methods 104
2.4.2 Cluster Job Management Systems 107
2.4.3 Load Sharing Facility (LSF) for Cluster Computing 109
2.4.4 MOSIX: An OS for Linux Clusters and Clouds 110
2.5 Case Studies of Top Supercomputer Systems 112
2.5.1 Tianhe-1A: The World Fastest Supercomputer in 2010 112
2.5.2 Cray XT5 Jaguar: The Top Supercomputer in 2009 116
2.5.3 IBM Roadrunner: The Top Supercomputer in 2008 119
2.6 Bibliographic Notes and Homework Problems 120
Acknowledgments 121
References 121
Homework Problems 122

CHAPTER 3 Virtual Machines and Virtualization of Clusters and Data Centers 129
Summary 130
3.1 Implementation Levels of Virtualization 130
3.1.1 Levels of Virtualization Implementation 130
3.1.2 VMM Design Requirements and Providers 133
3.1.3 Virtualization Support at the OS Level 135
3.1.4 Middleware Support for Virtualization 138
3.2 Virtualization Structures/Tools and Mechanisms 140
3.2.1 Hypervisor and Xen Architecture 140
3.2.2 Binary Translation with Full Virtualization 141
3.2.3 Para-Virtualization with Compiler Support 143

3.3 Virtualization of CPU, Memory, and I/O Devices ... 145
3.3.1 Hardware Support for Virtualization ... 145
3.3.2 CPU Virtualization ... 147
3.3.3 Memory Virtualization ... 148
3.3.4 I/O Virtualization ... 150
3.3.5 Virtualization in Multi-Core Processors ... 153
3.4 Virtual Clusters and Resource Management ... 155
3.4.1 Physical versus Virtual Clusters ... 156
3.4.2 Live VM Migration Steps and Performance Effects ... 159
3.4.3 Migration of Memory, Files, and Network Resources ... 162
3.4.4 Dynamic Deployment of Virtual Clusters ... 165
3.5 Virtualization for Data-Center Automation ... 169
3.5.1 Server Consolidation in Data Centers ... 169
3.5.2 Virtual Storage Management ... 171
3.5.3 Cloud OS for Virtualized Data Centers ... 172
3.5.4 Trust Management in Virtualized Data Centers ... 176
3.6 Bibliographic Notes and Homework Problems ... 179
Acknowledgments ... 179
References ... 180
Homework Problems ... 183

PART 2 COMPUTING CLOUDS, SERVICE-ORIENTED ARCHITECTURE, AND PROGRAMMING 189

CHAPTER 4 Cloud Platform Architecture over Virtualized Data Centers ... 191
Summary ... 192
4.1 Cloud Computing and Service Models ... 192
4.1.1 Public, Private, and Hybrid Clouds ... 192
4.1.2 Cloud Ecosystem and Enabling Technologies ... 196
4.1.3 Infrastructure-as-a-Service (IaaS) ... 200
4.1.4 Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) ... 203
4.2 Data-Center Design and Interconnection Networks ... 206
4.2.1 Warehouse-Scale Data-Center Design ... 206
4.2.2 Data-Center Interconnection Networks ... 208
4.2.3 Modular Data Center in Shipping Containers ... 211
4.2.4 Interconnection of Modular Data Centers ... 212
4.2.5 Data-Center Management Issues ... 213
4.3 Architectural Design of Compute and Storage Clouds ... 215
4.3.1 A Generic Cloud Architecture Design ... 215
4.3.2 Layered Cloud Architectural Development ... 218
4.3.3 Virtualization Support and Disaster Recovery ... 221
4.3.4 Architectural Design Challenges ... 225

4.4 Public Cloud Platforms: GAE, AWS, and Azure 227
4.4.1 Public Clouds and Service Offerings 227
4.4.2 Google App Engine (GAE) 229
4.4.3 Amazon Web Services (AWS) 231
4.4.4 Microsoft Windows Azure 233
4.5 Inter-cloud Resource Management 234
4.5.1 Extended Cloud Computing Services 235
4.5.2 Resource Provisioning and Platform Deployment 237
4.5.3 Virtual Machine Creation and Management 243
4.5.4 Global Exchange of Cloud Resources 246
4.6 Cloud Security and Trust Management 249
4.6.1 Cloud Security Defense Strategies 249
4.6.2 Distributed Intrusion/Anomaly Detection 253
4.6.3 Data and Software Protection Techniques 255
4.6.4 Reputation-Guided Protection of Data Centers 257
4.7 Bibliographic Notes and Homework Problems 261
Acknowledgements 261
References 261
Homework Problems 265

CHAPTER 5 Service-Oriented Architectures for Distributed Computing 271
Summary 272
5.1 Services and Service-Oriented Architecture 272
5.1.1 REST and Systems of Systems 273
5.1.2 Services and Web Services 277
5.1.3 Enterprise Multitier Architecture 282
5.1.4 Grid Services and OGSA 283
5.1.5 Other Service-Oriented Architectures and Systems 287
5.2 Message-Oriented Middleware 289
5.2.1 Enterprise Bus 289
5.2.2 Publish-Subscribe Model and Notification 291
5.2.3 Queuing and Messaging Systems 291
5.2.4 Cloud or Grid Middleware Applications 291
5.3 Portals and Science Gateways 294
5.3.1 Science Gateway Exemplars 295
5.3.2 HUBzero Platform for Scientific Collaboration 297
5.3.3 Open Gateway Computing Environments (OGCE) 301
5.4 Discovery, Registries, Metadata, and Databases 304
5.4.1 UDDI and Service Registries 304
5.4.2 Databases and Publish-Subscribe 307
5.4.3 Metadata Catalogs 308
5.4.4 Semantic Web and Grid 309
5.4.5 Job Execution Environments and Monitoring 312

5.5 Workflow in Service-Oriented Architectures ... 314
5.5.1 Basic Workflow Concepts ... 315
5.5.2 Workflow Standards ... 316
5.5.3 Workflow Architecture and Specification ... 317
5.5.4 Workflow Execution Engine ... 319
5.5.5 Scripting Workflow System Swift ... 321
5.6 Bibliographic Notes and Homework Problems ... 322
Acknowledgements ... 324
References ... 324
Homework Problems ... 331

CHAPTER 6 Cloud Programming and Software Environments ... 335
Summary ... 336
6.1 Features of Cloud and Grid Platforms ... 336
6.1.1 Cloud Capabilities and Platform Features ... 336
6.1.2 Traditional Features Common to Grids and Clouds ... 336
6.1.3 Data Features and Databases ... 340
6.1.4 Programming and Runtime Support ... 341
6.2 Parallel and Distributed Programming Paradigms ... 343
6.2.1 Parallel Computing and Programming Paradigms ... 344
6.2.2 MapReduce, Twister, and Iterative MapReduce ... 345
6.2.3 Hadoop Library from Apache ... 355
6.2.4 Dryad and DryadLINQ from Microsoft ... 359
6.2.5 Sawzall and Pig Latin High-Level Languages ... 365
6.2.6 Mapping Applications to Parallel and Distributed Systems ... 368
6.3 Programming Support of Google App Engine ... 370
6.3.1 Programming the Google App Engine ... 370
6.3.2 Google File System (GFS) ... 373
6.3.3 BigTable, Google's NOSQL System ... 376
6.3.4 Chubby, Google's Distributed Lock Service ... 379
6.4 Programming on Amazon AWS and Microsoft Azure ... 379
6.4.1 Programming on Amazon EC2 ... 380
6.4.2 Amazon Simple Storage Service (S3) ... 382
6.4.3 Amazon Elastic Block Store (EBS) and SimpleDB ... 383
6.4.4 Microsoft Azure Programming Support ... 384
6.5 Emerging Cloud Software Environments ... 387
6.5.1 Open Source Eucalyptus and Nimbus ... 387
6.5.2 OpenNebula, Sector/Sphere, and OpenStack ... 389
6.5.3 Manjrasoft Aneka Cloud and Appliances ... 393
6.6 Bibliographic Notes and Homework Problems ... 399
Acknowledgement ... 399
References ... 399
Homework Problems ... 405

PART 3 GRIDS, P2P, AND THE FUTURE INTERNET 413

CHAPTER 7 Grid Computing Systems and Resource Management 415

Summary 416

7.1 Grid Architecture and Service Modeling 416

7.1.1 Grid History and Service Families 416

7.1.2 CPU Scavenging and Virtual Supercomputers 419

7.1.3 Open Grid Services Architecture (OGSA) 422

7.1.4 Data-Intensive Grid Service Models 425

7.2 Grid Projects and Grid Systems Built 427

7.2.1 National Grids and International Projects 428

7.2.2 NSF TeraGrid in the United States 430

7.2.3 DataGrid in the European Union 431

7.2.4 The ChinaGrid Design Experiences 434

7.3 Grid Resource Management and Brokering 435

7.3.1 Resource Management and Job Scheduling 435

7.3.2 Grid Resource Monitoring with CGSP 437

7.3.3 Service Accounting and Economy Model 439

7.3.4 Resource Brokering with Gridbus 440

7.4 Software and Middleware for Grid Computing 443

7.4.1 Open Source Grid Middleware Packages 444

7.4.2 The Globus Toolkit Architecture (GT4) 446

7.4.3 Containers and Resources/Data Management 450

7.4.4 The ChinaGrid Support Platform (CGSP) 452

7.5 Grid Application Trends and Security Measures 455

7.5.1 Grid Applications and Technology Fusion 456

7.5.2 Grid Workload and Performance Prediction 457

7.5.3 Trust Models for Grid Security Enforcement 461

7.5.4 Authentication and Authorization Methods 464

7.5.5 Grid Security Infrastructure (GSI) 466

7.6 Bibliographic Notes and Homework Problems 470

Acknowledgments 471

References 471

Homework Problems 473

CHAPTER 8 Peer-to-Peer Computing and Overlay Networks 479

Summary 480

8.1 Peer-to-Peer Computing Systems 480

8.1.1 Basic Concepts of P2P Computing Systems 480

8.1.2 Fundamental Challenges in P2P Computing 486

8.1.3 Taxonomy of P2P Network Systems 490

8.2 P2P Overlay Networks and Properties 492

8.2.1 Unstructured P2P Overlay Networks 492

8.2.2 Distributed Hash Tables (DHTs) ... 496
8.2.3 Structured P2P Overlay Networks ... 498
8.2.4 Hierarchically Structured Overlay Networks ... 501
8.3 Routing, Proximity, and Fault Tolerance ... 505
8.3.1 Routing in P2P Overlay Networks ... 505
8.3.2 Network Proximity in P2P Overlays ... 507
8.3.3 Fault Tolerance and Failure Recovery ... 509
8.3.4 Churn Resilience against Failures ... 512
8.4 Trust, Reputation, and Security Management ... 514
8.4.1 Peer Trust and Reputation Systems ... 514
8.4.2 Trust Overlay and DHT Implementation ... 517
8.4.3 PowerTrust: A Scalable Reputation System ... 520
8.4.4 Securing Overlays to Prevent DDoS Attacks ... 522
8.5 P2P File Sharing and Copyright Protection ... 523
8.5.1 Fast Search, Replica, and Consistency ... 524
8.5.2 P2P Content Delivery Networks ... 529
8.5.3 Copyright Protection Issues and Solutions ... 533
8.5.4 Collusive Piracy Prevention in P2P Networks ... 535
8.6 Bibliographic Notes and Homework Problems ... 538
Acknowledgements ... 538
References ... 539
Homework Problems ... 541

CHAPTER 9 Ubiquitous Clouds and the Internet of Things ... 545
Summary ... 546
9.1 Cloud Trends in Supporting Ubiquitous Computing ... 546
9.1.1 Use of Clouds for HPC/HTC and Ubiquitous Computing ... 546
9.1.2 Large-Scale Private Clouds at NASA and CERN ... 552
9.1.3 Cloud Mashups for Agility and Scalability ... 555
9.1.4 Cloudlets for Mobile Cloud Computing ... 558
9.2 Performance of Distributed Systems and the Cloud ... 561
9.2.1 Review of Science and Research Clouds ... 562
9.2.2 Data-Intensive Scalable Computing (DISC) ... 566
9.2.3 Performance Metrics for HPC/HTC Systems ... 568
9.2.4 Quality of Service in Cloud Computing ... 572
9.2.5 Benchmarking MPI, Azure, EC2, MapReduce, and Hadoop ... 574
9.3 Enabling Technologies for the Internet of Things ... 576
9.3.1 The Internet of Things for Ubiquitous Computing ... 576
9.3.2 Radio-Frequency Identification (RFID) ... 580
9.3.3 Sensor Networks and ZigBee Technology ... 582
9.3.4 Global Positioning System (GPS) ... 587
9.4 Innovative Applications of the Internet of Things ... 590
9.4.1 Applications of the Internet of Things ... 591

9.4.2 Retailing and Supply-Chain Management 591
9.4.3 Smart Power Grid and Smart Buildings 594
9.4.4 Cyber-Physical System (CPS) 595
9.5 Online Social and Professional Networking 597
9.5.1 Online Social Networking Characteristics 597
9.5.2 Graph-Theoretic Analysis of Social Networks 600
9.5.3 Communities and Applications of Social Networks 603
9.5.4 Facebook: The World's Largest Social Network 608
9.5.5 Twitter for Microblogging, News, and Alert Services 611
9.6 Bibliographic Notes and Homework Problems 614
Acknowledgements 614
References 614
Homework Problems 618

Index **623**