

# Contents

Foreword .....	xiii
Preface to the Second Edition .....	xv
Abbreviations .....	xix
<b>CHAPTER 1 Introduction .....</b>	<b>1</b>
1.1 Cloud Computing .....	2
1.2 Network-Centric Computing and Network-Centric Content .....	4
1.3 Cloud Computing, an Old Idea Whose Time Has Come .....	6
1.4 Cloud Delivery Models and Defining Attributes .....	8
1.5 Ethical Issues in Cloud Computing .....	10
1.6 Cloud Vulnerabilities .....	11
<b>CHAPTER 2 Cloud Service Providers and the Cloud Ecosystem .....</b>	<b>13</b>
2.1 The Cloud Ecosystem .....	13
2.2 Cloud Computing Delivery Models and Services .....	15
2.3 Amazon Web Services .....	19
2.4 The Continuing Evolution of AWS .....	26
2.5 Google Clouds .....	30
2.6 Microsoft Windows Azure and Online Services .....	34
2.7 Cloud Storage Diversity and Vendor Lock-In .....	35
2.8 Cloud Computing Interoperability; the Intercloud .....	37
2.9 Service-Level Agreements and Compliance-Level Agreements .....	39
2.10 Responsibility Sharing Between a User and the CSP .....	40
2.11 User Experience .....	42
2.12 Software Licensing .....	44
2.13 Energy Use and Ecological Impact of Cloud Computing .....	45
2.14 Major Challenges Faced by Cloud Computing .....	46
2.15 Further Readings .....	47
2.16 Exercises and Problems .....	48

## SECTION 1

---

<b>CHAPTER 3 Concurrency in the Cloud .....</b>	<b>53</b>
3.1 Enduring Challenges; Concurrency and Cloud Computing .....	54
3.2 Communication and Concurrency in Computing .....	56
3.3 Computational Models; the BSP Model .....	61
3.4 A Model for Multicore Computing .....	64
3.5 Modeling Concurrency with Petri Nets .....	66
3.6 Process State; Global State of a Process or Thread Group .....	72
3.7 Communication Protocols and Process Coordination .....	76

3.8	Communication, Logical Clocks, and Message Delivery Rules	78
3.9	Runs and Cuts; Causal History	82
3.10	Threads and Activity Coordination	85
3.11	Critical Sections, Locks, Deadlocks, and Atomic Actions	93
3.12	Consensus Protocols	98
3.13	Load Balancing	100
3.14	Multithreading and Concurrency in Java; FlumeJava	106
3.15	History Notes and Further Readings	108
3.16	Exercises and Problems	109
<b>CHAPTER 4</b>	<b>Parallel and Distributed Systems</b>	<b>113</b>
4.1	Data, Thread-Level and Task-Level Parallelism	114
4.2	Parallel Architectures	116
4.3	SIMD Architectures; Vector Processing and Multimedia Extensions	120
4.4	Graphics Processing Units	122
4.5	Speedup, Amdahl's Law, and Scaled Speedup	124
4.6	Multicore Processor Speedup	126
4.7	Distributed Systems; System Modularity	129
4.8	Soft Modularity versus Enforced Modularity	131
4.9	Layering and Hierarchy	135
4.10	Virtualization; Layering and Virtualization	137
4.11	Peer-to-Peer Systems	140
4.12	Large-Scale Systems	143
4.13	Composability Bounds and Scalability	144
4.14	History Notes and Further Readings	146
4.15	Exercises and Problems	149
 <b>SECTION 2</b>		
<b>CHAPTER 5</b>	<b>Cloud Access and Cloud Interconnection Networks</b>	<b>153</b>
5.1	Packet-Switched Networks and the Internet	154
5.2	The Transformation of the Internet	159
5.3	Web Access and the TCP Congestion Control Window	163
5.4	Named Data Networks	166
5.5	Software Defined Networks	168
5.6	Interconnection Networks for Computer Clouds	169
5.7	Multistage Interconnection Networks	173
5.8	InfiniBand and Myrinet	175
5.9	Storage Area Networks and the Fibre Channel	178
5.10	Scalable Data Center Communication Architectures	180
5.11	Network Resource Management Algorithms	185
5.12	Content Delivery Networks	188
5.13	Vehicular Ad Hoc Networks	192

5.14	Further Readings	193
5.15	Exercises and Problems	193
<b>CHAPTER 6</b>	<b>Cloud Data Storage</b>	<b>195</b>
6.1	The Evolution of Storage Technologies	196
6.2	Storage Models, File Systems, and Databases	198
6.3	Distributed File Systems; the Precursors	202
6.4	General Parallel File System	208
6.5	Google File System	211
6.6	Locks; Chubby – a Locking Service	213
6.7	NoSQL Databases	218
6.8	Data Storage for Online Transaction Processing Systems	219
6.9	BigTable	220
6.10	Megastore	222
6.11	Storage Reliability at Scale	224
6.12	Disk Locality versus Data Locality in Computer Clouds	228
6.13	Database Provenance	229
6.14	History Notes and Further Readings	231
6.15	Exercises and Problems	233
 <b>SECTION 3</b>		
<b>CHAPTER 7</b>	<b>Cloud Applications</b>	<b>237</b>
7.1	Cloud Application Development and Architectural Styles	238
7.2	Coordination of Multiple Activities	241
7.3	Workflow Patterns	245
7.4	Coordination Based on a State Machine Model – the ZooKeeper	248
7.5	The MapReduce Programming Model	251
7.6	Case Study: the GrepTheWeb Application	254
7.7	Hadoop, Yarn, and Tez	257
7.8	SQL on Hadoop: Pig, Hive, and Impala	262
7.9	Current Cloud Applications and New Opportunities	266
7.10	Clouds for Science and Engineering	268
7.11	Cloud Computing for Biology Research	272
7.12	Social Computing, Digital Content, and Cloud Computing	275
7.13	Software Fault Isolation	277
7.14	Further Readings	278
7.15	Exercises and Problems	278
<b>CHAPTER 8</b>	<b>Cloud Hardware and Software</b>	<b>281</b>
8.1	Challenges; Virtual Machines and Containers	282
8.2	Cloud Hardware; Warehouse-Scale Computers	284
8.3	WSC Performance	286
8.4	Hypervisors	290
8.5	An Engine for Coarse-Grained Data-Parallel Applications	290

8.6	Fine-Grained Cluster Resource Sharing	292
8.7	Cluster Management with Borg	294
8.8	Shared State Cluster Management	296
8.9	QoS-Aware Cluster Management	298
8.10	Resource Isolation	301
8.11	In-Memory Cluster Computing for Big Data	306
8.12	Containers; Docker Containers	313
8.13	Kubernetes	315
8.14	Further Readings	317
8.15	Exercises and Problems	318
<b>CHAPTER 9</b>	<b>Cloud Resource Management and Scheduling</b>	<b>321</b>
9.1	Policies and Mechanisms for Resource Management	322
9.2	Cloud Resource Utilization and Energy Efficiency	324
9.3	Resource Management and Dynamic Application Scaling	327
9.4	Control Theory and Optimal Resource Management	328
9.5	Stability of a Two-Level Resource Allocation Architecture	332
9.6	Feedback Control Based on Dynamic Thresholds	334
9.7	Coordination of Autonomic Performance Managers	335
9.8	A Utility Model for Cloud-Based Web Services	337
9.9	Scheduling Algorithms for Computer Clouds	341
9.10	Delay Scheduling	343
9.11	Data-Aware Scheduling	348
9.12	Apache Capacity Scheduler	351
9.13	Start-Time Fair Queuing	352
9.14	Borrowed Virtual Time	356
9.15	Further Readings	359
9.16	Exercises and Problems	362
<b>CHAPTER 10</b>	<b>Cloud Resource Virtualization</b>	<b>365</b>
10.1	Performance and Security Isolation in Computer Clouds	366
10.2	Virtual Machines	367
10.3	Full Virtualization and Paravirtualization	370
10.4	Hardware Support for Virtualization	372
10.5	Xen – a Hypervisor Based on Paravirtualization	375
10.6	Optimization of Network Virtualization in Xen 2.0	380
10.7	Kernel-Based Virtual Machine	383
10.8	Nested Virtualization	383
10.9	A Trusted Kernel-Based Virtual Machine for ARMv8	388
10.10	Paravirtualization of Itanium Architecture	389
10.11	A Performance Comparison of Virtual Machines	392
10.12	Open-Source Software Platforms for Private Clouds	395
10.13	The Darker Side of Virtualization	398
10.14	Virtualization Software	399

10.15 History Notes and Further Readings ..... 400  
 10.16 Exercises and Problems ..... 401

**SECTION 4**

**CHAPTER 11 Cloud Security ..... 405**  
 11.1 Security, the Top Concern for Cloud Users ..... 406  
 11.2 Cloud Security Risks ..... 407  
 11.3 Privacy and Privacy Impact Assessment ..... 411  
 11.4 Trust ..... 413  
 11.5 Cloud Data Encryption ..... 415  
 11.6 Security of Database Services ..... 417  
 11.7 Operating System Security ..... 419  
 11.8 Virtual Machine Security ..... 420  
 11.9 Security of Virtualization ..... 422  
 11.10 Security Risks Posed by Shared Images ..... 425  
 11.11 Security Risks Posed by a Management OS ..... 428  
 11.12 Xoar – Breaking the Monolithic Design of the TCB ..... 431  
 11.13 A Trusted Hypervisor ..... 434  
 11.14 Mobile Devices and Cloud Security ..... 435  
 11.15 Further Readings ..... 436  
 11.16 Exercises and Problems ..... 437

**CHAPTER 12 Big Data, Data Streaming, and the Mobile Cloud ..... 439**  
 12.1 Big Data ..... 440  
 12.2 Data Warehouses and Google Databases for Big Data ..... 442  
 12.3 Bootstrapping Techniques for Data Analytics ..... 450  
 12.4 Approximate Query Processing ..... 453  
 12.5 Dynamic Data-Driven Applications ..... 457  
 12.6 Data Streaming ..... 460  
 12.7 A Dataflow Model for Data Streaming ..... 463  
 12.8 Joining Multiple Data Streams ..... 467  
 12.9 System Availability at Scale ..... 469  
 12.10 The Scale and the Latency ..... 471  
 12.11 Mobile Computing and Applications ..... 476  
 12.12 Energy Efficiency of Mobile Computing ..... 478  
 12.13 Alternative Mobile Cloud Computing Models ..... 480  
 12.14 Mobile Edge Cloud and Markov Decision Processes ..... 482  
 12.15 Further Readings ..... 485  
 12.16 Exercises and Problems ..... 486

**CHAPTER 13 Advanced Topics ..... 489**  
 13.1 A Glimpse at the Future ..... 489  
 13.2 Cloud Scheduling Subject to Deadlines ..... 491  
 13.3 Scheduling MapReduce Applications Subject to Deadlines ..... 496

13.4	Emergence and Self-Organization .....	497
13.5	Resource Bundling; Combinatorial Auctions for Cloud Resources .....	500
13.6	Cloud Interoperability and Super Clouds .....	503
13.7	In Search for Blooms Amid a Flurry of Challenges .....	505
<b>APPENDIX A</b>	<b>Cloud Application Development (online appendix) .....</b>	<b>507</b>
<b>APPENDIX B</b>	<b>Cloud Projects (online appendix) .....</b>	<b>509</b>
Literature .....		511
Glossary .....		533
Index .....		549

## ONLINE APPENDICES

---

<b>APPENDIX A</b>	<b>Cloud Application Development .....</b>	<b>e1</b>
A.1	AWS EC2 Instances .....	e2
A.2	Connecting Clients to Cloud Instances through Firewalls .....	e3
A.3	Security Rules for Application- and Transport-Layer Protocols in EC2 .....	e7
A.4	How to Launch an EC2 Linux Instance and Connect to It .....	e8
A.5	How to Use S3 in Java .....	e12
A.6	How to Manage AWS SQS Services in C# .....	e14
A.7	How to Install SNS on Ubuntu 10.04 .....	e15
A.8	How to Create an EC2 Placement Group and use MPI .....	e17
A.9	StarCluster – A Cluster Computing Toolkit for EC2 .....	e19
A.10	An Alternative Setting of an MPI Virtual Cluster .....	e19
A.11	How to Install Hadoop on Eclipse on a Windows System .....	e21
A.12	Exercises and Problems .....	e24
<b>APPENDIX B</b>	<b>Cloud Projects .....</b>	<b>e27</b>
B.1	Cloud-Based Simulation of a Distributed Trust Algorithm .....	e27
B.2	Simulation of Traffic Management in a Smart City .....	e31
B.3	A Cloud Trust Management Service .....	e37
B.4	A Cloud Service for Adaptive Data Streaming .....	e45
B.5	Cloud-Based Optimal FPGA Synthesis .....	e49
B.6	Tensor Network Contraction on AWS .....	e51