Contents

Acknowledgments xi
Preface xiii

PART 1 FOUNDATIONS

CHAPTER 1 Introduction 3
1.1 Cloud computing at a glance 3
1.1.1 The vision of cloud computing 5
1.1.2 Defining a cloud 7
1.1.3 A closer look 9
1.1.4 The cloud computing reference model 11
1.1.5 Characteristics and benefits 13
1.1.6 Challenges ahead 14
1.2 Historical developments 15
1.2.1 Distributed systems 15
1.2.2 Virtualization 18
1.2.3 Web 2.0 19
1.2.4 Service-oriented computing 20
1.2.5 Utility-oriented computing 21
1.3 Building cloud computing environments 22
1.3.1 Application development 22
1.3.2 Infrastructure and system development 23
1.3.3 Computing platforms and technologies 24
Summary 26
Review questions 27

CHAPTER 2 Principles of Parallel and Distributed Computing 29
2.1 Eras of computing 29
2.2 Parallel vs. distributed computing 29
2.3 Elements of parallel computing 31
2.3.1 What is parallel processing? 31
2.3.2 Hardware architectures for parallel processing 32
2.3.3 Approaches to parallel programming 36
2.3.4 Levels of parallelism 36
2.3.5 Laws of caution 37

2.4 Elements of distributed computing 39
2.4.1 General concepts and definitions 39
2.4.2 Components of a distributed system 39
2.4.3 Architectural styles for distributed computing 41
2.4.4 Models for interprocess communication 51
2.5 Technologies for distributed computing 54
2.5.1 Remote procedure call 54
2.5.2 Distributed object frameworks 56
2.5.3 Service-oriented computing 61
Summary 69
Review questions 70

CHAPTER 3 Virtualization 71
3.1 Introduction 71
3.2 Characteristics of virtualized environments 73
3.2.1 Increased security 74
3.2.2 Managed execution 75
3.2.3 Portability 77
3.3 Taxonomy of virtualization techniques 77
3.3.1 Execution virtualization 77
3.3.2 Other types of virtualization 89
3.4 Virtualization and cloud computing 91
3.5 Pros and cons of virtualization 93
3.5.1 Advantages of virtualization 93
3.5.2 The other side of the coin: disadvantages 94
3.6 Technology examples 95
3.6.1 Xen: paravirtualization 96
3.6.2 VMware: full virtualization 97
3.6.3 Microsoft Hyper-V 104
Summary 109
Review questions 109

CHAPTER 4 Cloud Computing Architecture 111
4.1 Introduction 111
4.2 The cloud reference model 112
4.2.1 Architecture 112
4.2.2 Infrastructure- and hardware-as-a-service 114

4.2.3 Platform as a service 117
4.2.4 Software as a service 121
4.3 Types of clouds 124
4.3.1 Public clouds 125
4.3.2 Private clouds 126
4.3.3 Hybrid clouds 128
4.3.4 Community clouds 131
4.4 Economics of the cloud 133
4.5 Open challenges 135
4.5.1 Cloud definition 135
4.5.2 Cloud interoperability and standards 136
4.5.3 Scalability and fault tolerance 137
4.5.4 Security, trust, and privacy 138
4.5.5 Organizational aspects 138
Summary 139
Review questions 139

PART 2 CLOUD APPLICATION PROGRAMMING AND THE ANEKA PLATFORM

CHAPTER 5 Aneka 143
5.1 Framework overview 143
5.2 Anatomy of the Aneka container 146
5.2.1 From the ground up: the platform abstraction layer 147
5.2.2 Fabric services 147
5.2.3 Foundation services 150
5.2.4 Application services 153
5.3 Building Aneka clouds 155
5.3.1 Infrastructure organization 155
5.3.2 Logical organization 155
5.3.3 Private cloud deployment mode 158
5.3.4 Public cloud deployment mode 158
5.3.5 Hybrid cloud deployment mode 160
5.4 Cloud programming and management 162
5.4.1 Aneka SDK 162
5.4.2 Management tools 167
Summary 168
Review questions 168

CHAPTER 6 Concurrent Computing 171
6.1 Introducing parallelism for single-machine computation 171
6.2 Programming applications with threads 173
6.2.1 What is a thread? 174
6.2.2 Thread APIs 174
6.2.3 Techniques for parallel computation with threads 177
6.3 Multithreading with Aneka 189
6.3.1 Introducing the thread programming model 190
6.3.2 Aneka thread vs. common threads 191
6.4 Programming applications with Aneka threads 195
6.4.1 Aneka threads application model 195
6.4.2 Domain decomposition: matrix multiplication 196
6.4.3 Functional decomposition: *Sine*, *Cosine*, and *Tangent* 203
Summary 203
Review questions 210

CHAPTER 7 High-Throughput Computing 211
7.1 Task computing 211
7.1.1 Characterizing a task 212
7.1.2 Computing categories 213
7.1.3 Frameworks for task computing 214
7.2 Task-based application models 216
7.2.1 Embarrassingly parallel applications 216
7.2.2 Parameter sweep applications 217
7.2.3 MPI applications 218
7.2.4 Workflow applications with task dependencies 222
7.3 Aneka task-based programming 225
7.3.1 Task programming model 226
7.3.2 Developing applications with the task model 227
7.3.3 Developing a parameter sweep application 243
7.3.4 Managing workflows 248
Summary 250
Review questions 251

CHAPTER 8 Data-Intensive Computing 253
8.1 What is data-intensive computing? 253
8.1.1 Characterizing data-intensive computations 254

8.1.2 Challenges ahead 254
8.1.3 Historical perspective 255
8.2 Technologies for data-intensive computing 260
8.2.1 Storage systems 260
8.2.2 Programming platforms 268
8.3 Aneka MapReduce programming 276
8.3.1 Introducing the MapReduce programming model 276
8.3.2 Example application 293
Summary 309
Review questions 310

PART 3 INDUSTRIAL PLATFORMS AND NEW DEVELOPMENTS

CHAPTER 9 Cloud Platforms in Industry 315
9.1 Amazon web services 315
9.1.1 Compute services 316
9.1.2 Storage services 321
9.1.3 Communication services 329
9.1.4 Additional services 332
9.2 Google AppEngine 332
9.2.1 Architecture and core concepts 333
9.2.2 Application life cycle 338
9.2.3 Cost model 340
9.2.4 Observations 341
9.3 Microsoft Azure 341
9.3.1 Azure core concepts 342
9.3.2 SQL Azure 347
9.3.3 Windows Azure platform appliance 349
9.3.4 Observations 349
Summary 350
Review questions 351

CHAPTER 10 Cloud Applications 353
10.1 Scientific applications 353
10.1.1 Healthcare: ECG analysis in the cloud 353
10.1.2 Biology: protein structure prediction 355
10.1.3 Biology: gene expression data analysis for cancer diagnosis 357
10.1.4 Geoscience: satellite image processing 358

10.2 Business and consumer applications 358
10.2.1 CRM and ERP 359
10.2.2 Productivity 362
10.2.3 Social networking 365
10.2.4 Media applications 366
10.2.5 Multiplayer online gaming 369
Summary 370
Review questions 371

CHAPTER 11 Advanced Topics in Cloud Computing 373
11.1 Energy efficiency in clouds 373
11.1.1 Energy-efficient and green cloud computing architecture 375
11.2 Market-based management of clouds 377
11.2.1 Market-oriented cloud computing 378
11.2.2 A reference model for MOCC 379
11.2.3 Technologies and initiatives supporting MOCC 384
11.2.4 Observations 389
11.3 Federated clouds/InterCloud 390
11.3.1 Characterization and definition 391
11.3.2 Cloud federation stack 392
11.3.3 Aspects of interest 399
11.3.4 Technologies for cloud federations 417
11.3.5 Observations 422
11.4 Third-party cloud services 422
11.4.1 MetaCDN 423
11.4.2 SpotCloud 425
Summary 425
Review questions 427
References 429
Index 439