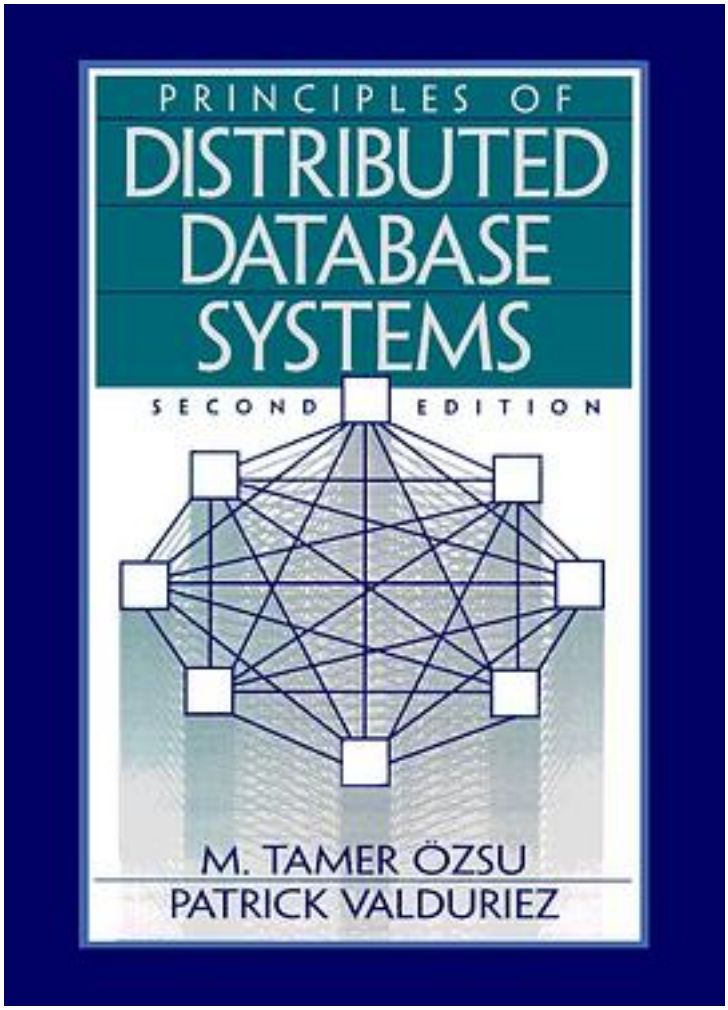
Principles of Distributed Database Systems



Principles of Distributed Database Systems_下载链接1_

著者:M. Tamer Özsu

出版者:Springer

出版时间:2020-1-4

装帧:Hardcover

isbn:9783030262525

The fourth edition of this classic textbook provides major updates. This edition has

completely new chapters on Big Data Platforms (distributed storage systems, MapReduce, Spark, data stream processing, graph analytics) and on NoSQL, NewSQL and polystore systems. It also includes an updated web data management chapter that includes RDF and semantic web discussion, an integrated database integration chapter focusing both on schema integration and querying over these systems. The peer-to-peer computing chapter has been updated with a discussion of blockchains. The chapters that describe classical distributed and parallel database technology have all been updated.

The new edition covers the breadth and depth of the field from a modern viewpoint. Graduate students, as well as senior undergraduate students studying computer science and other related fields will use this book as a primary textbook. Researchers working in computer science will also find this textbook useful.

This textbook has a companion web site that includes background information on relational database fundamentals, query processing, transaction management, and computer networks for those who might need this background. The web site also includes all the figures and presentation slides as well as solutions to exercises (restricted to instructors).

作者介绍:

M. Tamer Özsu is a University Professor at Cheriton School of Computer Science at University of Waterloo, Canada. He has been conducting research in distributed data management for thirty years. He is a Fellow of the Royal Society of Canada, American Association for the Advancement of Science (AAAS), Association for Computing Machinery (ACM) and Institute of Electrical and Electronics Engineers (IEEE). He is an elected member of the Science Academy, Turkey, and a member of Sigma Xi. He has received the CS-Can/Info-Can (Canadian Computer Science Society) Lifetime Achievement Award in 2019, ACM SIGMOD Test-of-Time Award in 2015, ACM SIGMOD Contributions Award in 2008 and the Ohio State University College of Engineering Distinguished Alumnus Award in 2008 and has two best paper awards and one honourable mention for his publications. He serves on the editorial boards of many journals and book series, and is also the co-editor-in-chief, with Ling Liu, of the Encyclopedia of Database Systems.

Patrick Valduriez is a senior scientist at Inria, France. He has also been a professor of computer science at University Pierre et Marie Curie (UPMC) in Paris (2000-2002) and a researcher at Microelectronics and Computer Technology Corp. in Austin, Texas (1985-1989). Since 2019, he is the scientific advisor of the LeanXcale startup.

He is currently the head of the Zenith team (between Inria and University of Montpellier, LIRMM) that focuses on data science, in particular data management in large-scale distributed and parallel systems and scientific data management. He currently serves as associate editor of several journals, including the VLDB Journal, Distributed and Parallel Databases, and Internet and Databases. He has served as PC chair of major conferences such as SIGMOD and VLDB. He was the general chair of SIGMOD 2004, EDBT 2008 and VLDB 2009 Conferences. He obtained several best paper awards, including at VLDB 2000. He was the recipient of the 1993 IBM scientific prize in Computer Science in France and the 2014 Innovation Award from Inria – French Academy of Science – Dassault Systems. He is an ACM Fellow.

目录: 1 Introduction 1
1.1 What Is a Distributed Database System?..... 1
1.2 History of Distributed DBMS 3
1.3 Data Delivery Alternatives..... 5
1.4 Promises of Distributed DBMSs 7
1.4.1 Transparent Management of Distributed and
Replicated Data..... 7
1.4.2 Reliability Through Distributed Transactions 10
1.4.3 Improved Performance..... 11
1.4.4 Scalability 13
1.5 Design Issues 13
1.5.1 Distributed Database Design 13
1.5.2 Distributed Data Control..... 14
1.5.3 Distributed Query Processing..... 14
1.5.4 Distributed Concurrency Control..... 14
1.5.5 Reliability of Distributed DBMS 15
1.5.6 Replication..... 15
1.5.7 Parallel DBMSs 16
1.5.8 Database Integration 16
1.5.9 Alternative Distribution Approaches 16
1.5.10 Big Data Processing and NoSQL..... 16
1.6 Distributed DBMS Architectures 17
1.6.1 Architectural Models for Distributed DBMSs 17
1.6.2 Client/Server Systems..... 20
1.6.3 Peer-to-Peer Systems..... 22
1.6.4 Multidatabase Systems..... 25
1.6.5 Cloud Computing 27
1.7 Bibliographic Notes 31

2 Distributed and Parallel Database Design 33
2.1 Data Fragmentation 35
2.1.1 Horizontal Fragmentation..... 37
2.1.2 Vertical Fragmentation 52
2.1.3 Hybrid Fragmentation..... 65
2.2 Allocation..... 66
2.2.1 Auxiliary Information..... 68
2.2.2 Allocation Model..... 69
2.2.3 Solution Methods 72
2.3 Combined Approaches 72
2.3.1 Workload-Agnostic Partitioning Techniques 73
2.3.2 Workload-Aware Partitioning Techniques 74
2.4 Adaptive Approaches 78
2.4.1 Detecting Workload Changes..... 79
2.4.2 Detecting Affected Items 79
2.4.3 Incremental Reconfiguration..... 80
2.5 Data Directory 82
2.6 Conclusion..... 83
2.7 Bibliographic Notes 84
3 Distributed Data Control 91
3.1 View Management 92
3.1.1 Views in Centralized DBMSs..... 92
3.1.2 Views in Distributed DBMSs 95
3.1.3 Maintenance of Materialized Views..... 96

3.2 Access Control 102
3.2.1 Discretionary Access Control..... 103
3.2.2 Mandatory Access Control 106
3.2.3 Distributed Access Control 108
3.3 Semantic
3.3.1 Centralized Semantic Integrity Control 111
3.3.2 Distributed Semantic Integrity Control..... 116
Integrity Control..... 110 3.4
Conclusion..... 123
3.5 Bibliographic Notes 123
4 Distributed Query Processing 129
4.1 Overview..... 130
4.1.1 Query Processing Problem..... 130
4.1.2 Query Optimization 133
4.1.3 Layers Of Query Processing 136
4.2 Data Localization..... 140
4.2.1 Reduction for Primary Horizontal Fragmentation 141
4.2.2 Reduction with Join 142
4.2.3 Reduction for Vertical Fragmentation 143
4.2.4 Reduction for Derived Fragmentation..... 145
4.2.5 Reduction for Hybrid Fragmentation..... 148
Contents
xiii
4.3 Join Ordering in Distributed Queries..... 149
4.3.1 Join Trees 149
4.3.2 Join Ordering 151
4.3.3 Semijoin-Based Algorithms 153
4.3.4 Join Versus Semijoin 156
4.4 Distributed Cost Model 157
4.4.1 Cost Functions..... 157
4.4.2 Database Statistics 159
4.5 Distributed Query Optimization 161
4.5.1 Dynamic Approach..... 161
4.5.2 Static Approach 165
4.5.3 Hybrid Approach 169
4.6 Adaptive Query Processing 173
4.6.1 Adaptive Query Processing Process..... 174
4.6.2 Eddy Approach 176
4.7 Conclusion..... 177
4.8 Bibliographic Notes 178
5 Distributed Transaction Processing 183
5.1 Background and Terminology 184
5.2 Distributed Concurrency Control 188
5.2.1 Locking-Based Algorithms 189
5.2.2 Timestamp-Based Algorithms 197
5.2.3 Multiversion Concurrency Control 203
5.2.4 Optimistic Algorithms 205
5.3 Distributed Concurrency Control Using Snapshot Isolation 206
5.4 Distributed DBMS Reliability 209
5.4.1 Two-Phase Commit Protocol 211
5.4.2 Variations of 2PC..... 217
5.4.3 Dealing with Site Failures 220
5.4.4 Network Partitioning 227
5.4.5 Paxos Consensus Protocol 231

5.4.6 Architectural Considerations 234
5.5 Modern
5.5.1 Spanner 237
5.5.2 LeanXcale 237
Approaches to Scaling Out Transaction Management 236 5.6
Conclusion..... 239
5.7 Bibliographic Notes 241
6 Data Replication 247
6.1 Consistency of Replicated Databases 249
6.1.1 Mutual Consistency 249
6.1.2 Mutual Consistency Versus Transaction Consistency . . . 251
6.2 Update Management Strategies..... 252
6.2.1 Eager Update Propagation 253
6.2.2 Lazy Update Propagation 254
xiv
Contents
7
6.3.1 Eager Centralized Protocols 256
6.3.2 Eager Distributed Protocols..... 262
6.3.3 Lazy Centralized Protocols 262
6.3.4 Lazy Distributed Protocols 268
6.4 Group Communication..... 269
6.5 Replication and Failures 272
6.5.1 Failures and Lazy Replication 273
6.5.2 Failures and Eager Replication 273
6.6 Conclusion..... 276
6.7 Bibliographic Notes 277
Database Integration—Multidatabase Systems 281
7.1 Database Integration 282
7.1.1 Bottom-Up Design Methodology 283
7.1.2 Schema Matching 287
7.1.3 Schema Integration..... 296
7.1.4 Schema Mapping 298
7.1.5 Data Cleaning 306
7.2 Multidatabase Query Processing 307
7.2.1 Issues in Multidatabase Query Processing 308
7.2.2 Multidatabase Query Processing Architecture 309
7.2.3 Query Rewriting Using Views 311
7.2.4 Query Optimization and Execution 317
7.2.5 Query Translation and Execution..... 329
7.3 Conclusion..... 332
7.4 Bibliographic Notes 334
Parallel Database Systems 349
8.1 Objectives..... 350
8.2 Parallel Architectures 352
8.2.1 General Architecture ..... 353 8.2.2 Shared-Memory
..... 355 8.2.3 Shared-Disk 357 8.2.4
Shared-Nothing..... 358
8.3 Data Placement 359
8.4 Parallel Query Processing..... 362
8
6.2.3
Centralized Techniques 254
Distributed Techniques..... 255

6.2.4
6.3 Replication Protocols 255
8.4.1
8.4.2 8.5 Load 8.5.1 8.5.2 8.5.3 8.5.4
Parallel Algorithms for Data Processing 362
Parallel Query Optimization 369
Balancing..... 374 Parallel Execution Problems
..... 374 Intraoperator Load Balancing..... 376 Interoperator
Load Balancing..... 378 Intraquery Load Balancing 378

8.6 Fault-Tolerance 383
8.7 Database Clusters 384
8.7.1 Database Cluster Architecture 385
8.7.2 Replication..... 386
8.7.3 Load Balancing..... 386
8.7.4 Query Processing..... 387
8.8 Conclusion..... 390
8.9 Bibliographic Notes 390
9 Peer-to-Peer Data Management..... 395
9.1 Infrastructure 398
9.1.1 Unstructured P2P Networks 399
9.1.2 Structured P2P Networks 402
9.1.3 Superpeer P2P Networks 406
9.1.4 Comparison of P2P Networks 408
9.2 Schema Mapping in P2P Systems 408
9.2.1 Pairwise Schema Mapping..... 408
9.2.2 Mapping Based on Machine Learning Techniques 409
9.2.3 Common Agreement Mapping 410
9.2.4 Schema Mapping Using IR Techniques..... 411
9.3 Querying
9.3.1 Top-k Queries 412
9.3.2 Join Queries 424
9.3.3 Range Queries 425
9.4 Replica Consistency..... 428
9.4.1 Basic Support in DHTs 429
9.4.2 Data Currency in DHTs..... 431
9.4.3 Replica Reconciliation 432
9.5 Blockchain..... 436
9.5.1 Blockchain Definition..... 437
9.5.2 Blockchain Infrastructure 438
9.5.3 Blockchain 2.0..... 442
9.5.4 Issues..... 443
9.6 Conclusion..... 444
9.7 Bibliographic Notes 445
10 Big Data Processing 449
10.1 Distributed Storage Systems 451
10.1.1 Google File System 453
10.1.2 Combining Object Storage and File Storage..... 454
10.2 Big Data Processing Frameworks 455
10.2.1 MapReduce Data Processing 456
10.2.2 Data Processing Using Spark 466
10.3 Stream Data Management 470
10.3.1 Stream Models, Languages, and Operators 472

10.3.2 Query Processing over Data Streams..... 476
10.3.3 DSS Fault-Tolerance 483
Over P2P Systems..... 411

11
10.4 Graph 10.4.1 10.4.2 10.4.3 10.4.4 10.4.5 10.4.6 10.4.7 10.4.8 10.4.9
10.4.10 10.4.11 10.4.12
10.5 Data Lakes 508
10.5.1 Data Lake Versus Data Warehouse 508
10.5.2 Architecture 510
10.5.3 Challenges 511
10.6 Conclusion..... 512
10.7 Bibliographic Notes 512
NoSQL, NewSQL, and Polystores 519
11.1 Motivations for NoSQL 520
11.2 Key-Value Stores 521
11.2.1 DynamoDB 522
11.2.2 OtherKey-ValueStores..... 524
11.3 Document Stores 525
11.3.1 MongoDB 525
11.3.2 Other Document Stores 528
11.4 Wide Column Stores ..... 529 11.4.1 Bigtable
..... 529 11.4.2 Other Wide Column Stores 531
11.5 Graph DBMSs..... 531 11.5.1
Neo4j..... 532 11.5.2 Other Graph Databases
..... 535
11.6 Hybrid Data Stores 535
11.6.1 Multimodel NoSQL Stores 536
11.6.2 NewSQL DBMSs 537
11.7 Polystores..... 540
11.7.1 Loosely Coupled Polystores 540
11.7.2 Tightly Coupled Polystores 544
11.7.3 Hybrid Systems 549
11.7.4 Concluding Remarks 553
11.8 Conclusion..... 554
11.9 Bibliographic Notes 555
Analytics Platforms..... 486 Graph Partitioning.....
489 MapReduce and Graph Analytics ..... 494 Special-Purpose Graph Analytics
Systems ..... 495 Vertex-Centric Block Synchronous..... 498 Vertex-Centric
Asynchronous ..... 501 Vertex-Centric Gather-Apply-Scatter 503
Partition-Centric Block Synchronous Processing . . . . . 504 Partition-Centric
Asynchronous ..... 506 Partition-Centric Gather-Apply-Scatter 506
Edge-Centric Block Synchronous Processing . . . . . 507 Edge-Centric
Asynchronous ..... 507 Edge-Centric Gather-Apply-Scatter 507

12 Web Data Management 559
12.1 Web Graph Management..... 560
12.2 Web Search 562
12.2.1 Web Crawling 563
12.2.2 Indexing 566
12.2.3 Ranking and Link Analysis 567
12.2.4 Evaluation of Keyword Search 568
12.3 Web Querying 569

12.3.1 Semistructured Data Approach 570
12.3.2 Web Query Language Approach 574
12.4 Question Answering Systems..... 580
12.5 Searching and Querying the Hidden Web 584
12.5.1 Crawling the Hidden Web 585
12.5.2 Metasearching 586
12.6 Web Data Integration..... 588
12.6.1 Web Tables/Fusion Tables 589
12.6.2 Semantic Web and Linked Open Data 590
12.6.3 Data Quality Issues in Web Data Integration 608
12.7 Bibliographic Notes 615
A Overview of Relational DBMS 619
B Centralized Query Processing..... 621
C Transaction Processing Fundamentals 623
D Review of Computer Networks..... 625
References..... 627
Index..... 663
・・・・・・(收起)

Principles of Distributed Database Systems_下载链接1_

标签

分布式系统

数据库

计算机科学

分布式

DistributedSystem

计算机

Database

软件工程

评论

Principles of Distributed Database Systems_下载链接1_

书评

Principles of Distributed Database Systems_下载链接1_