

Database System Concepts 6e

About the book

Now in its sixth edition, Silberschatz, Korth, and Sudarshan's *Database System Concepts* is one of the cornerstone texts of database education. It presents the fundamental concepts of database management in an intuitive manner geared toward allowing students to begin working with databases as quickly as possible. The text emphasizes practical issues, applications, and implementation, coupled with coverage of key theoretical concepts in a clear and well-motivated manner. To keep the focus manageable, concepts and algorithms are presented in a general setting that is not tied to one particular database system, but annotated where required with variations specific to particular database systems. Familiarity with basic data structures, computer organization, and a high-level programming language are the only prerequisites.

New In The Sixth Edition

- **Revised coverage of SQL**
 - Greater attention to variants of SQL in actual systems
 - puts SQL first, allowing earlier SQL projects
- **New running example of a university database**
 - more intuitive and motivating for students
 - illustrates more complex design tradeoffs
- **Revised ER notation more compatible with UML**
- **Update chapter on relational design**
 - more intuitive coverage followed by theory
 - examples illustrating real world issues in database design
- **Expanded coverage of Web applications and security**
- **Improved coverage of indexing and query optimization**
- **Revised coverage of transaction management**
 - expanded coverage of basics in chapter 14
 - detailed coverage of snapshot isolation
 - revised and simplified coverage of recovery
- **Expanded coverage of parallel and distributed databases**
 - database processing on multicore processors
 - cloud-based data storage systems
- **All chapters have been revised to reflect the latest advances in technology**
- **Updated website db-book.com**
 - new material for laboratory exercises
 - online appendices, slides for all chapters
 - solutions for exercises

Mc
Graw
Hill
Education

Higher
Education

ISBN-13: 978-93-3290-138-4
ISBN-10: 93-3290-138-4



9 789332 901384

6e Database System Concepts

Silberschatz
Korth
Sudarshan

Mc
Graw
Hill
Education

INDIAN
EDITION



Database System Concepts

For sale in
India, Pakistan,
Nepal, Bangladesh,
Sri Lanka and
Bhutan only.

Abraham Silberschatz
Henry F. Korth
S. Sudarshan

6e

274

COLLEGE OF AVIAN SCIENCES & MANAGEMENT
KVASU CAMPUS, THIRUVAZHAMKUNNU

Acc. No. 274

Call No.

This book should be returned on or before the
date last given below.

19 OCT 2018



E
IS

imited

Library TVK
005.74 SIL/D
Books
CASMTVK
274

McGraw Hill Education Offices
New Delhi New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto



McGraw Hill Education (India) Private Limited

DATABASE SYSTEM CONCEPTS, SIXTH EDITION

Copyright © 2011 by The McGraw Hill Companies, Inc. All rights reserved. Previous editions © 2006, 2002, and 1999. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of The McGraw-Hill Companies, Inc., including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

The Internet addresses listed in the text were accurate at the time of publication. The inclusion of a Web site does not indicate an endorsement by the authors of McGraw Hill, and McGraw Hill does not guarantee the accuracy of the information presented at these sites.

McGraw Hill Education (India) Edition 2013

Fourth reprint 2014
RYBLCRYUDQLAY

Reprinted in India by arrangement with The McGraw-Hill Companies, Inc., New York

Sales territories: India, Pakistan, Nepal, Bangladesh, Sri Lanka and Bhutan

Library of Congress Cataloging-in-Publication Data

Silberschatz, Abraham.

Database system concepts / Abraham Silberschatz. — 6th ed.
p. cm.

ISBN 978-0-07-352332-3 (alk. paper)

1. Database management. I. Title.

QA76.9.D3S5637 2011

005.74—dc22

2009039039

ISBN-13: 978-93-329-0138-4

ISBN-10: 93-329-0138-4

Published by McGraw Hill Education (India) Private Limited, P-24, Green Park Extension New Delhi 110 016,
and printed at Gopsons Papers Ltd., Noida.

*In memory of my father Joseph Silberschatz
my mother Vera Silberschatz
and my grandparents Stepha and Aaron Rosenblum*

Avi Silberschatz

*To my wife, Joan
my children, Abigail and Joseph
and my parents, Henry and Frances*

Hank Korth

*To my wife, Sita
my children, Madhur and Advait
and my mother, Indira*

S. Sudarshan



Contents

Chapter 1 Introduction

- | | | | |
|----------------------------------|----|--|----|
| 1.1 Database-System Applications | 1 | 1.10 Data Mining and Information Retrieval | 25 |
| 1.2 Purpose of Database Systems | 3 | 1.11 Specialty Databases | 26 |
| 1.3 View of Data | 6 | 1.12 Database Users and Administrators | 27 |
| 1.4 Database Languages | 9 | 1.13 History of Database Systems | 29 |
| 1.5 Relational Databases | 12 | 1.14 Summary | 31 |
| 1.6 Database Design | 15 | Exercises | 33 |
| 1.7 Data Storage and Querying | 20 | Bibliographical Notes | 35 |
| 1.8 Transaction Management | 22 | | |
| 1.9 Database Architecture | 23 | | |

PART ONE ■ RELATIONAL DATABASES

Chapter 2 Introduction to the Relational Model

- | | | | |
|---------------------------------------|----|---------------------------|----|
| 2.1 Structure of Relational Databases | 39 | 2.6 Relational Operations | 48 |
| 2.2 Database Schema | 42 | 2.7 Summary | 52 |
| 2.3 Keys | 45 | Exercises | 53 |
| 2.4 Schema Diagrams | 46 | Bibliographical Notes | 55 |
| 2.5 Relational Query Languages | 47 | | |

Chapter 3 Introduction to SQL

- | | | | |
|--|----|----------------------------------|-----|
| 3.1 Overview of the SQL Query Language | 57 | 3.7 Aggregate Functions | 84 |
| 3.2 SQL Data Definition | 58 | 3.8 Nested Subqueries | 90 |
| 3.3 Basic Structure of SQL Queries | 63 | 3.9 Modification of the Database | 98 |
| 3.4 Additional Basic Operations | 74 | 3.10 Summary | 104 |
| 3.5 Set Operations | 79 | Exercises | 105 |
| 3.6 Null Values | 83 | Bibliographical Notes | 112 |

Chapter 4 Intermediate SQL

- | | |
|------------------------------------|---------------------------|
| 4.1 Join Expressions 113 | 4.6 Authorization 143 |
| 4.2 Views 120 | 4.7 Summary 150 |
| 4.3 Transactions 127 | Exercises 152 |
| 4.4 Integrity Constraints 128 | Bibliographical Notes 156 |
| 4.5 SQL Data Types and Schemas 136 | |

Chapter 5 Advanced SQL

- | | |
|---|---|
| 5.1 Accessing SQL From a Programming Language 157 | 5.5 Advanced Aggregation Features** 192 |
| 5.2 Functions and Procedures 173 | 5.6 OLAP** 197 |
| 5.3 Triggers 180 | 5.7 Summary 209 |
| 5.4 Recursive Queries** 187 | Exercises 211 |
| | Bibliographical Notes 216 |

Chapter 6 Formal Relational Query Languages

- | | |
|--|---------------------------|
| 6.1 The Relational Algebra 217 | 6.4 Summary 248 |
| 6.2 The Tuple Relational Calculus 239 | Exercises 249 |
| 6.3 The Domain Relational Calculus 245 | Bibliographical Notes 254 |

PART TWO ■ DATABASE DESIGN**Chapter 7 Database Design and the E-R Model**

- | | |
|--|---|
| 7.1 Overview of the Design Process 259 | 7.8 Extended E-R Features 295 |
| 7.2 The Entity-Relationship Model 262 | 7.9 Alternative Notations for Modeling Data 304 |
| 7.3 Constraints 269 | 7.10 Other Aspects of Database Design 310 |
| 7.4 Removing Redundant Attributes in Entity Sets 272 | 7.11 Summary 313 |
| 7.5 Entity-Relationship Diagrams 274 | Exercises 315 |
| 7.6 Reduction to Relational Schemas 283 | Bibliographical Notes 321 |
| 7.7 Entity-Relationship Design Issues 290 | |

Chapter 8 Relational Database Design

- | | |
|---|--|
| 8.1 Features of Good Relational Designs 323 | 8.6 Decomposition Using Multivalued Dependencies 355 |
| 8.2 Atomic Domains and First Normal Form 327 | 8.7 More Normal Forms 360 |
| 8.3 Decomposition Using Functional Dependencies 329 | 8.8 Database-Design Process 361 |
| 8.4 Functional-Dependency Theory 338 | 8.9 Modeling Temporal Data 364 |
| 8.5 Algorithms for Decomposition 348 | 8.10 Summary 367 |
| | Exercises 368 |
| | Bibliographical Notes 374 |

Chapter 9 Application Design and Development

- | | |
|--|---|
| 9.1 Application Programs and User Interfaces 375 | 9.6 Application Performance 400 |
| 9.2 Web Fundamentals 377 | 9.7 Application Security 402 |
| 9.3 Servlets and JSP 383 | 9.8 Encryption and Its Applications 411 |
| 9.4 Application Architectures 391 | 9.9 Summary 417 |
| 9.5 Rapid Application Development 396 | Exercises 419 |
| | Bibliographical Notes 426 |

PART THREE ■ DATA STORAGE AND QUERYING**Chapter 10 Storage and File Structure**

- | | |
|---|---|
| 10.1 Overview of Physical Storage Media 429 | 10.6 Organization of Records in Files 457 |
| 10.2 Magnetic Disk and Flash Storage 432 | 10.7 Data-Dictionary Storage 462 |
| 10.3 RAID 441 | 10.8 Database Buffer 464 |
| 10.4 Tertiary Storage 449 | 10.9 Summary 468 |
| 10.5 File Organization 451 | Exercises 470 |
| | Bibliographical Notes 473 |

Chapter 11 Indexing and Hashing

- | | |
|------------------------------|---|
| 11.1 Basic Concepts 475 | 11.8 Comparison of Ordered Indexing and Hashing 523 |
| 11.2 Ordered Indices 476 | 11.9 Bitmap Indices 524 |
| 11.3 B+-Tree Index Files 485 | 11.10 Index Definition in SQL 528 |
| 11.4 B+-Tree Extensions 500 | 11.11 Summary 529 |
| 11.5 Multiple-Key Access 506 | Exercises 532 |
| 11.6 Static Hashing 509 | Bibliographical Notes 536 |
| 11.7 Dynamic Hashing 515 | |

Chapter 12 Query Processing

- 12.1 Overview 537
- 12.2 Measures of Query Cost 540
- 12.3 Selection Operation 541
- 12.4 Sorting 546
- 12.5 Join Operation 549
- 12.6 Other Operations 563
- 12.7 Evaluation of Expressions 567
- 12.8 Summary 572
 - Exercises 574
 - Bibliographical Notes 577

Chapter 13 Query Optimization

- 13.1 Overview 579
- 13.2 Transformation of Relational Expressions 582
- 13.3 Estimating Statistics of Expression Results 590
- 13.4 Choice of Evaluation Plans 598
- 13.5 Materialized Views** 607
- 13.6 Advanced Topics in Query Optimization** 612
- 13.7 Summary 615
 - Exercises 617
 - Bibliographical Notes 622

PART FOUR ■ TRANSACTION MANAGEMENT**Chapter 14 Transactions**

- 14.1 Transaction Concept 627
- 14.2 A Simple Transaction Model 629
- 14.3 Storage Structure 632
- 14.4 Transaction Atomicity and Durability 633
- 14.5 Transaction Isolation 635
- 14.6 Serializability 641
- 14.7 Transaction Isolation and Atomicity 646
- 14.8 Transaction Isolation Levels 648
- 14.9 Implementation of Isolation Levels 650
- 14.10 Transactions as SQL Statements 653
- 14.11 Summary 655
 - Exercises 657
 - Bibliographical Notes 660

Chapter 15 Concurrency Control

- 15.1 Lock-Based Protocols 661
- 15.2 Deadlock Handling 674
- 15.3 Multiple Granularity 679
- 15.4 Timestamp-Based Protocols 682
- 15.5 Validation-Based Protocols 686
- 15.6 Multiversion Schemes 689
- 15.7 Snapshot Isolation 692
- 15.8 Insert Operations, Delete Operations, and Predicate Reads 697
- 15.9 Weak Levels of Consistency in Practice 701
- 15.10 Concurrency in Index Structures** 704
- 15.11 Summary 708
 - Exercises 712
 - Bibliographical Notes 718

Chapter 16 Recovery System

- 16.1 Failure Classification 721
- 16.2 Storage 722
- 16.3 Recovery and Atomicity 726
- 16.4 Recovery Algorithm 735
- 16.5 Buffer Management 738
- 16.6 Failure with Loss of Nonvolatile Storage 743
- 16.7 Early Lock Release and Logical Undo Operations 744
- 16.8 ARIES** 750
- 16.9 Remote Backup Systems 756
- 16.10 Summary 759
 - Exercises 762
 - Bibliographical Notes 766

PART FIVE ■ SYSTEM ARCHITECTURE**Chapter 17 Database-System Architectures**

- 17.1 Centralized and Client-Server Architectures 769
- 17.2 Server System Architectures 772
- 17.3 Parallel Systems 777
- 17.4 Distributed Systems 784
- 17.5 Network Types 788
- 17.6 Summary 791
 - Exercises 793
 - Bibliographical Notes 794

Chapter 18 Parallel Databases

- 18.1 Introduction 797
- 18.2 I/O Parallelism 798
- 18.3 Interquery Parallelism 802
- 18.4 Intraquery Parallelism 803
- 18.5 Intraoperation Parallelism 804
- 18.6 Interoperation Parallelism 813
- 18.7 Query Optimization 814
- 18.8 Design of Parallel Systems 815
- 18.9 Parallelism on Multicore Processors 817
- 18.10 Summary 819
 - Exercises 821
 - Bibliographical Notes 824

Chapter 19 Distributed Databases

- 19.1 Homogeneous and Heterogeneous Databases 825
- 19.2 Distributed Data Storage 826
- 19.3 Distributed Transactions 830
- 19.4 Commit Protocols 832
- 19.5 Concurrency Control in Distributed Databases 839
- 19.6 Availability 847
- 19.7 Distributed Query Processing 854
- 19.8 Heterogeneous Distributed Databases 857
- 19.9 Cloud-Based Databases 861
- 19.10 Directory Systems 870
- 19.11 Summary 875
 - Exercises 879
 - Bibliographical Notes 883

PART SIX ■ DATA WAREHOUSING, DATA MINING, AND INFORMATION RETRIEVAL

Chapter 20 Data Warehousing and Mining

- | | |
|--------------------------------------|-------------------------------------|
| 20.1 Decision-Support Systems 887 | 20.7 Clustering 907 |
| 20.2 Data Warehousing 889 | 20.8 Other Forms of Data Mining 908 |
| 20.3 Data Mining 893 | 20.9 Summary 909 |
| 20.4 Classification 894 | Exercises 911 |
| 20.5 Association Rules 904 | Bibliographical Notes 914 |
| 20.6 Other Types of Associations 906 | |

Chapter 21 Information Retrieval

- | | |
|---|---|
| 21.1 Overview 915 | 21.7 Crawling and Indexing the Web 930 |
| 21.2 Relevance Ranking Using Terms 917 | 21.8 Information Retrieval: Beyond Ranking of Pages 931 |
| 21.3 Relevance Using Hyperlinks 920 | 21.9 Directories and Categories 935 |
| 21.4 Synonyms, Homonyms, and Ontologies 925 | 21.10 Summary 937 |
| 21.5 Indexing of Documents 927 | Exercises 939 |
| 21.6 Measuring Retrieval Effectiveness 929 | Bibliographical Notes 941 |

PART SEVEN ■ SPECIALTY DATABASES

Chapter 22 Object-Based Databases

- | | |
|---|--|
| 22.1 Overview 945 | 22.8 Persistent Programming Languages 964 |
| 22.2 Complex Data Types 946 | 22.9 Object-Relational Mapping 973 |
| 22.3 Structured Types and Inheritance in SQL 949 | 22.10 Object-Oriented versus Object-Relational 973 |
| 22.4 Table Inheritance 954 | 22.11 Summary 975 |
| 22.5 Array and Multiset Types in SQL 956 | Exercises 976 |
| 22.6 Object-Identity and Reference Types in SQL 961 | Bibliographical Notes 980 |
| 22.7 Implementing O-R Features 963 | |

Chapter 23 XML

- | | |
|---|-------------------------------|
| 23.1 Motivation 981 | 23.6 Storage of XML Data 1009 |
| 23.2 Structure of XML Data 986 | 23.7 XML Applications 1016 |
| 23.3 XML Document Schema 990 | 23.8 Summary 1019 |
| 23.4 Querying and Transformation 998 | Exercises 1021 |
| 23.5 Application Program Interfaces to XML 1008 | Bibliographical Notes 1024 |

PART EIGHT ■ ADVANCED TOPICS

Chapter 24 Advanced Application Development

- | | |
|---|----------------------------|
| 24.1 Performance Tuning 1029 | 24.4 Standardization 1051 |
| 24.2 Performance Benchmarks 1045 | 24.5 Summary 1056 |
| 24.3 Other Issues in Application Development 1048 | Exercises 1057 |
| | Bibliographical Notes 1059 |

Chapter 25 Spatial and Temporal Data and Mobility

- | | |
|---------------------------------------|---|
| 25.1 Motivation 1061 | 25.5 Mobility and Personal Databases 1079 |
| 25.2 Time in Databases 1062 | 25.6 Summary 1085 |
| 25.3 Spatial and Geographic Data 1064 | Exercises 1087 |
| 25.4 Multimedia Databases 1076 | Bibliographical Notes 1089 |

Chapter 26 Advanced Transaction Processing

- | | |
|---|--------------------------------------|
| 26.1 Transaction-Processing Monitors 1091 | 26.6 Long-Duration Transactions 1109 |
| 26.2 Transactional Workflows 1096 | 26.7 Summary 1115 |
| 26.3 E-Commerce 1102 | Exercises 1117 |
| 26.4 Main-Memory Databases 1105 | Bibliographical Notes 1119 |
| 26.5 Real-Time Transaction Systems 1108 | |

PART NINE ■ CASE STUDIES

Chapter 27 PostgreSQL

- | | |
|--|---|
| 27.1 Introduction 1123 | 27.5 Storage and Indexing 1146 |
| 27.2 User Interfaces 1124 | 27.6 Query Processing and Optimization 1151 |
| 27.3 SQL Variations and Extensions 1126 | 27.7 System Architecture 1154 |
| 27.4 Transaction Management in PostgreSQL 1137 | Bibliographical Notes 1155 |

Chapter 28 Oracle

- | | |
|--|--|
| 28.1 Database Design and Querying Tools 1157 | 28.6 System Architecture 1183 |
| 28.2 SQL Variations and Extensions 1158 | 28.7 Replication, Distribution, and External Data 1188 |
| 28.3 Storage and Indexing 1162 | 28.8 Database Administration Tools 1189 |
| 28.4 Query Processing and Optimization 1172 | 28.9 Data Mining 1191 |
| 28.5 Concurrency Control and Recovery 1180 | Bibliographical Notes 1191 |

Chapter 29 IBM DB2 Universal Database

- | | |
|---|---|
| 29.1 Overview 1193 | 29.9 Tools and Utilities 1215 |
| 29.2 Database-Design Tools 1194 | 29.10 Concurrency Control and Recovery 1217 |
| 29.3 SQL Variations and Extensions 1195 | 29.11 System Architecture 1219 |
| 29.4 Storage and Indexing 1200 | 29.12 Replication, Distribution, and External Data 1220 |
| 29.5 Multidimensional Clustering 1203 | 29.13 Business Intelligence Features 1221 |
| 29.6 Query Processing and Optimization 1207 | Bibliographical Notes 1222 |
| 29.7 Materialized Query Tables 1212 | |
| 29.8 Autonomic Features in DB2 1214 | |

Chapter 30 Microsoft SQL Server

- | | |
|--|--|
| 30.1 Management, Design, and Querying Tools 1223 | 30.8 Distributed Heterogeneous Query Processing 1250 |
| 30.2 SQL Variations and Extensions 1228 | 30.9 Replication 1251 |
| 30.3 Storage and Indexing 1233 | 30.10 Server Programming in .NET 1253 |
| 30.4 Query Processing and Optimization 1236 | 30.11 XML Support 1258 |
| 30.5 Concurrency and Recovery 1241 | 30.12 SQL Server Service Broker 1261 |
| 30.6 System Architecture 1246 | 30.13 Business Intelligence 1263 |
| 30.7 Data Access 1248 | Bibliographical Notes 1267 |

PART TEN ■ APPENDICES**Appendix A Detailed University Schema**

- | | |
|----------------------|----------------------|
| A.1 Full Schema 1271 | A.3 Sample Data 1276 |
| A.2 DDL 1272 | |

Appendix B Advanced Relational Design (contents online)

- | | |
|---------------------------------|---------------------------|
| B.1 Multivalued Dependencies B1 | Exercises B10 |
| B.3 Domain-Key Normal Form B8 | Bibliographical Notes B12 |
| B.4 Summary B10 | |

Appendix C Other Relational Query Languages (contents online)

- | | |
|-------------------------|---------------------------|
| C.1 Query-by-Example C1 | C.4 Summary C25 |
| C.2 Microsoft Access C9 | Exercises C26 |
| C.3 Datalog C11 | Bibliographical Notes C30 |

Appendix D Network Model (contents online)

- | | |
|--------------------------------------|--------------------------------------|
| D.1 Basic Concepts D1 | D.6 DBTG Set-Processing Facility D22 |
| D.2 Data-Structure Diagrams D2 | D.7 Mapping of Networks to Files D27 |
| D.3 The DBTG CODASYL Model D7 | D.8 Summary D31 |
| D.4 DBTG Data-Retrieval Facility D13 | Exercises D32 |
| D.5 DBTG Update Facility D20 | Bibliographical Notes D35 |

Appendix E Hierarchical Model (contents online)

- | | |
|---------------------------------|---|
| E.1 Basic Concepts E1 | E.6 Mapping of Hierarchies to Files E22 |
| E.2 Tree-Structure Diagrams E2 | E.7 The IMS Database System E24 |
| E.3 Data-Retrieval Facility E13 | E.8 Summary E25 |
| E.4 Update Facility E17 | Exercises E26 |
| E.5 Virtual Records E20 | Bibliographical Notes E29 |

Bibliography 1283**Index 1315**

- [X/Open 1991] X/Open Snapshot: X/Open DTP: XA Interface. X/Open Company, Ltd. (1991).
- [Yan and Larson 1995] W. P. Yan and P. A. Larson, "Eager Aggregation and Lazy Aggregation", In *Proc. of the International Conf. on Very Large Databases* (1995).
- [Yannakakis et al. 1979] M. Yannakakis, C. H. Papadimitriou, and H. T. Kung, "Locking Protocols: Safety and Freedom from Deadlock", In *Proc. of the IEEE Symposium on the Foundations of Computer Science* (1979), pages 286-297.
- [Zaharioudakis et al. 2000] M. Zaharioudakis, R. Cochrane, G. Lapis, H. Pirahesh, and M. Urata, "Answering Complex SQL Queries using Automatic Summary Tables", In *Proc. of the ACM SIGMOD Conf. on Management of Data* (2000), pages 105-116.
- [Zeller and Gray 1990] H. Zeller and J. Gray, "An Adaptive Hash Join Algorithm for Multiuser Environments", In *Proc. of the International Conf. on Very Large Databases* (1990), pages 186-197.
- [Zhang et al. 1996] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: An Efficient Data Clustering Method for Very Large Databases", In *Proc. of the ACM SIGMOD Conf. on Management of Data* (1996), pages 103-114.
- [Zhou and Ross 2004] J. Zhou and K. A. Ross, "Buffering Database Operations for Enhanced Instruction Cache Performance", In *Proc. of the ACM SIGMOD Conf. on Management of Data* (2004), pages 191-202.
- [Zhuge et al. 1995] Y. Zhuge, H. Garcia-Molina, J. Hammer, and J. Widom, "View maintenance in a warehousing environment", In *Proc. of the ACM SIGMOD Conf. on Management of Data* (1995), pages 316-327.
- [Ziauddin et al. 2008] M. Ziauddin, D. Das, H. Su, Y. Zhu, and K. Yagoub, "Optimizer plan change management: improved stability and performance in Oracle 11g", *Proceedings of the VLDB Endowment*, Volume 1, Number 2 (2008), pages 1346-1355.
- [Zikopoulos et al. 2004] P. Zikopoulos, G. Baklarz, D. deRoos, and R. B. Melnyk, *DB2 Version 8: The Official Guide*, IBM Press (2004).
- [Zikopoulos et al. 2007] P. Zikopoulos, G. Baklarz, L. Katsnelson, and C. Eaton, *IBM DB2 9 New Features*, McGraw Hill (2007).
- [Zikopoulos et al. 2009] P. Zikopoulos, B. Tassi, G. Baklarz, and C. Eaton, *Break Free with DB2 9.7*, McGraw Hill (2009).
- [Zilio et al. 2004] D. C. Zilio, J. Rao, S. Lightstone, G. M. Lohman, A. Storm, C. Garcia-Arellano, and S. Fadden, "DB2 Design Advisor: Integrated Automatic Physical Database Design", In *Proc. of the International Conf. on Very Large Databases* (2004), pages 1087-1097.
- [Zloof 1977] M. M. Zloof, "Query-by-Example: A Data Base Language", *IBM Systems Journal*, Volume 16, Number 4 (1977), pages 324-343.

Index

- 2PC. See two-phase commit
 3NF. See third normal form
 3PC. See three-phase commit
- abstract data types, 1127
 access paths, 542
 ACID properties. See
 atomicity; consistency;
 durability; isolation
 Active Server Pages (ASP), 397
 ADO.NET, 169, 395, 1249, 1253
 Advanced Encryption Standard
 (AES), 412-413
 agglomerative clustering,
 907-908
 aggregate functions
 basic, 85-86
 Boolean values and, 89-90
 SQL, 84
 fusion, 960
 with grouping, 86-88
 having clause, 88-89
 null values and, 89-90
 aggregation
 advanced features of, 192-197
 alternative notations for,
 304-310
 entity-relationship (E-R)
 model and, 301-302, 304
 IBM DB2 and, 1209-1210
 intraoperation parallelism
 and, 811
 .NET Common Language
 Runtime (CLR) and,
 1257-1258
 OLAP and, 197-209
 PostgreSQL and, 1153
- query optimization and, 597
 query processing and, 566-567
 ranking and, 192-195
 relational algebra and,
 235-239
 representation of, 304
 view maintenance and,
 610-611
 windowing, 195-197
 Ajax, 390-391, 398, 867
 aliases, 75, 355, 829, 872-873,
 1229
 alter table, 63, 129
 alter trigger, 185
 alter type, 140
 American National Standards
 Institute (ANSI), 57, 1051
 analysis pass, 753
 analytic workspaces, 1161
 and operation, 66, 83-84, 1174
 any keyword, 92n8
 Apache, 386, 399, 426, 980
 Apple Macintosh OS X, 1124
 application design, 418
 application architectures and,
 391-396
 authentication and, 405-407
 business-logic layer and,
 391-392
 client-server architecture and,
 376-377
 common gateway interface
 (CGI), 380-381
 cookies and, 382-385
 data access layer and, 391,
 393, 395
- disconnected operation and,
 395-396
 encryption and, 411-417
 HyperText Markup Language
 (HTML) and, 378-380
 HyperText Transfer Protocol
 (HTTP) and, 377-381, 383,
 395, 404-406, 417
 Java Server Pages (JSP) and,
 377, 383-391
 performance and, 400-402
 rapid application
 development and,
 396-400
 security and, 402-417
 servlets and, 383-391
 three-layer architecture and,
 318
 TP-monitors and, 1095-1096
 uniform resource locators
 (URLs), 377-378
 user interfaces and, 375-377
 World Wide Web and, 377-382
- application development
 performance benchmarks
 and, 1045-1048
 performance tuning and,
 1029-1045
 set orientation and, 1030-1031
 standardization and,
 1051-1056
 testing applications and,
 1048-1051
 updates and, 1030-1033
 application migration,
 1050-1051

application program interface (API)

ADO.NET, 169, 1054
 application design and, 383-386, 395
 C++, 1054
 customized maps and, 1068, 1070
 DOM, 1020
 IBM DB2, 1196
 Java, 158-166, 213, 383-386, 1018, 1030
 LDAP, 874
 Microsoft SQL Server, 1229, 1245, 1248-1250, 1253-1255, 1265, 1267
 ODBC, 166-169, 1053
 PostgreSQL, 1125
 SAX, 1020
 standards for, 1051-1056
 system architectures and, 772
 XML, 985, 1008-1009

apply operator, 1230-1231

architectures, 767
 business logic and, 25, 1158, 1221-1222, 1228, 1232, 1253, 1263-1267
 centralized, 769-771
 client-server, 771-772
 cloud-based, 777
 data server, 773, 775-777
 data warehouse, 889-891
 distributed systems, 784-788
 hierarchical, 781, 784
 hypercube, 781
 mesh, 780-781
 network types and, 788-791
 parallel databases, 797-820
 parallel systems, 777-784
 process monitor, 774
 server system, 772-777
 shared-disk, 781, 783, 789
 shared memory, 781-783
 shared-nothing, 781, 783-784, 803
 shared server, 1185
 single-user system, 770-771
 source-driven, 890
 storage-area network (SAN), 789
 thread pooling and, 1246
 three-tier, 25
 TP-monitor, 1092-1095
 transaction-server, 773-775

two-phase commit protocol (2PC), 786-788
 two-tier, 24-25
 wide-area networks (WANs), 788, 790-791

ARIES, 1146

analysis pass, 753
 compensation log records (CLRs), 751-752, 754
 data structure, 751-753
 dirty page table, 750-755
 fine-grained locking, 756
 fuzzy checkpoints, 750-752
 log sequence number (LSN), 750-755
 nested top actions, 755-756
 optimization and, 756
 physiological redo and, 750
 recovery and, 751, 753, 756
 redo pass, 754
 savepoints and, 756
 transaction rollback and, 754-755
 undo pass, 754-755

Arpanet, 790**array types, 956-961****asc expression, 77-78****as clause, 75-76****ASP.NET, 387****assertions, 11, 135-136****assignment operation, 176, 217, 232, 1052****associations, 17, 43**

entity sets and, 290 (see also entity sets)
 relationship sets and, 264-267, 308-309, 314
 rules for, 904-907

associative property, 584-585**Aster Data, 816****asymmetric-key encryption, 412****atomic domains, 42**

first normal form and, 327-329
 object-based databases and, 947

relational database design and, 327-329**atomicity, 4-5, 22-23, 104, 625**

cascadeless schedules and, 647-648

commit protocols and, 832-838

defined, 628

two-phase commit protocol (2PC), 786-788

distributed transactions and, 830-832
 isolation and, 646-648
 log records and, 726-728, 730-734

recoverable schedules and, 647

recovery systems and, 726-735
 storage structure and, 632-633
 workflows and, 1099-1100

attribute inheritance, 298-299**attributes**

atomic domains and, 327-329
 classifiers and, 896-897
 closure of attribute sets and, 340-342
 complex, 277-278, 284-285
 composite, 267
 continuous valued, 898
 decomposition and, 329-338, 348-360
 derived, 268
 design issues and, 290-291
 domain, 267
 entity-relationship diagrams and, 277-278
 entity-relationship (E-R) model and, 263, 267-269
 entity sets and, 283-286, 290-291
 multiple-key access and, 506-509
 multivalued, 267-268, 327-329
 naming of, 362-363
 nesting, 958-961
 null values and, 268-269
 partitioned, 896-897
 placement of, 294-295
 search key and, 476
 simple, 267, 283-284
 single-valued, 267-268
 Unified Modeling Language (UML) and, 308-310
 unifiers and, 498-499
 unnesting, 958-961
 value set of, 267
 xmlattributes and, 1015
 XML types and, 990-998

attribute-value skew, 800-801

audit trails, 409-410

augmentation rule, 339

authentication
 challenge-response system and, 415

digital certificates and, 416-417
 digital signatures and, 416-417
 encryption and, 415-417
 security and, 405-407 (see also security)

single sign-on system and, 406-407

smart cards and, 415-416
 two-factor, 405-407

authorization, 11, 21, 58

administrators and, 143
 application-level, 407-409
 check clause, 148
 database design and, 312
 end-user information and, 407-408
 granting/revoking privileges, 29, 143-145, 149-150
 lack of fine-grained, 408-409
 roles and, 145-146
 schema and, 147-148
 Security Assertion Markup Language (SAML) and, 407
 sql security invoker, 147
 transfer of privileges, 148-149
 update and, 147, 148
 views and, 146-147

autonomy, 858**availability**

CAP theorem and, 852-853
 consistency and, 852-853
 coordinator selection and, 850-852
 majority-based approach and, 848-849
 read one, write all approach, 849-850
 remote backup and, 850
 robustness and, 847
 site reintegration and, 850
 sequential records and, 524-525
 structure of, 525-527
 blind writes, 687
 blobs, 138, 166, 457, 502, 1013, 1198-1199, 1259
 block-interleaved parity organization, 445-446
 block-level striping, 443-4
 block nested-loop join, 551-552
 Boolean operations, 83, 89-90, 94, 161, 176, 873, 1256.

backup, 186-187

application design and, 415
 distributed databases and, 839, 877

IBM DB2 and, 1215, 1218
 Microsoft SQL Server and, 1223, 1227-1228, 1245, 1262

Oracle and, 1165, 1169, 1181-1183, 1187-1190
 parallel databases and, 816
 recovery systems and, 726-738 (see also recovery systems)

remote systems for, 756-759, 850, 1095-1096

system architectures and, 770
 transactions and, 632, 1115

backup coordinator, 851**balanced range-partitioning, 801****balanced tree, 486****BASE properties, 853****batch scaleup, 779****batch update, 1030-1031****Bayesian classifiers, 900-901****Bayes' theorem, 900****BCNF. See Boyce-Codd normal form**

begin atomic...end, 128, 176, 181, 183-185

best splits, 897-899

biased protocol, 841

big-bang approach, 1050-1051

Bigtable, 862-867

binary splits, 898

bit-interleaved parity organization, 445

bit-level striping, 442-444

bitmap indices, 507, 509, 531, 536

B+-trees and, 528

efficient operations of, 527-528

existence, 526-527

Oracle and, 1166-1167

scans, 1153

sequential records and, 524-525

structure of, 525-527

blind writes, 687

blobs, 138, 166, 457, 502, 1013, 1198-1199, 1259

block-interleaved parity organization, 445-446

block-level striping, 443-4

block nested-loop join, 551-552

Boolean operations, 83, 89-90, 94, 161, 176, 873, 1256.

See also specific operation**bottlenecks**

application design and, 402, 1029, 1033-1035, 1039
 file structure and, 468
 system architectures and, 782-783, 800, 816-819, 829, 839-840
 transactions and, 1106-1107, 1116

Boyce-Codd normal form (BCNF)

decomposition algorithms and, 349-356
 dependency preservation and, 334-336
 relational database design and, 333-336

broadcast data, 1082-1083**B-trees**

application development and, 1039

IBM DB2 and, 1205

indices and, 504-506, 530, 1039

Oracle and, 1159, 1164-1169, 1173

PostgreSQL and, 1135, 1148-1150

spatial data and, 1064, 1071-1072, 1076, 1086

B+-trees, 1075

balanced trees and, 486

bitmap indices and, 528

bulk loading of indices and, 503-504

deletion time and, 491, 495-497, 499-501

extensions and, 500-506

fanout and, 487

file organization and, 500-502

flash memory and, 506

indexing strings and, 502-503

insertion time and, 491-495, 499-501

on multiple keys, 508

nonleaf nodes of, 487

nonunique search keys and, 497-499

performance and, 485-486

pointers and, 486

queries on, 488-491, 538, 544

record relocation and, 502

secondary indices and, 502

structure of, 486-488

- updates and, 491-500
- buffer manager**, 21, 464-466
- buffer pools**, 1201-1202, 1220
- buffers**
 - database buffering and, 739-741
 - file structure and, 437-438
 - force/no-force policy and, 739-740
 - force output and, 725-726
 - IBM DB2 and, 1200-1203
 - log-record buffering and, 738-739
 - main-memory databases and, 1105-1108
 - management of, 738-743
 - multiple pool, 1220
 - operating system role in, 741-742
 - recovery systems and, 738-743
 - replacement policies and, 465-468
 - steal/no-steal policy and, 740
 - transaction servers and, 774
 - write-ahead logging (WAL) rule and, 739-741
- bugs**, 174n4
 - application design and, 404, 1048-1050
 - recovery systems and, 721-722
 - system architectures and, 787-788
 - transactions and, 1093, 1102
 - workflows and, 1101
- build input**, 558
- bulk export**, 1032
- bulk insert**, 1032
- bulk loads**, 503-504, 1031-1033
- bully algorithm**, 851, 852
- business logic**, 25, 173
 - application design and, 376, 383, 391-393, 396, 410, 418
 - IBM DB2 and, 1221-1222
 - Microsoft SQL Server and, 1228, 1232, 1253, 1263-1267
 - Oracle and, 1158
- business-logic layer**, 39, 391
- bus system**, 780
- BYNET**, 806
- byte-code**, 389
- C, 7, 14**,
 - advanced SQL and, 157, 169, 173, 180
- application design and, 387, 397-398
 - Microsoft SQL Server and, 1228, 1253
 - ODBC and, 167-168
 - Unified Modeling Language (UML) and, 308
- C++, 7n1, 14**
 - advanced SQL and, 169, 173
 - Microsoft SQL Server and, 1253
 - object-based databases and, 945
 - persistent systems and, 968-971
 - standards for, 1054
 - Unified Modeling Language (UML) and, 308
- caching**, 429
 - application design and, 400-401
 - coherency and, 776
 - data servers and, 776
 - locks and, 776
 - multithreading and, 817-818
 - Oracle and, 1184
 - query plans and, 605, 775
 - shared-memory architecture and, 783
- call**, 175
- callable statements**, 164
- call back**, 776
- Call Level Interface (CLI)**
 - standards, 1053
- candidate keys**, 45-46
- canonical cover**, 342-345
- CAP theorem**, 852-853
- Cartesian products**, 68
 - equivalence and, 584
 - join expressions and, 229-232
 - queries and, 573, 584, 589, 595-596, 606, 616
 - relational algebra and, 50-51, 217, 222-232
 - SQL and, 68-75, 120, 209
- cascade**, 133, 150, 186
- cascadeless schedules**, 647-648
- cascading stylesheet (CSS)**
 - standard, 380
- case construct**, 102-103
- CASE tools**, 1194-1195
- cast**, 136, 139-140
- catalogs**
 - application development and, 1053

- distributed databases and, 830, 848
- e-catalogs and, 1103-1104
- IBM DB2 and, 1195, 1220
- indices and, 476 (see also indices)
- information retrieval and, 915, 935
- Microsoft SQL Server and, 1235-1236, 1250, 1256, 1266
- PostgreSQL and, 1151, 1154
- query optimization and, 590-596, 1151
- SQL and, 142-143, 165, 168-169
- system, 462, 468, 801, 1132
- transaction processing and, 1116
- XML and, 1017
- categories**, 935-937
- centralized architectures**, 769-771
- centralized deadlock detection**, 845-846
- challenge-response system**, 415
- change isolation level**, 649
- check clause**, 130, 134
 - assertion and, 135-136
 - authorization and, 148
 - data constraints and, 310
 - dependency preservation and, 334-336
 - user-defined types and, 140
- check constraints**, 134, 148, 310, 317, 334, 628, 1130
- checkpoint log record**, 752
- checkpoints**
 - fuzzy, 750-752
 - Microsoft SQL Server and, 1246
 - Oracle and, 1185
 - recovery systems and, 734-735, 742-743
 - transaction servers and, 774
- checksums**, 434
- classification hierarchy**, 935-937
- classifiers**
 - Bayesian, 900-901, 1191, 1266
 - best splits and, 897-899
 - decision-tree, 895-900
 - entropy measure and, 897
 - Gini measure and, 897
 - information gain and, 897-898
 - kernel functions and, 901-902

- neural-net, 900
- partitions and, 896-897
- prediction and, 894-904
- purity and, 897
- regression and, 902-903
- Support Vector Machine, 900-902
- training instances and, 895
- validating, 903-904
- client-server systems**, 23, 32, 204
 - application design and, 376-377, 1031, 1053, 1056
 - Microsoft SQL Server and, 1228
 - recovery systems and, 756
 - system architecture and, 769-772, 777, 788, 791
 - transaction processing and, 1092-1096
- client-side scripting**, 389-391
- clobs**, 138, 166, 457, 502, 1010-1013, 1196-1199
- cloud computing**, 777
 - challenges with, 868-870
 - data representation and, 863-865
 - partitions and, 865-866
 - replication and, 866-868
 - retrieval and, 865-866
 - storage and, 862-863
 - traditional databases and, 868
 - transactions and, 866-868
- clusters**, 781
 - agglomerative, 907-908
 - cloud computing and, 867
 - data mining and, 894, 907-908
 - divisive, 907-908
 - hierarchical, 907-908
 - IBM DB2 and, 1203-1207
 - multidimensional, 1203-1207
 - Oracle and, 1173, 1186
 - Real Application Clusters (RAC) and, 1186
- coalescence**, 491, 706
- code breaking**. See **encryption**
- ColdFusion Markup Language (CFML)**, 387
- collect function**, 959
- collection volumes**, 957, 957-958
- column-oriented storage**, 892-893
- combinatorics**, 639
- commit protocols**, 832-838
- commit time**, 758
- commit work**, 127
- common gateway interface (CGI) standard**, 380-381
- Common Language Runtime (CLR)**, 180
- Common Object Request Broker Architecture (CORBA)**, 1054-1055
- common subexpression elimination**, 614
- commutative property**, 584-585
- compatibility function**, 662
- compensation log records (CLRs)**, 751-752, 754
- complex data types**, 28, 31, 138, 1061
 - entity-relationship (E-R) model and, 946-947
 - keywords and, 947-949
 - normal forms and, 947
 - object-based databases and, 945-949, 963, 970-974
 - system architecture and, 864
- component object model (COM)**, 1253
- compression**, 1077-1078
 - application development and, 1041
 - data warehouses and, 893
 - IBM DB2 and, 1194
 - Microsoft SQL Server and, 1236
 - Oracle and, 1165-1167
 - prefix, 503, 1166
- computer-aided design (CAD)**, 312, 1061, 1064-1068
- conceptual design**, 15-16, 260
- concurrency control**, 631, 636, 639, 709
 - access anomalies and, 5
 - blind writes and, 687
 - consistency and, 695, 701-704, 710
 - deadlock handling and, 674-679
 - delete operations and, 697-701
 - distributed databases and, 835-836, 839-847
 - DML commands and, 1138-1139
 - false cycles and, 846-847
 - IBM DB2 and, 1200-1203, 1217-1218
 - index structures and, 704-708
 - insert operations and, 697-701

- locks and, 661-686, 839-842
- logical undo and, 749-750
- long-duration transactions and, 1111-1112
- Microsoft SQL Server and, 1241-1246
- multiple granularity and, 679-682
- multiversion schemes and, 689-692, 1137-1146
- Oracle and, 1180-1183
- PostgreSQL and, 1137-1145
- predicate reads and, 697-701
- recovery systems and, 729-730
- rollbacks and, 667, 670, 674-679, 685, 689, 691, 709
- serializability and, 650, 662, 666-667, 671, 673, 681-690, 693-697, 701-704, 708
- snapshot isolation and, 692-697, 729-730
- timestamp-based protocol and, 682-686
- updates and, 867-868
- user interactions and, 702-704
- validation and, 686-689, 729-730
- Web crawlers and, 930-931
- condition-defined entity sets**, 299
- confidence**, 893, 905
- conformance levels**, 168-169
- conjunction**, 545-546, 594
- connect to**, 170
- consistency**, 11, 22, 104
 - availability and, 852-853
 - CAP theorem and, 852-853
 - concurrency control and, 695, 701-704, 710
 - cursor stability and, 702
 - deadlocks and, 665-666
 - degree-two, 701-702
 - distributed transactions and, 830-832
 - logical operations and, 746
 - mobile networks and, 1083-1085
 - replication with weak, 843-844
 - requirement of, 630
 - transactions and, 627-631, 635-636, 640, 648-650, 655
 - user interactions and, 702-704
 - weak levels of, 701-704

- constraints
 - condition-defined, 299
 - decomposition and, 354-355
 - dependency preservation and, 334-336
 - disjoint, 300
 - entity-relationship (E-R) model and, 269-272, 299-301
 - IBM DB2 and, 1199
 - integrity, 4 (see also integrity constraints)
 - keys, 271-272
 - mapping cardinalities and, 269-270, 276-277
 - overlapping, 300
 - partial, 300
 - participation, 270
 - PostgreSQL and, 1130-1131, 1153-1154
 - total, 300
 - transactions and, 628, 1108-1109
 - UML and, 309-310
 - user-defined, 299
- contains, 93
- context switch, 1092
- conversation group, 1262
- cookies, 382-385, 403-405
- cores, 770
- correlated subqueries, 93
- correlations, 906
- correlation variables, 605-607
- count, 84-86, 89, 566-567
- count-distinct, 236
- count value, 204
- crashes, 434, 467-468
 - actions after, 736-738
 - algorithms for, 735-738
 - ARIES and, 750-756
 - checkpoints and, 734-735
 - failure classification and, 721-722
 - recovery systems and, 736-738 (see also recovery systems)
 - transactions and, 628
- create assertion, 135
- create distinct type, 141, 1194-1195, 1197
- create domain, 140-141
- create function, 175, 177, 189
- create index, 528-529, 1150-1151
- create or replace, 174n4
- create procedure, 175, 179
- create recursive view, 192
- create role, 146
- create schema, 143
- create snapshot, 843-844
- create table, 60-63, 139, 161
 - with data, 141-142
 - extensions for, 141-142
 - integrity constraints and, 129
 - object-based databases and, 950, 961-962
- create table...as, 142
- create type, 139-141
- create unique index, 529
- create view, 121-123, 125, 142, 147, 1130-1131
- cross-site request forgery (XSRF), 403-404
- cross-site scripting (XSS), 403-405
- cross-tabulation, 199-203, 205, 210
- CRUD Web interfaces, 399
- Crystal Reports, 399-400
- cube by, 1221-1222
- cube construct, 206-210
- current date, 137
- cursor stability, 702
- curve fitting, 902-903
- Cyc project, 925, 927
- data abstraction, 6-8
- data access layer, 391, 393, 395
- data analysis
 - data mining and, 893-910
 - decision-support systems and, 887-891
 - information retrieval and, 915-938
 - warehousing and, 887-891
- database design, 257, 313-314
 - adapting for modern architectures and, 818-819
 - alternatives and, 261-262, 304-310
 - application, 375-418
 - authorization requirements and, 312
 - automated tuning and, 1040-1041
 - bottom-up, 297
 - buffers and, 464-468
 - client-server architecture and, 32, 204, 376-377, 756, 766-767, 769-772, 777, 788, 791
 - complexity of, 260
 - conceptual-design phase and, 15-16, 260
 - data constraints and, 310-311
 - direct design process and, 259-260
 - encryption and, 411-417
 - entity-relationship (E-R) model and, 17-18, 249-313
 - IBM DB2 and, 1194-1195
 - incompleteness and, 262
 - logical-design phase and, 16, 260-261
 - Microsoft SQL Server and, 1223-1228
 - normalization and, 18-20
 - Oracle and, 1157-1158
 - overview of process, 259-262
 - parallel systems and, 815-817
 - phases of, 259-261
 - physical-design phase and, 16, 261
 - redundancy and, 261-262
 - relational, 323-368
 - specification of functional requirements and, 16
 - top-down, 297
 - universities and, 16-17
 - user needs and, 260
 - user requirements and, 311-312
 - workflow and, 312-313
- database graph, 671-674
- database instance, 42
- databases
 - abstraction and, 6-8, 10
 - architecture and, 28-29, 767 (see also architectures)
 - buffering and, 739-741 (see also buffers)
 - concurrency control and, 661-710 (see also concurrency control)
 - distributed, 825-878, 1188
 - dumping and, 743-744
 - file-processing system and, 3-6
 - force output and, 725-726
 - history of, 29-31
 - indexing and, 475-531 (see also indices)
 - information retrieval and, 915-937

- locks and, 661-679 (see also locks)
- main-memory, 1105-1108
- mobile, 1079-1085
- modification and, 98-103, 728-729
- multimedia, 1076-1079
- normalization and, 18-20
- parallel, 797-820
- personal, 1079-1085
- recovery systems and, 631, 633, 721-761 (see also recovery systems)
- storage and, 20-22, 427 (see also storage)
- time in, 1062-1064
- databases administrator (DBA), 28-29, 149, 1152, 1214-1215, 1243
- database schemas. *See* schemas
- Database Task Group, 1052
- database writer process, 773
- data cleansing, 890
- data cube, 200, 206-210
- data-definition language (DDL), 9-12, 14, 32
 - authorization and, 58
 - basic types and, 59-60
 - concurrency control and, 1144-145
 - dumping and, 743-744
 - IBM DB2 and, 1194-1197, 1204
 - indices and, 58
 - integrity and, 58
 - Microsoft SQL Server and, 1225, 1228-1233, 1245, 1253, 1256, 1261
 - Oracle and, 1162, 1181
 - PostgreSQL and, 1144-1145, 1150
 - querying and, 21-22
 - schema definition and, 28, 58, 60-63
 - security and, 58
 - set of relations in, 58-61
 - SQL basics and, 57-63, 104
 - storage and, 58
- data dictionary, 12, 21, 462-464
- Data Encryption Standard (DES), 413
- DataGrid control, 398
- data guard, 1183
- data inconsistency. *See* consistency
- data isolation. *See* isolation

- DATALlegro, 816
- Datalog, 37
- data-manipulation language (DML), 12-14
 - authorization and, 143
 - compiler and, 21-22
 - concurrency control and, 1138-1139, 1145
 - declarative, 10
 - defined, 10, 32
 - host language and, 15
 - Microsoft SQL Server and, 1231-1233, 1245, 1261
 - Oracle and, 1161-1162, 1165, 1181
 - PostgreSQL and, 1137-1138
 - precompiler and, 15
 - procedural/nonprocedural, 10
 - querying and, 21-22
 - snapshot isolation and, 1137 (see also snapshot isolation)
 - storage manager and, 20-21
 - triggers and, 1161-1162
- data mediation, 1018-1019
- data mining, 25-26, 33, 771-772, 887-889
 - association rules and, 904-907
 - best splits and, 897-899
 - classification and, 894-904
 - clusters and, 894, 907-908
 - data-visualization, 909
 - defined, 893
 - descriptive patterns and, 894
 - entropy measure and, 897
 - Gini measure and, 897
 - information gain and, 897-898
 - Microsoft SQL Server and, 1266-1267
 - Oracle and, 1191
 - prediction and, 894-904
 - purity and, 897
 - rules for, 893-894
 - text, 908
- data models. *See* specific model
- data parallelism, 805
- data server systems, 773, 775-777, 782
- data storage and definition language, 11
- data striping, 444
- data-transfer rate, 435-436
- data types. *See* types
- data warehouses, 888
 - column-oriented storage and, 892-893
 - components of, 889-891
 - deduplication and, 890-891
 - defined, 889
 - fact tables and, 891-892
 - householding and, 891
 - IBM DB2 and, 1194, 1221-1222
 - materialized views and, 1171-1172
 - merger-purge operation and, 890-891
 - Microsoft SQL Server and, 1264
 - Oracle and, 1158, 1162, 1169-1172, 1178, 1189
 - transformations and, 891
 - updates and, 891
- Data Encryption Standard (DES), 413
- datetime, 201
- DB-Lib, 1249
- deadlines, 1108-1109
- deadlocks
 - consistency and, 665-666
 - distributed databases and, 839, 841, 844-847
 - handling of, 674-679
 - IBM DB2 and, 1217-1220
 - long-duration transactions and, 1110-1111
 - Microsoft SQL Server and, 1243-1244, 1246
 - PostgreSQL and, 1143-1145
 - prevention of, 675-676
 - rollback and, 678-679
 - starvation and, 679
 - victim selection and, 678
 - wait-for graph and, 676-677, 676-678
- decision support, 1047
- decision-support queries, 797
- decision-support systems, 887-891
- decision-tree classifiers, 895-900
- declare, 175-178
- decode, 208
- decomposition
 - algorithms for, 348-355
 - Boyce-Codd normal form and, 333-336, 349-355
 - dependency preservation and, 334-336

fourth normal form and, 358-360
 functional dependencies and, 329-338, 355-360
 higher normal forms and, 337-338
 keys and, 330-333
 lossless, 345-346
 lossless-join, 345-346
 lossy, 345-346
 multivalued dependencies and, 355-360
 relational database design and, 329-338, 348-360
 third normal form and, 336-337, 352-355
decomposition rule, 339
DEC Rdb, 30
deduplication, 890-891
Deep Web crawlers, 931
default values, 133, 137, 140, 144, 425, 899, 952, 991-992, 996, 1128
deferred integrity constraints, 134
degree-two consistency, 701-702
deletion, 61, 63, 98-100, 102, 161
 concurrency control and, 697-701
 EXEC SQL and, 171
 hashing and, 510, 513, 516, 523
 integrity constraints and, 133
 PostgreSQL and, 1130-1131
 privileges and, 143-145
 transactions and, 629, 653
 triggers and, 183
 views and, 125
delta relation, 186
demand-driven pipelining, 569-570
denormalization, 363-364
dependency preservation, 334-336, 346-348
desc, 77-78
descriptive attributes. *See* attributes
descriptive patterns, 894
deviation, 215, 906-907
dicing, 201
dictionary attacks, 414
digital certificates, 416-417
digital signatures, 416
direct-access storage, 431
directories, 935-937
directory information tree (DIT), 872-875
directory systems, 870-875
dirty blocks, 741
dirty page table, 750-756, 1244-1245
dirty read, 1137, 1181
dirty writes, 649
disable trigger, 185
disconnected operation, 395-396
disjoint entity sets, 300
disjoint specialization, 296-297
disjunction, 545-546, 594
disk-arm-scheduling, 437
disk controller, 434
distinct types, 84-86, 138-141
distinguished name (DN), 872
distributed databases, 876-878
 availability and, 847-853
 cloud-based, 861-870
 commit protocols and, 832-838
 concurrency control and, 835-836, 839-847
 deadlock handling and, 844-847
 directory systems and, 870-875
 failure and, 831-835
 fragmentation and, 826-829
 heterogeneous, 825-826, 857-861
 homogeneous, 825-826
 joins and, 855-857
 locks and, 839-847
 partitions and, 835
 persistent messaging and, 836-838
 query processing and, 854-860
 recovery and, 835-836
 replication and, 826, 829, 843-844
 storage and, 826-830
 timestamps and, 842-843
 transparency and, 829-830
 unified view of data and, 858-859
distributed-lock manager, 840
distributed systems
 autonomy and, 785
 availability and, 785
 example of, 786
 global transactions and, 784

greater potential for bugs in, 787-788
 implementation and, 786-788
 increased processing overhead of, 788
 local transactions and, 784
 nodes and, 784
 ready state and, 787
 replication and, 785
 sharing data and, 785
 sites and, 784
 software-development cost and, 787
 two-phase commit protocol (2PC) and, 786-788
distributor, 1252
divisive clustering, 907-908
Document Object Model (DOM), 390
document type definition (DTD), 990-994
domain, 42
domain constraints, 11
domain-key normal form (DKNF), 360
domain relational calculus, 249
 example queries, 246-247
 expressive power of languages, 248
 formal definition, 245
 safety of expressions, 247-248
double-pipelined hash-join, 571-572
drill down, 201
Driver-Manager class, 160
drop index, 529
drop table, 63, 164
drop trigger, 185
drop type, 140
dumping, 743-744
duplicate elimination, 563-564
durability, 22-23, 104, 625, 630-631
 defined, 628
 distributed transactions and, 830-832
 one-safe, 758
 remote backup systems and, 758
 storage structure and, 632-633
 two-safe, 758
 two-very-safe, 758
dynamic SQL, 58, 158, 175
e-catalogs, 1103

Eclipse, 386
E-commerce, 1102-1105
efficiency, 6-8
election algorithms, 851-852
embedded SQL, 58, 158, 169-173
empty relations, 93-94
encryption
 Advanced Encryption Standard (AES), 412-413
 applications of, 411-417
 asymmetric-key, 412
 authentication and, 415-417
 challenge-response system and, 415
 database support and, 414-415
 dictionary attacks and, 414
 digital certificates and, 416-417
 digital signatures and, 416
 nonrepudiation and, 416
 Oracle and, 1165-1166
 prime numbers and, 413
 public-key, 412-414
 Rijndael algorithm and, 412-413
 techniques of, 412-414
end-user information, 407-408
enterprise information, 1-2
Entity Data Model, 395
entity-relationship (E-R) diagram, 17-18
 alternative notations for, 304-310
 basic structure of, 274-275
 complex attributes, 277-278
 entity sets and, 279-281
 generalization and, 298
 identifying relationship and, 280
 mapping cardinality, 276
 nonbinary relationship sets, 272-279
 relationship sets, 278-279
 roles, 278
 university enterprise example, 282-283
 weak entity sets, 279-281
entity-relationship (E-R) model, 9, 17-18, 259, 313-314, 963
 aggregation and, 301-302, 304
 alternative modeling data notations and, 304-310
 atomic domains and, 327-329

attributes and, 263, 267-269, 290-291, 294-295, 298-299, 327-329
 complex data types and, 946-947
 constraints and, 269-272
 design issues and, 290-295
 enterprise schema and, 262
 entity sets and, 262-267, 272-274, 279-286, 290-291, 296-298
 extended features, 295-304
 generalization and, 297-304
 normalization and, 361-362
 object-oriented data model and, 27
 reduction to relational schemas and, 283-290
 redundancy and, 272-274
 relationship sets and, 264-267, 286-290, 291-295, 296-297
 specialization and, 295-297
 Unified Modeling Language (UML) and, 308-310
entity sets
 alternative notations for, 304-310
 attributes and, 263, 284-285, 290-291
 condition-defined, 299
 defined, 262-263
 design issues and, 290-292
 disjoint, 299
 extension of, 263
 generalization and, 297-304
 identifying relationship and, 280
 overlapping, 299
 properties of, 262-264
 relationship sets and, 264-267, 291-292
 removing redundancy in, 272-274
 role of, 264-265
 simple attributes and, 283-284
 strong, 283-285
 subclass, 298
 superclass, 298
 superclass-subclass relationship and, 296-297
 Unified Modeling Language (UML) and, 308-310
 user-defined, 299
 weak, 279-281, 285-286
Entity SQL, 395

enterprise resource planning (ERP) systems, 1101
entropy measure, 897-898
equi-joins, 549-559, 563, 566, 571, 807, 819
equivalence
 cost analysis and, 601-602
 join ordering and, 588-589
 relational algebra and, 582-590
 transformation examples for, 586-588
error-correcting code (ECC) organization, 444-445
ERWin, 1194
escape, 77
evaluation primitive, 539
every clause, 90
except clause, 82-83, 93, 188
exchange system, 1104
exclusive-mode locks, 661
EXEC SQL, 169-173
execute, 147
existence bitmap, 526-527
exists clause, 93
extensibility contracts, 1256-1258
external language routines, 179-180
external sort-merge algorithm, 548-549
Facebook, 31, 862
factorials, 639
fact tables, 891-892
fail-stop assumption, 722
false cycles, 846-847
false drops, 929
false negatives, 903, 929-930
false positives, 903, 929-930
false value, 90, 208
fanout, 487
fetching, 21, 138, 906, 1078, 1097
 advanced SQL and, 161, 166-173, 176, 180, 194
 application design and, 389-397, 1030, 1038
 IBM DB2 and, 1199, 1202, 1209, 1211, 1219
 information retrieval and, 921, 929, 936 (see also information retrieval)
 Microsoft SQL Server and, 1241, 1251

object-based databases and, 965, 969-972
 PostgreSQL and, 1137, 1146, 1151, 1153
 storage and, 437, 439, 444 (see also storage)
 Web crawlers and, 930-931
Fibre Channel interface, 434, 436
fifth normal forms, 360
file header, 454
file manager, 21
file organization, 3-4. *See also* storage
 B+-trees and, 500-502
 blobs, 138, 166, 457, 502, 1013, 1198-1199, 1259
 block-access time and, 438
 clobs, 138, 166, 457, 502, 1010-1013, 1196-1199
 fixed-length records and, 452-454
 hashing, 457
 heap file, 457
 indexing and, 475 (see also indices)
 journaling systems and, 439
 multitable clustering and, 458, 460-462
 pointers and, 454
 security and, 5-6 (see also security)
 sequential, 457-459
 structured, 451-468
 system structure and, 451-452
 variable-length records and, 454-457
file scan, 541-544, 550, 552, 570
final/not final expressions, 949, 953
fine-granularity parallelism, 771
FireWire interface, 434
first committer wins, 692-693
first updater wins, 693
flash storage, 403
 B+-trees and, 506
 cost of, 439
 erase speed and, 440
 hybrid disk drives and, 440-441
 NAND, 430, 439-440
 NOR, 430, 439
flash translation layer, 440
floppy disks, 430
flow-distinct, 1240-1241
FLWOR (for, let, where, order by, return) expressions, 1002-1003
forced output, 465, 725-726
force policy, 739-740
for each row clause, 181-184
for each statement clause, 68, 183
foreign keys, 46, 61-62, 131-133
fourth normal forms, 356, 358-360
fragmentation, 827-829
free space control record (FSCR), 1202
from statement
 aggregate functions and, 84-90
 basic SQL queries and, 63-74
 on multiple relations, 66-71
 natural join and, 71-74
 null values and, 83-84
 rename operation and, 74-75
 set operations and, 79-83
 on single relation, 63-66
 string operations and, 76-79
 subqueries and, 95-96
full outer joins, 117-120, 233-234, 565-566
functional dependencies, 18, 129
 attribute set closure and, 340-342
 augmentation rule and, 339
 BCNF algorithm and, 349-352
 Boyce-Codd normal form and, 333-336
 canonical cover and, 342-345
 closure of a set, 338-340
 decomposition rule, 339
 dependency preservation and, 334-336, 346-348
 extraneous, 342
 higher normal forms and, 337-338
 keys and, 330-333
 lossless decomposition and, 345-346
 multivalued, 355-360
 pseudotransitivity rule, 339
 reflexivity rule and, 339
 theory of, 338-348
 third normal form and, 336-337, 352-355
 transitivity rule and, 339

union rule, 339
functionally determined attributes, 340-342
function-based indices, 1167-1168
functions. *See also* specific function
 declaring, 174-175
 external language routines and, 179-180
 IBM DB2 and, 1197-1198
 language constructs for, 176-179
 polymorphic, 1128-1129
 PostgreSQL, 1133-1135
 state transition, 1134
 syntax and, 173-174, 178
 writing in SQL, 173-180
 XML and, 1006-1007
fuzzy checkpoints, 742-743, 750-752

generalization
 aggregation and, 301-302
 attribute inheritance and, 298-299
 bottom-up design and, 297
 condition-defined, 299
 constraints on, 299-301
 disjoint, 300
 entity-relationship (E-R) model and, 297-304
 overlapping, 300
 partial, 300
 representation of, 302-304
 subclass set and, 298
 superclass set and, 298
 top-down design and, 297
 total, 300
 user-defined, 299
Generalized Inverted Index (GIN), 1149
generalized-projection, 235
Generalized Search Tree (GiST), 1148-1149
geographic data, 1061
 applications of, 1068
 information systems and, 1065
 raster data and, 1069
 representations of, 1065-1066, 1069-1070
 spatial queries and, 1070-1071
 vector data and, 1069

getColumnCount method, 164-165
getConnection method, 160
GET method, 405
getString method, 161
Gini measure, 897
Glassfish, 386
global class identifier, 1055
global company identifier, 1055
Global Positioning System (GPS), 1068
global product identifier, 1055
global wait-for graph, 845-847
Google, 31
 application design and, 378-382, 396, 407
 distributed databases and, 862, 866-867
 information retrieval and, 933 (see also information retrieval)
 PageRank and, 922-925
grant, 144-150
granted by current role, 150
graph-based protocols, 671-674
Greenplum, 816
group by, 86-89, 96, 194, 203, 206-209
growing phase, 667-669

hackers. *See* security
handoff, 1081
hard disks, 29-30
hardware RAID, 448
hardware tuning, 1035-1038
harmonic mean, 1046
hash cluster access, 1173
hash functions, 457-458, 476, 530-531
 closed, 513
 data structure and, 515-516
 deletion and, 510, 513, 516, 523-524
 dynamic, 515-523
 extendable, 515
 indices and, 514-515, 523-524
 insertion and, 513, 516-524
 insufficient buckets and, 512
 joins and, 809-810
 lookup and, 516-518, 522, 524
 open, 513
 Oracle and, 1170
 overflows and, 512-513
 partitioning and, 807
 PostgreSQL and, 1148

queries and, 516-522
 skew and, 512
 static, 509-515, 522-523
 updates and, 516-522
hash join, 602
 basics of, 558-559
 build input and, 558
 cost of, 561-562
 double-pipelined, 571-572
 hybrid, 562
 overflows and, 560
 query processing and, 557-562
 recursive partitioning and, 539-540
 skewed partitioning and, 560
hash-table overflow, 560
having, 88-89, 96
heap file, 457, 523, 1147-1149, 1153
heuristics, 1075-1076
 data analysis and, 899, 910
 distributed databases and, 859
 greedy, 910
 IBM DB2 and, 1211
 information retrieval and, 934
 Microsoft SQL Server and, 1240
 Oracle and, 1176
 parallel databases and, 815
 query optimization and, 598-605, 615-616
Hibernate system, 393-395
hierarchical architecture, 781, 784
hierarchical classification, 935-937
hierarchical clustering, 907-908
hierarchical data model, 9
high availability, 756
HIPAA, 1248
histograms, 195, 591-596, 616, 801, 901, 1152, 1175, 1211, 1239
HITS algorithm, 925
home processor, 803
homonyms, 925-927
horizontal fragmentation, 827-828
horizontal partitioning, 798
hot-spare configuration, 758
hot swapping, 449
householding, 891
HP-UX, 1193
hubs, 924
hybrid disk drives, 440-441

hybrid hash join, 562
hybrid merge join, 557
hybrid OLAP (HOLAP), 204
hypercube architecture, 781
hyperlinks
 PageRank and, 922-923, 925
 popularity ranking and, 920-922
 search engine spamming and, 924-925
HyperText Markup Language (HTML), 378-380
 client-side scripting and, 388-391
 DataGrid and, 398
 embedded, 397
 information retrieval and, 915 (see also Information retrieval)
 Java Server Pages (JSP) and, 387-391
 rapid application development (RAD) and, 397
 security and, 402-417
 server-side scripting and, 386-388
 stylesheets and, 380
 Web application frameworks and, 398-400
 web sessions and, 380-382
 XML and, 981
HyperText Transfer Protocol (HTTP)
 application design and, 377-381, 383, 395, 404-406, 417
 as connectionless, 381
 digital certificates and, 417
 man-in-the-middle attacks and, 406
 Representation State Transfer (REST) and, 395
 Simple Object Access Protocol (SOAP) and, 1017-1018, 1249-1250
IBM AIX, 1193
IBM DB2, 30, 96, 141, 160n3, 172, 180, 184-185, 216, 1121
 administrative tools, 1215-1216
 autonomic features, 1214-1215
 buffer pools and, 1201-1202

business intelligence features, 1221-1222
 concurrency control and, 1200-1203, 1217-1218
 constraints and, 1199
 Control Center, 1195, 1215
 database-design tools and, 1194-1195
 data type support and, 1196-1197
 Data Warehouse Edition, 1221
 development of, 1193-1193
 distribution, 1220-1221
 external data and, 1220-1221
 indexing and, 1199-1205
 isolation and, 1217-1218
 joins and, 1209-1210
 large objects and, 1198-1199
 locks and, 1217-1220
 massively parallel processors (MPP) and, 1193
 materialized views and, 1212-1214
 multidimensional clustering and, 1203-1207
 query processing and, 593, 604, 612, 1207-1216
 recovery and, 1200-1203
 replication, 1220-1221
 rollback and, 1218
 set operations and, 1209-1210
 SQL variations and, 1195-1200
 storage and, 1200-1203
 system architecture, 1219-1220
 System R and, 1193
 Universal Database Server, 1193-1194
 user-defined functions and, 1197-1198
 utilities, 1215-1216
 Web services, 1199-1200
 XML and, 1195-1196
IBM MVS, 1193
IBM OS/400, 1193-1194
IBM VM, 1193-1194
IBM z/OS, 1194
identifiers, 546
 global, 1055
 OrdPath, 1260-1261
 standards and, 1055-1056
 tags and, 982-985
identifying relationship, 280
identity declaration, 1043
if clause, 184

Illustra Information Technologies, 1123-1124
immediate-modification technique, 729
incompleteness, 262
inconsistent state, 630. See also consistency
in construct, 91, 92n8
incremental view maintenance, 608-611
independent parallelism, 814
indexed nested-loop join, 552-553
index entry, 477
indexing strings, 502-503
index-organized tables (IOTs), 1164-1165
index record, 477
indices, 21, 137-138, 530-531
 access time and, 476, 479, 523
 access types and, 476
 bitmap, 507, 509, 524-528, 531, 536, 1166-1167
 block, 1205
 bulk loading of, 503-504
 clustering, 476-477, 483-485, 542
 comparisons and, 544-545
 composite, 545-546
 concurrency control and, 704-708
 construction of, 1150-1151
 covering, 509
 definition in SQL and, 528-529
 deletion time and, 476, 483, 491, 495-501, 523-524
 dense, 477-483
 domain, 1168-1169
 on expressions, 1149
 function-based, 1167-1168
 Generalized Inverted Index (GIN) and, 1149
 hashed, 476 (see also hash functions)
 IBM DB2 and, 1199-1205
 identifiers and, 546
 information retrieval and, 927-929
 insertion time and, 476, 482-483, 491-495, 499-501, 523-524
 inverted, 927-929
 join, 1168 (see also joins)
 linear search and, 541-542
 logical row-ids and, 1164-1165

materialized views and, 612
 Microsoft SQL Server and, 1231-1236
 multicolumn, 1149
 multilevel, 480-482
 multiple-key access and, 485, 506-509
 nonclustering, 477
 operator classes and, 1150
 Oracle and, 1162-1173
 ordered, 475-485, 523-524
 partial, 1150
 partitions and, 1169-1171
 performance tuning and, 1039-1041
 persistent programming languages and, 964-972
 pointers and, 546
 PostgreSQL and, 1135-1136, 1146-1151
 primary, 476-477, 542, 544
 query processing and, 541-544
 record relocation and, 502
 search key and, 476
 secondary, 477, 483-485, 502, 542, 544-545
 selection operation and, 541-544
 sequential, 485-486
 sorting and, 547-549
 space overhead and, 476, 479, 486, 522
 sparse, 477-480, 482-483
 spatial data and, 1071-1076
 support routines and, 1135-1136
 trees and, 1148-1149 (see also trees)
 unique, 1149
 updates and, 482-483
 XML, 1160
information-extraction systems, 932-933
information gain, 897-898
information retrieval, 25-26, 885, 938
 adjacency test and, 922-923
 applications of, 915-917, 931-935
 categories and, 935-937
 defined, 915
 development of field, 915
 directories and, 935-937
 false negatives and, 929-930
 false positives and, 929-930

homonyms and, 925-927
 indexing of documents and, 927-929
 information extraction and, 932-933
 keywords and, 916-927
 measuring effectiveness of, 929-930
 ontologies and, 925-927
 PageRank and, 922-923, 925
 precision and, 929-930
 query result diversity and, 932
 question answering and, 933-934
 recall and, 929-930
 relevance ranking using terms, 917-920
 relevance using hyperlinks, 920-925
 result diversity and, 932
 search engine spamming and, 924-925
 similarity-based, 919-920
 stop words and, 918
 structured data queries and, 934-935
 synonyms and, 925-927
 TF-IDF approach and, 917-925
 Web crawlers and, 930-931
Ingres, 30
inheritance, 298-299
 overriding method, 952
 SQL and, 949-956
 structured types and, 949-952
 tables and, 954-956
 types and, 952-953
initially deferred integrity constraints, 134
inner joins, 117-120, 601
inner relation, 550
insertion, 61, 100-102
 concurrency control and, 697-701
 EXEC SQL and, 171
 hashing and, 513, 516-523
 lookup and, 705
 phantom phenomenon and, 698-701
 PostgreSQL and, 1130-1131
 prepared statements and, 162-164
 privileges and, 143-145
 transactions and, 629, 653
 views and, 124-125
instances, 8, 904-905

instead of triggers, 1161-1162
integrated development environment (IDE), 111, 307, 386, 397, 426, 434, 932
integrity constraints, 4, 58
 add, 129
 alter table, 129
 assertion, 135-136
 check clause, 130, 134-136
 create table, 129, 130
 deferred, 134
 examples of, 128
 foreign key, 131-133
 functional dependencies and, 129
 hashing and, 809-810
 not null, 129-130, 133
 primary key, 130-131
 referential, 11, 46-47, 131-136, 151, 181-182, 628
 schema diagrams and, 46-47
 on single relation, 129
 unique, 130-131
 user-defined types and, 140
 violation during transaction, 133-134
 XML and, 1003-1004
integrity manager, 21
intention-exclusive (IX) mode, 680
intention-shard (IS) mode, 680
interconnection networks, 780-781
interesting sort order, 601
Interface Description Language (IDL), 1054-1055
interference, 780
internal nodes, 487
International Organization for Standardization (ISO), 57, 871, 1051
Internet, 31
 direct user access and, 2
 wireless, 1081-1082
interoperation parallelism, 804, 813-814
interquery parallelism, 802-803
intersect, 81-82, 585
intersect all, 81-82
intersection, 50
intersection set, 960
intervals, 1063-1064
intraoperation parallelism
 aggregation and, 811
 degree of parallelism and, 804

duplicate elimination and, 811
 operation evaluation costs and, 812
 parallel external sort-merge and, 806
 parallel join and, 806-811
 parallel sort and, 805-806
 projection and, 811
 range-partitioning sort and, 805
 selection and, 811
intraquery parallelism, 803-804
invalidation reports, 1083
inverse document frequency (IDF), 918
I/O parallelism
 hashing and, 799-800
 partitioning techniques and, 798-800
 range scheme and, 800
 round-robin scheme and, 799
 skew handling and, 800-802
is not null, 83
is not unknown, 84
is null, 83
isolation, 4, 1094
 atomicity and, 646-648
 cascadeless schedules and, 647-648
 concurrency control and, 631, 636-637, 639, 650, 1137-1138
 defined, 628
 dirty read and, 1137
 distributed transactions and, 830-832
 factorials and, 639
 improved throughput and, 635-636
 inconsistent state and, 631
 levels of, 648-653
 locking and, 651
 multiple versions and, 652-653, 1137-1138
 nonrepeatable read and, 1137
 Oracle and, 1181-1182
 phantom read and, 1137-1138
 PostgreSQL and, 1137-1138, 1142
 read committed, 649, 1042
 read uncommitted, 648, 649
 recoverable schedules and, 647
 repeated read, 649

resource allocation and, 635-636
 row versioning and, 1244
 serializability and, 640, 648-653
 snapshot, 652-653, 692-697, 704, 729-730, 1042, 1242, 1244
 timestamps and, 651-652
 transactions and, 628, 635-640, 646-653
 utilization and, 636
 wait time and, 636
 is unknown, 84
 item shipping, 776
 iteration, 176, 188-190

J++, 1228, 1253
 Jakarta Project, 386
 .jar files, 160
 Java, 14, 157, 169, 173, 387, 945
 DOM API, 1008-1009
 JDBC and, 158-166
 metadata and, 164-166
 persistent systems and, 971-972
 SQLJ and, 172
 Unified Modeling Language (UML) and, 308
 Java 2 Enterprise Edition (J2EE), 386, 1157-1158
 Java Database Objects (JDO), 971
 JavaScript
 application design and, 389-391, 398
 Representation State Transfer (REST) and, 395
 security and, 402-411
 JavaScript Object Notation (JSON), 395, 863-864
 JavaServer Faces (JSF) framework, 397
 Java Server Pages (JSP)
 application design and, 377, 383-391
 client-side scripting and, 389-391
 security and, 405
 server-side scripting and, 386-388
 servlets and, 383-391
 Web application frameworks and, 399
 JBoss, 386, 399

JDBC (Java Database Connectivity), 380, 1052, 1154
 advanced SQL and, 158-159
 blob column, 166
 caching and, 400-401
 callable statements and, 164
 clob column, 166
 connecting to database, 159-161
 E-R model and, 269, 275
 information protocol of, 160-161
 metadata features and, 164-166
 prepared statements and, 162-164
 query result retrieval and, 161-162
 shipping SQL statements to, 161
 updatable result sets and, 166
join dependencies, 360
joins
 complex, 563
 conditions and, 114-115
 cost analysis and, 555-557, 599-601
 distributed processing and, 855-857
 equi-joins, 549-559, 563, 566, 571, 807, 819
 filtering of, 1187
 fragment-and-replicate, 808-809
 full outer, 117-120, 233-234, 565-566
 hash join, 539-540, 557-562, 571-572, 602
 hybrid merge, 557
 IBM DB2 and, 1209-1210
 inner, 117-120, 601
 inner relation and, 550
 left outer, 116-120, 233-235, 565-566
 lossless decomposition and, 345-346
 merge-join, 553-555
 minimization and, 613
 natural, 71-74, 87, 113 (see also natural joins)
 nested-loop, 550-553 (see also nested-loop join)
 Oracle and, 1168, 1187
 ordering and, 588-589

outer, 115-120, 232-235, 565-566, 597
 outer relation, 550
 parallel, 806-811, 857
 partitioned, 539-540, 807-810
 PostgreSQL and, 1153
 prediction, 1267
 query processing and, 549-566, 855-857
 relational algebra and, 229-232, 239
 right outer, 117-120, 233-235, 565-566
 semijoin strategy and, 856-857
 size estimation and, 595-596
 sort-merge-join, 553
 theta, 584-585
 types and, 115-120
 view maintenance and, 609
join using, 74, 113-114
journaling file systems, 439
JPEG (Joint Picture Experts Group), 1077
 jukebox systems, 431

k-d trees, 1071-1072
 kernel functions, 901-902
keys, 45-46
 constraints and, 271-272
 decomposition and, 354-355
 encryption and, 412-418
 entity-relationship (E-R) model and, 271-272
 equality on, 542
 functional dependencies and, 330-333
 hashing and, 509-519, 524
 indexing and, 476-508, 476-509, 524, 529
 multiple access and, 506-509
 nonunique, 497-499
 smart cards and, 415-416
 storage and, 457-459
 USB, 430
 uniquifiers and, 498-499
keywords
 complex data types and, 947-949
 homonyms and, 925-927
 indices and, 927-929
 ontologies and, 925-927
 PostgreSQL and, 1130-1131
 query simplification and, 1237-1238
 ranking and, 915-925

search engine spamming and, 924-925
 stop words and, 918
 synonyms and, 925-927

language constructs, 176-179
Language Integrated Querying (LINQ), 1055, 1249
 large-object types, 138
 latent failure, 448
 lazy propagation, 844, 868
 lazywriter, 1246
LDAP Data Interchange Format (LDIF), 872
 least recently used (LRU) scheme, 465-467
 left outer join, 116-120, 233-235, 565-566
 legacy systems, 1050-1051
 lightweight directory access protocol (LDAP), 406, 871-875
 like, 76-77
 linear regression, 902
 linear search, 541-542
 linear speedup, 778-780
 Linux, 1124, 1193-1194, 1212
 local-area networks (LANs), 788-789, 1081
 localtime, 137
 local wait-for graph, 845
location-dependent queries, 1080
locking protocols, 666
 biased, 841
 distributed lock manager, 840
 graph-based, 671-674
 majority, 840-841
 primary copy, 840
 quorum consensus, 841-842
 single lock-manager, 839
 timestamping, 842-843
 two-phase, 667-669
lock manager, 670-671, 773
locks
 adaptive granularity and, 776
 caching and, 776
 call back and, 776
 compatibility function and, 662
 concurrency control and, 661-674
 deadlock and, 665-666, 674-679, 839, 841, 844-847, 1217-1220, 1243-1246

distributed databases and, 839-847
 dynamic, 1243
 exclusive, 651, 661-662, 668-669, 672-673, 679, 691, 698-702, 706-710, 729-730, 740-741, 803, 839, 841
 false cycles and, 846-847
 fine-grained, 756
 granting of, 666-667
 growing phase and, 667-669
 IBM DB2 and, 1217-1220
 implementation of, 670-671
 intention modes and, 680
 logical undo operations and, 744-750
 long-duration transactions and, 1110-1111
 lower/higher level, 745
 Microsoft SQL Server and, 1242-1244, 1246
 multiple granularity and, 679-682
 multiversion schemes and, 691-692
 PostgreSQL and, 1143-1145
 recovery systems and, 744-750
 request operation and, 662-671, 675-680, 709
 shared, 661, 841
 shrinking phase and, 667-669
 starvation and, 679
 timestamps and, 682-686
 transaction servers and, 773-775
 true matrix value and, 662
 wait-for graph and, 676-678, 845-847
log disk, 438-439
logical clock, 843
logical counter, 682
logical-design phase, 16, 260-261
logical error, 721
logical logging, 745-746, 1115
logical operations
 consistency and, 746
 early lock release and, 744-750
 rollback and, 746-749
 undo log records, 745-750
logical row-ids, 1164-1165
logical undo operation, 745-750
log records
 ARIES and, 750-756
 buffering and, 738-739

compensation log records (CLRs) and, 751-752, 754
 identifiers and, 727
 old/new values and, 727-728
 physical, 745
 recovery systems and, 726-728, 730-734
 redo and, 729-734
 steal/no-steal policy and, 740
 undo, 729-734, 745-746
 write-ahead logging (WAL) rule and, 739-741
log sequence number (LSN), 750-755
log writer process, 773-774
long-duration transactions
 compensation transactions and, 1113-1114
 concurrency control and, 1111-1112
 graph-based protocols and, 1110
 implementation issues, 1114-1115
 multilevel, 1111-1112
 nesting and, 1111-1112
 nonserializable executions and, 1110-1111
 operation logging and, 1115
 performance and, 1110
 recoverability and, 1110
 subtasks and, 1109
 timestamps and, 1110
 two-phase locking and, 1110
 uncommitted data and, 1109
lookup, 600, 1086
 concurrency control and, 700, 704-708
 distributed databases and, 865, 867, 870
 fuzzy, 890, 1266
 indices and, 482, 485-500, 505-513, 516-518, 522, 524
 Microsoft SQL Server and, 1238, 1241, 1266
 PostgreSQL and, 1148
 query processing and, 544, 552-553
lossless-join decomposition, 345-346
lossy decomposition, 345-346
lost update, 692
 machine learning, 25-26

magnetic disks, 430
 blocks and, 436-439
 buffering and, 437-438
 checksums and, 434
 crashes and, 434
 data-transfer rate and, 435-436
 disk controller and, 434
 failure classification and, 722
 hybrid, 440-441
 log disk and, 438-439
 mean time to failure and, 436
 optimization of disk-block access and, 436-439
 parallel systems and, 781-782
 performance measures of, 435-436
 physical characteristics of, 432-435
 read-ahead and, 437
 read-write heads and, 432-435
 recording density and, 433-434
 redundant arrays of independent disks (RAID) and, 441-449
 scheduling and, 437
 scrubbing and, 448
 sectors and, 432-434
 seek-time and, 435-436
 sizes of, 433
main-memory database systems, 724n1
majority protocol, 840-841, 848-849
man-in-the-middle attacks, 406
many server, many-router model, 1094
many-server, single-router model, 1093
many-to-many mapping, 270, 276-277
many-to-one mapping, 270, 276
mapping cardinalities, 269-270, 276-277
markup languages. See also specific language
 file processing and, 981-982
 structure of, 981-990
 tags and, 982-985
 transactions and, 983-985
master-slave replication, 843-844
master table, 1032-1033
materialization, 567-568

materialized query tables (MQTs), 1212-1214, 1221
materialized views, 123-124, 607
 aggregation and, 610-611
 IBM DB2 and, 1212-1214
 index selection and, 612
 join operation and, 609
 maintenance and, 608-611
 Oracle and, 1171-1172, 1174, 1188
 performance tuning and, 1039-1040
 projection and, 609-610
 query optimization and, 611-612
 replication and, 1251-1253
 selection and, 609-610
max, 84, 86, 96, 236, 566-567
mean time to failure (MTTF), 436
Media Access Control (MAC), 1129
mediators, 859-860, 1018-1019
memory. See also storage
 buffers and, 1184 (see also buffers)
 bulk loading of indices and, 503-504
 cache, 429, 817-818 (see also caching)
 data access and, 724-726
 flash, 403, 430, 439-441, 506
 force output and, 725-726
 magnetic-disk, 430, 432-439
 main, 429-430
 main-memory databases and, 1105-1108
 Microsoft SQL Server and, 1246-1247
 multitasking and, 1092-1095
 .NET Common Language Runtime (CLR) and, 1255-1256
 nonvolatile random-access, 438
 optical, 430
 Oracle structures and, 1183-1184
 overflows and, 560
 persistent programming languages and, 964-972
 query costs and, 544

recovery systems and, 724-726
 (see also Recovery systems)
 redo log buffer and, 1184
 shared pool, 1184
merge-join, 553
merge-purge operation, 890-891
merging
 complex, 1173-1174
 duplicate elimination and, 563-564
 Oracle and, 1173-1174
 parallel external sort-merge and, 806
 performance tuning and, 1033
 query processing and, 547-549, 553-555, 557
mesh system, 780-781
message delivery process, 838
metadata, 12, 164-166
Microsoft, 3, 31, 141
 advanced SQL and, 160n3, 169, 173, 180, 184, 197, 205
 application design and, 387, 395-401, 406-407
 distributed databases and, 863
 parallel databases and, 816
 query optimization and, 612
 storage and, 438
Microsoft Active Server Pages (ASP), 397
Microsoft Database Tuning Assistant, 1040
Microsoft Distributed Transaction Coordinator (MS DTC), 1242
Microsoft Office, 55, 399, 1016
Microsoft SQL Server, 1042, 1121
 business intelligence and, 1263-1267
 compilation and, 1236-1237
 compression and, 1236
 concurrency control and, 1241-1246
 data access and, 1248-1250
 database mirroring and, 1245-1246
 data mining and, 1266-1267
 data types and, 1229-1230
 design tools and, 1223-1228
 development of, 1223
 filegroups and, 1233-1234
 indexing and, 1231-1236
 locks and, 1242-1244

management tools and, 1223-1228
 memory management and, 1246-1247
 page units and, 1233-1234
 partitions and, 1235
 Query Editor, 1224-1225
 query processing and, 1223-1231, 1236-1241, 1250-1251
 read-ahead and, 1235-1236
 recovery and, 1241-1246
 reordering and, 1238-1239
 replication and, 1251-1253
 routines and, 1231
 security and, 1247-1248
 server programming in .NET, 1253-1258
 snapshot isolation and, 1242, 1244
 SQL Profiler and, 1225-1227
 SQL Server Broker and, 1261-1263
 SQL Server Management Studio and, 1223-1224, 1227-1228
 SQL variations and, 1228-1233
 storage and, 1233-1236
 system architecture, 1246-1248
 tables and, 1234
 thread pooling and, 1246
 triggers and, 1232-1233
 tuning and, 1224, 1227
 types and, 1257-1258
 updates and, 1232-1233, 1239
 Windows Mobile and, 1223
 XML support and, 1258-1261
Microsoft Transaction Server, 1091
Microsoft Windows, 195, 426, 1078
 IBM DB2 and, 1193-1194, 1212
 PostgreSQL and, 1124, 1155
 SQL Server and, 1223-1224, 1228, 1242, 1246-1248
 storage and, 438
min, 84, 86, 236, 566-567
minpctused, 1203
minus, 82n7
mirroring, 441-442, 444, 1245-1246
mobility, 1062, 1086
 broadcast data and, 1082-1083
 consistency and, 1083-1085

disconnectivity and, 1083-1085
 handoff and, 1081
 invalidation reports and, 1083
 mobile computer model and, 1080-1082
 queries and, 1082
 recoverability and, 1083
 routing and, 1082
 updates and, 1083-1084
 version-numbering schemes and, 1083-1084
 wireless communications and, 1080-1082
Model-View-Control design, 1157-1158
most recently used (MRU) scheme, 467
most-specific type, 953
MPEG (Moving Picture Experts Group), 1077-1078
multicore processors, 817-819
multidatabase system, 857-861
multidimensional data, 199
multimaster replication, 844
multimedia data, 1062
multimedia databases, 1076-1079
multiple granularity
 concurrency control and, 679-682
 hierarchy definition for, 679
 intention-exclusive (IX) mode and, 680
 intention-shared (IS) mode and, 680
 locking protocol and, 681-682
 shared and
 intention-exclusive (SIX) mode and, 680
 tree architecture and, 679-682
multiple-key access, 506-509
multiquery optimization, 614
multiset relational algebra, 238
multiset types, 956-961
multisystem applications, 1096
multitable cluster file
 organization, 458, 460-462
multitasking, 771, 1092-1095
multithreading, 817-818, 1093
multivalued attributes, 267-268, 327-329
multivalued dependencies, 355-360

multiversion concurrency control (MVCC)
 DDL commands and, 1144-1145
 DML commands and, 1138-1139
 implementation of, 1139-1143
 implications of using, 1143-1144
 indices and, 1145
 isolation levels and, 1137-1138
 locks and, 1145
 recovery and, 1145-1146
 schema for, 689-692
multiway splits, 898
MySQL, 31, 76, 111, 160n3, 1123, 1155
Naïve Bayesian classifiers, 901, 1191, 1266
naïve users, 27-28
name servers, 829
NAND flash memory, 430, 440-441
natural joins, 49-50, 87, 113
 conditions and, 114-115
 full outer, 117-120, 233-234, 565-566
 inner, 117-120, 601
 left outer, 116-120, 233-235, 565
 on condition and, 114-115
 outer, 115-120
 SQL queries and, 71-74
 relational algebra and, 229-232
 right outer, 117-120, 233-234, 565-566
 types and, 115-120
nearest-neighbor query, 1070-1071
negation, 595
nested-loop join, 1071
 IBM DB2 and, 1209-1210
 Oracle and, 1173
 parallel, 807, 810-811
 PostgreSQL and, 1152-1153
 query optimization and, 600n2, 602, 604
 query processing and, 550-555, 558-560, 565, 571, 573
nested subqueries
 application development and, 1031, 1047

duplicate tuples and, 94-95
 empty relations and, 93-94
 from clause and, 95-96
 optimization of, 605-607
 scalar, 97-98
 set operations and, 90-93
 with clause and, 97

nesting
 ARIES and, 755-756
 concurrency control and, 679, 709
 granularities and, 679
 IBM DB2 and, 1197, 1209-1210, 1218
 long-duration transactions and, 1112-1113
 object-based databases and, 945, 948, 958-961
 Oracle and, 1159, 1164, 1182
 queries and, 601-607, 1004-1007, 1013-1014, 1017
 transactions and, 1091, 1112-1113, 1116, 1218
 XML and, 27, 943, 984-998, 1001, 1004-1007, 1010

.NET, 169
NetBeans, 386, 397
.NET Common Language Runtime (CLR)
 aggregates and, 1257-1258
 basic concepts of, 1254
 extensibility contracts and, 1256-1258
 Microsoft SQL Server and, 1253-1258
 routines and, 1256-1257
 SQL hosting and, 1254-1256
 table functions and, 1256-1257
 types and, 1257-1258

Netezza, 816

networks
 data model and, 9, 1080-1082
 local area, 788-789, 1081
 mobility and, 1079-1085
 wide-area types and, 788, 790-791

nextval for, 1043

nodes. *See also storage*
 B+-trees and, 485-506
 coalescing, 491, 706
 distributed systems and, 784
 IBM DB2 and, 1200-1201
 mesh architecture and, 780-781

multiple granularity and, 679-682
 nonleaf, 487
 overfitting and, 899-900
 splitting of, 491, 706
 updates and, 491-500
 XML and, 998

no-force policy, 739-740

nonacceptable termination states, 1099

nonclustering, 477

nondeclarative actions, 158

nonleaf nodes, 487

nonprocedural languages, 47-48

nonrepeatable read, 1137

nonrepudiation, 416

nonunique search keys, 497-499

nonvolatile random-access memory (NVRAM), 438

nonvolatile storage, 432, 632, 722, 724-726, 743-744

nonvolatile write buffers, 438

NOR flash memory, 430, 439

normal forms, 18
 atomic domains and, 327-329
 Boyce-Codd, 333-336, 349-352, 354-356
 complex data types and, 947
 domain-key, 360
 fifth, 360
 first, 327-329
 fourth, 356, 358-360
 higher, 337-338
 join dependencies and, 360
 project-join, 360
 second, 361
 third, 336-337

normalization, 16, 18-20
 denormalization and, 363-364
 entity-relationship (E-R) model and, 361-362
 performance and, 363-364
 relational database design and, 361-362

no-steal policy, 740

not connective, 66

not exists, 93, 192

not in, 90-91, 92n8

not null, 61, 83, 129-130, 133, 140

not unique, 95

null bitmap, 456

null values, 19, 83-84
 aggregation with, 89-90
 attributes and, 268-269

decode and, 208

decorrelation and, 1174

file organization and, 451-468

integrity constraints and, 128-130, 133-134

left outer join, 233

OLAP and, 202

query simplification and, 1237-1238

right outer join, 234-235

temporal data and, 364-367

user-defined types and, 140

numeric, 59, 62

nvarchar, 60

N-way merge, 547

object-based databases, 975
 array types and, 956-961
 collection volumes and, 957-958
 complex data types and, 946-949
 correspondence and, 955
 feature implementation and, 963-964
 inheritance and, 949-956
 mapping and, 973
 multiset types and, 956-961
 nesting and, 945, 948, 958-961
 object-identity types and, 961-963
 object-oriented vs. object-relational approaches and, 973-974
 persistent programming languages and, 964-972, 974
 reference types and, 961-963
 relational data model and, 945
 structured types and, 949-953
 unnesting and, 958-961

Object Database Management Group (ODMG), 1054-1055

object-oriented databases, 393

object-oriented data model, 27

object-relational data model, 27

object-relational mapping, 393, 946, 973

observable external writes, 634-635

ODBC (Open Database Connectivity), 380, 1052
 advanced SQL and, 166-169
 API definition and, 166-167

caching and, 401

conformance levels and, 168-169

Microsoft SQL Server and, 1249

PostgreSQL and, 1154

standards for, 1053-1054

type definition and, 168

OLAP (online analytical processing), 1046
 all attribute and, 201-203, 205
 applications for, 197-201
 cross-tabulation and, 199-203, 205, 210
 data cube and, 200, 206-210
 decode function and, 208
 dicing and, 201
 drill down and, 201
 implementation of, 204
 Microsoft SQL Server and, 1223, 1266
 multidimensional data and, 199
 null value and, 202
 Oracle and, 1161
 order by clause and, 205
 pivot clause and, 205, 210
 relational tables and, 202-203, 205
 rollup and, 201, 206-210
 slicing and, 201
 in SQL, 205-209

OLE-DB, 1249

OLTP (online transaction processing), 1046-1047, 1165, 1186, 1264

on condition, 114-115

on delete cascade, 133, 185

one-to-many mapping, 269, 276

one-to-one mapping, 269, 276

ontologies, 925-927

OOXML (Office Open XML), 1016

Open Document Format (ODF), 1016

open statement, 170-171

operation logging, 1115

operator tree, 803-804

optical storage, 430-431, 449-450

optimistic concurrency control without read validation, 704

Oracle, 3, 30, 216, 1121
 access path selection and, 1174

analytic workspaces and, 1161

archiver and, 1185

caching and, 1179-1180, 1184

checkpoint and, 1185

clusters and, 1173, 1186

compression and, 1165

concurrency control and, 1180-1183

database administration tools and, 1189-1191

database design and, 355, 386, 396, 408n5, 409, 1157-1158

database writer, 1185

data guard, 1183

data mining and, 1191

data warehousing and, 1158

dedicated servers and, 1183-1185

dimensional modeling and, 1160, 1171

distribution and, 1188-1189

encryption and, 1165-1166

Exadata and, 1187-1188

external data and, 1188-1189

hashing and, 1170

indices and, 1162-1173

isolation levels and, 1181-1182

joins and, 1168, 1187

logical row-ids and, 1164-1165

log writer and, 1185

materialized views and, 1171-1172, 1174, 1188

memory structures and, 1183-1184

optimizer of, 1174-1176

parallel execution and, 1178

partitions and, 1169-1172, 1176

process monitor and, 1185

process structures and, 1184-1185

projection and, 1187

query optimization and, 582, 593, 603-604, 612, 1173-1178

query processing and, 1157-1158, 1162-1172

Real Application Clusters (RAC) and, 1186

recovery and, 1180-1183, 1185

replication and, 1188-1189

result caching and, 1179-1180

security and, 1165-1166

segments and, 1163

serializability and, 1181-1182

shared server and, 1185

as Software Development Laboratories, 1157

SQL basics and, 55, 75n4, 82n7, 96, 141, 160-161, 172-174, 178, 180, 184-185, 197, 205

SQL Loader and, 1189

SQL Plan Management and, 1177-1178

SQL Tuning Advisor, 1176-1177

SQL variations and, 1158-1162

subquery flattening and, 1174

system architecture, 795, 803, 843, 1183-1188

system monitor, 1185

tables and, 1163-1166, 1172-1173, 1187, 1189

transactions and, 649, 653, 692-693, 697, 710, 718

transformations and, 1173-1174

trees and, 1191

triggers and, 1161-1162

updates and, 1179-1180

virtual private database and, 1166

XML DB and, 1159-1160

Oracle Application Development Framework (ADF), 1157-1158

Oracle Automatic Storage Manager, 1186-1187

Oracle Automatic Workload Repository (AWR), 1190

Oracle Business Intelligence Suite (OBI), 1158

Oracle Database Resource Management, 1190-1191

Oracle Designer, 1158

Oracle Enterprise Manager (OEM), 1190

Oracle JDeveloper, 1158

Oracle Tuxedo, 1091

or connective, 66

order by, 77-78, 193

organize by dimensions, 1204

or operation, 83-84

outer-join, 115-120, 232-235, 565-566, 597

outer relation, 550

overfitting, 899-900

overflow avoidance, 560

overflow buckets, 512-514
 overflow resolution, 560
 overlapping entity sets, 300
 overlapping specialization,
 296-297
 overloading, 968

P + Q redundancy schema, 446
PageLSN, 751, 753, 754
PageRank, 922-925, 928
page shipping, 776
parallel databases

cache memory and, 817-818
 cost of, 797
 decision-support queries and,
 797
 failure rates and, 816
 increased use of, 797
 interoperability parallelism
 and, 813-814
 interquery parallelism and,
 802-803
 intraoperation parallelism
 and, 804-812
 intraquery parallelism and,
 803-804
 I/O parallelism and, 798-802
 massively parallel processors
 (MPP) and, 1193
 multicore processors and,
 817-819
 multithreading and, 817-818
 operator tree and, 803-804
 Oracle and, 1178-1179
 partitioning techniques and,
 798-799
 pipelines and, 814-815
 query optimization and,
 814-817
 raw speed and, 817
 skew and, 800-801, 805-808,
 812, 814, 819
 success of, 797
 system design and, 815-817
**parallel external sort-merge,
 806**
parallelism, 442-444
parallel joins, 806, 857
 fragment-and-replicate,
 808-809
 hash, 809-810
 nested-loop, 810-811
 partitioned, 807-810
parallel processing, 401-402

parallel systems
 coarse-grain, 777
 fine-grain, 777
 hierarchical, 781
 interconnection networks
 and, 780-781
 interference and, 780
 massively parallel, 777-778
 scaleup and, 778-780
 shared disk, 781
 shared memory, 781-783
 shared nothing, 781
 skew and, 780
 speedup and, 778-780
 start-up costs and, 780
 throughput and, 778

parameter style general, 179
**parametric query optimization,
 615**

parity bits, 444-446

parsing

application design and, 388
 bulk loads and, 1031-1033
 query processing and,
 537-539, 572, 1236-1237

participation constraints, 270
partitioning vector, 798-799

partitions

attributes and, 896-897
 availability and, 847-853
 balanced range, 801
 classifiers and, 896-897
 cloud computing and, 865-866
 composite, 1170-1171
 condition and, 896-897
 distributed databases and,
 832, 835
 hash, 798-799, 807, 1170
 joins and, 807-810
 list, 1170
 Microsoft SQL Server and,
 1235
 Oracle and, 1169-1172, 1176
 point queries and, 799
 pruning and, 1176
 query optimization and,
 814-815
 range, 798-800, 805, 1170
 reference, 1171
 round-robin, 798-801
 scanning a relation and, 799

**Partner Interface Processes
 (PIPs), 1055**

passwords. See also security
 application design and, 376,
 382, 385, 393, 405-407, 415
 dictionary attacks and, 414
 distributed databases and, 871
 leakage of, 405
 man-in-the-middle attacks
 and, 406
 one-time, 406
 single sign-on system and,
 406-407
 SQL and, 142, 160, 168, 170
 storage and, 463-464
PATA (parallel ATA), 434
pctfree, 1203
performance

access time and, 431-439, 447,
 450-451, 476, 479, 523,
 540-541, 817
 application design and,
 400-402
 B+-trees and, 485-486
 caching and, 400-401
 data-transfer rate and, 435-436
 denormalization and, 363-364
 magnetic disk storage and,
 435-436
 parallel processing and,
 401-402
 response time and, 400 (see
 also response time)
 seek times and, 433, 435-439,
 450-451, 540, 555
 sequential indices and,
 485-486
 transaction time and, 365n8,
 1062
 web applications and, 377-382

performance benchmarks
 database-application classes
 and, 1046
 suites of tasks, 1045-1046
 Transaction Processing
 Performance Council
 (TPC), 1046-1048

performance tuning

bottleneck locations and,
 1033-1035
 bulk loads and, 1031-1033
 concurrent transactions and,
 1041-1044
 hardware and, 1035-1038
 indices and, 1039-1041
 materialized views and,
 1039-1041

parameter adjustment and,
 1029-1030, 1035
 physical design and,
 1040-1041
 RAID choice and, 1037-1038
 of schema, 1038-1039
 set orientation and, 1030-1031
 simulation and, 1044-1045
 updates and, 1030-1033
Perl, 180, 387, 1154
persistent messaging, 836-837
**persistent programming
 languages, 974**
 approaches for, 966-967
 byte code enhancement and,
 971
 C++, 968-971
 class extents and, 969, 972
 database mapping and, 971
 defined, 965
 iterator interface and, 970
 Java, 971-972
 object-based databases and,
 964-972, 974
 object identity and, 967
 object persistence and,
 966-968
 overloading and, 968
 persistent objects and, 969
 pointers and, 967, 969, 972
 reachability and, 971
 relationships and, 969
 single reference types and, 972
 transactions and, 970
 updates and, 970
**person-in-the-middle attacks,
 1105**
phantom phenomenon, 698-701
**phantom read, 1137-1138, 1142,
 1217-1218**
PHP, 387-388
physical data independence, 6
physical-design phase, 16, 261
physiological redo, 750
pinned blocks, 465
pipelining, 539, 568
 demand-driven, 569-570
 double-pipelined hash-join
 and, 571-572
 parallel databases and,
 813-815
 producer-driven, 569-571
 pulling data and, 570-571
pivot clause, 205, 210, 1230
plan caching, 605

PL/SQL, 173, 178

pointers. See also indices

application design and, 409
 child nodes and, 1074
 concurrency control and,
 706-707
 IBM DB2 and, 1199, 1202-1203
 information retrieval and, 936
 main-memory databases and,
 1107
 multimedia databases and,
 1077
 Oracle and, 1165
 persistent programming
 languages and, 967, 969,
 972
 PostgreSQL and, 1134,
 1147-1148
 query optimization and, 612
 query processing and,
 544-546, 554
 recovery systems and, 727,
 754
 SQL basics and, 166, 179-180
 storage and, 439, 452-462

point queries, 799

polymorphic types, 1128-1129

popularity ranking, 920-925

PostgreSQL, 31, 1121

access methods and, 1153
 aggregation and, 1153
 command-line editing and,
 1124
 concurrency control and, 692,
 697, 701, 1137-1145
 constraints and, 1130-1131,
 1153-1154
 DML commands and,
 1138-1139
 extensibility, 1132
 functions, 1133-1135
 Generalized Inverted Index
 (GIN) and, 1149
 Generalized Search Tree
 (GIST) and, 1148-1149
 hashing and, 1148
 indices and, 1135-1136,
 1146-1151
 isolation levels and,
 1137-1138, 1142
 joins and, 1153
 locks and, 1143-1145
 major releases of, 1123-1124
 multiversion concurrency
 control (MVCC) and,
 1137-1146
 operator classes and, 1150
 operator statements and, 1136
 parallel databases and,
 816-817
 performance tuning and, 1042
 pointers and, 1134, 1147-1148
 procedural languages and,
 1136
 query optimization and, 582,
 593
 query processing and,
 1151-1154
 recovery and, 718
 rollbacks and, 1142-1144
 rules and, 1130-1131
 serializability and, 1142-1143
 server programming
 interface, 1136
 sort and, 1153
 SQL basics and, 140, 160, 173,
 180, 184
 state transition and, 1134
 storage and, 1146-1151
 system architecture, 1154-1155
 system catalogs and, 1132
 transaction management in,
 649, 653, 1137-1146
 trees and, 1148-1149
 triggers and, 1153-1154
 trusted/untrusted languages
 and, 1136
 tuple ID and, 1147-1148
 tuple visibility and, 1139
 types, 1126-1129, 1132-1133
 updates and, 1130, 1141-1144,
 1147-1148
 user interfaces, 1124-1126
 vacuum, 1143
precedence graph, 644
precision, 903
predicate reads, 697-701
prediction
 classifiers and, 894-904
 data mining and, 894-904
 joins and, 1267
prepared statements, 162-164
**presentation facilities,
 1094-1095**
presentation layer, 391
**prestige ranking, 920-925,
 930-931**
primary copy, 840

primary keys, 45-46, 60-62
 decomposition and, 354-355
 entity-relationship (E-R) model and, 271-272
 functional dependencies and, 330-333
 integrity constraints and, 130-131
primary site, 756
privacy, 402, 410-411, 418, 828, 869-870, 1104
privileges
 all, 143-144
 execute and, 147
 granting of, 143-145
 public, 144
 revoking of, 143-145, 149-150
 transfer of, 148-149
procedural DMLs, 10
procedural languages, 20
 advanced SQL and, 157-158, 173, 178
 IBM DB2 and, 1194
 Oracle and, 1160, 1191
 PostgreSQL and, 1130, 1133, 1136
 relational model and, 47-48
procedures
 declaring, 174-175
 external language routines and, 179-180
 language constructs for, 176-179
 syntax and, 173-174, 178
 writing in SQL, 173-180
producer-driven pipeline, 569-570
program global area (PGA), 1183
programming languages. *See also specific language*
 accessing SQL from, 157-173
 mismatch and, 158
 variable operations of, 158
projection
 intraoperation parallelism and, 811
 Oracle and, 1187
 queries and, 564, 597
 view maintenance and, 609-610
project-join normal form (PJNF), 360
project operation, 219
PR quadrees, 1073

pseudotransitivity rule, 339
public-key encryption, 412-414
publishing, 1013, 1251-1253
pulling data, 570-571
purity, 897
Python, 180, 377, 387, 1123, 1125, 1136
QBE, 37, 245, 770
quadratic split, 1075-1076
quadrees, 1069, 1072-1073
queries, 10
 ADO.NET and, 169
 availability and, 826-827
 B+-trees and, 488-491
 basic structure of SQL, 63-71
 caching and, 400-401
 Cartesian product and, 50-51, 68-69, 71-75, 120, 209, 217, 222-229, 232, 573, 584, 589, 595-596, 606, 616
 complex data types and, 946-949
 correlated subqueries and, 93
 data-definition language (DDL) and, 21-22
 data-manipulation language (DML) and, 21-22
 decision-support, 797
 delete and, 98-100
 distributed databases and, 825-878 (see also distributed databases)
 hashing and, 475, 516-522 (see also hash functions)
 indices and, 475 (see also indices)
 information retrieval and, 915-938
 insert and, 100-101
 intermediate SQL and, 113-151
 JDBC and, 158-166
 location-dependent, 1080
 metadata and, 164-166
 multiple-key access and, 506-509
 on multiple relations, 66-71
 natural joins and, 71-74, 87, 113-120 (see also joins)
 nearest-neighbor, 1070-1071
 nested subqueries, 90-98
 null values and, 83-84
 object-based databases and, 945-975

ODBC and, 166-169
 OLAP and, 197-209
 Oracle and, 1171-1172
 PageRank and, 922-923
 parallel databases and, 797-820
 persistent programming languages and, 964-972
 point, 799
 programming language
 access and, 157-173
 range, 799
 read only, 804
 recursive, 187-192
 result diversity and, 932
 ResultSet object and, 159, 161, 164-166, 393, 397-398, 490
 retrieving results, 161-162
 scalar subqueries and, 97-98
 security and, 402-417
 servlets and, 383-391
 set operations and, 79-83, 90-93
 on single relation, 63-66
 spatial data and, 1070-1071
 string operations and, 76-79
 transaction servers and, 775
 universal Turing machine and, 14
 user requirements and, 311-312
 views and, 120-128
 XML and, 998-1008
query cost
 Microsoft SQL Server and, 1237-1239
 optimization and, 580-581, 590-602
 processing and, 540-541, 544, 548, 555-557, 561
query evaluation engine, 22
query-evaluation plans, 537-539
 choice of, 598-607
 expressions and, 567-572
 materialization and, 567-568
 optimization and, 579-616
 pipelining and, 568-572
 response time and, 541
 set operations and, 564
 viewing, 582
query-execution engine, 539
query-execution plan, 539

query languages, 249. *See also specific language*
 accessing from a
 programming language, 157-173
 centralized systems and, 770-771
 domain relational calculus and, 245-248
 expressive power of
 languages, 244, 248
 formal relational, 217-248
 nonprocedural, 239-244
 procedural, 217-239
 relational algebra and, 217-239
 relational model and, 47-48, 50
 temporal, 1064
 tuple relational calculus and, 239-244
query optimization, 22, 537, 539, 552-553, 562, 616
 access path selection and, 1174-1176
 aggregation and, 597
 cost analysis and, 580-581, 590-602
 distributed databases and, 854-855
 equivalence and, 582-588
 estimating statistics of
 expression results, 590-598
 heuristics in, 602-605
 IBM DB2 and, 1211-1212
 join minimization, 613
 materialized views and, 607-612
 Microsoft SQL Server and, 1236-1241
 multiquery, 614
 nested subqueries and, 605-607
 Oracle and, 1173-1178
 parallel databases and, 814-817
 parametric, 615
 parallel execution and, 1178-1179
 partial search and, 1240
 partitions and, 1174-1176
 plan choice for, 598-607
 PostgreSQL and, 1151-1154
 process structure and, 1179

relational algebra and, 579-590
 result caching and, 1179-1180
 set operations and, 597
 shared scans and, 614
 simplification and, 1237-1238
 SQL Plan Management and, 1177-1178
 SQL Tuning Advisor and, 1176-1177
 top-K, 613
 transformations and, 582-590, 1173-1174
 updates and, 613-614
query processing, 21-22, 30, 32
 aggregation, 566-567
 basic steps of, 537
 binding and, 1236-1237
 comparisons and, 544-545
 compilation and, 1236-1237
 cost analysis of, 540-541, 544, 548, 555-557, 561
 CPU speeds and, 540
 distributed databases and, 854-857, 859-860
 distributed heterogeneous, 1250-1251
 duplicate elimination and, 563-564
 evaluation of expressions, 567-572
 executor module and, 1152-1153
 file scan and, 541-544, 550, 552, 570
 hashing and, 557-562
 IBM DB2 and, 1207-1216
 identifiers and, 546
 information retrieval and, 915-937
 join operation and, 549-566
 LINQ and, 1249
 materialization and, 567-568, 1212-1214
 Microsoft SQL Server and, 1223-1231, 1236-1241, 1250-1251
 mobile, 1082
 operation evaluation and, 538-539
 Oracle and, 1157-1158, 1172-1180
 parsing and, 537-539, 572-573, 1236-1237
 pipelining and, 568-572

PostgreSQL and, 1151-1154
 projection and, 563-564
 recursive partitioning and, 539-540
 relational algebra and, 537-539
 reordering and, 1238-1239
 selection operation and, 541-546
 set operations and, 564-565
 sorting and, 546-549
 SQL and, 537-538
 standard planner and, 1152
 syntax and, 537
 transformation and, 854-855
 triggers and, 1153-1154
 XML and, 1259-1260
question answering, 933-934
queueing systems, 1034-1035
quorum consensus protocol, 841-842
random access, 437
random samples, 593
random walk model, 922
range-partitioning sort, 805
range-partitioning vector, 801
range queries, 799
ranking, 192-195
rapid application development (RAD)
 functions library and, 396
 report generators and, 399-400
 user interface building tools and, 396-398
 Web application frameworks and, 398-399
raster data, 1069
Rational Rose, 1194
read-ahead, 437
read committed
 application development and, 1042
 Microsoft SQL Server and, 1242
 Oracle and, 1181
 PostgreSQL and, 1138, 1141-1142
 transaction management and, 649, 658, 685, 701-702
read one, write all available protocol, 849-850
read one, write all protocol, 849
read only queries, 804
read quorum, 841-842

read uncommitted, 648
 read-write contention, 1041-1042
 read/write operations, 653-654
 real, double precision, 59
 real-time transaction systems, 1108-1109
 recall, 903
 recovery interval, 1244-1245
 recovery manager, 22-23
 recovery systems, 186, 631, 760-761, 1083
 actions after crash, 736-738
 algorithm for, 735-738
 ARIES, 750-756
 atomicity and, 726-735
 buffer management and, 738-743
 checkpoints and, 734-735, 742-743
 concurrency control and, 729-730
 data access and, 724-726
 database mirroring and, 1245-1246
 database modification and, 728-729
 disk failure and, 722
 distributed databases and, 835-836
 early lock release and, 744-750
 fail-stop assumption and, 722
 failure and, 721-723, 743-744
 force/no-force policy and, 739-740
 IBM DB2 and, 1200-1203, 1217-1218
 logical undo operations and, 744-750
 log records and, 726-728, 730-734, 738-739
 log sequence number (LSN) and, 750
 long-duration transactions and, 1110
 Microsoft SQL Server and, 1241-1246
 Oracle and, 1180-1183
 partitions and, 1169-1172
 PostgreSQL and, 1145-1146
 redo and, 729-738
 remote backup, 723, 756-759, 850, 1095-1096
 rollback and, 729-734, 736

shadow-copy scheme and, 727
 snapshot isolation and, 729-730
 steal/no-steal policy and, 740
 storage and, 722-726, 734-735, 743-744
 successful completion and, 723
 undo and, 729-738
 workflows and, 1101
 write-ahead logging (WAL) rule and, 739-741, 1145-1146
 recovery time, 758
 recursive partitioning, 539-540
 recursive queries, 187
 recursive relationship sets, 265
 redo
 actions after crash, 736-738
 pass, 754
 phase, 736-738
 recovery systems and, 729-738
 redundancy, 4, 261-262, 272-274
 redundant arrays of independent disks (RAID), 435, 759, 1147
 bit-level striping, 442-444
 error-correcting-code (ECC) organization and, 444-445
 hardware issues, 448-449
 hot swapping and, 449
 levels, 444-448
 mirroring and, 441-442, 444
 parallelism and, 442-444
 parity bits and, 444-446
 performance reliability and, 442-444
 performance tuning and, 1037-1038
 recovery systems and, 723
 reliability improvement and, 441-442
 scrubbing and, 448
 software RAID and, 448
 striping data and, 442-444
 references, 131-133, 148
 referencing new row as, 181-182
 referencing new table as, 183
 referencing old row as, 182
 referencing old table as, 183
 referencing relation, 46

referential integrity, 11, 46-47, 131-136, 151, 181-182, 628
 referrals, 875
 reflexivity rule, 339
 region quadrees, 1073
 regression, 902-903, 1048-1049
 relational algebra, 51-52, 248-249, 427
 aggregate functions, 235-239
 assignment, 232
 avg, 236
 Cartesian-product, 222-226
 composition of relational operations and, 219-220
 count-distinct, 236
 equivalence and, 582-588, 601-602
 expression transformation and, 582-590
 expressive power of languages, 244
 formal definitions of, 228
 fundamental operations, 217-228
 generalized-projection, 235
 join expressions, 239
 max, 236
 min, 236
 multiset, 238
 natural-join, 229-232
 outer-join, 232-235
 project operation, 219
 query optimization and, 579-590
 query processing and, 537-539
 rename, 226-228
 select operation, 217-219
 semijoin strategy and, 856-857
 set-difference, 221-222
 set-intersection, 229
 SQL and, 219, 239
 sum, 235-236
 union operation, 220-221
 relational database design, 368
 atomic domains and, 327-329
 attribute naming, 362-363
 decomposition and, 329-338, 348-360
 design process and, 361-364
 features of good, 323-327
 first normal form and, 327-329
 fourth normal form and, 356, 358-360
 functional dependencies and, 329-348

larger schemas and, 324-325
 multivalued dependencies and, 355-360
 normalization and, 361-362
 relationship naming, 362-363
 second normal form and, 336n5, 361
 smaller schemas and, 325-327
 temporal data modeling and, 364-367
 third normal form and, 336-337
 relational databases
 access from application programs and, 14-15
 data-definition language and, 14
 data-manipulation language (DML) and, 13-14
 storage and, 1010-1014
 tables and, 12-13
 relational model, 9
 disadvantages of, 30
 domain and, 42
 keys and, 45-46
 natural joins and, 49-50
 operations and, 48-52
 query languages and, 47-48, 50
 referencing relation and, 46
 schema for, 42-47, 302-304, 1012
 structure of, 39-42
 tables for, 39-44, 49-51, 202-205
 tuples and, 40-42, 49-50
 relation instance, 42-45, 264
 relationship sets
 alternative notations for, 304-310
 atomic domains and, 327-329
 attribute placement and, 294-295
 binary vs. n-ary, 292-294
 descriptive attributes, 267
 design issues and, 291-295
 entity-relationship diagrams and, 278-279
 entity-relationship (E-R) model and, 264-267, 286-290, 296-297
 entity sets and, 291-292
 naming of, 362-363
 nonbinary, 278-279
 recursive, 265

redundancy and, 288
 representation of, 286-290
 schema combination and, 288-290
 superclass-subclass, 296-297
 Unified Modeling Language (UML) and, 308-310
 relative distinguished names (RDNs), 872
 relevance
 adjacency test and, 922-923
 hubs and, 924
 PageRank and, 922-923, 925
 popularity ranking and, 920-922
 ranking using TF-IDF, 917-920, 925
 search engine spamming and, 924-925
 similarity-based retrieval and, 919-920
 TF-IDF approach and, 917-925, 928-929
 using hyperlinks and, 3421
 Web crawlers and, 930-931
 relevance feedback, 919-920
 remote backup systems, 723, 756-759, 850, 1095-1096
 remote-procedure-call (RPC) mechanism, 1096
 rename operation, 75-76, 226-228
 repeat, 176
 repeatable read, 649
 repeat loop, 188, 341, 343, 490
 replication
 cloud computing and, 866-868
 distributed databases and, 843-844
 Microsoft SQL Server and, 1251-1253
 system architectures and, 785, 826, 829
 report generators, 399-400
 Representation State Transfer (REST), 395
 request forgery, 403-405
 request operation
 deadlock handling and, 675-679
 locks and, 662-671, 675-680, 709
 lookup and, 706
 multiple granularity and, 679-680

multiversion schemes and, 691
 snapshot isolation and, 693
 timestamps and, 682
 resource managers, 1095
 response time
 application design and, 400, 1037, 1046
 concurrency control and, 688
 E-R model and, 311
 Microsoft SQL Server and, 1261
 Oracle and, 1176-1177, 1190
 query evaluation plans and, 541
 query processing and, 541
 storage and, 444, 1106, 1109-1110
 transactions and, 636
 system architecture and, 778, 798, 800, 802
 restriction, 149-150, 347
 ResultSet object, 159, 161, 164-166, 393, 397-398, 490
 revoke, 145, 149
 right outer join, 117-120, 233-235, 565-566
 Rijndael algorithm, 412-413
 robustness, 847
 roles, 264-265
 authorization and, 145-146
 entity-relationship diagrams, 278
 Unified Modeling Language (UML) and, 308-310
 rollback, 173
 ARIES and, 754-755
 cascading, 667
 concurrency control and, 667, 670, 674-679, 685, 689, 691, 709
 IBM DB2 and, 1218
 logical operations and, 746-749
 PostgreSQL and, 1142-1144
 recovery systems and, 729-734, 736
 remote backup systems and, 758-759
 transactions and, 736
 timestamps and, 685-686
 undo and, 729-734
 rollback work, 127
 rollup, 201, 206-210, 1221-1222
 RosettaNet, 1055

row triggers, 1161-1162
R-trees, 1073-1076
Ruby on Rails, 387, 399
runstats, 593

SAS (Serial attached SCSI), 434
Sarbanes-Oxley Act, 1248
SATA (serial ATA), 434, 436
savepoints, 756
scalar subqueries, 97-98
scaleup, 778-780
scheduling

Microsoft SQL Server and, 1254-1255

PostgreSQL and, 1127
query optimization and, 814-815

storage and, 437
transactions and, 641, 1099-1100, 1108

schema definition, 28
schema diagrams, 46-47
schemas, 8

alternative notations for modeling data, 304-310
authorization on, 147-148
basic SQL query structures and, 63-74
canonical cover and, 342-345
catalogs and, 142-143
combination of, 288-290
concurrency control and, 661-710 (see also concurrency control)
data-definition language (DDL) and, 58, 60-63
data mining, 893-910
data warehouses, 889-893
entity-relationship (E-R) model and, 262-313
functional dependencies and, 329-348
generalization and, 297-304
larger, 324-325
locks and, 661-686
performance tuning of, 1038-1039
physical-organization modification and, 28
recovery systems and, 721-761
reduction to relational, 283-290
redundancy of, 288
relational algebra and, 217-239

relational database design and, 323-368
relational model and, 42-47
relationship sets and, 286-288
shadow-copy, 727
smaller, 325-327
strong entity sets and, 283-285
timestamps and, 682-686
tuple relational calculus, 239-244
version-numbering, 1083-1084
weak entity sets and, 285-286
XML documents, 990-998

scripting languages, 389

scrubbing, 448

search engine spamming, 924-925

search keys

hashing and, 509-519, 524
indexing and, 476-509, 524, 529

nonunique, 497-499

storage and, 457-459
uniquifiers and, 498-499

secondary site, 756

second normal form, 361

Secure Electronic Transaction (SET) protocol, 1105

security, 5, 147

abstraction and, 6-8, 10
application design and, 402-417

audit trails and, 409-410
authentication and, 405-407
authorization and, 11, 21, 407-409

concurrency control and, 661-710 (see also concurrency control)
cross site scripting and, 403-405

dictionary attacks and, 414
encryption and, 411-417, 1165-1166

end-user information and, 407-408

GET method and, 405

integrity manager and, 21
isolation and, 628, 635-640, 646-653

keys and, 45-46
locks and, 661-686 (see also locks)

long-duration transactions and, 1109-1115

man-in-the-middle attacks and, 406

Microsoft SQL Server and, 1247-1248

observable external writes and, 634-635

Oracle and, 1165-1166

passwords and, 142, 160, 168, 170, 376, 382, 385,

393, 405-407, 415, 463-464, 871

person-in-the-middle attacks and, 1105

physical data independence and, 6

privacy and, 402, 410-411, 418, 828, 869-870, 1104

remote backup systems and, 756-759

request forgery and, 403-405
single sign-on system and, 406-407

SQL injection and, 402-403
unique identification and, 410-411

virtual private database and, 1166

Security Assertion Markup Language (SAML), 407

seek times, 433, 435-439, 450-451, 540, 555

select, 363

aggregate functions and, 84-90

attribute specification, 77
basic SQL queries and, 63-74
on multiple relations, 66-71

natural join and, 71-74
null values and, 83-84

privileges and, 143-145, 148
ranking and, 194

rename operation and, 74-75
set membership and, 90-91

set operations and, 79-83
on single relation, 63-65, 63-66

string operations and, 76-79

select all, 65

select distinct, 64-65, 84-85, 91, 125

select-from-where

delete and, 98-100
function/procedure writing and, 174-180

inheritance and, 949-956
insert and, 100-101

join expressions and, 71-74, 87, 113-120

natural joins and, 71-74, 87, 113-120

nested subqueries and, 90-98
transactions and, 651-654

types handling and, 949-963
update and, 101-103

views and, 120-128

selection

comparisons and, 544-545
complex, 545-546

conjunctive, 545-546

disjunctive, 545-546

equivalence and, 582-588
file scans and, 541-544, 550,

552, 570

identifiers and, 546

indices and, 541-544

intraoperation parallelism and, 811

linear search and, 541-542

relational algebra and, 217-219

SQL and,

view maintenance and, 609-610

Semantic Web, 927

semistructured data models, 9, 27

sensitivity, 903

Sequel, 57

sequence associations, 906-907

sequence counters, 1043

sequential-access storage, 431, 436

sequential files, 459

sequential scans, 1153

serializability

blind writes and, 687

concurrency control and, 662, 666-667, 671, 673, 681-690, 693-697, 701-704, 708

conflict, 641-643

distributed databases and, 860-861

isolation and, 648-653

Oracle and, 1181-1182

order of, 644-646

performance tuning and, 1042

PostgreSQL and, 1142-1143

precedence graph and, 644

predicate reads and, 701

in the real world, 650

snapshot isolation and, 693-697

topological sorting and, 644-646

transactions and, 640-646, 648, 650-653

view, 687

serializable schedules, 640

server programming interface (SPI), 1136

server-side scripting, 386-388

server systems

categorization of, 772-773

client-server, 771-772

cloud-based, 777

data servers, 773, 775-777

transaction-server, 773-775

servlets

client-side scripting and, 389-391

example of, 383-384

life cycle and, 385-386

server-side scripting and, 386-388

sessions and, 384-385

support and, 385-386

set clause, 103

set default, 133

set difference, 50, 221-222, 585

set-intersection, 2229

set null, 133

set operations, 79, 83

IBM DB2 and, 1209-1210

intersect, 50, 81-82, 585, 960

nested subqueries and, 90-93

query optimization and, 597

query processing and, 564-565

set comparison and, 91-93

union, 80-81, 220-221, 339, 585

set role, 150

set transactions isolation level

serializable, 649

shadow-copy scheme, 727

shadowing, 441-442

shadow-paging, 727

shared and intention-exclusive (SIX) mode, 680

shared-disk architecture, 781,

783, 789

shared-memory architecture,

781-783

shared-mode locks, 661

shared-nothing architecture,

781, 783-784

shared scans, 614

Sherpa/PNUTS, 866-867

shredding, 1013, 1258-1259

similarity-based retrieval, 919-920, 1079

Simple API for XML (SAX), 1009

Simple Object Access Protocol (SOAP), 1017-1018, 1056, 1249-1250

single lock-manager, 839-840

single-server model, 1092-1093

single-valued attributes, 267-268

site reintegration, 850

skew, 512

attribute-value, 800-801

parallel databases and, 800-801, 805-808, 812, 814, 819

parallel systems and, 780
partitioning and, 560, 800-801

slicing, 201

small-computer-system interconnect (SCSI), 434

snapshot isolation, 652-653, 704, 1042

Microsoft SQL Server and, 1244

recovery systems and, 729-730
serializability and, 693-697

validation and, 692-693

snapshot replication, 1252-1253

snapshots

DML commands and,

1138-1139

Microsoft SQL Server and, 1242

multiversion concurrency

control (MVCC) and, 1137-1146

PostgreSQL and, 1137-1146

read committed, 1242

software RAID, 448

Solaris, 1193

solid-state drives, 430

some, 90, 92, 92n8

sorting, 546

cost analysis of, 548-549

duplicate elimination and,

563-564

external sort-merge algorithm

and, 547-549

parallel external sort-merge

and, 806

PostgreSQL and, 1153

range-partitioning, 805
 topological, 644-646
 XML and, 1106
sort-merge-join, 553
space overhead, 476, 479, 486, 522
spatial data
 computer-aided-design data and, 1061, 1064-1068
 geographic data and, 1061, 1064-1066
 indexing of, 1071-1076
 queries and, 1070-1071
 representation of geometric information and, 1065-1066
 topographical information and, 1070
 triangulation and, 1065
 vector data and, 1069
specialization
 entity-relationship (E-R) model and, 295-296
 partial, 300
 single entity set and, 298
 total, 300
specialty databases, 943
 object-based databases and, 945-975
 XML and, 981-1020
specification of functional requirements, 16, 260
specificity, 903
speedup, 778-780
spider traps, 930
SQL (Structured Query Language), 10, 13-14, 57, 151, 210, 582
 accessing from a programming language, 157-163
 advanced, 157-210
 aggregate functions, 84-90, 192-197
 application-level authorization and, 407-409
 application programs and, 14-15
 array types and, 956-961
 authorization and, 58, 143-150
 basic types and, 59-60
 blobs and, 138, 166, 457, 502, 1013, 1198-1199, 1259
 bulk loads and, 1031-1033

catalogs, 142-143
 clobs and, 138, 166, 457, 502, 1010-1013, 1196-1199
 CLR hosting and, 1254-1256
 create table, 60-63, 141-142
 database modification and, 98-103
 data-definition language (DDL) and, 57-63, 104
 data-manipulation language (DML) and, 57-58, 104
 data mining and, 26
 date/time types in, 136-137
 decision-support systems and, 887-889
 default values and, 137
 delete and, 98-100
 dumping and, 743-744
 dynamic, 58, 158
 embedded, 58, 158, 169-173, 773
 Entity, 395
 environments, 43
 function writing and, 173-180
 IBM DB2 and, 1195-1200, 1210
 index creation and, 137-138, 528-529
 inheritance and, 949-956
 injection and, 402-403
 insert and, 100-101
 integrity constraints and, 58, 128-136
 intermediate, 113-151
 isolation levels and, 648-653
 JDBC and, 158-166
 join expressions and, 71-120 (see also joins)
 lack of fine-grained authorization and, 408-409
 large-type objects, 138
 Management of External Data (MED) and, 1077
 Microsoft SQL Server and, 1223-1267
 multiset types and, 956-961
 MySQL and, 31, 76, 111, 160n3, 1123, 1155
 nested subqueries and, 90-98
 nonstandard syntax and, 178
 null values and, 83-84
 object-based databases and, 945-975
 ODBC and, 166-169
 OLAP and, 197-209

Oracle variations and, 1158-1162
 overview of, 57-58
 persistent programming languages and, 964-972
 PostgreSQL and, 31 (see also PostgreSQL)
 prepared statements and, 162-164
 procedure writing and, 173-180
 query processing and, 537-538 (see also query processing)
 rapid application development (RAD) and, 397
 relational algebra and, 219, 239
 rename operation and, 74-80
 report generators and, 399-400
 ResultSet object and, 159, 161, 164-166, 393, 397-398, 490
 revoking of privileges and, 149-150
 roles and, 145-146
 schemas and, 47, 58-63, 141-143, 147-148
 security and, 402-403
 select clause and, 77
 set operations and, 79-83
 as standard relational database language, 57
 standards for, 1052-1053
 string operations and, 76-77
 System R and, 30, 57
 time specification in, 1063-1064
 transactions and, 58, 127-128, 773 (see also transactions)
 transfer of privileges and, 148-149
 triggers and, 180-187
 tuples and, 77-78 (see also tuples)
 under privilege and, 956
 update and, 101-103
 user-defined types, 138-141
 views and, 58, 120-128, 146-147
 where clause predicates, 78-79
SQLLoader, 1032, 1189
SQL Access Group, 1053
SQL/DS, 30
SQL environment, 143

SQLJ, 172
SQL Plan Management, 1177-1178
SQL Profiler, 1225-1227
SQL Security Invoker, 147
SQL Server Analysis Services (SSAS), 1264, 1266-1267
SQL Server Broker, 1261-1263
SQL Server Integration Services (SSIS), 1263-1266
SQL Server Management Studio, 1223-1224, 1227-1228
SQL Server Reporting Services (SSRS), 1264, 1267
sqlstate, 179
SQL Transparent Data Encryption, 1248
SQL Tuning Advisor, 1176-1177
SQL/XML standard, 1014-1015
Standard Generalized Markup Language (SGML), 981
standards
 ANSI, 57, 1051
 anticipatory, 1051
 Call Level Interface (CLI), 1053
 database connectivity, 1053-1054
 data pump export/import and, 1189
 DBTG CODASYL, 1052
 ISO, 57, 871, 1051
 ODBC, 1053-1055
 reactionary, 1051
 SQL, 1052-1053
 Wi-Max, 1081
 XML, 1055-1056
 X/Open XA, 1053-1054
Starburst, 1193
start-up costs, 780
starvation, 679
Statement object, 161-164
statement triggers, 1161-1162
state transition, 1134
state value, 1134
statistics
 catalog information and, 590-592
 computing, 593
 join size estimation and, 595-596
 maintaining, 593

number of distinct values and, 597-598
 query optimization and, 590-598
 random samples and, 593
 selection size estimation and, 592-595
steal policy, 740
steps, 1096
stop words, 918
storage, 427
 archival, 431
 atomicity and, 632-633
 authorization and, 21
 Automatic Storage Manager and, 1186-1187
 backup, 431, 723, 756-759, 850, 1095-1096
 bit-level striping, 442-444
 buffer manager and, 21 (see also buffers)
 byte amount and, 20
 checkpoints and, 734-735, 742-743
 clob values and, 1010-1011
 cloud-based, 777, 862-863
 column-oriented, 892-893
 content dump and, 743
 cost per bit, 431
 crashes and, 467-468 (see also crashes)
 data access and, 724-726
 data-dictionary, 462-464
 data mining and, 25-26, 893-910
 data-transfer rate and, 435-436
 data warehouses and, 888
 decision-storage systems and, 887-889
 direct-access, 431
 distributed databases and, 826-830
 distributed systems and, 784-788
 dumping and, 743-744
 durability and, 632-633
 error-correcting-code (ECC) organization and, 444-445
 Exadata and, 1187-1188
 file manager and, 21
 file organization and, 451-462
 flash, 403, 430, 439-441, 506
 flat files and, 1009-1010
 force output and, 725-726
 fragmentation and, 826-829

hard disks and, 29-30
 IBM DB2 and, 1200-1203
 indices and, 21 (see also indices)
 information retrieval and, 915-937
 integrity manager and, 21
 jukebox, 431
 magnetic disk, 430, 432-439
 main memory and, 429-430
 Microsoft SQL Server and, 1233-1236
 mirroring and, 441-442, 1245-1246
 native, 1013-1014
 nonrelational data, 1009-1010
 nonvolatile, 432, 632, 722, 724-726, 743-744
 optical, 430-431, 449-450
 Oracle and, 1162-1172, 1186-1188
 parallel systems and, 777-784
 persistent programming languages and, 967-968
 physical media for, 429-432
 PostgreSQL and, 1146-1151
 publishing/shredding data and, 1013, 1258-1259
 punched cards and, 29
 query processor and, 21-22
 recovery systems and, 722-726 (see also recovery systems)
 redundant arrays of independent disks (RAID), 435, 441-449
 relational databases and, 1010-1014
 remote backup systems and, 723, 756-759, 850, 1095-1096
 replication and, 826, 829
 scrubbing and, 448
 seek times and, 433, 435-439, 450-451, 540, 555
 segments and, 1163
 sequential-access, 431, 436
 solid-state drives and, 430
 stable, 632, 722-724
 striping data and, 442-444
 tape, 431, 450-451
 tertiary, 431, 449-451
 transaction manager and, 21 (see also transactions)
 transparency and, 829-830

volatile, 431, 632, 722
 wallets and, 415
 XML and, 1009-1016
storage area network (SAN), 434-435, 789
storage manager, 20-21
string operations
 aggregate, 84
 attribute specification, 77
 escape, 77
 JDBC and, 158-166
 like, 76-77
 lower, 76
 query result retrieval and, 161-162
 similar to, 77
 trim, 76
 tuple display order, 77-78
 upper function, 76
 where predicates, 78-79
striping data, 442-444
structured types, 138-141, 949-952
stylesheets, 380
sublinear speedup, 778-780
submultiset, 960
suffix, 874
sum, 84, 123, 207, 235-236, 566-567, 1134
superclass-subclass relationship, 296-297
superkeys, 45-46, 271-272, 330-333
superuser, 143
Support Vector Machine (SVM), 900-901, 1191
swap space, 742
Swing, 399
Sybase, 1223
symmetric multiprocessors (SMPs), 1193
synonyms, 925-927
sysaux, 1172-1173
system architecture. *See* architectures
system catalogs, 462-464, 1132
system change number (SCN), 1180-1181
system error, 721
System R, 30, 57, 1193

table inheritance, 954-956
tables, 12-13
 filtering and, 1187
 IBM DB2 and, 1200-1203

materialized, 1212-1214
 Microsoft SQL Server and, 1230, 1234
 .NET Common Language Runtime (CLR) and, 1257-1258
 Oracle and, 1163-1166, 1187, 1189
 partitions and, 1169-1172
 relational model and, 39-44, 49-51
 SQL Server Broker and, 1262
tablespaces, 1146, 1172-1173
tag library, 388
tag
 application design and, 378-379, 388, 404
 information retrieval and, 916
 XML and, 982-986, 989, 994, 999, 1004, 1019
tape storage, 431, 450-451
Tapestry, 399
task flow. *See* workflows
Tcl, 180, 1123-1125, 1136
temporal data, 1061
 intervals and, 1063-1064
 query languages and, 1064
 relational databases and, 364-367
 time in databases and, 1062-1064
 timestamps and, 1063-1064
 transaction time and, 1062
temporal relation, 1062-1063
Teradata Purpose-Built Platform Family, 806
term frequency (TF), 918
termination states, 1099
tertiary storage, 431, 449-451
TF-IDF approach, 928-929
theta join, 584-585
third normal form (3NF)
 decomposition algorithms and, 352-355
 relational databases and, 336-337, 352-355
Thomas' write rule, 685-686
thread pooling, 1246
three-phase commit (3PC) protocol, 826
three-tier architecture, 25
throughput
 application development and, 1037, 1045-1046
 defined, 311

harmonic mean of, 1046
 improved, 635-636, 655
 log records and, 1106
 main memories and, 1116
 Microsoft SQL Server and, 1255
 Oracle and, 1159, 1184
 parallel systems and, 778
 performance and, 1110
 range partitioning and, 800
 storage and, 444, 468
 system architectures and, 771, 778, 800, 802, 819
 transactions and, 635-636, 655
timestamps, 136-167
 concurrency control and, 682-686, 703
 distributed databases and, 842-843
 logical counter and, 682
 long-duration transactions and, 1110
 multiversion schemes and, 690-691
 ordering scheme and, 682-685
 rollback and, 685-686
 temporal data and, 1063-1064
 Thomas' write rule and, 685-686
 transactions and, 651-652
 with time zone, 1063
time to completion, 1045
time with time zone, 1063
timezone, 136-137, 1063
Tomcat, 386
top-down design, 297
top-K optimization, 613
topographic information, 1070
topological sorting, 644-646
training instances, 895
transactional replication, 1252-1253
transaction control, 58
transaction coordinator, 830-831, 834-835, 850-852
transaction manager, 21, 23, 830-831
transaction-processing monitors, 1091
 application coordination using, 1095-1096
 architectures of, 1092-1095
 durable queue and, 1094
 many-server, many-router model and, 1094

many-server, single-router model and, 1093
 multitasking and, 1092-1095
 presentation facilities and, 1094-1095
 single-server model and, 1092-1093
 switching and, 1092
Transaction Processing Performance Council (TPC), 1046-1048
transactions, 32, 625, 655-656, 1116
 aborted, 633-634, 647
 actions after crash, 736-738
 active, 633
 advanced processing of, 1091-1116
 association rules and, 904-907
 atomicity and, 22-23, 628, 633-635, 646-648 (see also atomicity)
 availability and, 847-853
 begin/end operations and, 627
 cascadeless schedules and, 647-648
 check constraints and, 628
 cloud computing and, 866-868
 commit protocols and, 832-838
 committed, 127, 633-635, 639, 647, 692-693, 730, 758, 832-838, 1107, 1218
 compensating, 633, 1113-1114
 concept of, 627-629
 concurrency control and, 661-710, 1241-1246 (see also concurrency control)
 consistency and, 22, 627-631, 635-636, 640, 648-650, 655 (see also consistency)
 crashes and, 628
 data mining and, 893-910
 decision-storage systems and, 887-889
 defined, 22, 627
 distributed databases and, 830-832
 durability and, 22-23, 628, 633-635 (see also durability)
 E-commerce and, 1102-1105
 failure of, 633, 721-722

force/no-force policy and, 739-740
 global, 784, 830, 860-861
 integrity constraint violation and, 133-134
 isolation and, 628, 635-640, 646-653 (see also isolation)
 killed, 634
 local, 784, 830, 860-861
 locks and, 661-669, 661-686 (see also locks)
 log records and, 726-728, 730-734
 long-duration, 1109-1115
 main-memory databases and, 1105-1108
 multidatabases and, 860-861
 multilevel, 1112-1113
 multitasking and, 1092-1095
 multiversion concurrency control (MVCC) and, 1137-1146
 multiversion schemes and, 689-692
 object-based databases and, 945-975
 observable external writes and, 634-635
 parallel databases and, 797-820
 performance tuning and, 1041-1044
 persistent messaging and, 836-837
 persistent programming languages and, 970
 person-in-the-middle attacks and, 1105
 PostgreSQL and, 1137-1146
 read/write operations and, 653-654
 real-time systems and, 1108-1109
 recoverable schedules and, 647
 recovery manager and, 22-23
 recovery systems and, 631, 633 (see also recovery systems)
 remote backup systems and, 756-759
 restart of, 634
 rollback and, 127, 736, 746-749, 754-755

serializability and, 640-653
 shadow-copy scheme and, 727
 simple model for, 629-631
 SQL Server Broker and, 1261-1263
 as SQL statements, 653-654
 starved, 666
 states of, 633-635
 steal/no-steal policy and, 740
 storage structure and, 632-633
 timestamps and, 682-686
 two-phase commit protocol (2PC) and, 786-788
 uncommitted, 648
 as unit of program, 627
 validation and, 686-689
 wait-for graph and, 676-678
 workflows and, 836-838, 1096-1102
 write-ahead logging (WAL) rule and, 739-740, 739-741
transaction scaleup, 779
transactions-consistent snapshot, 843-844
transaction-server systems, 773-775
transactions per second (TPS), 1046-1047
transaction time, 365n8, 1062
TransactSQL, 173
transfer of control, 757
transfer of prestige, 921-922
transformations
 equivalence rules and, 583-586
 examples of, 586-588
 join ordering and, 588-589
 query optimization and, 582-590
 relational algebra and, 582-590
 XML and, 998-1008
transition tables, 183-184
transition variable, 181
transitive closure, 188-190
transitivity rule, 339-340
transparency, 829-830, 854-855
trees, 1086
 B, 504-506, 530, 1039, 1064, 1071-1072, 1076, 1086, 1135, 1148-1150, 1159, 1164-1169, 1173, 1205
 B+, 12-34-1235 (see also B+-trees)

decision-tree classifiers and, 895-900
 directory information (DIT), 872-875
 distributed directory, 874-875
 Generalized Search Tree (GiST) and, 1148-1149
 index-organized tables (IOTs) and, 1164-1165
 k-d, 1071-1072
 multiple granularity and, 679-682
 Oracle and, 1164-1165, 1191
 overfitting and, 899-900
 PostgreSQL and, 1148-1149
 quadratic split and, 1075-1076
 quadrees, 1069, 1072-1073
 query optimization and, 814-815 (see also query optimization)
 R, 1073-1076
 scheduling and, 814-815
 spatial data support and, 1064-1076
 XML, 998, 1011

triggers
 alter, 185
 disable, 185
 drop, 185
 IBM DB2 and, 1210
 Microsoft SQL Server and, 1232-1233
 need for, 180-181
 nonstandard syntax and, 184
 Oracle and, 1161-1162
 PostgreSQL and, 1153-1154
 recovery and, 186
 in SQL, 181-187
 transition tables and, 183-184
 when not to use, 186-187

true negatives, 903
true predicate, 67
true relation, 90, 93
tuple ID, 1147-1148
tuple relational calculus, 239, 249
 example queries, 240-242
 expressive power of languages, 244
 formal definition, 243
 safety of expressions, 244

tuples, 40-42
 aggregate functions and, 84-90
 Cartesian product and, 50

delete and, 98-100
 domain relational calculus and, 245-248
 duplicate, 94-95
 eager generation of, 569-570
 insert and, 100-101
 joins and, 550-553 (see also joins)
 lazy generation of, 570-571
 ordering display of, 77-78
 parallel databases and, 797-820
 pipelining and, 568-572
 PostgreSQL and, 1137-1146
 query structures and, 68
 query optimization and, 579-616
 query processing and, 537-573
 ranking and, 192-195
 relational algebra and, 217-239, 582-590
 set operations and, 79-83
 update and, 101-103
 views and, 120-128
 windowing and, 195-197

tuple visibility, 1139
two-factor authentication, 405-407
two-phase commit (2PC) protocol, 786-788, 832-836
two-tier architecture, 24-25

types, 1017, 1159
 abstract data, 1127
 array, 956-961
 base, 1127
 blob, 138, 166, 457, 502, 1013, 1198-1199, 1259
 clob, 138, 166, 457, 502, 1010-1013, 1196-1199
 complex data, 946-949 (see also complex data types)
 composite, 1127
 document type definition (DTD), 990-994
 enumerated, 1128
 IBM DB2 and, 1196-1197
 inheritance and, 949-956
 Microsoft SQL Server and, 1229-1230, 1257-1258
 most-specific, 953
 multiset, 956-961
 .NET Common Language Runtime (CLR) and, 1257-1258

nonstandard, 1129-1130
 object-based databases and, 949-963
 object-identity, 961-963
 Oracle and, 1158-1160
 performance tuning and, 1043
 polymorphic, 1128-1129
 PostgreSQL, 1126-1129, 1132-1133
 pseudotypes, 1128
 reference, 961-963
 user-defined, 138-141
 single reference, 972
 wide-area, 788-791
 XML, 990-998, 1006-1007

UDF. See user-defined functions

Ultra320 SCSI interface, 436

Ultrium format, 451

under privilege, 956

undo

concurrency control and, 749-750
 logical operations and, 745-750
 Oracle and, 1163
 recovery systems and, 729-738
 transaction rollback and, 746-749

undo pass, 754

undo phase, 737

Unified Modeling Language (UML), 17-18

associations and, 308-309
 cardinality constraints and, 309-310

components of, 308

relationship sets and, 308-309

uniform resource locators (URLs), 377-378

union, 80-81, 585, 220-221

union all, 80

union rule, 339

unique, 94-95

decomposition and, 354-355
 integrity constraints and, 130-131

uniquifier, 498-499

United States, 17, 45, 263, 267n3, 411, 788, 858, 869, 922

Universal Coordinated Time (UTC), 1063

Universal Description, Discovery, and Integration (UDDI), 1018
Universal Serial Bus (USB)
 slots, 430

universal Turing machine, 14
universities, 2

application design and, 375, 392, 407-409
 concurrency control and, 698
 database design and, 16-17
 databases for, 3-8, 11-12, 15-19, 27, 30

E-R model and, 261-274, 280, 282, 292, 294-299

indexing and, 477, 510, 529

query optimization and, 586, 589, 605

query processing and, 566

recovery system and, 724

relational database design and, 323-330, 334, 355, 364-365

relational model and, 41, 43-48

SQL and, 61-63, 70-72, 75, 99, 125-134, 145-150, 153, 170, 173, 187, 192-193, 197, 226-227

storage and, 452, 458, 460

system architecture and, 785, 828, 872

transactions and, 653

University of California, Berkeley, 30, 1123

Unix, 77, 438, 713, 727, 1124, 1154, 1193-1194, 1212, 1223

unknown, 83, 90

unnesting, 958-961

updatable result sets, 166

update-anywhere replication, 844

updates, 101-103

authorization and, 147, 148

B+-trees and, 491-497, 499-500

batch, 1030-1031

complexity of, 499-500

concurrency control and, 867-868

data warehouses and, 891

deletion time and, 491, 495-500

distributed databases and, 826-827

EXEC SQL and, 171

hashing and, 516-522

indices and, 482-483

insertion time and, 491-495, 499-500

log records and, 726-734

lost, 692

Microsoft SQL Server and, 1232-1233, 1239

mobile, 1083-1084

Oracle and, 1179-1180

performance tuning and, 1030-1033, 1043-1044

persistent programming languages and, 970

PostgreSQL and, 1130, 1141-1144, 1147-1148

privileges and, 143-145

query optimization and, 613-614

shipping SQL statements to database, 161

snapshot isolation and, 692-697

triggers and, 182, 184

views and, 124-128

XML and, 1259-1260

user-defined entity sets, 299

user-defined functions (UDFs), 1197-1198

user-defined types, 138-141

user interfaces, 27-28

application architectures and, 391-396

application programs and, 375-377

as back-end component, 376

business-logic layer and, 391-392

client-server architecture and, 32, 204, 376-377, 756-772, 777, 788, 791

client-side scripting and, 389-391

cloud computing and, 861-870

common gateway interface (CGI), 380-381

cookies and, 382-385, 403-405

CRUD, 399

data access layer and, 391, 393, 395

disconnected operation and, 395-396

HyperText Transfer Protocol (HTTP) and, 377-383, 395, 404-406, 417

IBM DB2, 1195

mobile, 1079-1085

persistent programming languages and, 970

PostgreSQL and, 1124-1126

presentation layer and, 391

report generators and, 399-400

security and, 402-417

storage and, 434 (see also storage)

tools for building, 396-398

Web services and, 395

World Wide Web and, 377-382

user requirements, 15-16, 27-28

E-R model and, 260, 298

performance and, 311-312

response time and, 311

throughput and, 311

using, 114

utilization, 636

vacuum, 1143

validation, 703-704

classifiers and, 903-904

concurrency control and, 686-689

first committer wins and, 692-693

first updater wins and, 693

long-duration transactions and, 1111

phases of, 688

recovery systems and, 729-730

snapshot isolation and, 692-693

view serializability and, 687

valid time, 365

varchar, 59-60, 62

VBScript, 387

vector data, 1069

vector space model, 919-920

version-numbering schemas, 1083-1084

version-vector scheme, 1084

vertical fragmentation, 828

video servers, 1078

view definition, 58

view maintenance, 608-611

views, 120

authorization on, 146-147

with check option, 126

complex merging and, 1173-1174
 create view and, 121-125
 cube, 1221-1222
 deferred maintenance and, 1039-1040
 definition, 121-122
 delete, 125
 immediate maintenance and, 1039-1040
 insert into, 124-125
 maintenance, 124
 materialized, 123-124 (see also materialized views)
 performance tuning and, 1039-1041
 SQL queries and, 122-123
 update of, 124-128
 view serializability, 687
 virtual machines, 777
 virtual processor, 801
 Virtual Reality Markup Language (VRML), 390-391
 Visual Basic, 169, 180, 397-398, 1228
 VisualWeb, 397
 volatile storage, 431, 632, 722

wait-for graph, 676-678, 845-847
 Web crawlers, 930-931
 Weblogic, 386
 WebObjects, 399
 Web servers, 380-382
 Web services, 395, 1199-1200
 Web Services Description Language (WSDL), 1018
 WebSphere, 386
 when clause, 181, 184
 where clause, 311
 aggregate functions and, 84-90
 basic SQL queries and, 63-74
 between, 78
 on multiple relations, 66-71
 natural join and, 71-74
 not between, 78
 null values and, 83-84
 query optimization and, 605-607
 rename operation and, 74-75
 security and, 409
 set operations and, 79-83
 on single relation, 63-65, 63-66
 string operations and, 76-79
 transactions and, 651-654
 while loop, 168, 171, 176
 wide-area networks (WANs), 788, 790-791
 Wi-Max, 1081
 windowing, 195-197
 Windows Mobile, 1223
 Wireless application protocol (WAP), 1081-1082
 wireless communications, 1079-1085
 with check option, 126
 with clause, 97, 190
 with data, 141-142
 with grant option, 148
 with recursive clause, 190
 with timezone, 136
 WordNet, 927
 workflows, 312-313, 836-838, 1017
 acceptable termination states and, 1099
 bugs and, 1101
 business-logic layer and, 391-392
 execution and, 1097-1101
 external variables and, 1098-1099
 failures and, 1099-1102
 management systems for, 1101-1102
 multisystem applications and, 1096
 nonacceptable termination states and, 1099
 performance and, 1029-1048
 recovery of, 1101
 specification and, 1097-1099
 steps and, 1096
 tasks and, 1096
 transactional, 1096-1102
 workload compression, 1041
 World Wide Web, 31, 885
 application design and, 377-382
 cookies and, 382-385, 403-405
 encryption and, 411-417
 HyperText Markup Language (HTML), 378-380
 HyperText Transfer Protocol (HTTP) and, 377-381, 383, 395, 404-406, 417
 information retrieval and, 915 (see also information retrieval)

security and, 402-417
 services processing and, 395
 Simple Object Access Protocol (SOAP) and, 1017-1018
 three-layer architecture and, 318
 uniform resource locators (URLs), 377-378
 Web application frameworks and, 398-400
 Web servers and, 380-382
 XML and, 1017-1018
 World Wide Web Consortium (W3C), 927, 1056
 wrapping, 1055-1056
 write-ahead logging (WAL), 739-7841, 1145-1146
 write once, read-many (WORM) disks, 431
 write quorum, 841-842
 write-write contention, 1042

X.500 directory access protocol, 871

XML (Extensible Markup Language), 31, 169, 386, 1020
 application program interfaces (APIs) to, 1008-1009
 applications, 1016-1019
 clob values and, 1010-1011
 data exchange formats and, 1016-1017
 data mediation and, 1018-1019
 data structure, 986-990
 document schema, 990-998
 document type definition (DTD), 990-994
 as dominant format, 985
 file processing and, 981-982
 format flexibility of, 985
 HTML and, 981
 IBM DB2 and, 1195-1196
 joins and, 1003-1004
 markup concept and, 981-985
 Microsoft SQL Server and, 1258-1261
 nesting and, 27, 943, 984-990, 995-998, 1001, 1004-1007, 1010
 Oracle XML DB and, 1159-1160

publishing/shredding data and, 1013
 queries and, 998-1008, 1259-1260
 relational databases and, 1010-1014
 relational maps and, 1012
 Simple Object Access Protocol (SOAP) and, 1017-1018
 sorting and, 1006
 SQL/XML standard and, 1014-1015
 standards for, 1055-1056
 storage and, 1009-1016
 tags and, 982-986, 989, 994, 999, 1004, 1019
 textual context and, 986
 transformation and, 998-1008
 tree model of, 998

updates and, 1259-1260
 web services and, 1017-1018
 wrapping and, 1055-1056
 xmlagg, 1015
 xmlattributes, 1015
 xmlconcat, 1015
 xmlelement, 1015
 xmlforest, 1015
 XMLIndex, 1160
 XML Schema, 994-998
 XMLType, 1159
 X/Open XA standards, 1053-1054
 XOR operation, 413
 XPath, 1160
 document schema and, 997
 queries and, 998-1002
 storage and, 1009-1015

XQuery, 31, 998
 FLWOR expressions and, 1002-1003
 functions and, 1006-1007
 joins and, 103-104
 Microsoft SQL Server and, 1260-1261
 nested queries and, 1004-1005
 Oracle and, 1160
 sorting of results and, 1006
 storage and, 1009-1015
 transformations and, 1002-1008
 types and, 1006-1007
 XSLT, 1160

Yahoo, 390, 863
 YUI library, 390