

4744

Database Management Systems



Raghu Ramakrishnan

DATABASE MANAGEMENT SYSTEMS

Raghu Ramakrishnan

University of Wisconsin

Madison, WI, USA



Boston, Massachusetts • Burr Ridge, Illinois
Dubuque, Iowa • Madison, Wisconsin
New York, New York • San Francisco, California
St. Louis, Missouri

CONTENTS

PREFACE	xix
1 INTRODUCTION TO DATABASE SYSTEMS	1
1.1 Overview	2
1.2 File Systems versus a DBMS	4
1.3 Describing and Storing Data in a DBMS	5
1.3.1 The Relational Model	5
1.3.2 Levels of Abstraction in a DBMS	7
1.3.3 Data Independence	9
1.4 Queries in a DBMS	10
1.5 Concurrent Access and Crash Recovery	11
1.5.1 Points to Note	13
1.6 Structure of a DBMS	14
1.7 Advantages of a DBMS	15
1.8 People Who Deal with Databases	16
1.9 Summary	18
2 THE RELATIONAL MODEL	21
2.1 Relations	22
2.1.1 Creating and Modifying Relations Using SQL-92	24
2.2 Integrity Constraints	25
2.2.1 Key Constraints	26
2.2.2 Foreign Key Constraints	28
2.2.3 General Constraints	30
2.3 Enforcing Integrity Constraints	31
2.4 Query Languages	32
2.5 Summary	34
3 STORING DATA: DISKS AND FILES	38
3.1 The Memory Hierarchy	39
3.1.1 Disks	40
3.1.2 Performance Implications of Disk Structure	41
3.2 Disk Space Management	42

3.2.1	Keeping Track of Free Blocks	43
3.2.2	Using OS File Systems to Manage Disk Space	43
3.3	Buffer Manager	44
3.3.1	Buffer Replacement Policies	46
3.3.2	Buffer Management in DBMS versus OS	48
3.4	Record Formats *	49
3.4.1	Fixed-Length Records	49
3.4.2	Variable-Length Records	50
3.5	Page Formats *	51
3.5.1	Fixed-Length Records	52
3.5.2	Variable-Length Records	53
3.6	Files and Indexes	55
3.6.1	Heap Files *	55
3.6.2	Introduction to Indexes	57
3.7	System Catalogs in a Relational DBMS	59
3.8	Summary	61
4	FILE ORGANIZATIONS AND INDEXES	67
4.1	Cost Model	68
4.2	Comparison of Three File Organizations	69
4.2.1	Heap Files	69
4.2.2	Sorted Files	70
4.2.3	Hashed Files	72
4.2.4	Choosing a File Organization	73
4.3	Overview of Indexes	74
4.3.1	Alternatives for Data Entries in an Index	75
4.4	Properties of Indexes	76
4.4.1	Clustered versus Unclustered Indexes	76
4.4.2	Dense versus Sparse Indexes	78
4.4.3	Primary and Secondary Indexes	79
4.4.4	Indexes Using Composite Search Keys	80
4.5	Index Specification in SQL-92	81
4.6	Summary	81
5	TREE-STRUCTURED INDEXING	84
5.1	Indexed Sequential Access Method (ISAM) *	85
5.2	B+ Trees: A Dynamic Index Structure	90
5.3	Format of a Node	92
5.4	Search	92
5.5	Insert	93
5.6	Delete *	97
5.7	Duplicates *	102
5.8	B+ Trees in Practice *	103

5.8.1	Key Compression	103
5.8.2	Bulk-Loading a B+ Tree	104
5.8.3	The Order Concept	107
5.8.4	The Effect of Inserts and Deletes on Rids	108
5.9	Multidimensional Indexes	108
5.10	Summary	110
6	HASH-BASED INDEXING	116
6.1	Static Hashing	117
6.1.1	Notation and Conventions	118
6.2	Extendible Hashing *	118
6.3	Linear Hashing *	124
6.4	Extendible Hashing Versus Linear Hashing *	130
6.5	Summary	131
7	EXTERNAL SORTING	137
7.1	A Simple Two-Way Merge Sort	138
7.2	External Merge Sort	139
7.2.1	Minimizing the Number of Runs *	143
7.3	Minimizing I/O Cost versus Number of I/Os	145
7.3.1	Blocked I/O *	145
7.3.2	Double Buffering *	147
7.4	Using B+ Trees for Sorting	148
7.4.1	Clustered Index	148
7.4.2	Unclustered Index	149
7.5	Summary	151
8	RELATIONAL ALGEBRA AND CALCULUS	154
8.1	Preliminaries	154
8.2	Relational Algebra	155
8.2.1	Selection and Projection	156
8.2.2	Set-Operations	157
8.2.3	Renaming	159
8.2.4	Joins	160
8.2.5	Division	162
8.2.6	More Examples of Relational Algebra Queries	163
8.3	Relational Calculus	167
8.3.1	Tuple Relational Calculus	168
8.3.2	Domain Relational Calculus	172
8.4	Expressive Power of Algebra and Calculus *	175
8.5	Summary	177
9	SQL: THE QUERY LANGUAGE	181

9.1	The Form of a Basic SQL Query	183
9.1.1	Examples of Basic SQL Queries	185
9.1.2	Expressions and Strings in the SELECT Command	187
9.2	UNION, INTERSECT and EXCEPT	188
9.3	Nested Queries	191
9.3.1	Introduction to Nested Queries	192
9.3.2	Correlated Nested Queries	193
9.3.3	Set-Comparison Operators	194
9.3.4	More Examples of Nested Queries	195
9.4	Aggregate Operators	196
9.4.1	The GROUP BY and HAVING Clauses	199
9.4.2	More Examples of Aggregate Queries	202
9.5	Null Values *	205
9.5.1	Comparisons Using Null Values	205
9.5.2	Logical Connectives AND, OR and NOT	206
9.5.3	Impact on SQL Constructs	206
9.5.4	Outer Joins	207
9.5.5	Disallowing Null Values	208
9.6	Embedded SQL *	208
9.6.1	Declaring Variables and Exceptions	209
9.6.2	Embedding SQL Statements	210
9.7	Cursors *	211
9.7.1	Basic Cursor Definition and Usage	211
9.7.2	Properties of Cursors	213
9.8	Dynamic SQL *	215
9.9	Queries in Complex Integrity Constraints *	216
9.9.1	Constraints over a Single Table	216
9.9.2	Domain Constraints	216
9.9.3	Assertions: ICs over Several Tables	217
9.10	Summary	218
10	SECURITY, VIEWS, AND SQL	226
10.1	Introduction to Database Security	227
10.2	Views	227
10.2.1	Destroying/Altering Tables and Views	228
10.2.2	Queries on Views	229
10.2.3	Updates on Views *	230
10.3	Access Control	232
10.4	Discretionary Access Control *	232
10.4.1	Grant and Revoke on Views and Integrity Constraints	240
10.5	Mandatory Access Control *	242
10.5.1	Multilevel Relations and Polyinstantiation	243
10.5.2	Covert Channels, DOD Security Levels	245
10.6	Additional Issues Related to Security *	246

10.6.1	Role of the Database Administrator	246
10.6.2	Security in Statistical Databases	246
10.6.3	Encryption	248
10.7	Summary	251
11	QUERY-BY-EXAMPLE (QBE)	255
11.1	Introduction	255
11.2	Basic QBE Queries	256
11.2.1	Other Features: Duplicates, Ordering Answers *	257
11.3	Queries Over Multiple Relations	258
11.4	Negation in the Relation-name Column	259
11.5	Aggregates	260
11.6	The Conditions Box	261
11.6.1	And/Or Queries *	262
11.7	Unnamed Columns	263
11.8	Updates	264
11.8.1	Restrictions on Update Commands	265
11.9	Division and Relational Completeness *	266
11.10	Summary	267
12	EVALUATION OF RELATIONAL OPERATORS	271
12.1	Introduction to Query Processing	272
12.1.1	Access Paths	272
12.1.2	Preliminaries: Examples and Cost Calculations	273
12.2	The Selection Operation	274
12.2.1	No Index, Unsorted Data	274
12.2.2	No Index, Sorted Data*	274
12.2.3	B+ Tree Index	275
12.2.4	Hash Index, Equality Selection	276
12.3	General Selection Conditions *	277
12.3.1	CNF and Index Matching	277
12.3.2	Selections without Disjunction	279
12.3.3	Selections with Disjunction	280
12.4	The Projection Operation *	280
12.4.1	Projection Based on Sorting	281
12.4.2	Projection Based on Hashing	282
12.4.3	Sorting versus Hashing for Projections	284
12.4.4	Use of Indexes for Projections	285
12.5	The Join Operation	285
12.5.1	Nested Loops Join	286
12.5.2	Sort-Merge Join *	291
12.5.3	Hash Join *	295
12.5.4	General Join Conditions *	300

12.6	The Set Operations *	301
12.6.1	Sorting for Union and Difference	301
12.6.2	Hashing for Union and Difference	302
12.7	Aggregate Operations *	302
12.7.1	Implementing Aggregation by Using an Index	303
12.8	The Impact of Buffering	304
12.9	Summary	305
13	RELATIONAL QUERY OPTIMIZATION	310
13.1	Overview of Relational Query Optimization	312
13.1.1	Query Evaluation Plans	312
13.1.2	Pipelined Evaluation	313
13.1.3	The Iterator Interface for Operators and Access Methods	315
13.1.4	The System R Optimizer	316
13.2	Translating SQL Queries into Algebra	316
13.2.1	Decomposition of a Query into Blocks	316
13.2.2	A Query Block as a Relational Algebra Expression	317
13.3	Alternative Plans: A Motivating Example	319
13.3.1	Pushing Selections	319
13.3.2	Using Indexes	321
13.4	Estimating the Cost of a Plan	324
13.4.1	Statistics Maintained by a DBMS *	325
13.4.2	Estimating Result Sizes *	326
13.5	Relational Algebra Equivalences *	327
13.5.1	Selections	328
13.5.2	Projections	328
13.5.3	Cross-Products and Joins	329
13.5.4	Selects, Projects and Joins	330
13.5.5	Others	331
13.6	Enumeration of Alternative Plans *	332
13.6.1	Single-Relation Queries	332
13.6.2	Multiple-Relation Queries	336
13.7	Nested Subqueries *	344
13.8	Other Approaches to Query Optimization *	346
13.9	Summary	347
14	CONCEPTUAL DESIGN AND THE ER MODEL	357
14.1	Overview of Database Design	357
14.2	The Entity-Relationship (ER) Data Model	358
14.2.1	Entities, Attributes and Entity Sets	359
14.2.2	Relationships and Relationship Sets	360
14.3	Additional Features of the ER Model	364
14.3.1	Key Constraints	364

14.3.2	Participation Constraints	367
14.3.3	Weak Entities *	370
14.3.4	Class Hierarchies *	372
14.3.5	Aggregation *	374
14.4	Conceptual Design Using the ER Model	376
14.4.1	Entity versus Attribute	377
14.4.2	Entity versus Relationship	379
14.4.3	Binary versus Ternary Relationships *	380
14.4.4	Aggregation versus Ternary Relationships *	384
14.4.5	Constraints beyond the ER Model *	384
14.4.6	Need for Further Refinement	386
14.5	Conceptual Design for Large Enterprises	387
14.6	Summary	388
15	SCHEMA REFINEMENT AND NORMAL FORMS	395
15.1	Introduction to Schema Refinement	396
15.1.1	Problems Caused by Redundancy	396
15.1.2	Use of Decompositions	398
15.1.3	Problems Related to Decomposition	399
15.2	Functional Dependencies	400
15.3	Examples Motivating Schema Refinement	401
15.3.1	Constraints on an Entity Set	401
15.3.2	Constraints on a Relationship Set	402
15.3.3	Identifying Attributes of Entities	402
15.3.4	Identifying Entity Sets	404
15.4	Reasoning About Functional Dependencies	405
15.4.1	Closure of a Set of FDs	406
15.4.2	Attribute Closure	407
15.5	Normal Forms	408
15.5.1	Boyce-Codd Normal Form	409
15.5.2	Third Normal Form	410
15.6	Decompositions	412
15.6.1	Lossless-Join Decomposition	413
15.6.2	Dependency-Preserving Decomposition	415
15.7	Normalization	416
15.7.1	Decomposition into BCNF	416
15.7.2	Decomposition into 3NF *	419
15.8	Other Kinds of Dependencies *	423
15.8.1	Multivalued Dependencies	423
15.8.2	4NF	426
15.8.3	Join Dependencies	428
15.8.4	5NF	428
15.8.5	Inclusion Dependencies	429
15.9	Summary	429

16	PHYSICAL DATABASE DESIGN AND TUNING	436
16.1	Introduction to Physical Database Design	437
16.1.1	Database Workloads	437
16.1.2	Physical Design and Tuning Decisions	438
16.1.3	Need for Database Tuning	439
16.2	Guidelines for Index Selection	439
16.3	Basic Examples of Index Selection	441
16.4	Clustering and Indexing *	444
16.4.1	Co-clustering Two Relations	447
16.5	Indexes on Multiple-Attribute Search Keys *	449
16.6	Indexes that Enable Index-Only Plans *	450
16.7	Overview of Database Tuning	453
16.7.1	Tuning Indexes	453
16.7.2	Tuning the Conceptual Schema	455
16.7.3	Tuning Queries and Views	456
16.8	Choices in Tuning the Conceptual Schema *	457
16.8.1	Settling for a Weaker Normal Form	457
16.8.2	Denormalization	458
16.8.3	Choice of Decompositions	458
16.8.4	Vertical Decomposition	460
16.8.5	Horizontal Decomposition	460
16.9	Choices in Tuning Queries and Views *	461
16.10	Impact of Concurrency *	463
16.11	DBMS Benchmarking *	465
16.11.1	Well-Known DBMS Benchmarks	465
16.11.2	Using a Benchmark	466
16.12	Summary	467
17	CONCURRENCY CONTROL	476
17.1	The Concept of a Transaction	477
17.1.1	Concurrent Execution and Database Consistency	477
17.1.2	Incomplete Transactions and Crash Recovery	478
17.2	Transactions and Schedules	479
17.3	Notions of Consistency	480
17.3.1	Serializability	480
17.3.2	Some Anomalies Associated with Interleaved Execution *	481
17.3.3	Schedules Involving Aborted Transactions	484
17.4	Lock-Based Concurrency Control	485
17.4.1	Strict Two-Phase Locking (Strict 2PL)	485
17.4.2	2PL, Serializability, and Recoverability	488
17.4.3	View Serializability *	490
17.5	Lock Management *	491
17.5.1	Implementing Lock and Unlock Requests	491
17.5.2	Deadlocks	493

17.5.3	Performance of Lock-Based Concurrency Control	496
17.6	Specialized Locking Techniques *	497
17.6.1	Dynamic Databases and the Phantom Problem	498
17.6.2	Concurrency Control in Tree Indexes	499
17.6.3	Multiple-Granularity Locking	502
17.7	Transaction Support in SQL-92 *	504
17.7.1	Transaction Characteristics	504
17.7.2	Transactions and Constraints	506
17.8	Concurrency Control Without Locking *	507
17.8.1	Optimistic Concurrency Control	507
17.8.2	Timestamp-Based Concurrency Control	509
17.8.3	Multiversion Concurrency Control	512
17.9	Summary	512
18	CRASH RECOVERY	519
18.1	Introduction to Crash Recovery	519
18.1.1	Stealing Frames and Forcing Pages	520
18.1.2	Recovery-Related Steps during Normal Execution	521
18.2	Overview of ARIES	522
18.2.1	The Log	523
18.2.2	Other Recovery-Related Data Structures	526
18.2.3	The Write-Ahead Log Protocol	527
18.2.4	Checkpointing	528
18.3	Recovering From a System Crash *	529
18.3.1	Analysis Phase	529
18.3.2	Redo Phase	531
18.3.3	Undo Phase	533
18.4	Media Recovery *	536
18.5	Other Algorithms and Interaction with Concurrency Control *	537
18.6	Summary	538
19	PARALLEL AND DISTRIBUTED DATABASES	544
19.1	Architectures for Parallel Databases	545
19.2	Parallel Query Evaluation	547
19.2.1	Data Partitioning	548
19.2.2	Parallelizing Sequential Operator Evaluation Code	548
19.3	Parallelizing Individual Operations	549
19.3.1	Bulk Loading and Scanning	549
19.3.2	Sorting	549
19.3.3	Joins	550
19.4	Parallel Query Optimization	553
19.5	Introduction to Distributed Databases	554
19.5.1	Types of Distributed Databases	554

19.6	Distributed DBMS Architectures	555
19.6.1	Client-Server Systems	555
19.6.2	Collaborating Server Systems	556
19.7	Storing Data in a Distributed DBMS	556
19.7.1	Fragmentation	556
19.7.2	Replication	558
19.8	Distributed Catalog Management	558
19.8.1	Naming Objects	558
19.8.2	Catalog Structure	559
19.8.3	Distributed Data Independence	560
19.9	Distributed Query Processing	561
19.9.1	Nonjoin Queries in a Distributed DBMS	561
19.9.2	Joins in a Distributed DBMS	562
19.9.3	Cost-Based Query Optimization	566
19.10	Updating Distributed Data	567
19.10.1	Synchronous Replication	567
19.10.2	Asynchronous Replication	568
19.11	Introduction to Distributed Transaction Management	571
19.12	Distributed Concurrency Control	572
19.12.1	Distributed Deadlock	573
19.13	Distributed Recovery	575
19.13.1	Normal Execution and Commit Protocols	575
19.13.2	Restart after a Failure	576
19.13.3	Two-Phase Commit Revisited	578
19.13.4	Three-Phase Commit	579
19.14	Summary	580
20	DEDUCTIVE AND ACTIVE DATABASES	589
20.1	Introduction to Recursive Queries	590
20.1.1	Datalog	591
20.1.2	The Fixpoint Operator	594
20.2	Recursive Queries with Negation	595
20.2.1	Stratification	596
20.2.2	Aggregate Operations	599
20.3	Efficient Evaluation of Recursive Queries	600
20.3.1	Fixpoint Evaluation without Repeated Inferences	600
20.3.2	Pushing Selections to Avoid Irrelevant Inferences	602
20.4	Introduction to Active Databases	604
20.4.1	Examples of Triggers in SQL	605
20.5	Designing Active Databases	607
20.5.1	Why Triggers Can Be Hard to Understand	607
20.5.2	Constraints versus Triggers	607
20.5.3	Other Uses of Triggers	608
20.6	Summary	609

21	OBJECT-DATABASE SYSTEMS	614
21.1	Motivating Example	616
21.1.1	New Data Types	616
21.1.2	Manipulating the New Kinds of Data	617
21.2	User-Defined Abstract Data Types	619
21.2.1	Defining Methods of an ADT	620
21.3	Constructed Types	622
21.3.1	Manipulating Data of Constructed Types	623
21.4	Objects, Object Identity, and Reference Types	626
21.4.1	Notions of Equality	626
21.4.2	Dereferencing Reference Types	627
21.5	Inheritance	627
21.5.1	Defining Types with Inheritance	628
21.5.2	Binding of Methods	628
21.5.3	Collection Hierarchies, Type Extents, and Queries	629
21.6	Database Design for an ORDBMS	630
21.6.1	Constructed Types and ADTs	630
21.6.2	Object Identity	633
21.6.3	Extending the ER Model	634
21.6.4	Using Nested Collections	636
21.7	New Challenges in Implementing an ORDBMS	636
21.7.1	Storage and Access Methods	637
21.7.2	Query Processing	639
21.7.3	Query Optimization	640
21.8	Comparing RDBMS with OODBMS and ORDBMS	642
21.8.1	OODBMS: ODL and OQL	642
21.8.2	RDBMS versus ORDBMS	643
21.8.3	OODBMS versus ORDBMS: Similarities	643
21.8.4	OODBMS versus ORDBMS: Differences	643
21.9	Summary	644
22	DECISION SUPPORT SYSTEMS	650
22.1	Introduction to Decision Support	650
22.2	Data Warehousing	651
22.2.1	Creating and Maintaining a Warehouse	652
22.3	OLAP	654
22.3.1	Multidimensional Data Model	654
22.3.2	OLAP Queries	657
22.3.3	Database Design for OLAP	661
22.4	Implementation Techniques for OLAP	662
22.4.1	Indexing Methods	662
22.4.2	File Organizations	664
22.4.3	Additional OLAP Implementation Issues	664
22.5	Data Mining	665

22.5.1	Introduction to Mining for Rules	666
22.5.2	The Use of Association Rules for Prediction	668
22.5.3	An Algorithm for Finding Association Rules	669
22.5.4	Generalized Association Rules	671
22.5.5	Sequential Patterns	673
22.5.6	Classification Rules	674
22.5.7	Additional Data Mining Tasks	675
22.6	Summary	677
23	ADDITIONAL TOPICS	682
23.1	Advanced Transaction Processing	682
23.1.1	Transaction Processing Monitors	682
23.1.2	New Transaction Models	683
23.1.3	Real-Time DBMSs	684
23.2	Networked Databases	684
23.2.1	World Wide Web	684
23.2.2	Integrated Access to Multiple Data Sources	686
23.2.3	Mobility and Databases	686
23.3	Main Memory Databases	687
23.4	Information Visualization	687
23.5	Domain-Specific DBMS Issues	688
23.5.1	Geographic Information Systems	688
23.5.2	Temporal and Sequence Databases	688
23.5.3	Image and Video Databases	689
23.5.4	Information Retrieval and Text Databases	689
23.6	Summary	690
A	THE MINIBASE SOFTWARE	692
A.1	What's Available	692
A.2	Overview of Minibase Assignments	693
A.2.1	Overview of Programming Projects	693
A.2.2	Overview of Nonprogramming Assignments	694
A.3	Acknowledgments	695
	REFERENCES	697
	INDEX	727