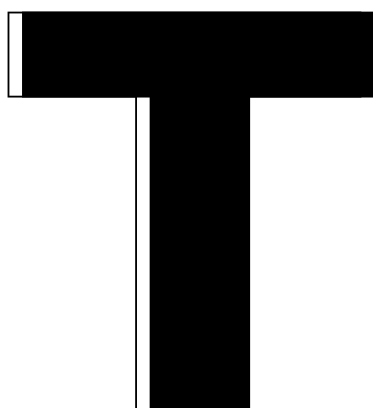


Official Textbooks for Huawei ICT Academy



DATABASE PRINCIPLES AND TECHNOLOGIES – BASED ON HUAWEI GAUSSDB

Huawei Technologies Co., Ltd.



 **人民邮电出版社**
POSTS & TELECOM PRESS

OPEN ACCESS

 Springer

Database Principles and Technologies – Based on Huawei GaussDB

Huawei Technologies Co., Ltd.

Database Principles and Technologies – Based on Huawei GaussDB

 Springer

 人民邮电出版社
POSTS & TELECOM PRESS

Huawei Technologies Co., Ltd.
Hangzhou, China



This work was supported by Huawei Technologies Co., Ltd.

ISBN 978-981-19-3031-7 ISBN 978-981-19-3032-4 (eBook)
<https://doi.org/10.1007/978-981-19-3032-4>

Jointly published with Posts & Telecom Press, Beijing, China

The print edition is not for sale in China (Mainland). Customers from China (Mainland) please order the print book from: Posts & Telecom Press.

© Posts & Telecom Press 2023. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this book or parts of it.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

This work is subject to copyright. All commercial rights are reserved by the author(s), whether the whole or part of the material is concerned, specifically the rights of reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Regarding these commercial rights a non-exclusive license has been granted to the publisher.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publishers, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Nowadays, database technology has developed from the early stage of simply saving and processing data files to a rich, comprehensive discipline with data modeling and database management system as the core, as the foundation and core of modern computer application system. Entering the Internet era, the traditional database system began to show decadence in response to the storage needs of big data, and enterprise customers urgently need a new generation of database products, that is, products with dynamic expansion and contraction capacity, high throughput, low cost, and other characteristics. As a result, cloud computing-based databases have emerged and risen, showing the future-oriented trend of cloud-based, distributed, and multi-mode processing.

Based on Huawei's GaussDB (for MySQL) cloud computing-based database, this book focuses on various cloud computing-based features and application scenarios of cloud computing-based databases. The division of the book's eight chapters is as follows:

Chapter 1 mainly introduces databases, including database technology overview, database technology history, relational database architecture, and mainstream application scenarios of relational databases.

Chapter 2 mainly teaches database basics, including the main responsibilities and contents of database management, and introduces some common and important basic concepts of databases.

Chapter 3 introduces SQL syntax, including GaussDB (for MySQL) data types, system functions and operators, which aims to help beginners master get started with SQL syntax.

Chapter 4 focuses on SQL syntax classification and further explains SQL statements accordingly, covering data query, data update, data definition, and data control.

Chapter 5 focuses on database security fundamentals, including basic security management techniques for databases, such as access control, user management, permission management, object permissions, and cloud auditing services, which will be elaborated from basic concepts, usages, and application scenarios.

Chapter 6 focuses on the database development environment, including the use of all the tools of GaussDB (for MySQL), for the convenience of users to learn and view.

Chapter 7 mainly teaches database design fundamentals, detailing the specific work of requirements analysis, conceptual design, logical design and physical design in accordance with the New Orleans design methodology, and finally introducing the specific means of database design implementation with relevant cases.

Chapter 8 mainly introduces the features of GaussDB database, involving Huawei relational database and Huawei NoSQL database.

This book is edited by Huawei Technologies Co., Ltd., thanks to Ma Ruixin for the specific writing and final compilation of the whole book. We welcome readers' criticism and correction if there are any shortcomings in the book, due to the limited time for compilation.

Hangzhou, China
December 2021

Huawei Technologies Co., Ltd.

Contents

1	Introduction to Databases	1
1.1	Overview of Database Technology	1
1.1.1	Data	1
1.1.2	Database	2
1.1.3	Database Management System	3
1.1.4	Database System	4
1.2	History of Database Technology	5
1.2.1	Emergence and Development of Database Technology	5
1.2.2	Comparison of the Three Stages of Data Management	6
1.2.3	Benefits of Database	8
1.2.4	Development Characteristics of the Database	9
1.2.5	Hierarchical Model, Mesh Model and Relational Model	10
1.2.6	Structured Query Language	14
1.2.7	Characteristics of Relational Databases	14
1.2.8	Historical Review of Relational Database Products	15
1.2.9	Other Data Models	17
1.2.10	New Challenges for Data Management Technologies	18
1.2.11	NoSQL Database	19
1.2.12	NewSQL Database	21
1.2.13	Database Ranking	23
1.3	Architecture of Relational Databases	24
1.3.1	Development of Database Architecture	24
1.3.2	Single-Host Architecture	24
1.3.3	Group Architecture: Master-Standby Architecture	26
1.3.4	Group Architecture: Master-Slave Architecture	27
1.3.5	Group Architecture: Multi-Master Architecture	28
1.3.6	Shared Disk Architecture	28
1.3.7	Sharding Architecture	29
1.3.8	Shared-Nothing Architecture	30

1.3.9	Massively Parallel Processing Architecture	31
1.3.10	Comparison of the Characteristics of Database Architectures	32
1.4	Mainstream Applications of Relational Databases	34
1.4.1	Online Transaction Processing	34
1.4.2	Online Analytical Processing	34
1.4.3	Database Performance Measurement Indicators	35
1.5	Summary	37
1.6	Exercises	37
2	Basic Knowledge of Database	41
2.1	Overview of Database Management	41
2.1.1	Database Management and Its Scope of Work	41
2.1.2	Object Management	43
2.1.3	Backup and Recovery Management	44
2.1.4	Security Management	49
2.1.5	Performance Management	53
2.1.6	O&M Management	56
2.2	Key Concepts of Database	60
2.2.1	Database and Database Instance	60
2.2.2	Database Connection and Session	61
2.2.3	Schema	63
2.2.4	Tablespace	64
2.2.5	Table	65
2.2.6	How the Table Is Stored	66
2.2.7	Partition	68
2.2.8	Data Distribution	71
2.2.9	Data Types	72
2.2.10	View	74
2.2.11	Index	76
2.2.12	Constraints	77
2.2.13	Transaction	80
2.3	Summary	84
2.4	Exercises	85
3	Getting Started with SQL Syntax	87
3.1	Overview of SQL Statements	88
3.1.1	What is an SQL Statement	88
3.1.2	Comprehensive Application of SQL Statements	89
3.2	Data Types	89
3.2.1	Common Data Types	90
3.2.2	Uncommon Data Types	92
3.2.3	Cases of Data Types	93
3.3	System Functions	94
3.3.1	Numeric Calculation Functions	94

- 3.3.2 Character Processing Functions 97
- 3.3.3 Time and Date Functions 100
- 3.3.4 Type Conversion Functions 101
- 3.3.5 System Information Functions 104
- 3.4 Operators 104
 - 3.4.1 Logical Operators 104
 - 3.4.2 Comparison Operators 105
 - 3.4.3 Arithmetic Operators 106
 - 3.4.4 Test Operators 107
 - 3.4.5 Other Operators 110
- 3.5 Summary 111
- 3.6 Exercises 111
- 4 SQL Syntax Categories 115**
 - 4.1 Data Query 115
 - 4.1.1 Simple Query 115
 - 4.1.2 Removing Duplicate Values 117
 - 4.1.3 Query Column Selection 118
 - 4.1.4 Conditional Query 120
 - 4.1.5 Join Query 123
 - 4.1.6 Subquery 128
 - 4.1.7 Merging Result Sets 131
 - 4.1.8 Difference Result Sets 133
 - 4.1.9 Data Grouping 133
 - 4.1.10 Data Sorting 135
 - 4.1.11 Data Restriction 136
 - 4.2 Data Update 137
 - 4.2.1 Data Insertion 138
 - 4.2.2 Data Modification 140
 - 4.2.3 Data Deletion 142
 - 4.3 Data Definition 144
 - 4.3.1 Database Objects 144
 - 4.3.2 Creating a Table 145
 - 4.3.3 Modifying Table Properties 148
 - 4.3.4 Deleting a Table 149
 - 4.3.5 Index 149
 - 4.3.6 View 153
 - 4.4 Data Control 155
 - 4.4.1 Transaction Control 155
 - 4.4.2 Committing a Transaction 155
 - 4.4.3 Rolling Back a Transaction 156
 - 4.4.4 Transaction Save Points 157
 - 4.5 Others 159
 - 4.5.1 SHOW Command 159
 - 4.5.2 SET Command 161

4.6	Summary	162
4.7	Exercises	162
5	Database Security Fundamentals	167
5.1	Overview of Database Security Features	167
5.1.1	What Is Database Security Management	167
5.1.2	Database Security Framework	167
5.1.3	Database Security Features	168
5.2	Access Control	168
5.2.1	What Is IAM	168
5.2.2	IAM Features	169
5.2.3	IAM Authorization	171
5.2.4	Relationship Between IAM and GaussDB (for MySQL) usage	172
5.2.5	How to Use GaussDB(for MySQL) with IAM	172
5.2.6	Detailed Explanation of SSL	173
5.3	User Permission Control	174
5.3.1	Permission Concept	174
5.3.2	Users	175
5.3.3	Modifying a User	176
5.3.4	Deleting a User	177
5.3.5	Roles	178
5.3.6	Authorization	179
5.3.7	Permission Recovery	181
5.4	Cloud Audit Services	183
5.4.1	What Are Cloud Audit Services	183
5.4.2	Key Operations to Support Cloud Audit Services	184
5.5	Summary	186
5.6	Exercises	186
6	Database Development Environment	189
6.1	GaussDB Database Driver	189
6.1.1	What Is a Driver	189
6.1.2	JDBC	190
6.1.3	ODBC	194
6.1.4	Others	200
6.2	Database Tools	202
6.2.1	DDM	202
6.2.2	DRS	208
6.2.3	DAS	215
6.3	Client Tools	223
6.3.1	zsql	225
6.3.2	gsql	235
6.3.3	Data Studio	238
6.3.4	MySQL Workbench	240

- 6.4 Summary 242
- 6.5 Exercises 242
- 7 Database Design Fundamentals 245**
 - 7.1 Database Design Overview 245
 - 7.1.1 Difficulties of Database Design 246
 - 7.1.2 Goal of Database Design 246
 - 7.1.3 Methods of Database Design 247
 - 7.2 Requirements Analysis 247
 - 7.2.1 Significance of Requirement Analysis 247
 - 7.2.2 Tasks of the Requirement Analysis Stage 248
 - 7.2.3 Methods of Requirement Analysis 249
 - 7.2.4 Data Dictionary 249
 - 7.3 Conceptual Design 250
 - 7.3.1 Conceptual Design and Conceptual Model 250
 - 7.3.2 E-R Approach 251
 - 7.4 Logical Design 253
 - 7.4.1 Logical Design and Logical Models 253
 - 7.4.2 IDEFIX Method 253
 - 7.4.3 Entities and Attributes in the Logic Model 254
 - 7.4.4 NF Theory 260
 - 7.4.5 Logic Design Considerations 265
 - 7.5 Physical Design 267
 - 7.5.1 Physical Design and Physical Models 267
 - 7.5.2 Denormalization of the Physical Model 268
 - 7.5.3 Maintaining Data Integrity 270
 - 7.5.4 Establishing a Physicalized Naming Convention 271
 - 7.5.5 Physicalizing Tables and Fields 272
 - 7.5.6 Using Modeling Software 274
 - 7.5.7 Physical Model Products 275
 - 7.6 Database Design Case 275
 - 7.6.1 Scenario Description 275
 - 7.6.2 Regularization Processing 275
 - 7.6.3 Data Types and Length 279
 - 7.6.4 Denormalization 280
 - 7.6.5 Index Selection 281
 - 7.7 Summary 282
 - 7.8 Exercises 282
- 8 Introduction to Huawei Cloud Database GaussDB 287**
 - 8.1 GaussDB Database Overview 287
 - 8.1.1 GaussDB Database Family 287
 - 8.1.2 Typical OLTP and OLAP Databases 289
 - 8.2 Relational Database Products and Related Tools 290
 - 8.2.1 GaussDB (for MySQL) 290

- 8.2.2 GaussDB (openGauss) 296
- 8.2.3 GaussDB (DWS) 299
- 8.2.4 Data Studio 304
- 8.3 NoSQL Databases 307
 - 8.3.1 GaussDB (for Mongo) 307
- 8.4 Summary 309
- 8.5 Exercises 311
- Index 313**