



高等职业院校计算机类规划教材

MySQL SHUJUKU XITONG YUANLI

MySQL数据库系统原理

(Oracle Linux 8.3 + MySQL 8.0.23)

李爱武 编著



北京邮电大学出版社
www.buptpress.com

内 容 简 介

本书包括关系型数据库基本理论以及 MySQL 数据库技术两部分。第 1~4 章为数据库基本理论部分, 主要包括关系模型基本理论、ER 图及范式理论; 第 5~20 章为 MySQL 数据库技术部分, 主要包括在 Oracle Linux 上安装和运行 MySQL 服务器软件, SQL 语言, 表空间, 索引原理, MySQL 数据库体系结构, 事务处理及锁, 备份恢复, 用户及权限管理等内容。

本书可作为高校计算机相关专业的数据库课程教材或参考书, 也可供从事 MySQL 数据库相关技术工作的人员参考。

图书在版编目(CIP)数据

MySQL 数据库系统原理: Oracle Linux 8.3 + MySQL 8.0.23 / 李爱武编著. -- 北京: 北京邮电大学出版社, 2021.7

ISBN 978-7-5635-6417-0

I. ①M… II. ①李… III. ①Linux 操作系统—高等学校—教材②关系数据库系统—高等学校—教材
IV. ①TP316.85②TP311.138

中国版本图书馆 CIP 数据核字 (2021) 第 140167 号

策划编辑: 马晓仟 责任编辑: 马晓仟 封面设计: 七星博纳

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路 10 号

邮政编码: 100876

发行部: 电话: 010-62282185 传真: 010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 唐山玺诚印务有限公司

开 本: 787 mm×1 092 mm 1/16

印 张: 18.25

字 数: 478 千字

版 次: 2021 年 7 月第 1 版

印 次: 2021 年 7 月第 1 次印刷

ISBN 978-7-5635-6417-0

定价: 49.00 元

· 如有印装质量问题, 请与北京邮电大学出版社发行部联系 ·

前 言

数据库技术经过近 60 年的发展,关系型数据库逐渐一统天下。可以预见,关系型数据库在未来的几十年内还会居于主导地位。20 世纪的最后 20 年,数据库市场主要被 Oracle、微软、IBM 三家公司控制。

进入 21 世纪后,Linux 在服务器市场的份额越来越大,MySQL 和 PostgreSQL 等开源数据库产品也迎来了黄金时代。考虑数据安全和成本等因素,国内越来越多的公司使用开源数据库产品作为服务器。几个大型科技公司,如阿里、腾讯、网易等,对 MySQL 的开源代码进行了二次定制开发,以满足企业自身的特殊要求。经过定制开发的 MySQL 甚至进入了一直由 Oracle 主导的金融行业,改变了人们对 MySQL 只部署于论坛、博客等非关键应用的印象。随着云计算技术的不断普及,以 MySQL 为代表的开源数据库将会迎来更广阔的市场。

本书包括关系型数据库基本理论和 MySQL 数据库技术两部分内容。在基本理论方面,本书力争用简短的篇幅和通俗的语言讲清其本质。在 MySQL 数据库技术方面,本书以简单的实验介绍相关知识、验证各种结论。在内容选取上,本书注重介绍数据库通用原理方面的知识,详细地讲解了 SQL 语言、索引、执行计划、统计信息、事务处理、锁、数据的导出与导入、重做日志等内容,为读者进一步学习 MySQL 和其他数据库产品知识打下基础。

本书的实验环境由 Windows 10 上的 VMware 16.1.0 虚拟机搭建,并在虚拟机上安装 Oracle Linux 8.3 和 MySQL 8.0.23。SSH 远程管理工具采用 MobaXterm 20.5,MySQL 图形管理工具采用 MySQL Workbench 8.0.23。以上所有软件都可以安装在一台笔记本式计算机上,以方便模拟生产环境。与 CentOS 一样,Oracle Linux 是 Red Hat Enterprise Linux (RHEL)的“克隆”,本书内容同样适用于 RHEL 和 CentOS。

目 录

第 1 章 数据库技术基础	1
1.1 数据库应用的场合	1
1.2 常用术语	1
1.3 数据库系统的构成	1
1.3.1 硬件	1
1.3.2 软件	2
1.3.3 人员	2
第 2 章 关系模型理论及主要产品	3
2.1 数据处理的历史	3
2.2 数据模型的概念	3
2.3 网状模型与层次模型	4
2.3.1 网状模型	4
2.3.2 层次模型	4
2.3.3 网状模型和层次模型的贡献及缺陷	5
2.4 关系模型的提出和成熟	5
2.4.1 关系模型要解决的问题	5
2.4.2 关系模型的提出与完善	6
2.4.3 IBM 的 System R 项目	6
2.4.4 加州大学伯克利分校的 Ingres 项目	6
2.5 关系模型的三个要素	7
2.5.1 关系模型的数据结构	7
2.5.2 关系模型的数据操作方式	8
2.5.3 关系模型中的完整性约束	8
2.5.4 关系型数据库的特点	8
2.6 主要关系型数据库产品介绍	9
2.6.1 传统商业数据库产品	9
2.6.2 开源数据库与 MySQL	9

第 3 章 ER 模型	10
3.1 数据库设计的主要步骤	10
3.2 ER 模型的主要概念	10
3.3 联系的映射约束	11
3.4 ER 图转化为表	11
3.5 使用 MySQL Workbench 进行数据库设计	11
第 4 章 规范化理论	17
4.1 引入范式理论的原因	17
4.1.1 存在数据冗余的概念设计举例	17
4.1.2 Insertion 异常	18
4.1.3 Deletion 异常	18
4.1.4 Update 异常	18
4.2 第一范式	18
4.3 第二范式	19
4.4 第三范式	20
第 5 章 安装和使用 MySQL	21
5.1 支持 MySQL 8.0 的操作系统	21
5.2 在 VMware 16.1 虚拟机中安装 Oracle Linux 8.3	21
5.2.1 下载 VMware 和 Oracle Linux	21
5.2.2 安装 Oracle Linux	22
5.3 使用 Windows SSH 工具管理 Oracle Linux 服务器	24
5.3.1 使用 Windows 10 的 SSH 工具操作 Oracle Linux 服务器	25
5.3.2 使用 Windows 10 的 scp 工具远程复制文件	25
5.4 使用 MobaXterm 操作 Oracle Linux 服务器	25
5.4.1 建立连接及基本 SSH 配置	26
5.4.2 MobaXterm 基本配置	26
5.4.3 在 Windows 和 Linux 之间拖动传输文件	27
5.4.4 显示 X Window 图形	27
5.5 使用 dnf 工具	28
5.5.1 repo 配置文件	28
5.5.2 管理 yum 源	29
5.5.3 管理软件包	30
5.5.4 dnf 模块管理	30

5.6 安装与删除 MySQL 服务器软件	32
5.6.1 检查 MySQL 软件是否已安装	32
5.6.2 使用 Oracle Linux 系统附带软件包安装	32
5.6.3 管理 MySQL 服务	32
5.6.4 删除 MySQL 软件	33
5.6.5 使用 MySQL 官方 yum 源安装	34
5.7 升级 MySQL	34
5.8 目录结构	35
5.9 连接至 mysqld 服务	35
5.9.1 连接至本地 mysqld 服务	35
5.9.2 连接至远端 mysqld 服务	35
5.10 mysql 工具执行常见操作	36
5.11 创建测试数据	37
第 6 章 SQL 查询语句	40
6.1 SQL 概述	40
6.1.1 SQL 语言的历史	40
6.1.2 SQL 的发音	41
6.1.3 SQL 查询的特点	41
6.1.4 SQL 标准	41
6.2 SQL 语言的主要类型	41
6.3 常用数据类型	42
6.3.1 数值类型	42
6.3.2 字符串类型	42
6.3.3 日期时间类型	42
6.4 简单的 SQL 查询语句	43
6.4.1 最简单的查询——只指定表	43
6.4.2 指定列	43
6.4.3 指定列别名	43
6.4.4 用 where 子句指定查询条件	44
6.4.5 使用 order by 子句给查询结果排序	44
6.4.6 分页查询	45
6.5 常用数值运算符及函数	45
6.5.1 数值运算符	45
6.5.2 常用数值函数	45
6.6 字符数据的处理	46

6.6.1	字符串常量	46
6.6.2	字符串模糊查询	46
6.6.3	处理字符串中的特殊字符	47
6.6.4	常用字符串函数	47
6.6.5	利用正则表达式搜索字符串	51
6.7	处理日期型数据	53
6.7.1	获得当前日期时间	53
6.7.2	日期型常量	53
6.7.3	指定格式显示日期型列值	53
6.7.4	抽取日期的指定部分	54
6.7.5	获取时间差	54
6.8	空值的处理	55
6.9	分组汇总	56
6.9.1	单独使用分组函数	56
6.9.2	使用 group by 子句执行分组汇总	56
6.9.3	having 子句	57
6.9.4	order by 子句	57
6.9.5	分组汇总查询小结	57
6.10	子查询	58
6.10.1	where 子句中使用子查询	58
6.10.2	select 子句中使用子查询	59
6.10.3	from 子句中使用子查询	59
6.10.4	非相关子查询与相关子查询	59
6.10.5	in 与 not in	62
6.10.6	exists 与 not exists	64
6.11	集合运算	65
6.12	多表连接查询	65
6.12.1	交叉连接	65
6.12.2	内连接	66
6.12.3	两种连接标准:SQL-86 与 SQL-92	67
6.12.4	外连接	68
6.13	构造复杂的查询语句	70
6.14	SQL 查询的等效转换	72
6.14.1	内连接与子查询	72
6.14.2	in,exists,内连接	72
6.14.3	not in,not exist,外连接	73

第 7 章 窗口函数	74
7.1 over()子句构造窗口	74
7.1.1 partition by 子句构造窗口	74
7.1.2 order by 子句设置窗口内排序	75
7.1.3 窗口内划分框架	76
7.1.4 命名窗口	79
7.2 窗口函数分类与示例	79
7.2.1 汇总函数	79
7.2.2 排名函数	80
第 8 章 数据修改语句	83
8.1 delete 语句	83
8.1.1 简单的 delete 语句	83
8.1.2 delete 语句使用 limit 子句	83
8.1.3 delete 语句同时删除多个表的行	83
8.2 update 语句	84
8.2.1 简单的 update 语句	84
8.2.2 update 语句修改多个表	84
8.3 insert 语句	85
8.4 replace 语句	85
第 9 章 表及约束	87
9.1 创建简单的表	87
9.2 字符集及排序规则	87
9.2.1 MySQL 与 Unicode 字符集	87
9.2.2 排序规则与中文排序	89
9.2.3 建表时设置字符集及排序规则	89
9.3 建表时指定存储引擎	90
9.4 使用 auto_increment 自增列	90
9.5 自填充时间列	91
9.6 约束	92
9.6.1 约束的种类	92
9.6.2 主键、非空及默认约束	92
9.6.3 唯一、检查及外键约束	93
9.6.4 对表增加约束	95

9.6.5	删除约束	96
9.6.6	查询约束的信息	96
9.7	复制表	97
9.8	修改表的结构	97
9.8.1	修改列的数据类型	97
9.8.2	添加或删除列	98
9.8.3	修改列名	98
9.8.4	修改表名	98
9.8.5	清空表:truncate table	98
9.8.6	删除表	99
9.9	查看表定义	99
第 10 章	分区表	100
10.1	分区类别	100
10.2	范围分区	100
10.2.1	单列范围分区	100
10.2.2	多列范围分区	102
10.2.3	增删范围分区	104
10.2.4	重组分区	105
10.3	列表分区	106
10.3.1	单列列表分区	106
10.3.2	多列列表分区	106
10.3.3	增删列表分区	108
10.3.4	重组列表分区	109
10.4	散列分区	109
10.4.1	普通散列分区	110
10.4.2	线性散列分区	110
10.4.3	键分区	111
10.4.4	重组散列分区	112
10.5	子分区	112
10.6	查询分区信息	113
10.7	改变表的分区类型	115
10.8	在 SQL 命令中直接操作分区	116
第 11 章	程序设计	117
11.1	用户变量	117

11.2	存储过程	118
11.2.1	存储过程的创建和执行	118
11.2.2	使用变量	118
11.2.3	使用 if 语句	119
11.2.4	使用 while 循环语句	120
11.2.5	使用输入参数及 SQL 语句	120
11.2.6	使用输出参数	121
11.3	函数	122
11.4	触发器	123
11.4.1	创建触发器的语法	123
11.4.2	old 和 new 的用法	124
11.4.3	模拟外键级联删除	125
11.4.4	约束检查	125
11.4.5	审计	127
11.4.6	查看触发器信息	128
11.5	查看可编程对象系统信息	128
11.5.1	使用 show 命令查看程序定义	129
11.5.2	使用 information_schema 系统库的 routines 和 triggers 视图	129
11.6	删除可编程对象	130
第 12 章	服务器体系结构	131
12.1	总体结构	131
12.2	内存结构	132
12.2.1	内存数据缓冲区	132
12.2.2	内存日志缓冲区	133
12.2.3	排序缓冲区和连接缓冲区	134
12.2.4	内部临时表内存	134
12.2.5	内存的自动设置	134
12.3	配置服务器和客户端参数	134
12.3.1	设置方式	135
12.3.2	使用命令行参数	135
12.3.3	使用配置文件	135
12.3.4	使用 set 命令	136
12.3.5	查看系统参数	136
12.4	事件日志文件	138
12.4.1	设置日志的时区和输出目标	138

12.4.2	错误日志	139
12.4.3	通用查询日志	140
12.4.4	慢查询日志	141
12.5	重做文件	141
12.5.1	binary log	141
12.5.2	redo log	145
12.6	系统数据库	146
12.6.1	mysql	146
12.6.2	sys	147
12.6.3	performance_schema	150
12.6.4	information_schema	150
第 13 章	表空间和数据文件	151
13.1	表空间的概念	151
13.1.1	数据目录	151
13.1.2	MySQL 的表空间分类	151
13.2	system 表空间	152
13.2.1	change buffer	152
13.2.2	doublewrite buffer	152
13.2.3	system 表空间的数据文件	152
13.3	temporary 表空间	153
13.3.1	会话临时表空间	153
13.3.2	全局临时表空间	154
13.4	undo 表空间	155
13.4.1	默认 undo 表空间	155
13.4.2	创建 undo 表空间	155
13.4.3	查询 undo 表空间信息	155
13.4.4	截断 undo 表空间	156
13.4.5	删除 undo 表空间	157
13.5	file-per-table 表空间	157
13.6	general 表空间	159
13.7	移动表所属的表空间	161
第 14 章	B 树索引	162
14.1	B 树索引能把查询速度提高多少	162
14.2	一个使用索引的例子	163

14.3	主键索引的结构	165
14.4	普通索引的结构	168
14.5	索引能够提高查询速度的原因	172
14.6	需要创建索引的情况	172
14.7	如何知道一个查询是否使用了索引	172
14.8	不使用索引的情况	174
14.9	DML 语句对索引的影响	174
14.9.1	insert 语句对索引的影响	174
14.9.2	delete 语句对索引的影响	174
14.9.3	update 语句对索引的影响	174
14.10	基于函数的索引	174
14.11	设置索引的可见性	175
14.12	多列索引	176
14.13	约束与索引	177
14.14	查询索引的系统信息	177
14.14.1	show create table	177
14.14.2	show index from	178
14.14.3	查询 information_schema.innodb_indexes	178
第 15 章	执行计划	180
15.1	执行计划简单示例	180
15.2	执行计划的 select_type 属性	181
15.3	执行计划的 type 属性	186
15.4	执行计划的 ref 属性	189
15.5	执行计划的 Extra 属性	191
第 16 章	统计信息	195
16.1	统计信息的内容	195
16.2	统计信息的分类和收集	195
16.2.1	永久统计信息	195
16.2.2	临时统计信息	196
16.3	统计信息的存储和监控	197
16.3.1	innodb_index_stats 系统表	197
16.3.2	innodb_table_stats 系统表	198
16.3.3	information_schema 中的统计信息视图	198
16.3.4	show 命令	201

16.3.5	information_schema_stats_expiry 参数	203
16.4	统计信息的更新	203
16.4.1	自动更新	203
16.4.2	手动更新	204
16.5	列的直方图	204
16.5.1	直方图的作用	205
16.5.2	单值直方图和等高直方图	205
16.5.3	计算直方图	205
16.5.4	应用直方图实例	207
第 17 章	事务处理	210
17.1	事务的概念及应用实例	210
17.1.1	事务应用实例 1: 银行转账	210
17.1.2	事务应用实例 2: 超市收银	210
17.2	事务的 ACID 属性	211
17.2.1	原子性	211
17.2.2	一致性	211
17.2.3	隔离性	212
17.2.4	持久性	212
17.3	事务隔离级别	212
17.3.1	read uncommitted	212
17.3.2	read committed	212
17.3.3	repeatable read	213
17.3.4	serializable	214
17.4	事务控制命令	214
17.4.1	commit 和 rollback 命令	214
17.4.2	设置事务模式	215
17.4.3	设置事务隔离级别	215
17.4.4	设置事务只读性	216
17.5	并发控制要解决的问题	217
17.5.1	丢失更新	217
17.5.2	脏读	218
17.5.3	不可重复读	219
第 18 章	锁	220
18.1	MySQL 的锁类型和锁模式	220

18.2	表锁	221
18.2.1	表级 S 锁和表级 X 锁	221
18.2.2	DDL 语句产生的元数据锁	222
18.2.3	表级 IS 锁和表级 IX 锁	223
18.3	行锁	224
18.3.1	S, REC_NOT_GAP 锁和 X, REC_NOT_GAP 锁	224
18.3.2	S, GAP 锁和 X, GAP 锁	226
18.3.3	next-key 锁	227
18.3.4	insert intention 锁	230
18.4	insert 操作产生的隐式锁	233
18.5	外键对锁的影响	234
18.6	不同隔离级别下的加锁方式	235
18.6.1	read uncommitted 与 read committed 隔离级别下的锁	235
18.6.2	repeatable read 隔离级别下的锁	240
18.6.3	serializable 隔离级别下的锁	241
18.7	死锁	243
第 19 章	备份恢复	246
19.1	备份种类	246
19.1.1	热备份与冷备份	246
19.1.2	逻辑备份与物理备份	246
19.1.3	全备份与增量备份	246
19.2	恢复种类	246
19.3	备份恢复工具	247
19.4	导出导入数据行	248
19.4.1	一个简单示例	248
19.4.2	使用 select into outfile 及 load data 执行导出导入	249
19.4.3	使用 mysqlimport 导入	249
19.4.4	使用 mysqldump 导出	250
19.5	mysqlpump 执行逻辑备份	250
19.5.1	mysqlpump 执行导出	250
19.5.2	mysqlpump 的常用选项	251
19.6	使用 mysqlbinlog 导出二进制日志文件	252
19.7	基于时点的恢复	253
第 20 章	用户和权限管理	255
20.1	MySQL 用户的特点	255

20.2	预置用户	255
20.3	一个关于用户及权限的简单示例	256
20.4	用户管理	256
20.4.1	MySQL 的用户名标识	257
20.4.2	创建和删除用户	257
20.5	密码管理	257
20.5.1	密码加密算法	257
20.5.2	设置密码过期和重用规则	258
20.5.3	设置密码试错次数上限及锁定天数	258
20.5.4	查询用户的密码相关属性	259
20.5.5	设置密码相关系统参数	259
20.5.6	设置密码复杂度规则	260
20.6	对用户设置资源限制	261
20.7	修改用户属性	261
20.8	权限管理	262
20.8.1	MySQL 中的权限层次和名称	262
20.8.2	赋予和撤销权限	262
20.8.3	全局权限的部分撤销	264
20.8.4	查询用户的权限信息	266
20.9	角色	270
20.9.1	创建角色并对其赋权	270
20.9.2	角色的赋予及撤销	271
20.9.3	激活角色	272
20.9.4	用户的默认角色	273
20.9.5	设置公共角色	274
20.9.6	角色相关信息查询	274
	参考文献	276

第 1 章 数据库技术基础

随着计算机技术的快速发展,数据库应用的范围越来越广泛,日常生活对数据库应用的依赖越来越紧密。本章对数据库技术的相关背景和概念做简单介绍。

本章主要包括:

- 数据库应用的场合;
- 常用术语;
- 数据库系统的构成。

1.1 数据库应用的场合

随着移动互联网技术的飞速发展,人们可以使用网上银行、手机银行进行各种金融活动,使用淘宝、京东等网站或手机客户端软件购物,使用美团、饿了么点外卖,使用微信、QQ 等即时通信软件与朋友交流,使用滴滴打车、共享单车等出行。以上这些系统都离不开数据库的支撑,近年兴起的大数据技术更离不开数据库的支持。就业市场对数据库技术人员,如数据库管理、应用开发等方面人才的需求越来越大。

1.2 常用术语

数据库(Database, DB)是长期存储在计算机内的、有组织、有结构、可共享的数据集合。数据按一定数据模型组织和存储,具有较小冗余度、较高数据独立性和易扩展性。

数据库管理系统(Database Management System, DBMS)是位于用户和操作系统之间的一类系统软件。它用于科学、有效地组织和存储数据,高效地获取和维护数据。Oracle、DB2、SQL Server、MySQL、PostgreSQL 都是典型的 DBMS 产品。

数据库系统(Database System, DBS)是引入了数据库的计算机应用系统。

1.3 数据库系统的构成

数据库系统一般由硬件、软件及各种人员构成,下面分别介绍。

1.3.1 硬件

数据库系统的数据量和访问量一般很大,而且要保持常年 24 小时运转,整个数据库系统对硬件资源的要求很高。能够用作数据库服务器的计算机一般由 Oracle、IBM、HP 等著名计算机厂商生产,使用专用的 CPU、内存,数据存储采用磁盘阵列,性能和价格远远高于 PC 上的各种硬件。昂贵的服务器硬件限制了数据库应用系统的使用范围。

1.3.2 软件

数据库系统的软件主要包括操作系统、DBMS、前端应用软件。

大型数据库服务器一般采用专用服务器,安装专用的 Unix 操作系统,如 Oracle 的 SPARC 系列服务器产品只能安装 Solaris,而 IBM 公司的高端服务器只能安装 AIX。中小型应用对服务器硬件的要求稍低,一般使用 PC 服务器,安装 Windows 或 Linux 操作系统。

应用不同,DBMS 也不同。单用户应用使用桌面数据库产品就够了,如微软公司的 Access 产品,而大型、关键应用(指涉及金融或财务的应用)一般采用大型数据库产品,如 Oracle、DB2 或 SQL Server。

前端应用软件一般由另外的软件公司根据企业的具体要求进行定制开发。

随着开源软件和云计算技术的发展,越来越多的应用采用 Linux 系统或云架构运行其数据库服务器,这大大降低了软硬件购置和维护费用。典型的云平台包括亚马逊 AWS、微软 Azure 以及阿里云。当前很多数据库服务器由云上运行的 Linux+MySQL 等开源软件搭建。

1.3.3 人员

数据库系统涉及的人员主要包括系统分析和数据库设计人员、应用程序开发人员,数据库管理人员和终端用户。

系统分析人员负责应用系统的需求分析和规范说明,与用户及数据库管理人员讨论,确定系统的软硬件配置,并参与数据库系统的概要设计。系统分析人员要具备全面的软件工程知识及丰富的系统设计、开发经验。

数据库设计人员负责数据库中数据的确定,参加用户需求调查和系统分析,根据分析结果设计出应用所需的表结构。

应用程序开发人员负责设计和编写应用系统的程序模块及前端界面,需要掌握 SQL 语言及存储过程编写方法,并要掌握所使用的开发语言。

数据库管理人员负责管理和维护数据库系统,具体职责包括以下几个方面:决定数据库的存储结构和存取策略;确定数据库的安全性要求,完成用户和权限控制;负责数据库的备份和恢复;监控数据库运行,如空间使用情况,以及确定操作低效的成因并解决。

以上人员都要精通数据库理论及 DBMS 产品的使用,另外,数据库管理人员还要熟悉相关 DBMS 和操作系统的原理及操作,并具备一定的程序开发知识。

终端用户指数据库应用系统要服务的对象,这些人一般不需要具备计算机和数据库技术知识,例如使用银行自动取款机取款的人,就是银行数据库系统的终端用户;使用淘宝网购物的人就是淘宝数据库系统的终端用户。