

第 1 章 数据库概述

数据管理技术经过多年的发展，已经发展到数据库系统阶段。在该阶段会把数据存储到数据库（DataBase，DB）中，即数据库相当于存储数据的仓库。为了便于用户组织和管理数据，其还专门提供了数据库管理系统（DataBase Management System，DBMS），可以有效管理存储在数据库中的数据。本书所要讲的 MySQL 软件，就是一种非常优秀的数据库管理系统。本章抛开 MySQL 讲解关系数据库设计的相关知识，以“选课系统”为例，讲解“选课系统”数据库的设计流程。简单地说，数据库（DataBase 或 DB）是存储、管理数据的容器；严格地说，数据库是“按照某种数据结构对数据进行组织、存储和管理的容器”。

通过本章的学习，读者可以掌握如下内容：

- 数据管理技术。
- 数据库相关概念和知识。
- MySQL 数据库基本概念和知识。

1.1 数据库基础

1.1.1 数据库基本概念

- 数据（Data）。
- 数据库（DataBase）。
- 数据库管理系统（DBMS）。
- 数据库系统（DBS）。

1. 数据（Data）的定义

对客观事物的符号表示，如图形符号、数字、字母等，数据是数据库中存储的基本对象。

在日常生活中，人们直接用语言来描述事物；在计算机中，为了存储和处理这些事物，就要将事物的特征抽象出来组成一条记录来描述。

- （1）数据的种类：文字、图形、图像、声音。
- （2）数据的特点：数据与其语义是不可分的。
- （3）数据举例。

- 学生档案中的学生记录：（单光庆，男，1974，重庆，信息工程，1993）。
- 数据的形式不能完全表达其内容。
- 数据的解释。

语义：学生姓名、性别、出生年月、籍贯、所在系别、入学时间。

解释：单光庆是个大学生，1974 年出生，重庆人，1993 年考入信息工程学院。

2. 数据库 (Database , DB) 的定义

数据库是“按照数据结构来组织、存储和管理数据的仓库”。J.Martin 给数据库下了一个比较完整的定义：数据库是存储在一起的相关数据的集合，这些数据是结构化的，无有害的或不必要的冗余，并为多种应用服务。

3. 数据库管理系统

(1) 数据库管理系统 (Database Management System , DBMS) 是一种操纵和管理数据库的大型软件 ,用于建立、使用和维护数据库 ,简称 DBMS。关系型数据库管理系统称为 RDBMS , R 指 Relation。

(2) DBMS 的作用。对数据库进行统一管理和控制，以保证数据库的安全性和完整性。

(3) DBMS 的主要功能。

① 数据定义功能。

- 提供数据定义语言 (DDL)。
- 定义数据库中的数据对象。

② 数据操纵功能。

- 提供数据操纵语言 (DML)。
- 操纵数据实现对数据库的基本操作：查询、插入、删除和修改。

③ 数据库的运行管理。

- 保证数据的安全性、完整性。
- 多用户对数据的并发使用。
- 发生故障后的系统恢复。

④ 数据库的建立和维护功能 (实用程序)。

- 数据库数据批量装载。
- 数据库转储。
- 介质故障恢复。
- 数据库的重组织。
- 性能监视等。

4. 数据库系统

数据库系统 (Database System , DBS) 是一个实际可运行的存储、维护和应用系统提供数据的软件系统。

数据库系统构成：DBMS；DB；应用软件；数据库管理员；用户。

1.1.2 数据库的发展史

数据库的发展史分为如下四个阶段：

- (1) 人工管理阶段。手工整理存储数据。
- (2) 文件系统阶段。使用磁盘文件来存储数据。
- (3) 数据库系统阶段。关系型数据库。
- (4) 高级数据库阶段。“关系-对象”型数据库。

当前数据库产品：

- Oracle：甲骨文公司开发。
- DB2：BM 公司开发。
- SQL Server：微软公司开发。
- Sybase：赛贝斯公司开发。
- MySQL：甲骨文公司开发。

1.1.3 数据库的类型

数据库的类型如下：

(1) 纯文本数据库。纯文本数据库是只用空格符、制表符和换行符来分割信息的文本文件。适用于小型应用，对于大中型应用来说它存在诸多限制：

只能顺序访问，不能进行随机访问。

查找数据和数据关系或多用户同时访问进行写操作时非常困难。

(2) 关系数据库。由于纯文本数据库存在诸多局限，因此人们开始研究数据模型，设计各种类型的使用方便的数据库。在数据库的发展史上，最具影响的数据库模型有：层次模型、网状模型和关系模型。其中，关系模型是目前应用最广泛和最有发展前途的一种数据模型，其数据结构简单，当前主流的数据库系统几乎都采用关系模型。

关系数据库中所谓的“关系”，实质上是一张二维表，如表 1.1 所示。

作为数据库中最为重要的数据库对象，数据库表的设计过程并非一蹴而就，上述课程表根本无法满足“选课系统”的功能需求。

表 1.1 课程表

	第1个字段 (列)	第2个字段 (列)	第3个字段 (列)	第4个字段 (列)
字段名 (列名)	课程号	课程名	人数上限	任课教师
第1条记录	1	java语言程序设计	60	张老师
第2条记录	2	MySQL数据库	150	李老师
第3条记录	3	c语言程序设计	60	王老师
第4条记录	4	英语	230	马老师
第5条记录	5	数学	230	田老师

事实上，数据库表的设计过程并非如此简单，本章的重点就是讨论如何设计结构良好的数据库表。

1.1.4 数据库的优点

数据库的优点如下：

- 数据按一定的数据模型组织、描述和储存。
- 可为各种用户共享。

- 冗余度较小，节省存储空间。
- 易扩展，编写有关数据库应用程序。

1.1.5 关系数据库管理系统

- Oracle：应用广泛、功能强大，分布式数据库系统；“关系-对象”型数据库。
- MySQL：快捷、可靠；开源、免费、与 PHP 组成经典的 LAMP 组合。
- SQL Server：针对不同用户群体的五个特殊的版本；易用性好。
- DB2：应用于大型应用系统，具有较好的可伸缩性。

通过“数据库管理系统”，数据库用户可以轻松地实现对数据库容器中各种数据库对象的访问（增、删、改、查等操作），并可以轻松地完成数据库的维护工作（备份、恢复、修复等操作），如图 1.1 所示。

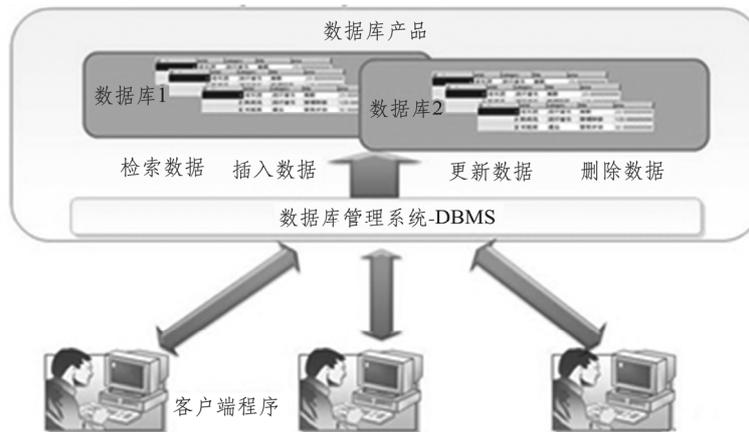


图 1.1 用户通过 DBMS 访问数据库

数据库用户无法直接通过操作系统获取数据库文件中的具体内容；数据库管理系统通过调用操作系统的进程管理、内存管理、设备管理以及文件管理等服务，为数据库用户提供管理、控制数据库容器中各种数据库对象、数据库文件的接口，如图 1.2 所示。

常用的数据库模型如图 1.3 所示。基于“关系模型”的数据库管理系统称为关系数据库管理系统（RDBMS）。

随着关系数据库管理系统的日臻完善，目前关系数据库管理系统已占据主导地位，如图 1.4 所示。

截至目前，MySQL 已经成功逆袭，如表 1.2 所示。

MySQL 逆袭原因如下：

- 第一，开源。MySQL 源代码免费下载。
- 第二，简单。MySQL 体积小，便于安装。
- 第三，性能优越。MySQL 性能足够与商业数据库媲美。
- 第四，功能强大。MySQL 提供的功能足够与商业数据库媲美。

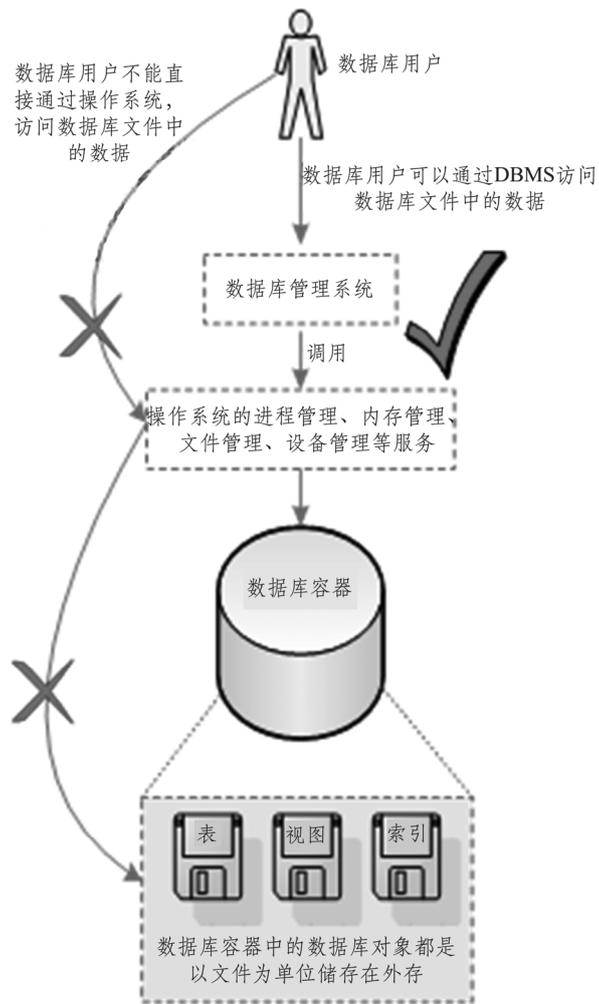


图 1.2 数据库管理系统调用操作系统的进程管理、内存管理、设备管理以及文件管理服务

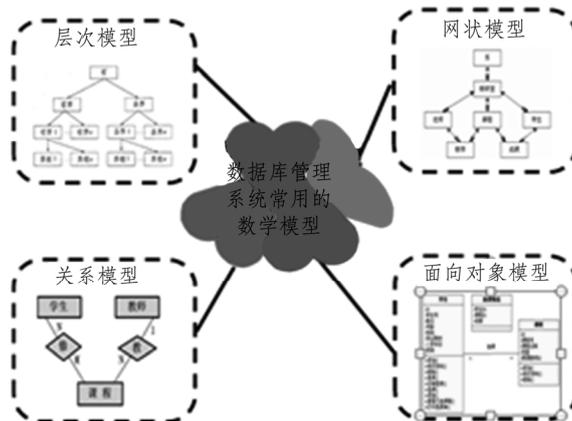


图 1.3 常用的数据库模型

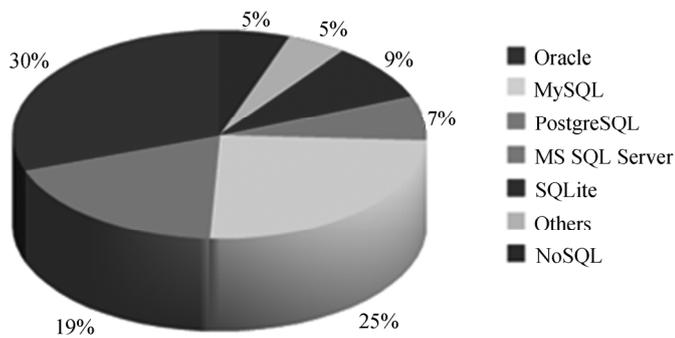


图 1.4 关系数据库管理系统占据市场份额

表 1.2 MySQL 市场占据逆袭数据

214 systems in ranking, February 2014

Rank	Last Month	DBMS	Database Model	Score	Changes
1.	1.	Oracle #	Relational DBMS	1500.23	+32.43
2.	2.	MySQL #	Relational DBMS	1288.39	-8.53
3.	3.	Microsoft SQL Server #	Relational DBMS	1214.27	-11.75
4.	4.	PostgreSQL #	Relational DBMS	230.45	+2.20
5.	↑ 6.	MongoDB #	Document store	195.17	+16.94
6.	↓ 5.	DB2 #	Relational DBMS	188.46	+0.15
7.	7.	Microsoft Access #	Relational DBMS	152.88	-22.11
8.	8.	SQLite #	Relational DBMS	93.00	-4.29
9.	9.	Sybase ASE #	Relational DBMS	87.88	-6.62
10.	10.	Cassandra #	Wide column store	80.31	-0.87
11.	11.	Teradata #	Relational DBMS	63.81	+2.36
12.	12.	Solr #	Search engine	62.70	+2.37
13.	13.	Redis #	Key-value store	55.81	+3.32
14.	14.	FileMaker #	Relational DBMS	51.90	+2.27
15.	↑ 17.	Informix #	Relational DBMS	35.67	+0.53

1.2 MySQL 基础

1.2.1 MySQL 介绍

1. MySQL 概念

MySQL 是一个小型关系型数据库管理系统,开发者为瑞典 MySQL AB 公司。目前 MySQL 被广泛地应用在 Internet 上的中小型网站中。由于其体积小、速度快、总体拥有成本低,尤其是开放源码这一特点,许多中小型网站为了降低网站总体拥有成本而选择了 MySQL 作为网站数据库。

2. MySQL 特征

- 性能快捷、优化 SQL 语言。
- 容易使用。
- 多线程和可靠性。
- 多用户支持。
- 可移植性和开放源代码。
- 遵循国际标准和国际化支持。
- 为多种编程语言提供 API。

3. MySQL5 特性

- 子查询。
- 视图。
- 存储过程。
- 触发器。
- 事务处理。
- 热备份。
- 二进制 Bit 类型。

4. MySQL 不足

- 不能直接处理 XML 数据。
- 一些功能上支持得不够完善和成熟。
- 不能提供任何 OLAP (实时分析系统) 功能。

5. MySQL 应用

MySQL 的官方网站引述 MySQL 是“世界上最受欢迎的开放源代码数据库”。这不是狂妄之语，数字可以证明它：目前，有超过 1 000 万份的 MySQL 被安装用于支付高负荷的网站和其他关键商业应用，包括像阿尔卡特、爱立信、朗讯、亚马逊、谷歌、纽约证券交易所、迪斯尼、雅虎、美国宇航局等这样的产业领袖。在下述网页你还能查看到 MySQL 和它竞争对手进行了短兵相接的比较。

<http://www.mysql.com/information/crash-me.php>

<http://www.mysql.com/information/benchmarks.html>

1.2.2 结构化查询语言 SQL

结构化查询语言 (Structured Query Language , SQL) 是一种应用最为广泛的关系数据库语言。该语言定义了操作关系数据库的标准语法，几乎所有的关系数据库管理系统都支持 SQL，如图 1.5 所示。

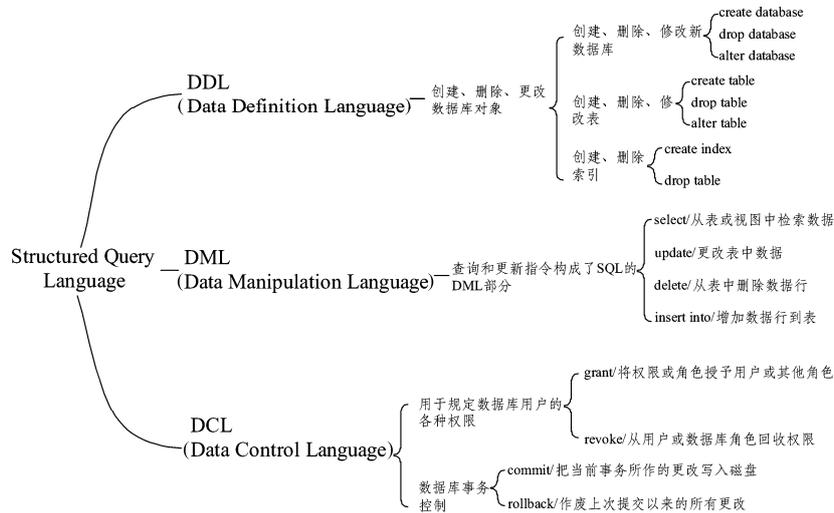


图 1.5 结构化查询语言构成