

# UNIX 数据库 实用教程

施雅民 倪琳 李惠群 等编  
胡传国 金荣得 审

上海科学普及出版社

# UNIX 数据库实用教程

---

施稚民 倪琳 李惠群 等编

胡传国 金荣得 审

上海科学普及出版社

(沪)新登字第 305 号

责任编辑: 胡名正 刘瑞莲

封面设计: 毛增南

# UNIX 数据库实用教程

施稚民 倪琳 李惠群 等编

胡传国 金荣得 审

上海科学普及出版社出版

(上海曹杨路 500 号 邮政编码 200063)

---

新华书店上海发行所发行 上海市委党校印刷厂印刷

开本 787×1092 1/16 印张 17 字数 407000

1994 年 9 月第 1 版 1994 年 9 月第 1 次印刷

---

ISBN 7-5427-0861-9/TP·213 定价: 16.50 元

## 内 容 提 要

本书是一本关于 UNIX 环境下数据库管理系统 (DBMS) 的实践性教程, 介绍了 UNIX 环境下开发 DBMS 应用程序的基础知识, 讨论了 DBMS 用户面临的实践性问题, 还叙述了 DBMS 在 MS-DOS、OS/2 以及 UNIX 环境下的区别。本书共分五部分。第一部分重点讨论关系模型。第二部分讨论 DBMS 如何与 UNIX 进行交互, 以及 DBMS 应用程序的开发和运用。第三部分综述 INFORMIX、INGRES、ORACLE 和 UNIFY 在 UNIX 环境下的编程基础知识。第四部分介绍如何选择适合应用需求的 DBMS 系统。第五部分展望 DBMS 的进展。

读者对象: UNIX 用户, UNIX 数据库应用程序开发人员, 大专院校计算机专业师生。

# 前 言

近年来, UNIX 在数据处理的领域中崭露头角, 推向市场的应用程序及软件包的数量日益增长。现在, 用户已可以使用各种工具, 特别是数据库管理系统 (DBMS) 来开发他们的各种应用软件。

在开发应用软件时, UNIX DBMS 的用户应清楚地了解操作系统在他们使用的应用程序中的作用, 也需要了解哪些因素会限制所选择的 DBMS 的进一步开发。应用程序的设计人员和开发人员应该知道, 在令人眼花缭乱的各种 DBMS 之间进行选择时, 如何才能选择好一个适合于他们应用需求的、更切实际的 DBMS。

本书不专门介绍 UNIX 操作系统。当然, 读者需要有一些数据处理的 UNIX 方面的知识。本书是一本 UNIX DBMS 方面的实践性指南。书中介绍了在实现一个应用程序时所要了解的必要基础知识, 也讨论了在 UNIX 环境中摆在 DBMS 用户面前的一些问题; 凡有需要和有可能的地方, 还对 DBMS 在 MS-DOS、OS/2 以及 UNIX 环境下的区别作了叙述, 以便读者了解 DBMS 在不同环境下的差异及其原因。

本书共分五个部分, 分别简介如下。

## 第一部分: 理论基础

这一部分简单回顾数据库管理系统的发展历史。因为 UNIX DBMS 的主体是基于关系模型的, 因此我们的重点放在讨论关系模型上。此外, 还叙述了一些基本原理和这一领域中的专业术语。

## 第二部分: UNIX 和 DBMS 应用程序

这一部分讨论数据库管理系统如何与基本操作系统进行交互。UNIX 会对 DBMS 可能提供哪些设施, 以及如何实现这些设施产生影响; 同时还会影响程序开发者设计一个应用程序时所使用的方法。这一部分的内容对应用程序设计人员、开发人员和最终用户都是很有帮助的。

## 第三部分: 四个 UNIX 数据库管理系统

我们在这一部分中综述了四个主要的 DBMS 产品所提供的设施。这四个 DBMS 产品是: INFORMIX、INGRES、ORACLE 和 UNIFY。大多数读者都会在这一部分中找到他们感兴趣的内容。

## 第四部分: 选择数据库管理系统

这一部分为如何选择一个最能适合你的应用需求的 DBMS 产品提供了许多实践经验。各种 DBMS 都有许多独特的、优异的特性, 但这并不意味着每个具体的应用都需要用到这些特性。这一部分的中心内容就是讨论如何确定哪些 DBMS 的设施和特性对特定的应用而言才是最重要的。

## 第五部分: 发展方向

展望了 DBMS 的进展。这些发展将会对今后应用程序的走向起到有意义的引导作用。

目前有许多软件包可在 UNIX 下运行。本书选择了四个最常用的 DBMS 产品进行讨论。这四个产品都支持工业标准 SQL 语言。其中的几个比较适合小型到中型的数据库，另外几个更适合大型的数据库。INFORMIX 在小型 UNIX 系统上已有一个广泛使用的产品，它的类似 PC 用户界面的技术改进了传统的 UNIX 方法。INGRES 提供了良好的产品化改进工具，开发人员可以自由选用；这些工具对于非技术型的用户来说也是比较容易掌握的。ORACLE 和 ACCELL (UNIFY) 都使用了 UNIX 中独特的“原始盘 (raw disk)”机制，这对于借助多卷标数据库来达到优异的性能是必要的。两者的不同之处在于，ACCELL 是 UNIX 下产生的，而 ORACLE 则源于其它操作系统，然后再移植到 UNIX 上。

本书对于对 UNIX DBMS 应用程序有兴趣的用户，已在使用或打算使用 UNIX DBMS 的应用程序开发人员，数据处理方面的管理人员，希望了解 UNIX (以及其它操作系统下，例如 MS-DOS) DBMS 有关知识的读者，以及大、中专院校师生和科研单位研究人员，都将具有实用的参考价值。

参加本书编、审工作的有施稚民、胡传国、金荣得、倪琳、朱伟华、张剑、李惠群、邱江等同志。

编者

# 目 录

|                        |      |
|------------------------|------|
| 第一部分 理论基础 .....        | (1)  |
| 第一章 什么是 DBMS .....     | (2)  |
| 1.1 简单的发展历史 .....      | (2)  |
| 1.2 DBMS 的目录 .....     | (4)  |
| 1.3 DBMS 模型 .....      | (5)  |
| 1.4 分层模型 .....         | (7)  |
| 1.5 网络模型 .....         | (8)  |
| 1.6 关系模型 .....         | (9)  |
| 1.7 实体关系和其它模型 .....    | (10) |
| 1.8 小结 .....           | (12) |
| 第二章 关系概念 .....         | (13) |
| 2.1 关系术语 .....         | (13) |
| 2.2 什么是规格化? .....      | (14) |
| 2.3 范式 .....           | (16) |
| 2.3.1 第一范式 (1NF) ..... | (16) |
| 2.3.2 第二范式 (2NF) ..... | (16) |
| 2.3.3 第三范式 (3NF) ..... | (17) |
| 2.3.4 第四范式 (4NF) ..... | (18) |
| 2.3.5 第五范式 (5NF) ..... | (19) |
| 2.3.6 规格化的程度 .....     | (20) |
| 2.4 关系型运算 .....        | (21) |
| 2.5 小结 .....           | (24) |
| 第三章 为什么要使用 DBMS? ..... | (25) |
| 3.1 数据控制 .....         | (25) |
| 3.1.1 数据字典 .....       | (25) |
| 3.1.2 数据类型 .....       | (26) |
| 3.1.3 数据存贮的管理 .....    | (28) |
| 3.1.4 存取方法 .....       | (29) |
| 3.1.5 用户视图 .....       | (32) |
| 3.1.6 存取保密控制 .....     | (32) |
| 3.1.7 并发控制 .....       | (33) |
| 3.1.8 事务控制和恢复 .....    | (35) |
| 3.1.9 数据完整性控制 .....    | (36) |

|                                    |             |
|------------------------------------|-------------|
| 3.2 实用程序包.....                     | (37)        |
| 3.2.1 菜单生成实用程序 .....               | (37)        |
| 3.2.2 屏幕表格实用程序 .....               | (38)        |
| 3.2.3 报表生成实用程序 .....               | (41)        |
| 3.2.4 特别询问设施 .....                 | (44)        |
| 3.2.5 主机语言接口 .....                 | (44)        |
| 3.2.6 第四代语言 .....                  | (45)        |
| 3.3 小结 .....                       | (46)        |
| <b>第四章 SQL 询问语言 .....</b>          | <b>(48)</b> |
| 4.1 为什么 SQL 很重要? .....             | (48)        |
| 4.2 SQL 数据定义语言 .....               | (49)        |
| 4.2.1 数据库定义 .....                  | (49)        |
| 4.2.2 表定义 .....                    | (50)        |
| 4.2.3 索引定义 .....                   | (51)        |
| 4.2.4 视图定义 .....                   | (52)        |
| 4.3 SQL 检索语句 .....                 | (52)        |
| 4.3.1 单表询问 .....                   | (53)        |
| 4.3.2 多表询问 .....                   | (54)        |
| 4.3.3 其它条件运算符 .....                | (56)        |
| 4.3.4 嵌套的 SQL 语句 .....             | (57)        |
| 4.4 SQL 数据管理语言 .....               | (59)        |
| 4.4.1 数据插入 .....                   | (59)        |
| 4.4.2 数据更新 .....                   | (60)        |
| 4.4.3 数据删除 .....                   | (60)        |
| 4.5 SQL 的扩展 .....                  | (61)        |
| 4.6 嵌入式 SQL 接口 .....               | (62)        |
| 4.7 小结 .....                       | (63)        |
| <b>第二部分 UNIX 和 DBMS 应用系统 .....</b> | <b>(65)</b> |
| <b>第五章 UNIX 功能和限制 .....</b>        | <b>(66)</b> |
| 5.1 UNIX 进程的基本结构 .....             | (66)        |
| 5.1.1 调度和交换 .....                  | (66)        |
| 5.1.2 用户进程 .....                   | (67)        |
| 5.1.3 进程间的通讯 .....                 | (68)        |
| 5.2 UNIX 下的数据存贮 .....              | (69)        |
| 5.2.1 文件系统 .....                   | (70)        |
| 5.2.2 原始盘 .....                    | (71)        |
| 5.3 UNIX 终端接口 .....                | (71)        |
| 5.3.1 与终端无关的用户接口 .....             | (71)        |

|                                 |             |
|---------------------------------|-------------|
| 5.3.2 终端驱动器功能 .....             | (72)        |
| 5.4 安全控制功能 .....                | (72)        |
| 5.5 并行控制工具 .....                | (74)        |
| 5.5.1 锁定文件 .....                | (74)        |
| 5.5.2 锁定信号 .....                | (75)        |
| 5.5.3 SYSTEM V 文件和记录锁定 .....    | (75)        |
| 5.5.4 连续处理事务的工具 .....           | (75)        |
| 5.6 网络功能 .....                  | (76)        |
| 5.6.1 UNIX 系统间的通讯 .....         | (76)        |
| 5.6.2 与其它操作系统的通讯 .....          | (77)        |
| 5.7 实时功能 .....                  | (77)        |
| 5.8 UNIX 调整 .....               | (78)        |
| 5.8.1 数据访问 .....                | (78)        |
| 5.8.2 终端激活 .....                | (79)        |
| 5.8.3 用户装入 .....                | (79)        |
| 5.9 应用系统开发工具 .....              | (79)        |
| 5.9.1 编程语言 .....                | (79)        |
| 5.9.2 程序开发工具 .....              | (80)        |
| 5.9.3 方案控制实用程序 .....            | (80)        |
| 5.10 小结 .....                   | (80)        |
| <b>第六章 开发一个 DBMS 应用系统 .....</b> | <b>(82)</b> |
| 6.1 数据存贮方法的选择 .....             | (82)        |
| 6.1.1 一个关系表一个文件 .....           | (82)        |
| 6.1.2 多个关系表一个文件 .....           | (84)        |
| 6.2 使用哪种访问方式? .....             | (85)        |
| 6.2.1 散列与 B 树访问的比较 .....        | (85)        |
| 6.2.2 参考连接的使用 .....             | (86)        |
| 6.2.3 数据集的使用 .....              | (87)        |
| 6.3 保持数据的一致性 .....              | (88)        |
| 6.4 并行控制要点 .....                | (89)        |
| 6.5 建立用户界面的功能 .....             | (90)        |
| 6.5.1 菜单界面的研究 .....             | (90)        |
| 6.5.2 表格界面 .....                | (92)        |
| 6.5.3 用一种第四代生成语言驱动表格 .....      | (93)        |
| 6.5.4 显示报表的界面 .....             | (94)        |
| 6.5.5 打印报表 .....                | (94)        |
| 6.6 如何用一个宿主程序接口进行开发 .....       | (95)        |
| 6.6.1 函数调用接口 .....              | (95)        |
| 6.6.2 嵌入式查询语言接口 .....           | (96)        |

|                                   |              |
|-----------------------------------|--------------|
| 6.7 宿主语言和第四代生成语言的比较.....          | (97)         |
| 6.8 小结.....                       | (97)         |
| <b>第七章 运用一个 DBMS 应用程序 .....</b>   | <b>(99)</b>  |
| 7.1 对 UNIX 系统需了解多少? .....         | (99)         |
| 7.1.1 集成化的菜单启动 .....              | (100)        |
| 7.1.2 暂回 SHELL 机制 .....           | (100)        |
| 7.1.3 打印队列界面 .....                | (101)        |
| 7.1.4 批作业界面 .....                 | (102)        |
| 7.2 用户的性能需求 .....                 | (102)        |
| 7.2.1 交互程序的性能 .....               | (103)        |
| 7.2.2 后台作业的性能 .....               | (103)        |
| 7.3 系统管理需求 .....                  | (104)        |
| 7.3.1 用户标识管理 .....                | (104)        |
| 7.3.2 数据库的备份与恢复 .....             | (106)        |
| 7.3.3 事务日志功能 .....                | (107)        |
| 7.4 安全访问控制 .....                  | (108)        |
| 7.4.1 登录和口令安全 .....               | (108)        |
| 7.4.2 数据访问控制 .....                | (108)        |
| 7.4.3 程序访问控制 .....                | (109)        |
| 7.4.4 目录和文件访问控制 .....             | (109)        |
| 7.5 适应性 .....                     | (110)        |
| 7.5.1 适应数据库结构的改变 .....            | (110)        |
| 7.5.2 适应修改功能 .....                | (111)        |
| 7.5.3 适应 DBMS 和 UNIX 升级 .....     | (111)        |
| 7.6 增长和移植 .....                   | (112)        |
| 7.7 小结 .....                      | (112)        |
| <b>第三部分 四个 UNIX 数据库管理系统 .....</b> | <b>(114)</b> |
| <b>第八章 Informix 数据库管理系统 .....</b> | <b>(115)</b> |
| 8.1 引言 .....                      | (115)        |
| 8.2 程序包和组件 .....                  | (115)        |
| 8.2.1 Informix 软件产品 .....         | (115)        |
| 8.2.2 开发组件 .....                  | (118)        |
| 8.3 数据控制 .....                    | (120)        |
| 8.3.1 数据类型 .....                  | (120)        |
| 8.3.2 存储结构 .....                  | (120)        |
| 8.3.3 访问方法 .....                  | (124)        |
| 8.3.4 访问安全性控制 .....               | (125)        |
| 8.3.5 并行性控制选择 .....               | (127)        |

|                                 |              |
|---------------------------------|--------------|
| 8.3.6 数据完整性度量 .....             | (128)        |
| 8.4 实用程序包 .....                 | (128)        |
| 8.4.1 菜单子系统 .....               | (128)        |
| 8.4.2 屏幕表格设施 .....              | (129)        |
| 8.4.3 报表生成 .....                | (132)        |
| 8.4.4 查询语言 .....                | (133)        |
| 8.4.5 主机语言接口 .....              | (134)        |
| 8.4.6 第四代语言 .....               | (135)        |
| 8.5 与 UNIX 的结合 .....            | (136)        |
| 8.5.1 注册访问用法 .....              | (137)        |
| 8.5.2 文件访问和管理 .....             | (137)        |
| 8.5.3 进程结构 .....                | (137)        |
| 8.5.4 SHFLL 访问 .....            | (138)        |
| 8.6 小结 .....                    | (138)        |
| <b>第九章 INGRES 数据库管理系统 .....</b> | <b>(139)</b> |
| 9.1 引言 .....                    | (139)        |
| 9.2 程序包和组件 .....                | (139)        |
| 9.2.1 INGRES 软件产品 .....         | (139)        |
| 9.2.2 开发组件 .....                | (141)        |
| 9.3 数据控制 .....                  | (142)        |
| 9.3.1 数据类型 .....                | (143)        |
| 9.3.2 存储器结构 .....               | (143)        |
| 9.3.3 访问方法 .....                | (145)        |
| 9.3.4 访问安全性控制 .....             | (145)        |
| 9.3.5 并行性控制选择 .....             | (147)        |
| 9.3.6 数据完整性测量 .....             | (147)        |
| 9.4 实用程序包 .....                 | (148)        |
| 9.4.1 菜单子系统 .....               | (148)        |
| 9.4.2 屏幕表格设施 .....              | (148)        |
| 9.4.3 报表生成 .....                | (150)        |
| 9.4.4 查询语言 .....                | (152)        |
| 9.4.5 主机语言接口 .....              | (153)        |
| 9.4.6 第四代语言 .....               | (154)        |
| 9.5 与 UNIX 组合 .....             | (156)        |
| 9.5.1 逻辑访问用法 .....              | (156)        |
| 9.5.2 文件访问及管理 .....             | (157)        |
| 9.5.3 进程体系结构 .....              | (157)        |
| 9.5.4 SHELL 访问 .....            | (157)        |
| 9.6 小结 .....                    | (158)        |

|                                     |                  |
|-------------------------------------|------------------|
| <b>第十章 ORACLE 数据库管理系统 .....</b>     | <b>(159)</b>     |
| 10.1 引言 .....                       | (159)            |
| 10.2 软件包和组件 .....                   | (159)            |
| 10.2.1 ORACLE 软件产品 .....            | (159)            |
| 10.2.2 开发组件 .....                   | (161)            |
| 10.3 数据控制 .....                     | (162)            |
| 10.3.1 数据类型 .....                   | (162)            |
| 10.3.2 存贮结构 .....                   | (163)            |
| 10.3.3 访问方法 .....                   | (165)            |
| 10.3.4 访问安全性控制 .....                | (166)            |
| 10.3.5 并行性控制任选项 .....               | (167)            |
| 10.3.6 数据完整性的考虑 .....               | (168)            |
| 10.4 程序包 .....                      | (169)            |
| 10.4.1 菜单子系统 .....                  | (169)            |
| 10.4.2 屏幕表格功能 .....                 | (169)            |
| 10.4.3 报表生成程序 .....                 | (172)            |
| 10.4.4 查询语言 .....                   | (175)            |
| 10.4.5 主机语言接口 .....                 | (177)            |
| 10.4.6 第四代语言 .....                  | (178)            |
| 10.5 与 UNIX 相组合 .....               | (180)            |
| 10.5.1 注册访问用法 .....                 | (180)            |
| 10.5.2 文件访问和管理 .....                | (180)            |
| 10.5.3 进程体系结构 .....                 | (181)            |
| 10.5.4 SHELL 访问 .....               | (182)            |
| 10.6 小结 .....                       | (182)            |
| <br><b>第十一章 ACCELL 应用开发系统 .....</b> | <br><b>(184)</b> |
| 11.1 引言 .....                       | (184)            |
| 11.2 软件包和组件 .....                   | (184)            |
| 11.2.1 Unify 软件产品 .....             | (184)            |
| 11.3 数据控制 .....                     | (188)            |
| 11.3.1 数据类型 .....                   | (188)            |
| 11.3.2 存贮结构 .....                   | (189)            |
| 11.3.3 访问方法 .....                   | (190)            |
| 11.3.4 访问安全性控制 .....                | (191)            |
| 11.3.5 并行控制的可选项 .....               | (192)            |
| 11.3.6 数据完整性度量 .....                | (193)            |
| 11.4 实用程序包 .....                    | (194)            |
| 11.4.1 菜单子系统 .....                  | (194)            |
| 11.4.2 屏幕表格设施 .....                 | (195)            |

|                               |              |
|-------------------------------|--------------|
| 11.4.3 报表生成 .....             | (196)        |
| 11.4.4 查询语言 .....             | (197)        |
| 11.4.5 主机语言接口 .....           | (198)        |
| 11.4.6 第四代语言 .....            | (199)        |
| 11.5 与 UNIX 结合 .....          | (201)        |
| 11.5.1 注册访问用法 .....           | (202)        |
| 11.5.2 文件访问和管理 .....          | (202)        |
| 11.5.3 进程体系结构 .....           | (202)        |
| 11.5.4 SHELL 访问 .....         | (203)        |
| 11.6 小结 .....                 | (203)        |
| <b>第四部分 选择数据库管理系统 .....</b>   | <b>(204)</b> |
| <b>第十二章 确定你的需求 .....</b>      | <b>(205)</b> |
| 12.1 数据容量 .....               | (205)        |
| 12.2 事务容量 .....               | (207)        |
| 12.3 性能需求 .....               | (208)        |
| 12.4 安全性需求 .....              | (209)        |
| 12.5 例程管理 .....               | (210)        |
| 12.6 应用功能的增强 .....            | (211)        |
| 12.7 未来的可移植性 .....            | (212)        |
| 12.8 小结 .....                 | (213)        |
| <b>第十三章 权衡评估 .....</b>        | <b>(216)</b> |
| 13.1 哪些数据库管理系统设施是至关重要的? ..... | (216)        |
| 13.2 成批运行还是交互运行? .....        | (218)        |
| 13.3 开发制约 .....               | (219)        |
| 13.4 程序修改的频繁程度和简易程度 .....     | (221)        |
| 13.5 管理控制 .....               | (221)        |
| 13.6 小结 .....                 | (222)        |
| <b>第十四章 性能的基准测试 .....</b>     | <b>(224)</b> |
| 14.1 开始操作 .....               | (224)        |
| 14.2 UNIX 工具 .....            | (225)        |
| 14.3 设计数据库管理系统基准测试 .....      | (226)        |
| 14.3.1 多少数据? .....            | (227)        |
| 14.3.2 多少用户? .....            | (227)        |
| 14.3.3 生成测试原本 .....           | (228)        |
| 14.4 带有逼真的输入/输出的测试 .....      | (229)        |
| 14.5 运行基准测试 .....             | (233)        |
| 14.6 解释测试结果 .....             | (235)        |

|                                 |              |
|---------------------------------|--------------|
| 14.6.1 输入 / 输出极限指示 .....        | (235)        |
| 14.6.2 CPU 极限指示 .....           | (236)        |
| 14.7 小结 .....                   | (237)        |
| <b>第五部分 发展方向 .....</b>          | <b>(239)</b> |
| <b>第十五章 其他问题 .....</b>          | <b>(240)</b> |
| 15.1 为什么要使用多台机器? .....          | (240)        |
| 15.1.1 地理位置 .....               | (240)        |
| 15.1.2 处理工作量的分布 .....           | (241)        |
| 15.1.3 部门分布 .....               | (242)        |
| 15.2 分布式系统 .....                | (242)        |
| 15.2.1 网络数据库系统 .....            | (243)        |
| 15.2.2 分布式数据库系统 .....           | (244)        |
| 15.2.3 异质分布式系统 .....            | (245)        |
| 15.3 分析员的工具 .....               | (246)        |
| 15.4 自然语言接口 .....               | (247)        |
| 15.5 专家系统 .....                 | (248)        |
| 15.6 小结 .....                   | (249)        |
| <b>附录 A 数据库管理系统评估检查清单 .....</b> | <b>(250)</b> |
| A.1 应用的特性 .....                 | (250)        |
| A.2 数据类型 .....                  | (250)        |
| A.3 交互表格功能 .....                | (251)        |
| A.4 批量更新功能 .....                | (252)        |
| A.5 报表生成设施 .....                | (252)        |
| A.6 安全性需求 .....                 | (253)        |
| A.7 管理功能 .....                  | (253)        |
| <b>附录 B 应用开发检查清单 .....</b>      | <b>(254)</b> |
| B.1 数据库设计 .....                 | (254)        |
| B.2 交互接口设计 .....                | (254)        |
| B.3 打印报表设计 .....                | (255)        |
| B.4 管理功能设计 .....                | (256)        |
| B.5 计划管理 .....                  | (256)        |

# 第一部分 理论 基础

这一部分简单地介绍一下数据库管理系统 (DBMS) 技术的一些基本概念和基础理论, 其目的并不是作深层次的技术性叙述, 而只是向读者提供必要的、有关 DBMS 的背景知识, 便于读者阅读本书后面那些更专业性和更技术性的章节。然而, 本部分的那些章节中也包含了一些非常重要的基础知识。这些知识将告诉你, DBMS 能为你的应用系统开发工作做些什么, 以及为什么你在开发应用系统时需要用到 DBMS。

第一章讨论了如果使用 DBMS 的数据存贮方法时, 开发人员将面临的一些问题。

第二章介绍关系数据库管理系统, 以及一些数据设计技术中用到的基本术语。

另外两章着重从商业产品的角度来描述某些设施类型, 目的在于建立一些使 UNIX 操作系统能作用于这些设施的基础规则, 本书中提到的设施是 DBMS 产品所提供的, 也可能是以后开发的目标。

# 第一章 什么是 DBMS

数据库是一些数据的贮存,是以某种形式关联起来的文件的集合。而数据库管理系统则是随着计算机技术的发展而发展起来的、用于管理数据库的一种工具。数据库管理系统的开发成功和迅速发展,打开了数据库管理的崭新局面,在各行各业中得到了广泛的应用。

本章简单地介绍一下数据库管理系统的作用和基本模型。所作的讨论是概念性的,与任何特定的操作系统都没有直接的关系。

## 1.1 简单的发展历史

当人们开始保存书写的记录时,实际上就开始有了数据库。无论这些书写记录是刻在石板上,写在纸上,保存在公文柜里,或者存贮在计算机中,这些贮存起来的数据都构成了数据库。在计算机发展的早期,或即使是在今天,我们都使用“文件”这个词来表示某些数据的集合。由于早期计算机硬件上的限制,这些文件大多数都贮存在卡片或纸带上。随着磁性存储载体的出现,我们开始使用磁带、磁鼓来存贮数据,然后发展到将文件联机贮存到磁盘上。

当我们开始联机保存越来越多的文件时,就出现了数据库。数据库是指以某种方式关联的文件的集合。例如,我们所说的帐务数据库、商场数据库或个人数据库。一个帐务数据库可能包括如总帐文件、卖主文件和顾客文件等几个文件。

但是,既然数据库已存在这么长时间,而且我们也根据自己的习惯管理它们,那为什么我们还需要数据库管理系统呢?问题是这些年来软件开发的方法能否提高开发效率。我们必须验证一下,如果不用 DBMS,将如何建立一个应用系统,以及为什么 DBMS 能帮助改善系统开发的效率。

最早的软件是用计算机能识别的基本机器码“0”和“1”写成的。然后,我们发展了汇编语言使软件写起来简单些。我们把机器码时代的程序称为第一代语言,汇编语言是第二代语言。有了汇编语言,我们可以编写软件,并且工作起来比用机器码时快得多。换句话说,我们提高了软件开发的生产效率。

接着发展为高级语言,如 COBOL、FORTRAN 及目前许多其它语言,我们称之为第三代程序设计语言,再次使开发生产效率提高了一个数量级。因为这些语言比汇编语言更易于编写。另外,与能理解汇编的人相比有更多的人能学会并有效地运用这些语言。通过把开发工作逻辑地推向下一步,DBMS 又一次提高了开发生产效率。与 COBOL 这样的第三代语言相比,第四代语言提供了更高级的开发接口。

DBMS 提供了许多加速屏幕接口开发的工具,例如需要好几页第三代程序语言代码才能完成的菜单和交互式表格。通过提供标准工具,它们不仅减少了程序中的重复代码,而且强制了用户接口的一致性。DBMS 为你开发典型的以表格为基础的交互式程序提供了更简便快捷的方法。同样,它提供了报表生成程序,你可以只用几条命令就建立一个典型的报表,而无需像 COBOL 之类的语言那样编写一页页的代码。这些工具也像 COBOL 语言提

高汇编语言的生产效率一样，通过更高级的开发接口进一步提高了生产效率。

DBMS 使人们认识到能通过终端用户询问来操作数据。非专业人员只需少量训练就能使用特设询问工具。你仍需要程序员编写程序、设置好报表或那些常用部分，但是，如果你现在需要数据，不必忙于安排，特设询问工具能帮你完成这项工作。

在改善开发语言的同时，为了免得编写代码去存取存贮器和外围设备，操作系统也随之有所改善。操作系统管理着外围设备和程序的运行。从数据管理角度来说，那些改进同样意味着我们不再需要关心某个文件起始和结束的磁盘区域。我们可以通过有意义的名字特指各个文件，并且按指定大小确定文件中的记录。操作系统的文件管理部分将这些文件名和记录说明转换成磁盘上的一些数据块。

操作系统的文件管理部分进一步发展为能提供不同类型的文件：顺序存取文件、索引存取文件和随机存取文件。有了这些文件，文件中各记录的快速存取就变得非常简单，因为我们可以利用关键字在索引中进行查找，而不必在整个文件中顺序查找。所以，我们又提高了程序设计的生产效率。（遗憾的是，UNIX 只开发了一种类型的文件：字节流文件。它所支持的唯一一种记录是在包括可打印字符的文件中把 ASCII 的新行（NL）字符当作记录终止符的任意记录。对于包含二进制数据的文件来说，没有这种约定。）

即使有了所有这些发展，程序设计仍面临许多费时的任务。例如，操作系统的文件管理设施实际上并不知道一个记录会包含什么类型的数据，它们只是简单地把记录当作一组字节，直到访问程序时才解释一个记录中每个数据条目（字段）在什么地方起始和终止。

一个 DBMS 能辨别各个数据项（字段），于是提供了比操作系统数据管理部分更高水平的抽取手段。所以，我们不再需要存取由字节组成的记录并把它们解释成有意义的字段。记录中字段的顺序和位置变得不重要了，并且在大多 DBMS 产品中也变得无关紧要了。

当用普通操作系统文件时，即使只用少量字段，程序也必须读入整个记录，每个程序必须知道这个记录中每个字段的确切地址。假设我们想在文件的记录中添加一个字段（在任何应用中都经常会出现这种情况），这个添加的字段可能会增加记录的大小，为了说明新的字段必须修改所有存取这个记录的程序，包括不用它的那些程序。于是，我们必须修改比实际需要的多得多的程序，而且当然还要对每个程序进行测试。所有这些行动消耗了程序设计资源中非常重要的部分。

针对这些不必要的修改，DBMS 提供给我们一种方法，它的接口使我们不用处理整个记录，而是直接处理字段，并且仅仅是那些在每个程序中都需要字段。于是，在记录中添加一个新字段并不影响操作这个记录的所有程序，而只影响那些需要用到这些新字段的程序。

使用普通操作系统文件时，开发者们还面临其他问题，例如，考虑一下商业应用中存贮的像日期这样的普通类型的数据，在数据库应用中，我们经常存贮如订货日期、租用日期、发票日期等等。我们开发了一些例行程序来解释日期、处理日期，并将它们用于报表中，比方说将报表限制在一段时间内。不幸的是，我们重复地开发了这些例行程序，有时甚至是在同一应用中。

DBMS 提供了处理常用数据的标准设施。它将数据类型同每个字段结合起来，所以它能用一种适当的方法操作像日期和钱这样的数据。这组普通的工具允许你指定像“今天”和“本周”这样的数据检索条件，我们不再需要为了这些普通功能开发和维护用户例行程序了。