



北京市高等教育精品教材立项项目

数据库

原理及应用

张晋连 等编著



3



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

北京市高等教育精品教材立项项目

TP311.13
Z123

数据库原理及应用

张晋连等 编著

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书较为详细地介绍数据库基本原理,但不在理论性太强的问题上花过多篇幅;结合当前流行软件,理论联系实际。本书分为两大部分,第一部分系统地介绍了数据库系统的基本原理,以及关系数据库实现所依据的数学基础,还较为详细地介绍了关系数据库设计的方法和标准的数据库查询语言 SQL。第二部分介绍了目前流行的 SQL Server 2000 数据库管理系统的使用和 SQL Server 2000 系统中使用的 Transact-SQL 语言,还介绍了数据库保护的内容,最后结合一个实际应用系统,介绍了如何进行数据库设计。

本书既可作为普通高校或成人院校本科非计算机专业的专业课教材,也可作为相关领域技术人员的参考用书或培训教材。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

数据库原理及应用/张晋连等编著. —北京:电子工业出版社,2004.7
北京市高等教育精品教材立项项目
ISBN 7-121-00080-6

I. 数… II. 张… III. 数据库系统—高等学校—教材 IV.TP311.13

中国版本图书馆 CIP 数据核字(2004)第 065156 号

责任编辑:李 岩

印 刷:北京牛山世兴印刷厂

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

经 销:各地新华书店

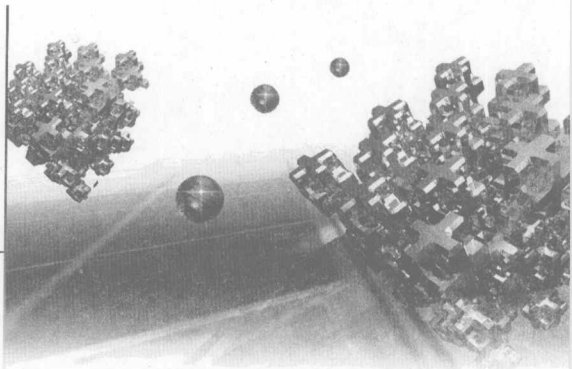
开 本:787×980 1/16 印张:16.5 字数:358.5 千字

印 次:2004 年 7 月第 1 次印刷

印 数:5 000 册 定价:21.00 元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。联系电话:(010) 68279077。质量投诉请发邮件至 zlts@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

前言



数据库技术是目前计算机领域发展最快、应用最广泛的技术。它的应用遍及各行各业，大到操作系统程序，全国联网的飞机票、火车票订票系统，银行业务系统；小到小型企事业单位的管理信息系统，家庭理财系统；互联网中日渐流行的动态网站都能看到数据库应用的身影。因此，数据库技术是计算机专业必须掌握的一门课程。

现有数据库类教材中所涉及的理论部分都比较全面，但所涉及的具体数据库管理系统，或者太过时，如 FoxBASE；或者对实验环境要求太高，不便于学生上机实际操作，如 Oracle。而大量流行的各种数据库管理系统的培训教程只是教授如何使用，理论联系实际不够，只适用于短期培训，不适合作为成人院校的教材使用。因此，需要一类既要比较全面地介绍数据库原理方面的知识，又能结合当前流行数据库系统的使用、对实验条件要求又不是很高的教材。编者正是基于这种考虑编写了本教材。

本书分为两大部分。前 5 章的内容作为第一部分。第一部分系统地介绍了数据库系统的基本原理，以及关系数据库实现所依据的数学基础，还较为详细地介绍了关系数据库设计的方法和标准的数据库查询语言——SQL。第 6~8 章是第二部分。第二部分介绍了目前流行的 SQL Server 2000 数据库管理系统的使用，内容包括：SQL Server 2000 系统的基本使用方法、Transact-SQL 语言、数据库的保护及一个数据库设计实例。使学生在学完这门课程之后，既能掌握一定的数据库理论，又能够联系实际独立开发具有一定实用功能的数据库应用系统。

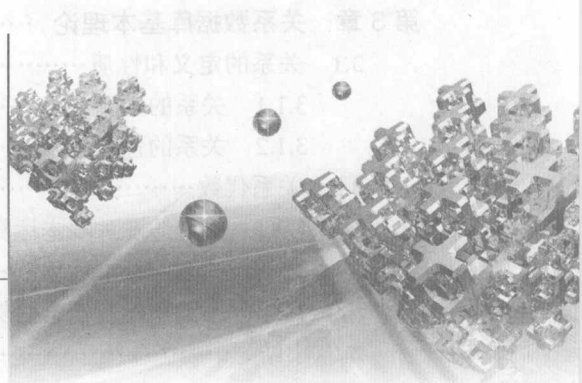
本书第 1, 2, 3 章和第 4 章的部分内容及全书的审定统稿由张晋连编写，第 4 章的大部分内容 by 刘振清编写，第 5 章由尚颖编写，第 6 章和第 8 章由柏浩编写，第 7 章由庞辉编写。在本书的编写过程中，得到了电子工业出版社刘宪兰、张毅和李岩编辑的大力支持与帮助，还得到了中国科学院管理干部学院领导的支持，在此表示由衷的感谢。

由于编者水平有限，书中难免出现问题与缺陷，恳请广大读者批评指正。

编者

2003 年 7 月于北京

目 录



第一部分 数据库基本原理	1
第 1 章 数据库系统概述	2
1.1 数据库技术的术语	2
1.1.1 信息、数据和数据处理	2
1.1.2 数据库、数据库管理系统和数据库系统	3
1.2 数据处理技术的发展过程	5
1.2.1 人工管理阶段	5
1.2.2 文件系统阶段	6
1.2.3 数据库系统阶段	7
1.2.4 数据库技术发展新趋势	8
1.3 数据库系统的体系结构	9
1.3.1 数据库体系结构	9
1.3.2 数据库系统体系结构	12
本章小结	14
习题 1	14
第 2 章 数据模型	15
2.1 现实世界的信息化过程	15
2.1.1 概念模型及其表示	16
2.1.2 概念模型的表示方法	19
2.2 常见的三种数据模型	21
2.2.1 层次模型	21
2.2.2 网状模型	23
2.2.3 关系模型	24
本章小结	27
习题 2	28

第3章 关系数据库基本理论	29
3.1 关系的定义和性质	29
3.1.1 关系的定义	29
3.1.2 关系的性质	30
3.2 关系代数	31
3.2.1 集合的三种基本运算——并、交、差	31
3.2.2 关系的基本运算	33
3.2.3 用关系代数检索的例子	37
本章小结	37
习题3	38
第4章 数据库设计	40
4.1 数据库设计概述	40
4.1.1 数据库设计的目标	40
4.1.2 数据库设计的方法	40
4.1.3 数据库设计的步骤	41
4.2 需求分析	42
4.2.1 需求分析的任务和重点	42
4.2.2 需求分析的步骤和常用方法	43
4.2.3 数据字典	44
4.3 概念结构设计	46
4.3.1 概念结构设计的方法与步骤	46
4.3.2 数据抽象与局部视图设计	47
4.3.3 视图的集成	48
4.4 逻辑结构设计	49
4.4.1 概念模型向关系模型的转换	49
4.4.2 数据模型的优化	50
4.4.3 用户子模式的设计	54
4.5 数据库的物理设计	55
4.5.1 确定数据库的物理结构	55
4.5.2 评价物理结构	57
4.6 数据库的实施和维护	57
4.6.1 数据库的实施	57
4.6.2 数据库维护	58
4.7 数据库设计举例	59
4.7.1 需求分析	60
4.7.2 概念结构设计	66

4.7.3	总体概念设计	67
4.7.4	逻辑结构设计	67
4.7.5	数据库物理设计和实施	68
	本章小结	70
	习题 4	71
第 5 章	关系数据库标准语言——SQL	72
5.1	SQL 概述及特点	72
5.1.1	SQL 概述	72
5.1.2	SQL 的特点	72
5.1.3	SQL 的基本概念	74
5.2	SQL 的数据定义	74
5.2.1	基本表的定义	74
5.2.2	基本表的修改和删除	76
5.2.3	索引的建立和删除	77
5.3	SQL 的数据查询	78
5.3.1	简单查询	79
5.3.2	连接查询	82
5.3.3	嵌套查询	84
5.3.4	常用库函数、Group By 和 Order By 的应用	88
5.4	SQL 的数据操纵	90
5.4.1	插入语句	90
5.4.2	删除语句	91
5.4.3	更新语句	91
5.5	视图	93
5.5.1	视图的定义和删除	93
5.5.2	视图查询	94
5.5.3	视图更新	95
5.5.4	视图的优点	97
5.6	数据控制	98
5.6.1	授权	99
5.6.2	收回权限	99
5.7	嵌入式 SQL	100
5.7.1	不用游标的 DML 语句	100
5.7.2	使用带游标的 DML 语句	102
	本章小结	104
	习题 5	104

第二部分 数据库应用	107
第 6 章 SQL Server 2000	108
6.1 SQL Server 2000 的安装与配置	108
6.1.1 系统需求	108
6.1.2 安装的过程	109
6.1.3 停止和启动 SQL Server 服务	111
6.1.4 SQL Server 2000 的安装目录	113
6.1.5 SQL Server 系统数据库和系统表	117
6.2 管理初步	119
6.2.1 使用 SQL Server 企业管理器	119
6.2.2 用企业管理器注册服务器	121
6.2.3 连接和断开服务器	123
6.3 使用企业管理器创建数据库	124
6.3.1 在企业管理器中创建表	126
6.3.2 在企业管理器中创建约束	127
6.3.3 在企业管理器中修改表	130
6.3.4 在企业管理器中查看表的属性	130
6.3.5 在企业管理器中查看表中的数据	130
6.3.6 在企业管理器中删除表	131
6.3.7 新建索引	132
6.4 浏览数据库	132
6.4.1 查看数据库	132
6.4.2 删除数据库	133
6.5 使用企业管理器创建存储过程	133
6.6 创建和管理视图	134
6.6.1 用企业管理器创建视图	134
6.6.2 修改视图	135
6.7 数据库备份	136
6.7.1 如何创建数据库备份 (使用企业管理器)	136
6.7.2 创建备份设备	136
6.7.3 创建数据库备份	137
6.7.4 还原数据库备份	139
6.7.5 查看备份设备	139
6.7.6 删除一个备份设备	140
6.7.7 调度备份	140
6.7.8 备份系统数据库	140

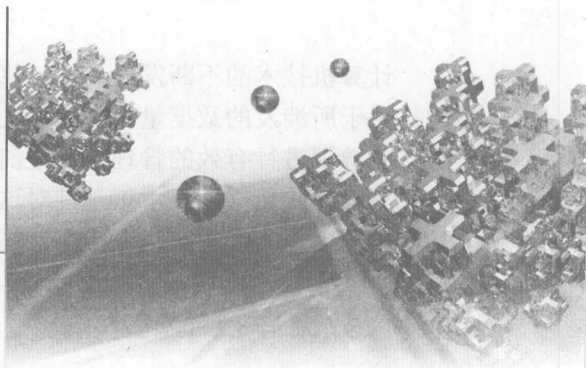
6.8	恢复数据库	141
6.9	安全配置	142
6.9.1	身份验证模式	142
6.9.2	使用企业管理器设置 Windows 身份验证模式的安全性	144
6.9.3	使用企业管理器设置混合模式的安全性	145
6.9.4	使用企业管理器添加 SQL Server 登录的方法	145
	本章小结	146
	习题 6	146
第 7 章	Transact-SQL 语言	147
7.1	Transact-SQL 的数据定义与操纵	147
7.1.1	数据库	147
7.1.2	数据库表	150
7.1.3	视图	152
7.1.4	数据库查询	153
7.2	Transact-SQL 的数据类型	153
7.2.1	系统数据类型	153
7.2.2	用户定义数据类型	157
7.3	Transact-SQL 的函数	159
7.3.1	集合函数	159
7.3.2	数学函数	160
7.3.3	字符串函数	161
7.3.4	日期函数	165
7.3.5	转换函数	165
7.3.6	系统函数	166
7.3.7	其他函数	167
7.4	SQL Server 编程结构	168
7.4.1	注释语句	168
7.4.2	批处理	168
7.4.3	变量	169
7.4.4	显示信息	172
7.4.5	流控制语句	173
7.4.6	事务	180
7.5	游标的使用	183
7.5.1	申明游标	183
7.5.2	打开游标	186
7.5.3	存取游标	186

7.5.4	定位游标	188
7.5.5	关闭游标	188
7.5.6	释放游标	188
7.5.7	游标的使用方法	189
7.6	存储过程	191
7.6.1	创建存储过程	191
7.6.2	带参数的存储过程	194
7.6.3	修改存储过程	201
7.6.4	删除存储过程	201
7.6.5	自动执行存储过程	201
7.6.6	系统存储过程	202
7.7	触发器的创建和管理	203
7.7.1	创建触发器	203
7.7.2	使用触发器	206
7.7.3	INSTEAD 函数	212
7.7.4	触发器组合	214
7.7.5	修改和删除触发器	217
	本章小结	218
	习题 7	218
第 8 章	数据库的保护	220
8.1	数据库的安全性	220
8.1.1	安全措施	220
8.1.2	特权和角色	221
8.1.3	数据的安全性策略	223
8.2	数据库完整性	224
8.2.1	数据完整性的分类	224
8.2.2	完整性约束的检查和处理	228
8.3	并发控制	230
8.3.1	事务	230
8.3.2	事务的串行调度、并发调度及其可串行化	233
8.3.3	事务并发执行所带来的问题	234
8.3.4	封锁	236
8.4	数据库备份和恢复	242
8.4.1	数据库系统的故障	242
8.4.2	恢复数据库所使用的结构	242
8.4.3	基于日志文件的数据恢复技术	243

本章小结	249
习题 8	249
参考文献	250

数据库系统教程 第一章

第一部分



数据库基本原理

第1章 数据库系统概述

计算机技术的不断发展,使计算机的应用从早期的数值计算扩展到了数值处理领域中。由于所涉及的数据量大,它们之间的联系又错综复杂,所以主要考虑的问题是如何对这些数据进行有效的管理,以便能对它们方便迅速地进行查找、插入、删除及修改等一系列操作。数据库技术正是针对这一目的而发展起来的一项重要技术。目前,数据库技术的应用十分广泛,几乎涉及到人类社会的所有领域。

1.1 数据库技术的术语

在数据库技术中,有一些常用的术语,这些术语是我们经常听到和看到的,但真正了解它们的含义很有必要。

1.1.1 信息、数据和数据处理

信息和数据是两个比较容易混淆的概念,通过实例可以较好地理解这两个概念的基本含义。

1. 信息和数据

信息 (Information): 信息是用于向人们(或机器)提供关于现实世界新的事实的知识。

数据 (Data): 数据是用于载荷信息的物理符号。在日常生活中,数据涉及的面很广、种类也很多,像数字、文字、图形、声音及学生档案记录等都属于数据。

【例 1.1】 气象台每天测量气温的变化,记录下来的一系列温度值即是数据。分析这些数据即可得到未来的气温将是多少度,是高了还是低了,分析的结果就是我们希望了解的信息。

【例 1.2】 人口普查。每个人所填写的人口普查表中的内容即是数据,分析这些数据即可得到全国人口总数,男、女人口比率,各年龄段人口比率,受不同教育层次的人口比率等信息。

从采集的数据到最终获得有用的信息,要经过一个过程,这个过程就是数据处理。

2. 数据处理

数据处理泛指对原始数据进行收集、整理、存储、分类、排序、检索、维护、加工、统计和传输等一系列活动的总称。其目的是获得人们所需要的数据和提取有用的信息作为决策的依据。

数据经过数据处理后就得到了有用的信息。即

数据 $\xrightarrow{\text{数据处理}}$ 信息

1.1.2 数据库、数据库管理系统和数据库系统

数据库、数据库管理系统和数据库系统是数据库技术中常用的三大术语，常常容易混淆，它们之间既有区别又有联系。

1. 数据库 (DB, DataBase)

通俗地说，数据库即是存放数据的仓库。但所存放的数据并不是简单地堆放在一起，而是相互之间有联系的，并按某种存储模式组织管理。

数据库是以一定组织方式存储在一起的、能为多个用户共享的、与应用程序彼此独立的相关数据的集合。它有如下特点：

- (1) 数据的共享性。数据库中的数据能为多个用户服务。
- (2) 数据的独立性。用户的应用程序与数据的逻辑组织和物理存储方式无关。
- (3) 数据的完整性。数据库中的数据在操作和维护过程中可以保持正确无误。
- (4) 数据库中的数据冗余（重复）少。

例如，一个单位的全体职工信息可作为一个数据库，各个部门都可以根据自己的需要使用其中的部分或全部信息，而没有必要自己单独存储这些信息（无不必要的冗余）。如人事部门、财务部门都可以根据自己的需要使用这些信息（为多种应用服务，即共享性）。当数据库中数据的逻辑结构或物理结构发生改变时，各部门处理这些数据的应用程序可以不变（数据的独立性）。另外，对这个数据库的管理，如对数据库中的数据进行增加、删除、修改等工作不能随意进行，必须按公用的可控制的方法进行，以保证数据的完整性。

2. 数据库管理系统 (DBMS, DataBase Management System)

数据库管理系统是一个数据库管理软件，位于用户和操作系统之间，其职能是维护数据库，接受和完成用户程序或命令提出的访问数据库的各种请求。

其功能可以概括为以下三个方面。

1) 数据库的定义和建立

数据库管理系统能够提供数据定义语言 (DDL, Data Definition Language)，并提供相应的建库机制。用户利用 DDL 可以方便地建立数据库。当需要时，用户将系统的数据及结构情况用 DDL 描述，数据库管理系统能够根据其描述执行建库操作。

2) 数据库的操纵

实现数据的插入、修改、删除、查询、统计等数据存取及处理功能统称为数据操纵功能。数据操纵功能是数据库的基本操作功能，数据库管理系统通过提供数据库操纵语

言 (DML, Data Manipulation Language) 实现其数据操纵功能。

DML 有以下两种形式:

(1) 宿主型 DML。宿主型 DML 只能嵌入在其他高级语言中使用, 不能单独使用。被 DML 嵌入的计算机语言称为主语言, 常用的主语言有 C, FORTRAN, COBOL 等。在由宿主型 DML 和主语言混合在一起设计的程序中, DML 语句只完成有关数据库的数据存取操作功能, 而其他功能由主语言的语句完成。

(2) 自主型 DML。既可以嵌入到主语言中使用, 也可以单独使用的 DML 称为自主型 DML。自主型 DML 可以作为交互式命令与用户对话, 执行其独立的单条语句功能, 也可以用其编程完成数据存取和其他功能。

3) 数据库的运行控制

数据库的运行控制包括完整性控制、安全性控制和并发控制。

完整性控制: 完整性包括数据的正确性、有效性和相容性。DBMS 应提供有效措施, 以保证数据处于约束范围内, 由 DBMS 而不是由用户程序自动检查数据的完整性。

安全性控制: 安全性主要指保密方式、保密级别和数据使用权。不是任何人都可以存取数据库的数据, 也不是每个合法用户可以存取相同的数据范围, 经常采用的安全性控制方法有口令、密码或通过系统直接给用户赋予某种存取权限。在 DBMS 的管理下, 只有合法用户才能访问数据库, 才能访问他有权访问的数据库, 才能对数据进行他有权进行的操作。

并发控制: 由 DBMS 提供并发控制手段, 使多个用户同时对数据库操作时, 可以保证共享、并发操作, 且不破坏数据库。

例如, Oracle, SQL Server, Access 等都是数据库管理系统。

3. 数据库系统 (DBS, DataBase System)

数据库系统并不单指数据库和数据库管理系统本身, 而是指计算机系统引进数据库技术后的整个系统, 数据库系统的基本结构如图 1.1 所示。

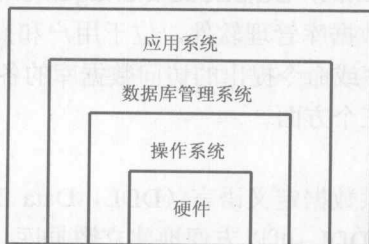


图 1.1 数据库系统的基本结构

数据库系统通常由以下部分组成:

(1) 硬件 (Hardware)。由五个组成部分构成的计算机系统 (CPU、运算器、内外

存储器、输入和输出设备)。

(2) 软件 (Software)。它包括操作系统、DBMS、主语言 (高级语言, 如 C, FORTRAN, COBOL 等)、应用开发工具软件 (如 Visual Basic, Delphi, PowerBuilder 等)、应用系统 (用户或专业软件开发人员根据需要编写的应用程序) 和数据库 (DB)。

(3) 数据库管理员 (DBA, DataBase Administrator), 其职责为:

- ① DBA 应参与数据库和应用系统的设计;
- ② DBA 应参与决定数据库的存储结构和存取策略的工作;
- ③ DBA 要负责定义数据的安全性和完整性条件;
- ④ DBA 负责监视和控制数据库系统的运行及系统的维护和数据恢复工作;
- ⑤ DBA 负责数据库的改进和重组。

(4) 用户 (User): 使用数据库应用系统的人员。

说明: 数据库管理员工作在 DBMS 之上, 用户则是应用系统的使用者。

1.2 数据处理技术的发展过程

人类活动的整个历史, 离不开对数据的收集、处理、保存和利用。初始时, 人类仅能用语言、绘画和火光传递信息, 通过用绳子打结和在树上刻画记录数据, 以供人们处理和利用。自从发明了文字、纸张和印刷术后, 就以纸张为介质, 通过汇编字典、登记账目与构造各类型的表、部、册等方法来收集和保存数据。出现电子技术后, 人们才利用磁性材料作为介质来保存和收集数据, 使得利用和加工数据进入更高级的阶段。

自 20 世纪 60 年代以来, 社会生产力高速发展, 新技术层出不穷, 信息量急剧膨胀, 使整个人类社会正成为信息化的社会。人们对数据和信息的利用和处理已进入自动化、网络化和国际化的阶段。例如, 查找情报资料、档案, 分析银行账目、资金往来, 进行人口统计分析, 办理航空订票、旅馆房间登记, 管理交通运输、生产科研, 订计划、做决策等。这些任务需要利用大量数据, 又要求快速处理和及时得出有用的信息, 如果仅靠人工翻阅文件和查找原始记录是难以完成的, 甚至是不可能的, 所以迫切需要借助于电子计算机的高速度和大容量来完成处理。

由于有了计算机这样一种雄厚的物质基础, 数据处理技术得到了很快的发展。目前, 它已经历了人工管理阶段、文件系统阶段和数据库系统阶段。

1.2.1 人工管理阶段

20 世纪 50 年代以前, 计算机主要用于数值计算, 外存储器只有纸带、磁带和卡片等。当时也没有操作系统。这一时期数据管理的特点是:

- (1) 由于计算机主要用于科学计算, 一般不需将数据长期保存起来供查询。因此,

在计算某一课题时才将数据输入，用完后就撤走，甚至系统软件也这样处理。

(2) 由于没有软件进行管理，程序员不仅要规定数据的逻辑结构，还要在程序中设计存储结构与输入/输出方法。在这样的程序设计中，数据的物理结构与逻辑结构密切相关，即数据与程序不具有独立性。程序员不仅要在物理布置上耗费精力，而且若在存储上有改变时，必须改变程序。

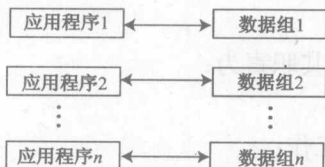


图 1.2 人工管理阶段的数据管理特点

(3) 这一时期基本无文件概念，即使有也大多是顺序文件。

(4) 一组数据对应一个程序，即数据是面向应用的。这样程序与程序之间就会有重复数据，即数据的冗余很大，浪费存储空间。

这一阶段的程序与数据的关系如图 1.2 所示。

1.2.2 文件系统阶段

从 20 世纪 50 年代后期到 60 年代中期，计算机大量用于管理领域。这时外存储器有了磁鼓、磁盘等可以直接存取的存储设备，在系统软件方面也出现了包含文件系统在内的操作系统。文件系统是专门管理数据的软件，用它来进行数据管理，其特点可归纳为：

(1) 由于计算机大量用于数据处理等方面，数据可以以文件形式长期保留在外存上，并经常反复地进行处置，如对文件进行查询、修改、增删等操作。

(2) 由于有专用软件进行管理，程序和数据间有公共存取方法进行转换，逻辑结构和物理结构可以互有区别。因此，程序和数据之间有一定的独立性。程序不必过多地考虑物理布置的细节，数据存储上的改变不一定都反映到程序上，从而减少了修改程序的工作量。

(3) 除了顺序文件外，又有了能直接存取的索引文件、链文件和直接存取文件等。

(4) 这一阶段的数据管理较第一阶段虽有改进，但仍存在很大缺点。如：

① 各数据文件基本上是根据程序员自身的程序需要建立的，数据和程序之间仍然存在着较密切的依赖关系。

② 由于文件是由程序员各自建立的，尽管各文件可以被共享，但存在于各文件中不必要的数据冗余依然存在。

③ 当用户希望将若干个文件中的部分数据组成一个新的文件，且按他自己所需要的逻辑结构而加以应用时，就必须在自己的程序中对所涉及到的有关文件重新进行加工和组合。这样，一方面要存取许多对他无用的数据，另一方面，这种本来是属于数据本身内在联系的问题，却要放在应用程序中来处理，势必增加了用户的额外负担。