# 第 1 章　绪 论

数据库技术是数据管理的最新技术，是计算机科学与技术的重要分支。信息技术的发展极大地推动了社会的进步和发展，信息资源已成为社会各行各业的重要资源和财富。作为实施有效信息处理的信息系统已成为一个企业或组织生存和发展的重要基础条件。由于数据库技术是信息系统的核心和基础，因而得到快速的发展和越来越广泛的应用。数据库技术可以为各种不同用户提供及时的、准确的、相关的信息，满足这些用户的各种不同的需要。

本章 1-1 节对数据库技术进行了概述，1-2 节和 1-3 节介绍了数据库系统的体系结构与组成，1-4 节从模型、技术与应用等方面介绍了数据库技术的发展趋势，1-5 节以一个实例介绍了一个数据库系统的概貌。

## 1-1　数据库技术概述

在系统地介绍数据库技术的基本概念之前，首先介绍一些与数据库技术相关的常用术语。

### 1-1-1　常用术语

#### 1. 数据

数据（Data）是关于人、地点、行为或事件的事实。数据的种类很多，如文字、图形、图像、声音、语言等，这些都是数据。因此，也有人把数据定义为描述事物的符号记录。

数据的形式有时还不能完全表达其内容，需要经过解释。因而数据与数据的解释是不可分的，数据的解释是指对数据含义的说明，数据的含义定义为数据语义。

#### 2. 信息

信息（Information）通常定义为经过加工处理后的数据。信息与原始资料同样重要，原始资料需要加工处理才能成为有用的信息。由于信息在决策过程中具有决定性作用，因而它的产生、使用和管理需要一个精心建造的环境，这个环境就是信息系统，由人、硬件、软件组成，如图 1-1 所示。

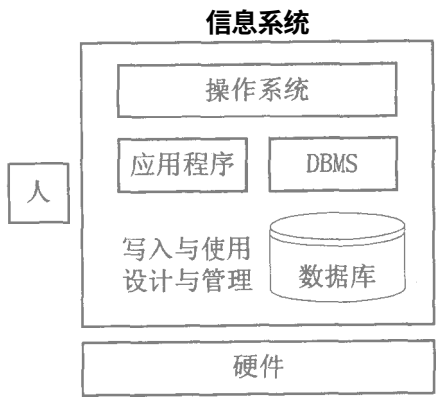



图 1-1　信息系统

信息是构成一定含义的一组数据，所以，数据和信息都是客观事物的反映，反映了人们对事物的认识。数据是信息的载体，信息则是数据加工的结果，是对数据的解释。信息与数据间的关系如图 1-2 所示。

需要全本请访问 www.ertongbook.com

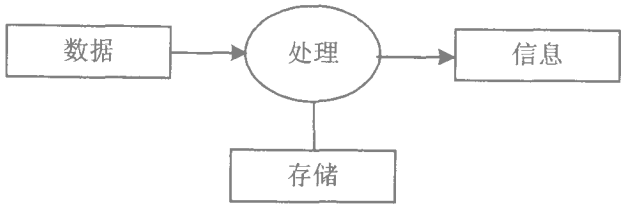



图 1-2 数据与信息的关系

计算机系统的每项操作，均是对数据进行某种处理。数据输入计算机后，经存储、传送、排序、计算、转换、检索、制表及仿真等操作，输出人们需要的结果，亦即产生信息。

### 3. 数据库

数据库（DataBase，简称 DB）是长期存储在计算机外存上的、有结构的、可共享的数据集合。数据库中的数据按一定的数据模型描述、组织和储存，具有较小的冗余度、较高的数据独立性和易扩展性，可为各种用户共享。

### 4. 数据库管理系统

数据库管理系统（DataBase Management System，简称 DBMS）是科学地组织和存储数据、有效地获取和维护数据的系统软件。它是位于用户和操作系统之间的一层数据管理软件，其主要功能包括以下几个方面：

- 数据定义功能——DBMS 提供数据定义语言（Data Definition Language，简称 DDL），用户可用 DDL 对数据库中的数据对象进行定义。
- 数据操纵功能——DBMS 提供数据操纵语言（Data Manipulation Language，简称 DML）以供用户实现对数据库的基本操作，如查询、插入、删除和修改等。
- 数据库的运行管理功能——为了保证数据的安全性、完整性、多用户对数据的并发使用等，数据库的建立、运行与维护必须在 DBMS 的统一管理和统一控制下进行。
- 数据库的建立与维护功能——包括初始数据的输入与转换功能，数据库的转存与恢复功能；数据库的重组织功能和性能监视与分析功能等。

数据库管理系统的主要成份如图 1-3所示。

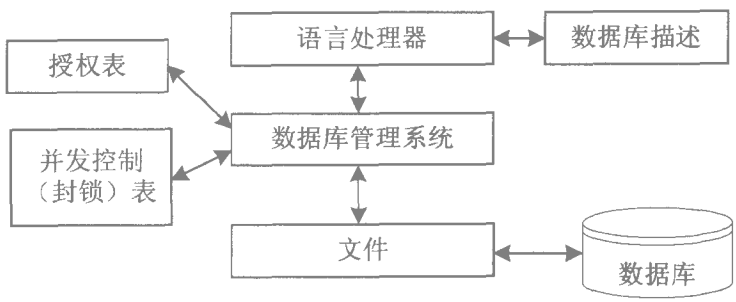



图 1-3 数据库管理系统的主要成份

### 5. 数据库系统

数据库系统（DataBase System，简称 DBS）是由数据库、数据库管理系统、应用系统、

数据库管理员和用户构成的人一机系统。数据库管理员( DataBase Administrator,简称 DBA)是专门从事数据库建立、使用和维护的工作人员。数据库系统的组成及其作用将在下一节详细讨论。

## 1-1-2 数据库技术的产生与发展

数据库技术是应数据管理任务的需求而产生的。而数据管理技术经历了人工管理、文件管理和数据库系统等三个阶段，表 1-1 给出了数据管理三个阶段的比较。

表 1-1 数据管理三个阶段的比较

| | | 人工管理 | 文件管理 | 数据库管理 |
|---|---|---|---|---|
| 背景 | 应用背景 | 科学计算 | 科学计算、管理 | 大规模管理 |
| | 硬件背景 | 无直接存取存储设备 | 磁盘、磁鼓 | 大容量磁盘 |
| | 软件背景 | 没有操作系统 | 有文件系统 | 有数据库管理系统 |
| | 处理方式 | 批处理 | 联机实时处理、批处理 | 联机实时处理、分布处理、批处理 |
| 特点 | 数据的管理者 | 用户 | 文件系统 | 数据库管理系统 |
| | 数据面向的对象 | 某一应用程序 | 某一应用 | 现实世界 |
| | 数据的共享程度 | 无共享，冗余度极大 | 共享性差，冗余度大 | 共享性高，冗余度小 |
| | 数据的独立性 | 不独立，完全依赖于程序 | 独立性差 | 具有高度的物理独立性和一定的逻辑独立性 |
| | 数据的结构化 | 无结构 | 记录内有结构，整体无结构 | 整体结构化，用数据模型描述 |
| | 数据控制能力 | 应用程序自己控制 | 应用程序自己控制 | 由数据库管理系统提供数据安全性、完整性、并发控制和恢复能力 |

### 1. 人工管理阶段

20 世纪 50 年代中期以前，计算机的软硬件还很不完善。硬件存储设备只有磁带、卡片和纸带，软件方面还没有操作系统，因而当时的计算机主要用于科学计算。在这个阶段，由于没有进行数据管理的软件系统，程序员在程序设计时不仅要规定数据的逻辑结构，还要设计其存取、输入/输出等物理结构。当数据的物理组织或计算机存储设备改变时，用户程序就必须重新编制。由于数据的组织面向应用，不同的应用程序之间无法共享数据，因此，使得不同的应用程序之间存在大量的重复数据，应用程序间的数据一致性很难保证与维护。总之，人工管理数据具有如下特点：

- 数据不保存。
- 应用程序管理数据。
- 数据不共享。

- 数据不具有独立性。
- 数据处理方式是批处理。

在人工数据管理阶段，应用程序与数据间的关系如图 1-4 所示。

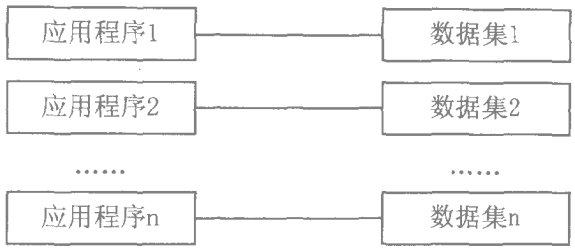



图 1-4 人工数据管理阶段应用程序与数据间的对应关系

## 2. 文件系统阶段

20 世纪 50 年代中期至 60 年代中期，在硬件方面，由于大容量存储设备（如磁盘、磁带等）的出现，推动了软件技术的发展；在软件方面，操作系统的出现标志着数据管理进入了一个新的阶段。数据以文件为单位存储在外存上，由操作系统统一管理，操作系统为用户使用文件提供了友好的界面。在这一阶段，由于操作系统的文件管理功能，使得文件的逻辑结构与物理结构脱钩、程序与数据分离，这样，数据与程序就有了一定的独立性。用户的应用程序与数据文件可分别存放在外存储器上，不同应用程序可以共享一组数据，实现了数据以文件为单位的共享。

但是，由于数据的组织仍然是面向应用程序，所以存在大量的数据冗余，而且数据的逻辑结构很难修改与扩充。数据逻辑结构的任何一点微小改变，都会影响到应用程序。另外，由于数据文件间相互独立，因而它们不能反映现实世界中数据间的联系。总之，用文件系统管理数据具有如下特点：

- 数据可以长期保存。
- 由文件系统管理数据。
- 数据共享性差，冗余度大。
- 数据独立性差。

文件系统中应用程序与数据的关系如图 1-5 所示。

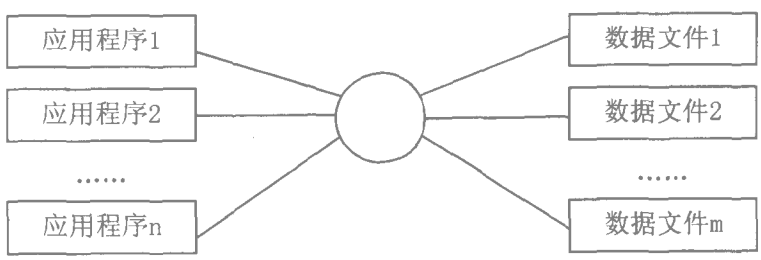



图 1-5 文件系统中应用程序与数据的关系

### 3. 数据库系统阶段

20世纪60年代末，在硬件方面，开始出现了具有数百兆字节容量、价格低廉的磁盘；在软件方面，操作系统已开始成熟，程序设计语言的功能更加强大，操作和使用更加方便。这些硬件与软件技术为数据库技术的发展提供了良好的基础。从现实需求来看，数据量急剧增加，对数据的管理和分析需求力度加大，人们对数据管理技术提出了更高的要求，希望面向企业或部门以数据为中心组织数据，减少数据的冗余，提供更高的数据共享能力。同时要求程序和数据具有较高的独立性，当数据的逻辑结构改变时，不至于影响数据的物理结构，也不影响应用程序，以降低应用程序的开发与维护费用。数据库技术正是在这样一个环境下发展起来的。

文件系统阶段向数据库系统阶段转变以下列三件大事为标志：

（1）1968年，IBM公司推出了商品化的基于层次模型的信息管理系统（Information Management System，IMS）。IMS是一种宿主语言系统。某种宿主语言加上数据操纵语言就组成了IMS的应用系统。

（2）1969年，美国CODASYL（Conference on Data System Language，数据系统语言协商会）组织下属的DBTG（DataBase Task Group，数据库任务组）发布了一系列研究数据库方法的DBTG报告，该报告奠定了网状数据模型的基础。

（3）1970年，IBM公司的研究人员E.F.Codd连续发表论文，提出了关系模型，奠定了关系型数据库管理系统的基础。目前，广为流行的关系型数据库系统的理论基础仍然是关系理论。

数据库管理系统克服了传统的文件管理方式的缺陷，提高了数据的一致性、完整性，减少了数据冗余。数据库系统阶段应用程序与数据间的关系如图1-6所示。在这种管理方式下，许多应用程序可以在数据库管理系统的控制下共享数据库中的数据。

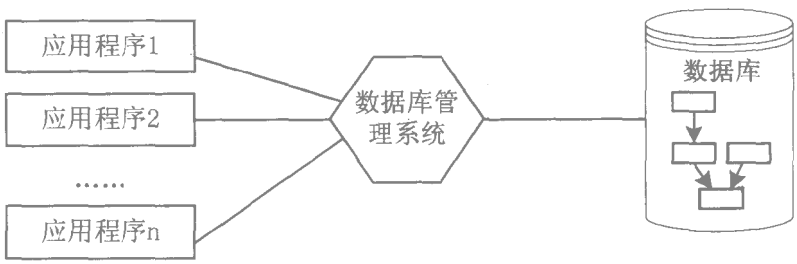



图1-6 数据库系统阶段应用程序与数据的关系

与传统的文件系统相比，数据库系统阶段具有明显的优点，从文件系统到数据库系统，标志着数据管理技术的飞跃。下面将对数据库系统的特点进行详细讨论。

## 1-1-3 数据库系统的特点

与人工管理和文件管理相比，数据库系统具有下列一些特点：

（1）面向企业或部门，以数据为中心组织数据，形成综合性的数据库为各应用共享。

（2）使用复杂的数据模型表示结构。在数据库系统中，数据模型不仅描述数据本身的

需要全本请访问 www.ertongbook.com

特征，而且还要描述数据之间的联系。这种联系通过存取路径来实现。通过所有存取路径表示自然的数据联系是数据库系统与传统的文件系统间的本质差别。这样，所要管理的数据不再面向特定的某个或某些应用，而是面向整个应用系统，因此大大降低了数据冗余度，实现了数据共享。

（3）程序和数据具有较高的独立性。数据库系统通过提供两个层次的映像功能，使数据的物理结构独立于全局逻辑结构。使数据的全局逻辑结构独立于应用程序。当数据的物理组织或存储设备改变时，不影响全局逻辑结构，而当数据的全局逻辑结构改变时，有可能不用改变应用程序。

（4）为用户提供了方便的接口。在数据库系统中，用户可以非常方便地使用查询语言，如 SQL，或实用程序命令操作数据库中的数据，也可以使用编程方式，如在高级程序设计语言中嵌入查询语言，操作数据库。

（5）提供了完整的数据管理与控制功能。这些功能包括并发性、完整性、可恢复性、安全性和审计性等。并发性就是允许多个用户或应用程序同时操作数据库中的数据，而数据库依然保证为这些用户或应用程序提供正确的数据。完整性就是始终包含正确的数据，例如，通过完成完整性的规则可以使数据值限制在指定的范围内。可恢复性是指在数据库遭到破坏后，系统有能力把数据库恢复到最近某个时刻的正确状态。安全性是指只有指定的用户才能使用数据库中的数据和执行允许的操作。审计性是指系统可以自动记录所有对数据库和数据的操作，以便跟踪和审计数据库系统的所有操作。

（6）提高了系统的灵活性。对数据库中数据的操作既可以以记录为单位，也可以以记录中的数据项为单位。例如在 SQL 语言中可以使用 SELECT语句指定记录或记录中的数据项。

数据库系统的出现使得信息系统从以加工数据的程序为中心转向围绕共享的数据库为中心的新阶段。这样既有利于数据的集中管理，又有利于应用程序的开发与维护，从而提高了数据的利用率与相容性，提高了决策的可靠性。目前，数据库已经成为现代信息系统不可分离的重要组成部分。具有数以百万甚至数百亿字节信息的数据库已经普遍存在于科学技术、工业、农业、商业、服务业和政府部门的信息系统中。

### 1-1-4 数据库技术的应用示例

如前所述，数据库技术已广泛应用于各行各业，可以毫不夸张地讲，几乎所有的计算机应用系统均在使用数据库技术。下面通过介绍几个数据库系统应用的典型示例，使读者对数据库技术有一个感性认识。

#### 1. 超市的销售业务系统

由于其种类繁多的商品、较低的价格和选物的便利，超市已逐渐成为人们购物的最佳选择。数据库技术作为超市的重要技术支持，发挥着非常重要的作用。在超市的销售业务系统中，主要的数据有如下几项。

- 销售信息：连锁店、顾客、商品、数量、总价、时间等。
- 商品信息：商品名称、类型、摆放位置、供应商、数量、单价等。

- 供应商信息：供应商名称、地点、商品、信誉、电话等。

对超市销售业务系统的主要操作包括记录顾客的购买信息、查询超市现存商品的情况、确定进货品种及商品的摆放位置、分析销售情况等。

### 2. 教学管理系统

高等学校教学管理系统主要涉及学生、教室、教师、课程、排课等信息的管理，其主要数据项如下。

- 学生信息：学号、姓名、性别、年级、系专业、班级、年龄、宿舍、电话、 E-mail 等。
- 教室信息：教室号、楼号、座位、类型等。
- 教师信息：工作证号、姓名、性别、年龄、职称、学历、学位、系、教研室、住址、电话、E-mail 等。
- 课程信息：课程名称、教材、学时、学分、开设时间等。
- 排课信息：课程名称、教室、班级、教师等。

当然，还应包括学生选课、考试成绩、实验室等信息。典型的操作包括学生成绩统计、教师工作量计算、教室安排、自动排课、考试成绩登记等。教学管理系统的关键在于保证大量教务数据的正确存储与处理，为学校的各项工作安排，特别是教务工作，提供安全的数据支持。

### 3. 图书管理系统

图书管理系统属于数据库技术的一个典型应用领域。在图书管理系统中，其主要数据项有如下几个。

- 图书信息：书号、书名、作者、出版时间、类型、价格、出版社等。
- 作者信息：姓名、身份证号、单位、性别、年龄、职称、学历、住址、电话、 E-mail 等。
- 出版社信息：名称、地址、法人、成立日期、信誉等。
- 读者信息：姓名、借书号、类型、书名、借书日期、借书数量等。

典型的查询操作有查看某种类型的图书、浏览指定出版社的图书、检索指定作者的图书、查看某一读者的借书情况、检索指定的图书等等；典型的更新操作有登记新书信息、作者信息、读者信息、出版社信息等。作为一个存书量达几百万的大学图书馆，没有管理图书的数据库系统，很难想像它将如何运行。

## 1-1-5 数据库应用软件的开发过程

一般来讲，用于开发数据库应用软件的方法包括结构化生命周期法和原型化法。结构化生命周期方法要求把信息系统的开发方法从初始到结束划分成若干个预先规定好任务的阶段，然后按照一定的准则进行设计与实现。原型化法是一种考虑到人本身的灵活、多变、依经验行事的特点而产生的信息系统开发方法，其基本思想是并非所有的需求都能预先明确地定义，反复修改是必要的和不可避免的。在结构化生命周期方法中，数据库应用软件的开发包括系统调研、系统分析、系统设计、系统实施与系统评价等五个阶段。

系统调研主要是对现行系统的调查研究和新系统开发的可行性研究。对现行系统的调

查研究是为新系统的开发进行原始资料的准备，并且使系统开发人员获得对现行系统的感性与理性认识。

系统分析包括目标分析、需求分析与功能分析等，是数据库系统开发中的一个重要阶段。在系统分析的基础上，设计信息系统的逻辑模型，完成系统分析说明书。

系统设计又称为物理设计。系统设计是根据一定的原则完成系统物理设计的全部内部，包括系统的体系结构设计、计算机系统设计、数据通信网络设计、数据库设计、输入 输出设计、用户界面设计等并完成系统设计说明书。

系统实施是指对系统设计阶段完成的系统物理设计的实现，需投入大量的人力、物力与财力。系统实施的主要内容包括系统实施的组织、软硬件环境的选型与建立、软件编制与系统集成、人员培训、系统调试与修改等。

系统评价阶段主要包括系统维护与评价。数据库系统往往是一个复杂的大系统，由于系统内外环境的变化，如组织机构的变化、应用需求的变化等，要求系统能够适应这种变化并且不断完善，这就需要系统的维护。当系统完成与运行之后，应该对系统设计的目标是否实现以及实现的程度进行系统的评价，以便于下一步系统的升级与更新。

## 1-2 数据库系统的体系结构

可以从不同的层次或不同的角度来分析数据库系统的体系结构，从 DBMS 角度来看，数据库系统通常采用三级模式结构，这是数据库管理系统内部的系统结构，通常称作数据库体系结构；从数据库最终用户角度看，数据库系统的结构可以分为集中式结构、分布式结构、C/S 结构、B/S 结构等，这就是数据库系统外部的体系结构，简称数据库系统体系结构。下面就对上述概念进行系统的介绍。

### 1-2-1 数据库体系结构

数据库体系结构是数据库的一个总的框架。虽然目前市场上流行的数据库系统软件产品多种多样、支持不同的数据模型、使用不同的数据库语言和应用系统开发工具，建立在不同的操作系统之上等，但绝大多数数据库系统在总的体系结构上都具有三级模式的结构特征。数据库的三级模式为外模式、概念模式和内模式，如图 1-7 所示。

概念模式又称数据库模式，简称模式（ Schema），它是数据库中全部数据的逻辑结构和特征的描述，它以数据模型为基础，采用数据库系统提供的模式描述语言（模式 DDL）进行定义。概念模式可以被看作是现实世界中一个组织或部门中的实体及其联系的抽象模型在具体数据库系统中的实现。

概念模式不同于外模式，与具体的应用程序无关；也不同于内模式，与数据的存储细节和硬件环境无关。

外模式又称用户模式或子模式，由概念模式导出，属于概念模式的子集。外模式定义了允许用户操作的数据库数据，通常由 DDL 语言描述。对最终用户来讲，所看到的数据视图（也就是与某一应用相关的数据的逻辑表示）即是外模式。由于不同用户需求相差很大，看待数据的方式与所使用的数据内容各不相同，对数据的保密性要求也有差异，因此，不同用户的外模式也不相同。

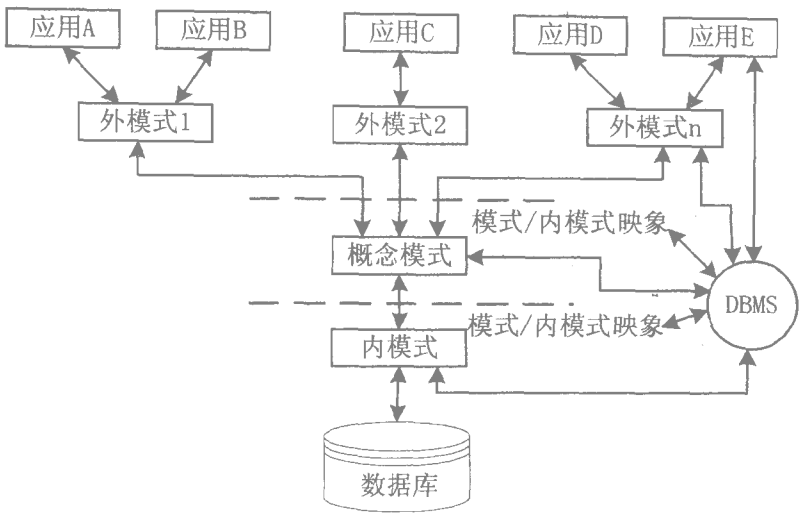



图 1-7　数据库体系结构——三级模式

概念模式与外模式不仅要定义数据的逻辑结构，而且要定义与数据有关的安全性和完整性；不仅要定义数据记录的内部结构，还要定义这些数据间的联系。

内模式又称存储模式，用来描述数据的物理结构与存储方式。通常，内模式也由 DDL 语言描述。

数据库体系结构的三级模式结构在数据的三个抽象层次上提供了两个映象，即内模式到概念模式的映象与概念模式到外模式的映象。

内模式到概念模式的映象提供了数据的物理独立性，即当数据的物理结构发生变化时，如存储设备的改变、数据存储位置或存储组织方式的改变等，不影响数据的逻辑结构。例如，为了提高应用程序的存取效率， DBA 和数据库的设计人员依据各应用程序对数据的存取要求，可以对数据的物理组织进行一定形式或程度的优化。而不需要重新定义概念模式与外模式，也不需要修改应用程序。

概念模式到外模式的映象提供了数据的逻辑独立性，即当数据的整体逻辑结构发生变化时，如为原有记录增加新的数据项、在概念模式中增加新的数据类型、在原有记录类型间增加新的联系等，不影响外模式。例如，在教务管理数据库系统中，随着需求的变化，需要增加双学位课程的选修与授予学位的信息，增加学生毕业就业去向的信息等。当根据新的功能要求对原模式进行修改或扩充新的结构时，这种修改不必重新编写应用程序，也不需要重新生成外模式，而仅需对概念模式作部分修改或扩充，对外模式的定义作某些调整。当然，为了实现新的处理功能，就需编写新的应用程序，或对原有的应用程序进行修改。

由于数据库的这种三级模式体系结构，提供了高度的数据独立性。其中，数据库的全局逻辑描述是独立于其他所有结构描述的，在定义数据库结构时，概念模式应首先定义。内模式则是将概念模式中所定义的数据进行适当的组织并加以存储，以实现较好的时空运行效率。

总之，数据库三级模式体系结构是数据管理的结构框架，依照这些数据框架组织的数据才是数据库的内容。在数据库设计时，主要是定义数据库的各级模式；而在用户使用数

据库时，关心的才是数据库的内容。数据库的模式通常是相对稳定的，而数据库的数据则是经常变化的，特别是一些工业过程的实时数据库，其数据的变化是连续不断的。

## 1-2-2 数据库系统体系结构

一个数据库应用系统通常包括数据存储层、应用层与用户界面层等三个层次。数据存储层一般由数据库系统来承担，完成对数据的各种维护操作；应用层是应用数据库处理的与用户相关的各种业务操作，使用有关的程序设计语言来完成；而用户界面层则是提供给用户的可视化图形操作界面，用于数据库系统与用户之间的交互。

数据库系统体系结构是指数据库系统中数据存储层、应用层、用户界面层以及网络通讯之间的布局与分布关系。依据目前数据库系统的应用与发展，可以分为单机结构、集中式结构、C/S 结构、B/S 结构与多层结构等。

### 1. 单机结构

随着 PC 机的速度与存储容量等性能指标的不断提高，出现了适合于 PC 的单用户数据库系统。这种可以运行在 PC 机上的数据库系统称为桌面型数据库管理系统（Desktop DataBase Management System）。这些桌面型 DBMS 虽然在数据的完整性、安全性、并发性等方面存在有许多缺陷，但已基本上实现 DBMS 应具备的功能。目前比较流行的桌面型 DBMS 有 Microsoft Access 、Visual Foxpro 等。

在这种桌面型 DBMS 中，数据存储层、应用层和用户界面层的所有功能都存在于单台 PC 机上，因而非常适合未联网用户，个人用户及移动用户等使用。

### 2. 集中式数据库系统结构

集中式数据库系统结构是一种采用大型主机和多个终端相结合的系统。这种结构将操作系统、应用程序、数据库系统等数据和资源均放在大型主机上，而连接在主机上的许多终端，只是作为主机的一种输入输出设备。在这种数据库系统结构中，数据存储层和应用层放在主机上，而用户界面层放在各个终端上。

在集中式数据库系统结构中，由于所有的处理均由主机完成，因而对主机的性能要求很高，这是数据库系统初期最流行的结构。随着计算机网络的兴起和 PC 机性能的大幅度提高而价格又大幅度下跌，这种传统的集中式数据库系统结构已被 C/S 数据库系统结构所代替。

### 3. 客户机/服务器结构

客户机/服务器（Client/Server，简称 C/S）结构是当前非常流行的数据库系统结构。在这种体系结构中，客户机提出请求，服务器对客户机的服务请求作出回应。客户机/服务器结构最早起源于计算机局域网中对打印机等外部设备资源的共享服务要求，即把文件打印和存取作为一种通用的服务功能，让局域网中的某些特定节点来完成，而其它节点在需要这些服务时，可以通过网络向特定节点发出服务请求，以得到相应的服务。这种外设共享处理结构在服务功能上的自然拓展就形成了今天的客户机/服务器数据库系统结构。

C/S 结构的本质在于通过对服务功能的分布实现分工服务。每一个服务器都为整个局

域网系统提供自己最擅长的服务，让所有客户机来分享；客户机上的应用程序借助于服务器的服务功能实现复杂的应用功能。在 C/S 结构中，数据存储层处于服务器上，应用层和用户界面层处于客户机上。在 C/S 数据库系统结构中，客户机负责管理用户界面，接收用户数据、处理应用逻辑、生成数据库服务请求，并将服务请求发送给数据库服务器，同时接收数据库服务器返回的结果，最后再将返回的结果按照一定的格式或方式显示给用户。数据库服务器接收客户机的请求，对服务请求进行处理，并返回处理结果给客户机。

C/S 系统结构使应用程序的处理更加接近用户，其好处在于使整个系统具有较好的性能。此外，C/S结构的通信成本也比较低，其主要原因在于：

（1）降低了数据传输量，数据库服务器返回给客户机的仅是执行数据操作后的结果数据。

（2）由于许多应用逻辑处理由客户机来完成，因而减少了许多的不必要的与服务器的通信开销。

### 4. 浏览器/服务器结构

浏览器/服务器( Browser/Server,简称 B/S）结构是随着计算机网络技术，特别是 Internet 技术的迅速发展与应用，而产生的一种数据库系统结构。 B/S 结构是针对 C/S 结构的不足而提出的。

如前所述，基于 C/S 结构的数据库应用系统把许多应用逻辑处理功能分散在客户机上完成，这样对客户机提出了较高的要求，客户机必须拥有足够的能力运行客户端应用程序与用户界面软件，客户机必须针对每种要连接的数据库安装客户端软件，造成客户机臃肿的局面。另一方面，由于应用程序运行在客户机端，当客户机上的应用程序修改之后，就必须在所有安装该应用程序的客户机上重新安装此应用程序，维护非常困难。

在 B/S 结构的数据库系统中，在客户机端仅安装通用的浏览器软件实现同用户的输入／输出，而应用程序在服务器端安装和运行。在服务器端，除了要有数据库服务器保存数据并运行基本的数据库操作外，还要有另外的称作为应用服务器的服务器处理客户端提交的处理要求。也就是说，C/S 结构中客户端运行的程序转移到了应用服务器中。应用服务器充当了客户机与数据库服务器的中介，架起了用户界面同数据库之间的桥梁，所以也称作为三层结构。

B/S 结构有效地克服了 C/S 结构的缺陷，客户机只要能够运行浏览器即可，能够有效地节省投资，同时客户机的配置和维护也变得异常轻松。

B/S 的典型应用用在 Internet 中，应用服务器即 Web 服务器，这样就可以利用数据库为网民提供功能强大的信息服务。由三层结构还扩展出来了多层结构，通过增加中间的服务器的层数来增强系统功能、优化系统配置、简化系统管理。相关的内容在后续章节中介绍应用系统开发和 Web 数据库时将会给出更详细的介绍。

## 1-3 数据库系统的组成

如前所述，数据库系统是由数据库、数据库管理系统、应用系统、数据库管理员和用户构成的人-机系统。在上一节已对数据库、数据库管理系统等进行了介绍，下面对数据库系统的软、硬件支持环境及数据库管理员与相关人员等概念作一简要介绍。

需要全本请访问 www.ertongbook.com

### 1. 硬件支持系统

硬件支持系统主要是指计算机。由于数据库应用系统的需求，要求数据库服务器应该具有足够大的外存、高效的 I/O 存取效率、大的吞吐量和强大的作业处理能力。对于分布式数据库系统，计算机网络也应该是其必不可少的硬件支持环境。

### 2. 软件支持系统

数据库系统的软件除数据库管理系统 DBMS 之外，还需要有一个软件支撑环境，包括操作系统、应用系统开发工具、各种宿主语言与实用程序等。 DBMS 是在操作系统的文件系统基础上发展起来的，而且一般是在操作系统的支持下工作的。

应用系统开发工具是为数据库系统应用开发人员提供的一组开发工具集，常见的有：面向表格应用的自动生成工具，面向报表的自动生成工具，绘制统计图表的自动生成工具等。这些开发工具基本上都是可视化的第四代语言开发工具，具有友好的图形用户界面、支持 C/S 运行模式，具有较高的应用系统开发效率。

### 3. 数据库管理员及相关人员

数据库系统涉及的有关人员包括数据库管理员、系统分析员和数据库设计人员、应用程序员和最终用户。不同人员涉及不同级别的数据抽象，具有不同的数据视图。图 1-8 给出了数据系统所涉及人员的数据视图。

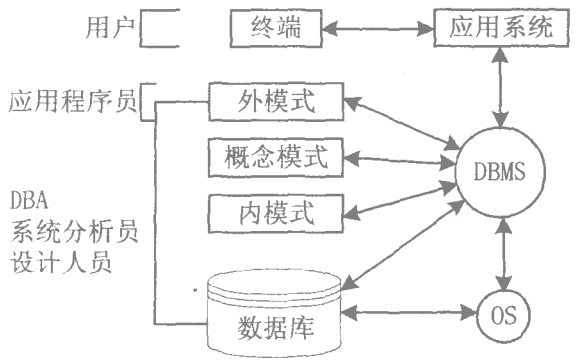



图 1-8 数据库人员的数据视图

最终用户是指数据库的最终使用者，他们通过应用程序的操作界面使用数据库，完成日常的业务处理。最终用户可以分为下列三类：

（1）偶然用户。这类用户一般是企业或组织机构的高级管理人员，他们不经常访问数据库，但每次访问则往往需要不同的数据库信息。

（2）简单用户。数据库的大多数用户都属于简单用户，其主要工作是通过应用程序员设计的具有友好界面的应用程序查询与存取数据库。

（3）复杂用户。这类用户比较熟悉数据库管理系统的各种功能，能够使用数据库语言访问数据库，甚至能够通过数据库管理系统提供的应用程序接口 API 编制自己的应用程序。这类用户一般都具有较高的科技背景，如科学家、工程师等。

应用程序员是负责设计、开发应用系统功能模块的软件编程人员，他们根据数据库的

外模式编写程序，同时负责应用系统的安装与调试。

系统分析员负责应用系统的需求分析和规范说明，要根据用户的需求，与 DBA 相结合确定系统的软、硬件配置，并参与数据库系统三级模式的概要设计。

数据库设计人员负责数据库中数据的确定、数据库各级模式的设计。数据库设计人员必须参加需求与系统分析，然后进行数据库设计。数据库设计人员可以由数据库管理员担任。

数据库管理员（DBA）负责管理和监控数据库系统的运行，负责为用户解决应用中出现的系统问题。为了保证数据库能正常稳定地运行，一般大型数据库系统都设有专人负责数据库系统的管理与维护。

数据库管理员一般应对他们所管理和维护的数据库的总体结构及实现细节具有十分透彻的了解。例如，数据库中有多少用户，他们存取和访问哪些数据，数据使用的频率多大，这些数据如何分布，如何使系统的性能最优等等。数据库管理员的主要任务包括：

（1）负责数据库系统软件的安装和维护。

（2）参与数据库设计。

（3）监控数据库的使用与运行。

（4）权限管理，DBA 的重要职责是保证数据库的安全性与完整性，确保多个用户能方便地共享数据，又要防止非法用户获取系统信息。

（5）数据库的日常维护，如定期对数据库进行备份，系统故障时，能尽快地将数据库恢复到正确状态等。

（6）数据库的改进与重组重构。DBA 在数据库系统运行期间，通过对空间利用率，处理效率等性能指标的记录与统计分析，不断改进数据库的设计；另外，在系统运行过程中，大量数据要不断地插入、删除、修改等，这样时间一长，会影响系统的性能。因此，DBA 应定期对数据库进行重组以提高系统的效率。当用户的需求改变时，DBA 就需要对数据库进行较大改造，即数据库的重构。

（7）数据库有关文档的管理。

### 4. 数据库工作流程

图 1-9 给出了引入数据库后计算机系统的层次图。

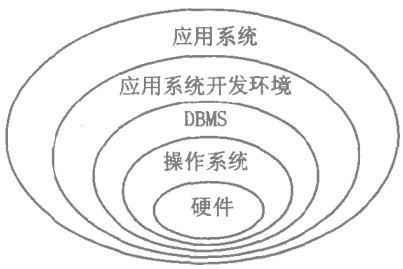



图 1-9　数据库系统层次图

下面通过一个用户对数据库的访问过程，向读者简要地介绍一下数据库系统的工作流程及各组成部分间的相互关系，如图 1-10 所示。

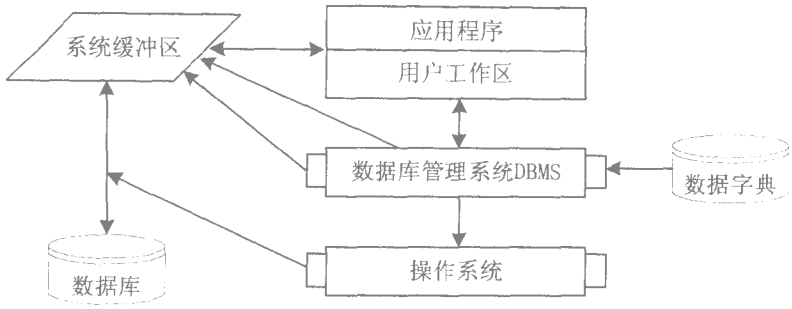



图 1-10　用户访问数据库的工作流程

假设用户提交了一条 DML 语句来查询数据库中的一组数据，则：

（1）DBMS 首先对该 DML 语句进行语法检查，然后从数据字典中寻找该用户对应的外模式进行权限检查，若有错误，会将信息返回给用户。

（2）DBMS 从数据字典读取相应的模式定义，并从外模式映像到概念模式，从中导出概念模式应该读入哪些记录。

（3）DBMS 读入内模式描述，并从概念模式映像到内模式，也就是将概念模式下的记录映像到内模式的记录格式，确定应读入的物理记录及其地址。

（4）DBMS 向操作系统发出读取指定地址记录的命令。

（5）操作系统执行读取记录的命令，将物理数据库的数据从硬盘读入系统缓冲区，并将执行结果通知 DBMS。

（6）DBMS 将系统缓冲区中的内容转换成用户所需的记录格式。

（7）DBMS 将数据记录从系统缓冲区传送到用户工作区。

从上述流程可以看出，在数据库系统中， DBMS 处于核心地位，它的工作需随时参照数据字典中的定义信息。

## 1-4　数据库技术的发展趋势

在许多新的数据库技术应用领域，传统的数据库技术与系统显得力不从心，对传统的数据库技术的研究工作提出了许多新挑战。数据库技术面临的挑战主要表现在以下几个方面：

- 环境的变化。数据库系统的应用环境由可控制的环境变成多变的异构信息集成环境与 Internet 环境。
- 数据类型的变化。数据库中的数据类型由结构化扩大至半结构化、非结构化和多媒体数据类型。
- 数据来源的变化。大量数据将来源于实时和动态的传感器或监测设备，需要处理的数据成倍剧增。
- 数据管理要求的变化。许多新型应用需要支持协同工作和工作流管理。

针对这些新的挑战，许多数据库技术研究与开发人员分别在数据模型、数据库与其他计算机新技术的结合及数据库技术的应用领域等三个方面对数据库新技术展开了一系列研究工作。下面分别就这三个方面对数据库技术的进一步发展趋势作一简要介绍。

## 1-4-1 数据模型的发展

有关数据模型的概念将在后续章节中详细介绍。

数据模型是数据特征的抽象，描述的是数据的共性。在数据库系统中针对不同的使用对象和应用目的，一般采取逐步抽象的方法，在不同层次采用不同的数据模型，一般分为物理层、逻辑层与概念层，如图 1-11 所示。

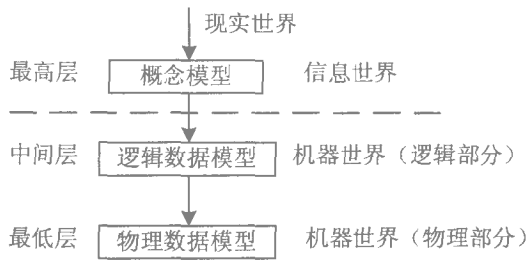



图 1-11 数据抽象的三个层次

物理层是数据抽象的最低层，这一层的数据抽象称为物理数据模型，用来描述数据的物理存储结构与存储方法。它不但由 DBMS 的设计决定，而且与操作系统、计算机硬件密切相关，一般都向用户隐蔽，用户不必了解其细节。

逻辑层是数据抽象的中间层，这一层的数据抽象称为逻辑数据模型（简称数据模型），用于描述数据库数据整体的逻辑结构。它是用户通过 DBMS 看到的数据的系统表示。传统的数据模型有层次、网状、关系模型等，非传统的数据模型有面向对象数据模型（简称 OO 模型）。

概念层是数据抽象的最高层，这一层的数据抽象称为概念数据模型（简称概念模型）。目的是按照用户的观点对现实世界进行建模，应该具有较强的语义表达能力、易于用户理解、独立于 DBMS 以及容易向 DBMS 所支持的数据模型转换等特点。概念模型的典型例子是实体联系模型（Entity-Relational Model，简称 E-R 模型）。

数据模型从最初的层次、网状数据模型发展到关系数据模型，促使数据库技术产生了巨大的飞跃。关系数据模型的提出是数据库发展史上具有划时代意义的重大成果，关系数据库管理系统的研究成功极大地促进了关系数据库的发展，使数据库技术的应用变成现实，因而，关系数据模型仍是目前具有统治地位的数据模型。 20 世纪 80 年代以来，几乎所有的数据库系统都是关系的，其应用遍布各个领域。

随着数据库应用领域的进一步扩大，传统的关系数据模型开始显示出其不足之处，如对复杂数据对象的表达能力较差、语义表达能力较弱以及缺乏灵活丰富的建模能力等。为此，人们提出并发展了许多新的数据模型。对数据模型的研究主要在以下三个方向进行：

（1）对传统的关系模型进行扩充。一种是对结构的扩充，如嵌套关系模型，它能表达表中表，且表中的一个域可以是一个函数。另一种是对语义的扩充，如 U.C. Berkeley 大学的 POSTGRES 系统，它支持关系间的继承，也支持在关系上定义函数与运算符。但关系的结构仍是一张平面表，表中表只能通过关系上定义的函数来模拟。

（2）提出和发展全新的数据构造器和数据处理原语，以表达复杂的结构和丰富的语义。

在这一方面比较有代表性的模型有函数数据模型（FDM），RM/T 模型、E-R 模型、语义数据模型（SDM）等，它们的共同特点是引入了丰富的语义关联，更能自然、恰当地表达现实世界中实体间的关系，同时也具有很强的结构表达能力。但由于这些模型比较复杂，至今在程序设计语言和技术方面仍没有相应的支持，至多被当作数据库设计中概念建模的一种工具。

（3）面向对象数据模型。这是将语义数据模型与面向对象程序设计方法相结合的产物，它是用面向对象的观点来描述现实世界中实体对象的逻辑组织、对象间关系、限制等的模型。基于面向对象数据模型的数据库系统称为面向对象数据库系统。

## 1-4-2 数据库技术与其他相关技术的结合

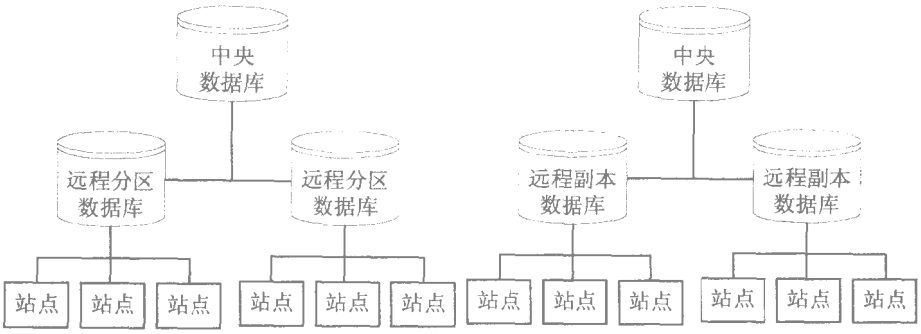
数据库技术与计算机或相近领域的其他新技术的结合，是数据库技术的重要发展趋势，已出现了许多新的数据库系统，如，数据库技术与分布处理技术相结合，产生了分布式数据库系统；数据库技术与并行处理技术相结合，产生了并行数据库系统；数据库技术与人工智能技术相结合，产生了知识库系统和主动数据库系统；数据库技术与多媒体技术相结合，产生了多媒体数据库系统；数据库技术与面向对象技术相结合，产生了面向对象数据库系统；数据库技术与工作流技术相结合，产生了工作流数据库系统；数据库系统与模糊技术相结合，产生了模糊数据库系统等。下面仅对几个典型的新型数据库系统作一简要介绍，以说明数据库技术是如何吸收、结合其他计算机技术的。

### 1. 分布式数据库系统

分布式数据库是一个数据在多个不同的地理位置存储的数据库。数据库的某一部分在一个位置存储和处理，数据库的其他部分在另外一个或多个位置存储和处理。有两种典型的分布式数据库，一种是中央数据库，包括分区数据库和副本式数据库；另一种是中央索引数据库，包括中央索引数据库和网络请求分布式数据库。

在分布式数据库中，中央数据库存储了所有数据，而在各分布式站点上只存储其中的某一部分数据，也就是说，各分站点数据库为中央数据库的分区。分区数据库主要是存储在该站点上经常需要处理的数据。在某一站点上，若需要其他数据，则需向中央数据库发出请求。副本分布式数据库是在各站点上存储了中央数据库的副本，这样，各站点就可以直接使用副本数据库中的数据。副本分布式数据库的关键技术是如何保证各站点副本数据库和中央数据库的一致性。图 1-12 给出了中央分布数据库的两种结构。

在中央索引数据库中，当用户使用数据时，先将数据服务请求发送到中央索引数据库中，中央索引数据库根据索引指针寻找存储这些数据的有关站点数据库，最后再把回应返送给发出请求的用户。在这种中央索引数据库中，数据存储在各站点上，由各站点负责维护自己的数据，而中央索引数据库只存储数据的索引，而不存储具体的数据。在网络请求分布式数据库中，没有中央索引数据库，依靠网络轮询来完成用户的数据请求。




（a）分区分布式数据库　　（b）分区副本式数据库

图 1-12　中央分布式数据库系统结构

### 2. 面向对象数据库

随着数据库技术应用领域的进一步拓宽，要求数据库不仅能方便地存储和检索结构化的数字和字符等信息，而且可以方便地存储和检索诸如图形、图像等复杂信息对象。传统的 DBMS 很难处理这些复杂对象，如包括了复杂关系、数据类型多种多样的 CAD 数据库中的设计数据。若在关系型数据库系统中要处理这些复杂数据对象，则需使用专门的应用程序把这些复杂的数据对象分解成适合于在二维表中存储的数据。面向对象数据库则可以像对待一般对象一样存储这些数据与过程，这些对象可以方便地被系统检索。

在面向对象数据库中，存储的对象除具有简单数据类型的对象外，还包括具有非常复杂的数据类型的对象，如树、文件、图形、图像等，这些复杂数据类型可能由其他基本数据类型组成。在面向对象的数据库系统中，可以从诸如整型、实型、布尔型、字符串型等基本类型开始，使用诸如记录结构、聚集类型、引用类型等类型构造符构造新的数据类型。

有关面向对象数据库系统的研究在数据库研究领域是近几年来的热点领域之一，主要包括以下几个主要方向：一是以关系数据库与 SQL 为基础研究其扩展关系模型。目前，Informix，DB2，Oracle，Sybase 等关系数据库厂商，都在不同程度上扩展了关系模型，推出了对象关系数据库产品。二是以面向对象的程序设计语言为基础，研究持久的程序设计语言,支持面向对象模型,例如 Servialogic 公司的 Gemstone 是以面向对象语言 Smalltalk 为基础的。三是研究开发新的面向对象数据库系统，支持面向对象数据模型，如美国 Itasca System 公司的 Itasca 产品等。

### 3. 多媒体数据库

传统的数据库管理系统在处理大字节的数据类型时，采取了复杂的方法。但对于要求处理大量图形、图像、音频、视频等多媒体数据时，这些方法就显得无能为力了。因此，如何存储和使用这些具有海量数据量的多媒体数据就成为摆在数据库研究与开发人员面前的重要课题。从技术角度讲，多媒体数据库涉及了诸如图像处理技术、音频处理技术、视频处理技术、三维动画技术、海量数据存储与检索技术等多方面的技术，如何综合处理这些技术是多媒体数据库技术需要解决的问题。

多媒体数据库是实现对格式化和非格式化的多媒体数据的存储、管理与检索，其主要

需要全本请访问 www.ertongbook.com