

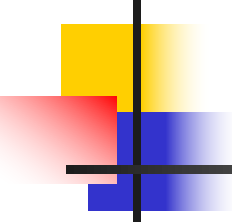


# 数据库原理

---

倪巍伟

[wni@seu.edu.cn](mailto:wni@seu.edu.cn)

- 
- 
- **数据库管理已经成为现代计算平台的核心成分之一。**
  - **本课程从平台对应用的支持出发，讨论DB系统的基本概念、方法、原理和技术，以提高设计、开发和系统分析能力。**



# 第1章 引论

---

## 1.1 数据管理的发展

**(1) 数据库技术是应数据管理的任务的需要而产生的。  
数据管理是指如何对数据进行分类、组织、编码、  
存储、检索和维护，它是数据处理的中心问题。**

**(2) 数据管理经历了三个阶段**

- .人工管理**
- .文件系统**
- .数据库系统**

背景		人工管理	文件系统	数据库系统
	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
特点	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
	数据的管理者	人	文件系统	数据库管理系统
	数据面向的对象	某一应用程序	某一应用程序	整个应用系统
	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



# 数据密集型应用（**data intensive application**）

---

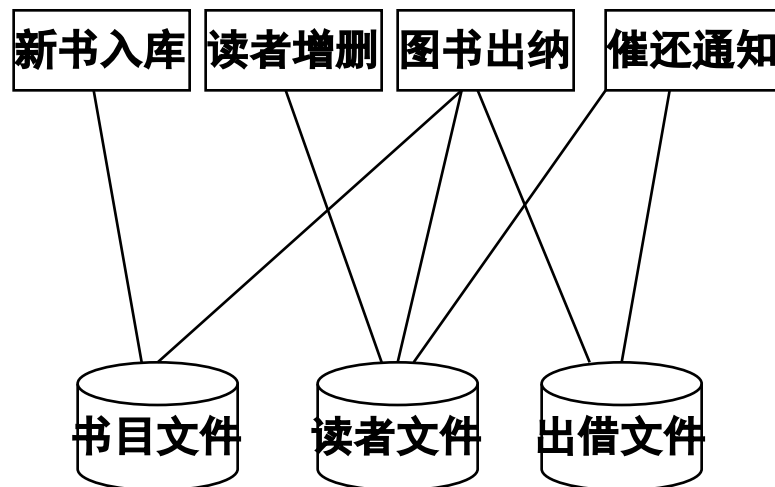
## ■ 数据密集型应用的特点

- 1.涉及数据量大（TB），一般内存容纳不下，需放在辅存；
- 2.数据为持久数据（persistent data）；
- 3.数据为多个应用程序所共享,MIS、银行、办公系统、网络数据服务等。

# 数据管理

## ■ 文件系统的缺点

- 1.编写应用程序不方便;
- 2.数据冗余;
- 3.文件结构的每一修改  
将导致应用程序的修改,  
应用程序维护工作量大;
- 4.文件系统不支持对文件的并发访问;
- 5.数据的规范化与标准化较难达到。





## 数据库系统（database system）

**数据与单位（enterprise）对应，看成一个单位的共享资源，由数据库管理系统软件（DataBase Management System，简称DBMS）统一管理。**

**数据库系统以统一管理和共享数据为主要特征**

### 数据结构化

**数据结构化是数据库与文件系统的根本区别。在描述数据时不仅要描述数据本身，还要描述数据之间的联系。**

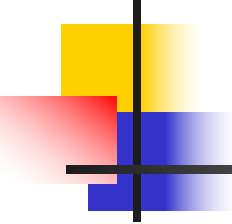


## DBMS的发展

---

- 1964年, IDS (Integrated Data Store) 网状, 通用电气, Bachman (图灵奖, 1973)
- 60年代末, IMS (Information Management System) 层次, IBM
- 1970 Codd提出了关系模型 (relational data model) (图灵奖, 1981)
- 1977 IBM System R和Berkeley分校Ingres运行, 关系模型
- 80年代, 商品化关系DBMS成为主流, DB2, Oracle, Sybase, Informix, MS SQL Server, Access等





---

**数据管理将继续发展，应用是动力，新的软硬件技术是基础，数据库是主要形式。**

## 1.2 数据库系统

数据库系统一般由数据库、数据库管理系统(及其开发工具)、应用系统、数据库管理员和用户构成。下面分别介绍这几个部分的内容。

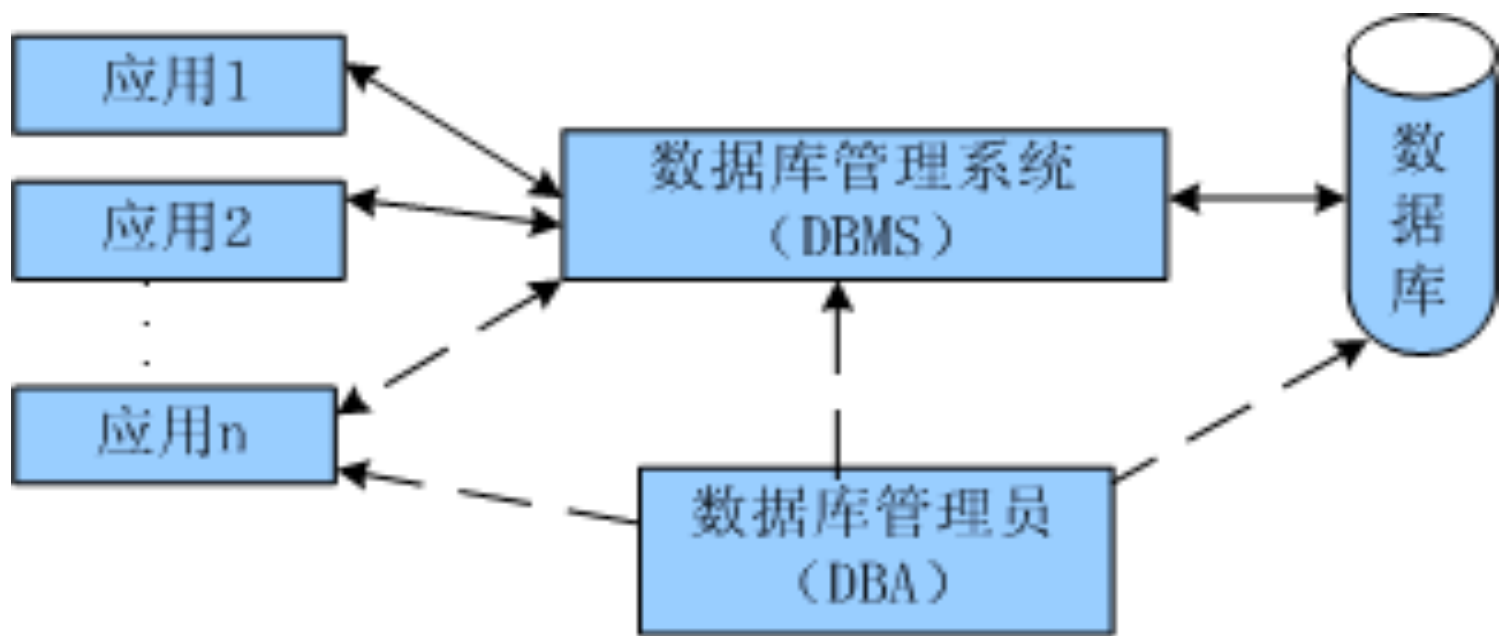


图1-2 数据库系统



## DBMS的主要功能

---

- 提供高级的用户接口
- 查询处理和优化
- 数据目录管理
- 并发控制
- 恢复功能
- 完整性约束检查
- 访问控制



## 1. 硬件平台及数据库

由于数据库系统数据量都很大，加之DBMS丰富的功能使得自身的规模也很大，因此，整个数据库系统对硬件资源提出了较高的要求，这些要求是：

(1)要有足够大的内存，存放操作系统、DBMS的核心模块、数据缓冲区和应用程序。

(2)有足够大的磁盘等直接存取设备存放数据库和备份数据。

(3)要求系统有较高的通道能力，以提高数据传送率。



## 2、软件

---

数据库系统的软件主要包括：

(1)**DBMS**。DBMS是为数据库的建立、使用和维护配置的软件。

(2)**支持DBMS运行的操作系统。**

(3)**具有数据库接口的高级语言及其编译系统，**  
便于开发应用程序。



---

**(4)以DBMS为核心的应用开发工具。**

**应用开发工具是系统为应用开发人员和最终用户提供的高效率、多功能的应用生成器、第四代语言等各种软件工具。它们为数据库系统的开发和应用提供了良好的环境。**

**(5)为特定应用环境开发的数据库应用系统。**



### 3. 人员

开发、管理和使用数据库系统的人员主要是：  
**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。**不同的人员涉及不同的数据抽象级别，具有不同的数据视图，有各自的职责。




## **(1)数据库管理员(DataBase Administrator, DBA)**

---

**具体职责包括:**

- 1)决定数据库中的信息内容和结构**
- 2)决定数据库的存储结构和存取策略**
- 3)定义数据的安全性要求和完整性约束条件**
- 4)监控数据库的使用和运行**





---

另外，在数据运行过程中，大量数据不断插入、删除、修改，时间一长，会影响系统的性能。因此，DBA要定期对数据库进行重组织，以提高系统的性能。

当用户的需求增加或者改变时，DBA还要对数据库进行较大的改造，包括修改部分设计，即数据库的重构造。



## **(2)系统分析员和数据库设计人员**

**系统分析员负责应用系统的需求分析和规范说明，要和用户及DBA相结合，确定系统的软硬件配置，并参与数据库系统的概要设计。**

**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用户需求调查和系统分析，然后进行数据库设计。在很多情况下，数据库设计人员就由数据库管理员担任。**



### **(3)应用程序员**

**应用程序员负责设计和编写应用系统的程序模块，并进行调试和安装。**

### **(4)用户**

**这里用户是指最终用户(EndUser)。最终用户通过应用系统的用户接口使用数据库。常用的接口方式有浏览器、菜单驱动、表格操作、图形显示、报表书写等。**




## 1.3 数据、数据模型和数据模式

---

**对事物描述的符号记录称为数据。**

**数据是信息存在的一种形式，只有通过解释或处理才能成为有用的信息。**

**模型，特别是具体的实物模型，人们并不陌生。例如，一张地图，一组建筑设计沙盘，一架精致的航模飞机，都是具体的模型。一眼望去，就会使人联想到真实生活中的事物。模型是现实世界特征的模拟和抽象。数据模型(DataModel)也是一种模型，它是现实世界数据特征的抽象。**



**数据库是某个企业、组织或部门所涉及的数据的综合，它不仅要反映数据本身的内容，而且要反映数据之间的联系。由于计算机不可能直接处理现实世界中的具体事物，所以人们必须事先把具体事物转换成计算机能够处理的数据。在数据库中用数据模型这个工具来抽象、表示和处理现实世界中的数据和信息。通俗地讲数据模型就是现实世界的模拟。**

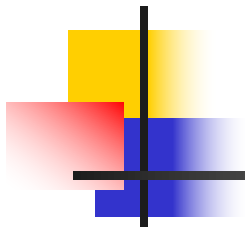


现有的数据库系统均是基于某种数据模型的。因此，了解数据模型的基本概念是学习数据库的基础。

**数据模型应满足三方面要求：**

- 一是能比较真实地模拟现实世界；**
- 二是容易为人所理解；**
- 三是便于在计算机上实现。**

一种数据模型要很好地满足这三方面的要求在目前还很困难。在数据库系统中针对不同的使用对象和应用目的，采用不同的数据模型。



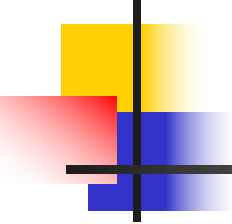
**不同的数据模型实际上是提供给我们模型化数据和信息的不同工具。**



**数据模型**是数据库系统的**核心**和**基础**。各种机器上实现的DBMS软件都是基于某种数据模型的。

为了把现实世界中的具体事物抽象、组织为某一DBMS支持的数据模型，人们常常首先将现实世界抽象为信息世界，然后将信息世界转换为机器世界。





---

**也就是说，首先把现实世界中的客观对象抽象为某一种信息结构，这种信息结构并不依赖于具体的计算机系统，不是某一个DBMS支持的数据模型，而是概念级的模型；然后再把概念模型转换为计算机上某一DBMS支持的数据模型。**



## 数据模型的组成要素

---

一般地讲，数据模型是严格定义的一组概念的集合。这些概念精确地描述了系统的静态特性、动态特性和完整性约束条件。因此数据模型通常由**数据结构**、**数据操作**和**完整性约束**三部分组成。



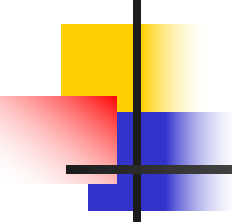
## 1. 数据结构

---

**数据结构是所研究的对象类型的集合。这些对象是数据库的组成成分，它们包括两类：**

**一类是与数据类型、内容、性质有关的对象，例如网状模型中的数据项、记录，关系模型中的域、属性、关系等；**

**一类是与数据之间联系有关的对象，例如网状模型中的系型(Set type)。**



---

数据结构是刻画一个数据模型性质最重要的方面，是对系统**静态特性**的描述。因此，在**数据库系统中**，通常按照其**数据结构的类型**来命名**数据模型**。例如层次结构、网状结构和关系结构的数据模型分别命名为**层次模型**、**网状模型**和**关系模型**。



## 2. 数据操作

---

**数据操作是指对数据库中各种对象(型)的实例(值)允许执行的操作的集合，包括操作及有关的操作规则。**

**数据库主要有检索和更新(包括插入、删除、修改)两大类操作。数据模型必须定义这些操作的确切含义、操作符号、操作规则(如优先级)以及实现操作的语言。**

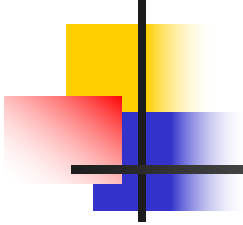
**数据操作是对系统动态特性的描述。**



### 3. 数据的约束条件

---

**数据的约束条件是一组完整性规则的集合。完整性规则是给定的数据模型中数据及其联系所具有的制约和依存规则，用以限定符合数据模型的数据库状态以及状态的变化，以保证数据的正确、有效、相容。**



**数据模型是用来描述现实世界的，因此，  
越自然越好（易于理解）。**



在数据库中，针对不同的使用对象和应用目的，采用多级数据模型，一般可分为三级：

**1. 概念数据模型（与DBMS无关）**

主要用来描述一个单位的概念化结构，面向用户、面向现实世界，是一种语义模型，常用E-R图表示，接近人的思维，容易理解。

**2. 逻辑数据模型（与DBMS有关）**

用概念数据模型表示的数据必须转化为逻辑数据模型表示的数据，才能在DBMS中实现，因此，逻辑模型既要面向用户，也要面向实现。





---

### **3.物理数据模型（与DBMS、OS硬件有关）**

它反映数据的存储结构，如物理块、指针、索引等。物理数据模型不但与**DBMS**有关，还与**OS**和**硬件**有关。

## 什么叫数据模式？

以一定的数据模型对一个单位的类型、结构及其相互间的关系所进行的描述。

数据模式有型与值之分：  
型——框架  
值——实例

例：学生记录

型：

姓名	学号	性别	出生年份	籍贯	系别	入学时间
----	----	----	------	----	----	------

(a)

一个值：

王彤	9098135	女	1980	江苏	计算机	1998
----	---------	---	------	----	-----	------

(b)

图1-3 数据模式及其实例



## 数据模型和数据模式区别：

**数据模型**是描述现实世界数据的**手段、工具**。  
**数据模式**是利用这个手段和工具对具体对象相互间的关系所进行的描述。是关于型的描述。它与DBMS和OS硬件无关。

例：C语言和用C语言编写的程序的区别。



## **★ 数据模式通常分为三级：**

### **1. 概念模式**

**是用逻辑数据模型对一个单位的数据的描述。是一个单位的全局模式，是整个数据库的逻辑结构。**

### **2. 外模式**

**是用逻辑数据模型，对用户所用到的那部分数据的描述。它是概念模式的一部分或由前者推导而来的。（这是每一个用户看到的模式。每一个用户的外模式不一定相同，外模式是一种局部模式）**



### 3. 内模式

是用物理数据模型对数据的描述。表示数据库存储结构，即物理结构。

概念模式、外模式和内模式都存于数据目录中，是数据目录的最基本内容。DBMS通过数据目录，管理和访问数据模式。



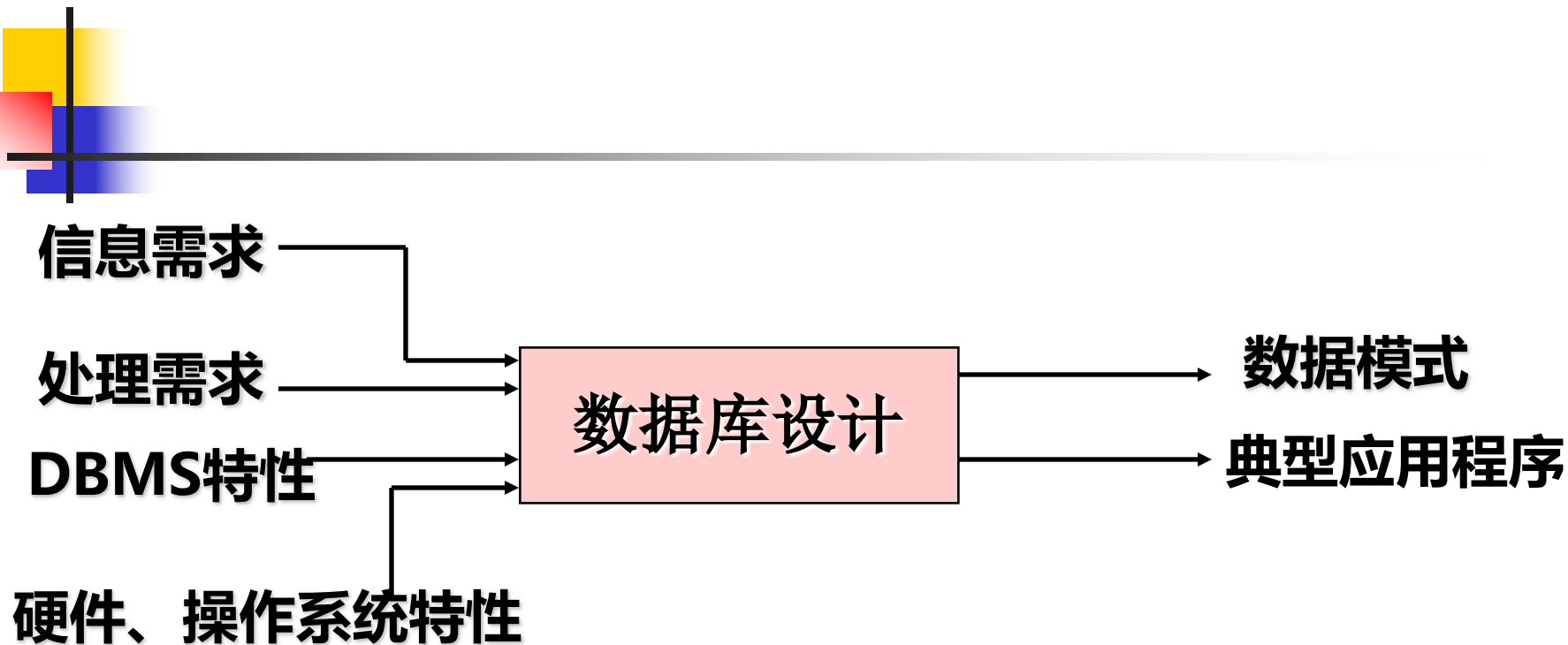
## 1.4 数据库应用

---

**数据库的应用就是选择合适的DBMS，设计、建立维护和管理数据库系统，为用户服务。**

### **1.数据库系统的规划**

### **2.数据库设计**



**图 1--5 数据库设计的基本任务**



---

**3.数据库的建立**

**4.数据库的运行、管理和维护**

**5.数据库的扩充和维护**





# 数据库技术的发展\*

---

- ❖ **第一代数据库系统**

- 20世纪70年代 层次和网状数据库系统**

- ❖ **第二代数据库系统**

- 关系数据库系统**

- ❖ **第三代数据库系统**

- **面向对象数据模型**

- **数据库技术和其他学科内容相结合（分布式数据库、工程数据库、演绎数据库、知识数据库、时态数据库、空间数据库）**



## ❖ Internet时代的数据库技术

**Internet技术的发展促进新的数据库应用，诸如支持高层决策的数据仓库、OLAP分析、数据挖掘、电子商务等。**

**出现了Web信息检索技术与系统、Web数据集成与共享技术与系统、数据流技术与系统等具有Internet时代特征的数据库技术和数据库管理系统。**