

UNIT 1 数据管理技术概述



本讲主要目标



学完本讲后, 你应该能够了解:

- 1、数据管理技术的发展历史阶段以及各阶段数据管理技术的优劣;
 - 2、文件系统到数据库系统的演化过程;
 - 3、数据库系统组成包括硬件、数据库、软件与人;
- 4、数据库系统的主要特点有:数据的高度结构化、数据的高共享性、程序与数据的高独立性、数据控制力度大;
- 5、数据库管理系统的控制功能包括:安全性保护、 完整性、并发与数据库恢复;
- 6、数据库系统的结构:内部三级模式、外部体系结构;
 - 7、概念模型和数据模型的概念和作用;
 - 8、目前常见的数据库管理系统产品。



本讲主要内容

- 一、数据管理技术的发展
- 二、数据库系统的组成
- 三. 数据库系统的特点
- 四.数据库的历史
- 五、数据库系统的结构
- 六. 数据库管理系统产品





1、考虑下面三个数据管理问题:

问题一: 在一个系,有一张学生选课表(学号,课程号,选修次数,成绩),记载着该系学生的选课情况。该系的教学秘书需要查询和统计某些学生的修课情况、各门课程的成绩、平均成绩;需要查询和统计选修了某门课程的学生、课程的平均成绩、成绩的分布情况…

问题二: 在一个大学,有一张学生选课表(学号,课程 号, 选修次数, 成绩), 记载着该校学生的选课情况, 学 生基本信息表(学号,姓名,性别,所属系,电话号码), 记载该校学生的信息,开设课程表(课程号,课程名,课 程类别, 学分), 记载该校开设的课程信息。学生工作处 需要查询和统计各系、各性别学生人数的分布情况、平 均成绩的分布情况, 以及重修学生的人数和比例; 学校 教务处需要查询和统计各系开设的课程总门数,不同类 别的课程门数、总学分数、各课程选修的学生总数和平 均成绩;各系的教学秘书需要对于本系学生完成问题一 中的功能…



问题三: 一个大学有多个管理部门: 学生工作处、 人事处、财务处、教务处、房管处、设备处、校长办公 室…, 负责对全校的学生、职工、财务、教学、房产、 设备、校日常事务…进行管理; 有多个学院: 计算机学 院、数学学院、生命科学学院、法学院、商学院…, 分 别负责对本院学生、教学、教师、设备等进行管理…



2、你如何解决这三个问题的计算机信息管理?







问题二







3、应该采用适当的工具来解决问题

从问题一至问题三, 为什么感觉越来越困难?

- ◇ 问题越来越复杂:用户越来越多,数据量和范围越来越大。需求越来越复杂
- * 解决问题的方法限于同样的技术

—— 文件方式存储数据 + 高级程序设计语言编制应用程序

针对问题的特点和规模, 应该采用适当的工具(技术)来解决

数据管理技术



4、数据管理技术就是管理数据的工具

数据管理对数据进行分类、组织、编码、存储、 检索和维护,是数据处理的中心问题。

数据管理技术的目标就是合理地组织数据、有效 地存取数据,为此必须解决下列问题:

- ◆存储介质
- ◆数据的表示
- ◆数据的操作
- ◆数据的管理



5、数据管理技术的发展阶段

在应用的驱动下, 在计算机硬件、软件发展的基础

- 上,数据管理技术经历了下面几个阶段(见教材P3):
 - 1)人工管理阶段
 - 2)文件系统阶段
 - 3)数据库系统阶段





1)人工管理阶段

- 》 硬件状况:没有直接存取的存储设备
- 》 软件状况:没有操作系统,没有管理数据的软件
- 》 数据处理方式: 批处理
- > 特点
 - ❖ 数据不保存
 - ◆ 数据没有结构
 - ◆ 应用程序管理数据
 - ◆ 数据不共享、冗余度极大
 - ※ 数据不具有独立性



2)文件阶段

- 》 硬件状况: 磁盘, 磁鼓
- 》 软件状况:有文件系统
- 》 数据处理方式: 联机实时处理, 批处理
- > 特点
 - ◆ 数据可以长期保存
 - 数据有一定的结构: 记录内有结构, 整体 无结构
 - ◆ 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



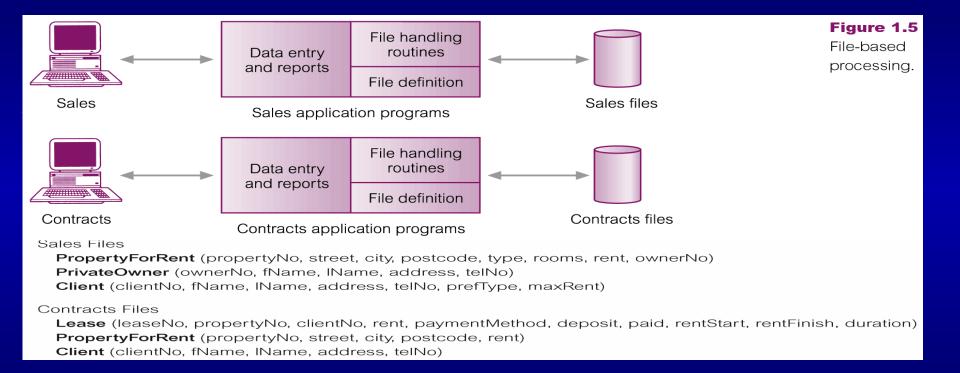


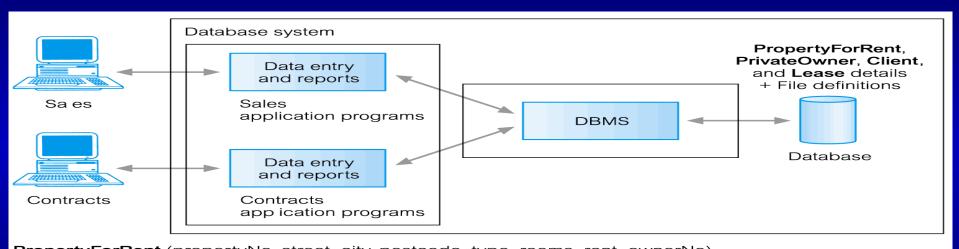
数据管理技术的发展

3)数据库系统阶段

- 硬件状况:大容量磁盘
- 软件状况:有数据库管理系统
- 数据处理方式: 联机实时处理, 批处理, 分布处 理
- 特点
 - 数据可以长期保存
 - 数据高度结构化,用数据模型描述
 - 由数据库管理系统处理数据。提供数据安全性、 完整性、并发控制与故障恢复能力
 - ◆ 数据的共享性高. 冗余度小
 - ❖ 具有高度的数据物理独立性和一定的数据逻辑 独立性

13

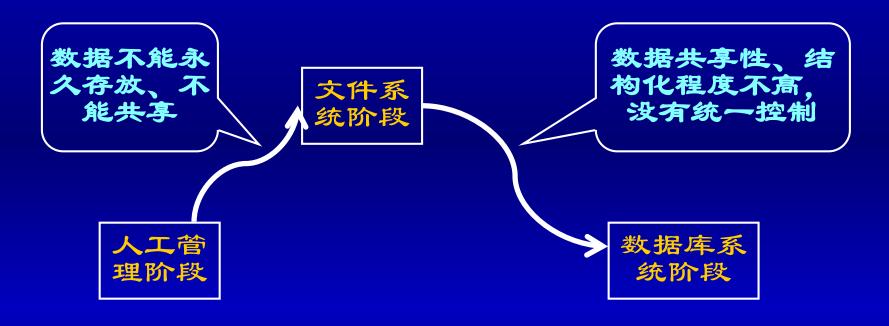




PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)
PrivateOwner (ownerNo, fName, IName, address, telNo)
Client (clientNo, fName, IName, address, telNo, prefType, maxRent)
Lease (leaseNo, propertyNo, clientNo, paymentMethod, deposit, paid, rentStart, rentF nish)



几个阶段的根本区别:



从文件系统到数据库系统,标志着 数据管理技术的飞跃。



本讲主要内容

- 一、数据管理技术的发展
- 二. 数据库系统的组成
- 三. 数据库系统的特点
- 四.数据库的历史
- 五、数据库系统的结构
- 六. 数据库管理系统产品

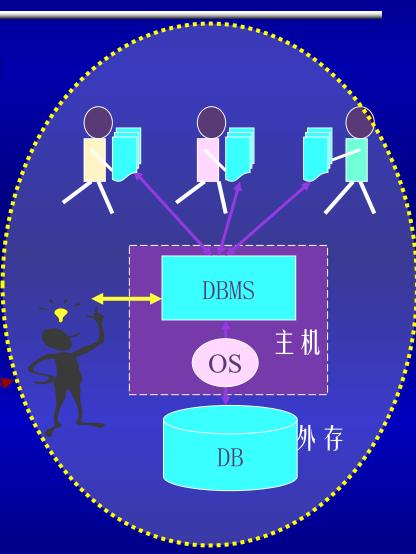




数据库系统的基本概念: (见教材P3)

①数据库系统框架:

- ◆ 数据库(Database, DB)
- ◆ 数据库管理系统(Database Management System, DBMS)
- ◆ 数据库系统(Database System, DBS)





数据库系统的基本概念:

②数据库(DB):

定义:数据库是长期存储在计算机内、有组织的、可 共享的大量数据的有机集合。

数据库用科学、规范的数据模型组织数据,为准确、有效地存取数据提供了基础。

基本的数据模型包括: 层次模型、网状模型、关系模型



数据库系统的基本概念:

③数据库管理系统(DBMS):

定义: DBMS是管理和维护数据库的系统软件。位于用户和操作系统之间, 对数据库进行统一管理。

功能: DBMS 能定义数据存储结构, 提供数据的操作机制, 维护数据库的安全性、完整性和可靠性。

数据库管理系统是统一管理数据的专用软件系统。 大大減少了应用程序的编程量。



数据库管理系统的主要功能

数据

控制

功能

防止非法用户

防止数据泄露、数据库的安全性

数据库中的数据从现实 世界抽象转换而来

的非法操作

更改或破坏

保证数据库中的数据真实反映现实世界的现状

数据库中的故障不可避免

能够将数据库从错误状态恢复到已知正确状态

保证事务的原子性 与持久性

对并发操作进行控制

保证事务的隔离性 与一致性 数据库的完整性

数据控制语言 (DCL)

数据库的可恢复性

并发控制

定义 功能

数据库

管理系统

数据定义语言(DDL)

对数据对象的定义

删除Drop

创建Create

修改Alter

数据 操纵 功能

数据操纵语言(DML)

对数据库中的 数据进行操作 查询Select

插入Insert

修改Update

删除Delete



数据库系统的基本概念:

④数据库系统(DBS):

定义:数据库系统是指在计算机系统中引入数据库 后的系统构成。

一个完整的数据库系统一般由数据库、数据库管理系统(及其应用开发工具)、 应用系统、数据库管理员和用户组成。

注: 在不引起混淆的情况下人们常常把数据库系统简称为数据库。





数据库系统的组成



1、数据库系统的组成成分(见教材P33)

- 硬件(足够的内存、外存、数据传输率)
- > 数据库
- 文件(OS、DBMS、以DBMS为核心的应用开发工具、高级语言及编译系统、数据库应用系统)
- 人(数据库管理员、应用程序员、最终用户)
 - ❖ 最终用户
 - 临时用户——用SQL访问DBMS的用户
 - · 初级用户——通过菜单访问DBMS的用户

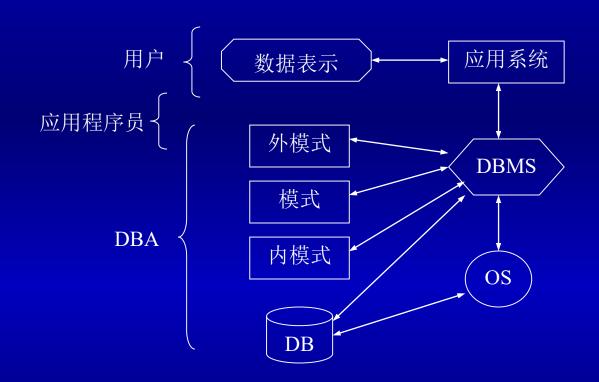


2、各种人员的职责

- ◆ **应用程序员** 直接与DBMS打交道,为初级用户编写菜单程序
- ◆ 最终用户 通过菜单或SQL使用数据库,以满足日常的事务处理要求
- ◆ 数据库管理员 (DBA) 的职能
 - 1) 决定DB中的信息内容和结构;
 - 2) 决定DB中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;
 - 5) 数据库的改进和重组重构。



3、各种人员的数据视图





本讲主要内容

- 一、数据管理技术的发展
- 二、数据库系统的组成
- 三. 数据库系统的特点
- 四.数据库的历史
- 五、数据库系统的结构
- 六. 数据库管理系统产品





1、数据的高度结构化(见教材P5)

students

sid	fname	lname	class	telephone
1	Jones	Allan	2	555-1234
2	Smith	John	3	555-4321
3	Brown	Harry	2	555-1122
5	White	Edward	3	555-3344

course

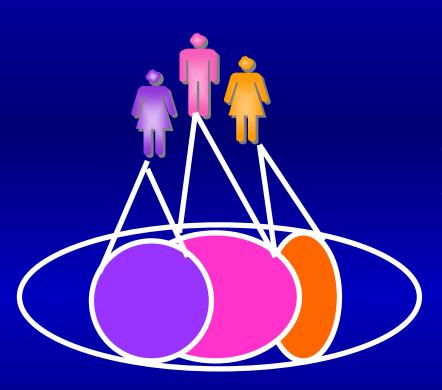
cno	cname	croom	time
101	French I	2-104	MW2
102	French II	2-113	MW3
105	Algebra	3-105	MW2
108	Calculus	2-113	MW4

enrollment

sid	cno	major
1	101	No
1	108	Yes
2	105	No
3	101	Yes
3	108	No
5	102	No
5	105	No



2、数据的高共享性



例如:校人事处、 财务处、房产处共 享学校职工的基本 信息,但共享的列 有一些区别



3、程序与数据的高独立性

程序与数据的独立性

逻辑独立性

物理独立性

- ◆ 逻辑独立性 ---- 用户的应用程序与数据库 的逻辑结构是相互独立的
- ◆ 物理独立性 ---- 用户的应用程序与存储在 磁盘上的数据库中的数据是相互独立的

数据库系统的三级模式结构和两级映象

---- 程序与数据的独立性



- 4、数据控制力度大
- -----DBMS统一管理和控制数据

数据库的共享是并发的共享,即多个用户可以同时存取数据库中的数据。

DBMS提供以下4个方面的数据控制功能:

- ◆数据的安全性(Security)保护
- ◆数据的完整性(Integrity)检查
- ◆并发(Concurrency)控制
- ◆数据库恢复 (Recovery)



本讲主要内容

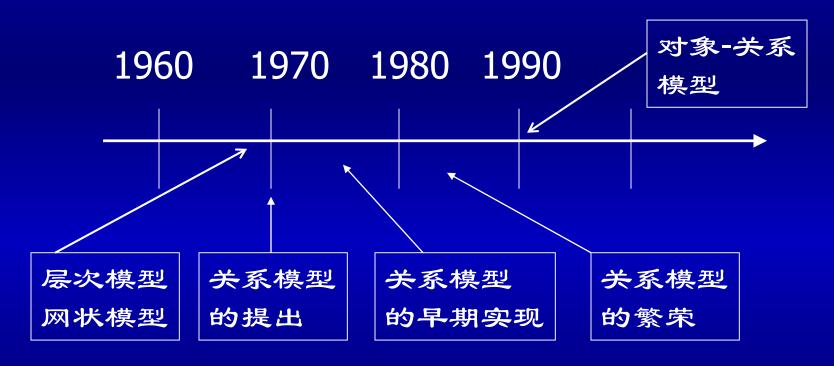
- 一、数据管理技术的发展
- 二、数据库系统的组成
- 三. 数据库系统的特点
- 四.数据库的历史
- 五、数据库系统的结构
- 六. 数据库管理系统产品





1、数据模型的发展历史(见教村P7-P22)

数据库的历史就是<mark>数据模型</mark>的发展历史, 典型的数据库模型有:





- 1) 层次模型 把数据组织成一棵根在上、叶在下的有向树
- 2) 网状模型 把数据组织成无环有向图
- 4) 对象-关系模型 用对象的方法组织数据



2、关系模型的提出

- ◆ 层次模型和网状模型的问题
 - ◆ 必须编写复杂的程序处理数据
 - ◆ 数据独立性差
 - ❖ 缺乏被广泛接受的理论基础
- ◆ 1970年. Ted Codd提出关系模型
 - ❖ 数据独立性
 - ◆ 关系代数
- ◆ 许多人认为关系模型无法实现,太慢,比不上 手工的程序快,人们不会接受



3、关系模型的实现

- ◆ 有些人看到它可以10倍地提高编程效率
 - ❖ 好像FORTRAN最初的时候
- ◆ 关系数据库的实现成为热门的研究课题
 - * 如何优化
 - 如何做事多处理
- ◆ 最早的研究系统出现在1970年代中期,IBM的 System R,和Berkeley的INGRES
- ◆ 1980年代初期出现了许多的商业产品

层次模型、网状模型、关系模型比较

	层次、网状模型	关系模型
数据结构	复杂	简单
联系表示	指针链接	二维表
出现年代	68, 69	70
理论基础	无	关系规范化理论
查询效率	较高	较低
数据独立性	较差	较高
DML操作方式	过程式	非过程式



本讲主要内容

- 一、数据管理技术的发展
- 二、数据库系统的组成
- 三. 数据库系统的特点
- 四.数据库的历史
- 五. 数据库系统的结构
- 六. 数据库管理系统产品





(见教村P22-P28)

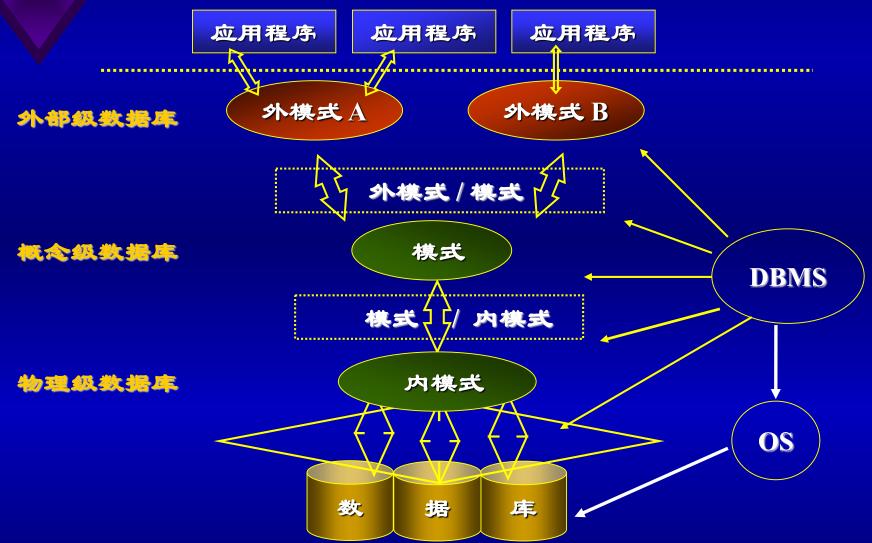
从DBMS角度看(DBS内部的三级模式结构):

- 外部级 (外模式)
- 概念级(模式)
- 内部级(内模式)

从DB用户角度看(DBS外部的体系结构)

- 单用户结构
- 主从结构
- 客户/服务器结构
- 分布式结构
- 云结构

1、三级模式结构





(1) 三级模式

- 1) 模式
- ◆模式 (Schema, 也称逻辑模式)
 - 数据库中全体数据的逻辑结构和特征的描述
 - 所有用户的公共数据视图,是数据库数据在逻辑级上的视图
- ◆模式的地位: 是数据库系统模式结构的中间层
 - 与数据的物理存储细节和硬件环境无关
 - 与具体的应用程序、开发工具及高级程序设计 语言无关



◆定义模式

- 数据的逻辑结构(数据项的名字、类型、取值范围等)
- 数据之间的联系
- 数据有关的安全性、完整性要求



2) 外模式

- ◆外模式 (External Schema, 也称子模式或用户模式)
 - 是数据库用户的数据视图数据库用户(包括应用程序员和最终用户)能够看见和使用的局部数据的逻辑结构和特征的描述
 - 数据库用户的数据视图,是与某一应用有关的数据的逻辑表示

◆外模式的用途

- 保证数据库安全性的一个有力措施。
- 每个用户只能看见和访问所对应的外模式中的数据



- ◆外模式的地位:介于模式与应用之间
 - 模式与外模式的关系: 一对多
 - 外模式通常是模式的子集
 - 一个数据库可以有多个外模式。反映了不同的用户的 应用需求、看待数据的方式、对数据保密的要求
 - 对模式中同一数据,在外模式中的结构、类型、长度、 保密级别等都可以不同
 - 外模式与应用的关系: 一对多
 - 同一外模式也可以为某一用户的多个应用系统所使用,
 - 但一个应用程序只能使用一个外模式。

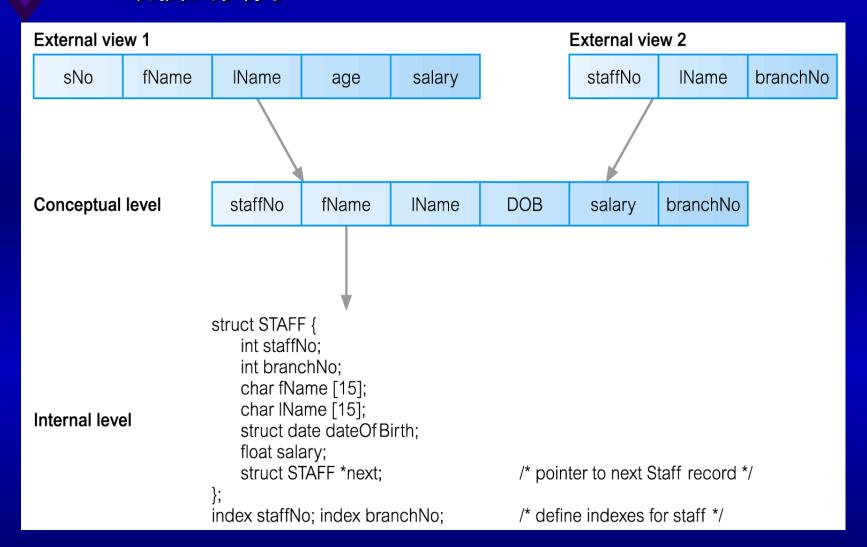


3) 内模式

- ◆内模式 (Internal Schema, 也称存储模式 Storage Schema)
 - 是数据物理结构和存储方式的描述
 - 是数据在数据库内部的表示方式
 - •记录的存储方式(顺序存储,按照B树结构 存储,按hash方法存储)
 - 索引的组织方式
 - 数据是否压缩存储
 - 数据是否加密
 - 数据存储记录结构的规定



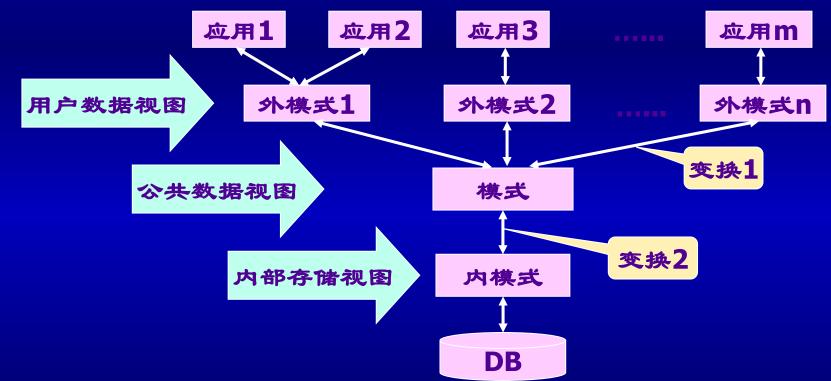
三级模式实例





(2) 三级模式间的联系

模式是内模式的逻辑表示;内模式是模式的物理实现;外模式是模式的逻辑子集。





(3) 数据库的二级映像:

- 外模式/模式映像(映像1)
- 模式/内模式映像(映像2)

二级映像作用:

- 实现逻辑数据到物理数据的转换
- 实现数据独立性

外模式/模式映像:实现数

模式/内模式映像:实现数

数据独立性有哪几种?



1) 外模式 / 模式映象(映像1)

- ◆外模式/模式映象定义外模式与模式之间的对应关系
 - 外模式/模式映象不唯一,每一个外模式都对应一个外模式/模式映象
 - 映象定义通常包含在各自外模式的描述中

◆外模式/模式映象用途—保证数据的逻辑独立性

- 当模式改变时,数据库管理员修改有关的外模式/模式映象,使 外模式保持不变
- 应用程序是依据数据的外模式编写的,从而应用程序不必修改, 保证了数据与程序的逻辑独立性,简称数据的逻辑独立性。



- 2) 模式 / 内模式映象(映像2)
- ◆模式 / 内模式映象
 - 模式/内模式映象定义了数据全局逻辑结构与 存储结构之间的对应关系。例如,说明逻辑记 录和字段在内部是如何表示的
 - 数据库中模式 / 内模式映象是唯一的
 - 该映象定义通常包含在模式描述中



◆模式/内模式映象的用途—保证数据的物理独立性

- 当数据库的存储结构改变了(例如选用了另一种存储结构),数据库管理员修改模式/内模式映象,使模式保持不变
- 应用程序不受影响。保证了数据与程序的物理独立性,简称数据的物理独立性。



2. DBS体系结构

从最终用户角度看,数据库系统分为:

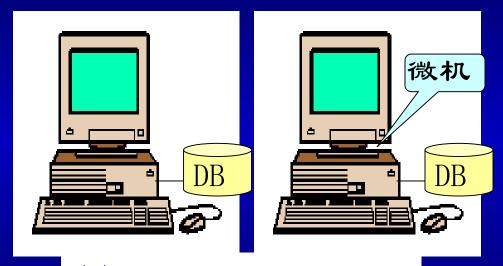
(1) 单用户结构

所有应用程序、DBMS、数据都装在一台计算机

上,由一个用户独占。

缺点:数据冗余大,

数据不共享。



(1) 单用户数据库系统



(2) 主从式结构

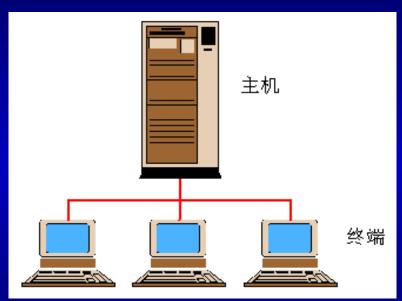
一个主机带多个终端的多用户系统。

主机上:应用程序、DBMS、数据库

终端: 各用户通过主机的终端并发存取数据库,

共享数据资源。

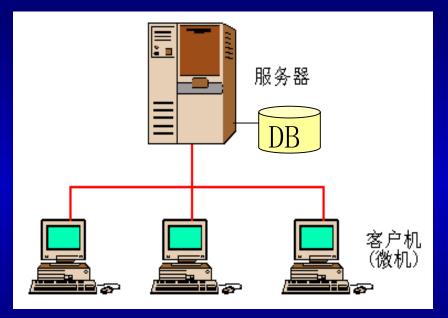
缺点: 终端数目增加时, 主机负担过重, 系统性能下降





(3)客户/服务器结构(Client/Server)

C/S体系结构的关键在于功能的分布上。从数据库角度看,把原主机负担的非数据库系统的工作交给客户机承担,而数据库服务器则专司DBMS的功能。





功能划分:

客户端:由一些应用程序构成,包括:格式处理、数据输入、报表输出、图形界面、用户界面等。

服务器端:数据库、完成事务处理和数据控制(并发、恢复、安全性、完整性)等方面的功能。



① 两层C/S结构

由服务器、客户机在局部范围内建立局域网。数据库设置在服务器中,客户机可存放其备份或临时表。

工作模式:应用程序运行在客户机上。当需要对数据库进行操作时,就向数据库服务器发一个请求,数据库服务器收到请求后执行相应的数据库操作,并将结果返回给客户机上的应用程序。



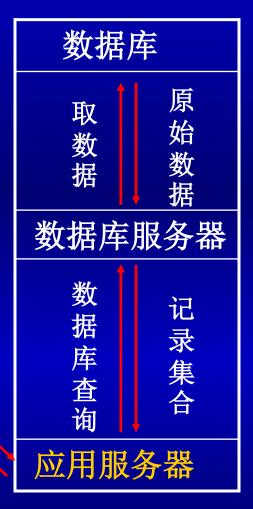
② 三层C/S结构(B/S结构)

Web浏览器

业务请求:查 找8G TF卡

客户机

HTML格式的结 果页面







客户端:

借助Web浏览器处理简单的客户端处理请求,显示用户界面及服务器端运行结果。

Web服务器:

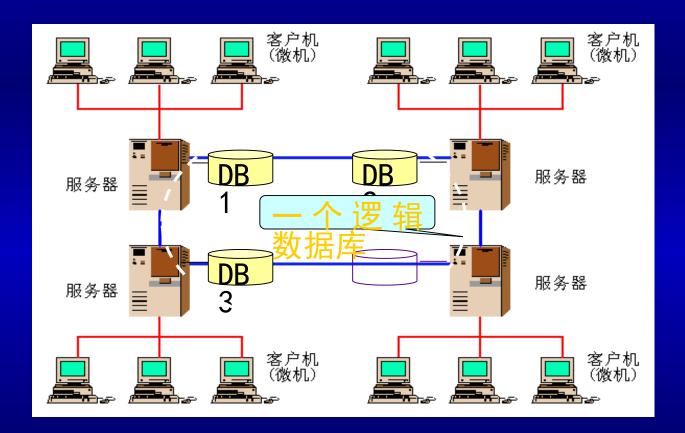
负责接收远程或本地的数据查询请求,然后运行服务器脚本,借助中间件把查询发送到数据库服务器上以获取相关数据,再把结果数据传回客户的Browser。

数据库服务器:

负责管理数据库,处理数据更新及完成查询要求, 运行存储过程。

(4) 分布式系统 (DDBS—Distributed Data Base System)

数据库中的数据在逻辑上是一个整体,但物理地分布在计算机网络的不同节点上。





DDBS的特点:

- ① 分布性:数据库中的数据是分布存储在多个场地上。
- ② 逻辑整体性: 数据在逻辑上是相互联系的整体, 支持全局应用。
- ③ 自治性:每个场地是独立的DBS,有自己的数据库、用户和CPU,运行自己的DBMS,执行局部应用,具有高度的自治性。
 - 4 协作性:各个场地的DBS又相互协作组成一个整体

优点: 更大范围内的数据共享

缺点:数据的分布存放,给数据的处理、管理与维护

带来困难;系统的效率受网络通信的制约。

60



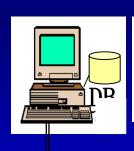
(5)并行结构(详见教材P240)

运行在并行机上的具有并行处理能力的数据库系统

优点: 并行技术能提高系统效率

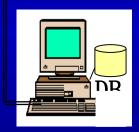
缺点:价格高、实现技术复杂



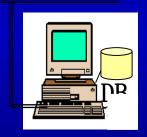












云结构的DBS是 把数据库在云环境下 ,以服务的形式提供 数据库的功能。现在 的云数据库是运行在 机群上的具有并行处 理能力的DBS



本讲主要内容

- 一、数据管理技术的发展
- 二、数据库系统的组成
- 三. 数据库系统的特点
- 四、数据库的历史
- 五、数据库系统的结构
- 六. 数据库管理系统产品





六、数据库管理系统产品

- > ORACLE
- > DB2
- > SQL Server
- > INGRES
- > SYBASE
- > Informix
- Foxpro

- > MySQL
- PostgreSQL
- > SQLite





六、数据库管理系统产品

- 》 武汉达梦
- 〉 人大金仓
- > 天津神州通用
- > 天津南大通用

- openGauss
- > TiDB
- > OceanBase
- > GaussDB



Questions?





本讲主要目标



学完本讲后, 你应该能够了解:

- 1、数据管理技术的发展历史阶段以及各阶段数据管理技术的优劣;
 - 2、文件系统到数据库系统的演化过程;
 - 3、数据库系统组成包括硬件、数据库、软件与人;
- 4、数据库系统的主要特点有:数据的高度结构化、数据的高共享性、程序与数据的高独立性、数据控制力度大;
- 5、数据库管理系统的控制功能包括:安全性保护、 完整性、并发与数据库恢复;
- 6、数据库系统的结构:内部三级模式、外部体系结构;
 - 7、概念模型和数据模型的概念和作用;
 - 8、目前常见的数据库管理系统产品。



问题讨论

- 1、什么是计算机文件系统?它存在哪些问题?
- 2、DBMS的控制功能有哪些?说明它们的重要性
- 3、为什么数据库设计很重要?
- 4、数据库系统中有哪些人员?他们的职责分别是什么?
- 5、针对一个目前用手工方式或 文件系统方式处理数据的 具体应用环境,设想用 数据库技术处理数据的 情况下,系统将如何?





练习

- 1、什么是数据冗余?数据库系统与文件系统相比怎样减少冗余?
- 2、数据库系统的三级模式描述了什么?为什么在 三级模式之间提供两级映射?
- 3、数据库系统(DBS)由哪几个部分组成?其特点是什么?
 - 4、什么是DBMS? DBMS包括哪些功能?

