

数据库系统原理

教程：数据库系统概论（第5版）

结合：CMU 15-445/645 INTRO TO DATABASE SYSTEMS

华中科技大学 计算机学院

左琼

第一章 绪论

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统的三级模式结构

1.4 数据库系统的组成

1.5 数据库系统的体系结构

1.3 数据库系统的三级模式结构

□ 数据库中数据如何存储？

- 数据库不仅存放数据的值，也需存放数据的语义。
- 数据的语义是通过存储数据的结构来实现的，库中数据有值与型两部分，分别存放在数据文件和字典中。

□ 数据模型中 “型” 和 “值” 的区别：

- 型(Type)：对某一类数据的结构和属性的说明。
- 值(Value)：是型的一个具体赋值。
- 例：

□ 学生记录型：（学号，姓名，性别，系别，年龄，籍贯）

□ 一个记录值：（900201，李明，男，计算机，22，江苏）

1.3.1 数据库系统模式的概念

- 从数据库应用开发人员角度看，数据库系统通常采用三级模式结构，是数据库系统内部的系统结构。
- 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。
 - 仅仅是型的描述，与具体的值无关。
 - 反映的是数据的结构及其联系。
 - 模式是相对稳定的。
- 实例 (Instance)
 - 模式的一个具体值；反映数据库某一时刻的状态；
 - 同一个模式可以有多个实例；
 - 实例随数据库中数据的更新而变动。

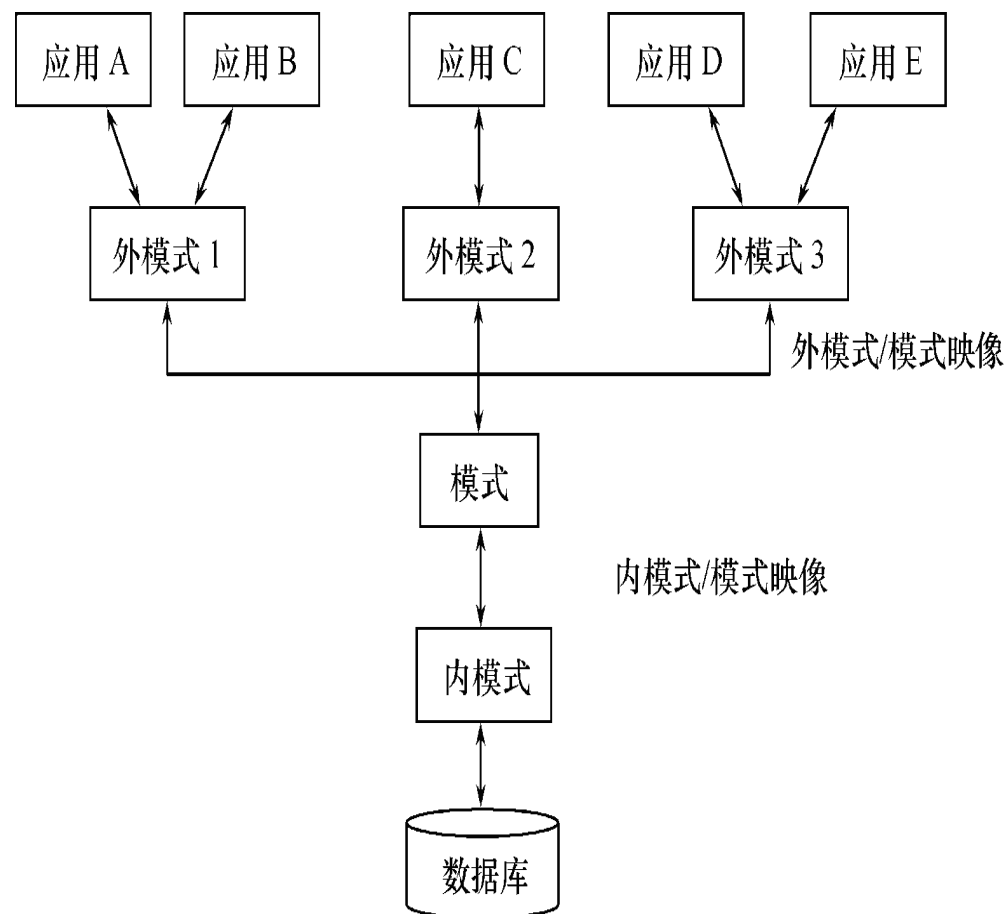
1.3.2数据库系统的三级模式

□ 数据库系统三级模式结构

- CODASYL (Conference On Data System Language) 提出**模式、外模式、内模式三级模式**的概念，三级模式之间有**两级映象**。

□ 模式的分级

- 从数据库用户观点，即用户看到的数据库，与数据库的物理方面，即实际存储的数据库区分开来，**数据库系统的模式是分级的**。
- 提高了**数据的逻辑独立性和物理独立性**。



1. 模式 (Schema)

□ 模式 (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——全局逻辑视图, 综合了所有用户的需求。
- 一个数据库只有一个模式。
- 模式的地位: 是数据库系统模式结构的中间层, 独立于数据库的其他层次。
 - 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。

1. 模式 (Schema)

- 模式的定义:
 - 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**。
- 模式一般由多个 **“记录”** 组成, 包含数据库的所有信息。
 - 如: 某课程选修信息系统的数据库模式包括下列记录:
 - 学生 (学号, 姓名, 性别, 系别, 年龄)
 - 课程 (课程号, 课程名)
 - 选修 (学号, 课程号, 成绩)
- 设计数据库模式结构时应**首先**确定数据库的逻辑模式。
- 模式的作用是为了支持数据的**少冗余共享**。

2. 外模式(external schema)

- 外模式（也称子模式subschema或用户模式）：
 - 数据库用户（包括应用程序员和最终用户）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，面向具体的应用程序，是与某一应用有关的数据的逻辑表示；
 - 定义在逻辑模式之上；
 - 独立于存储模式和存储设备；
 - 当应用需求发生较大变化，相应外模式不能满足其视图要求时，该外模式就得做相应改动；
 - 设计外模式时应充分考虑到应用的扩充性。

2. 外模式

□ 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。

1.3.2 数据库的三级模式结构

外模式

劳资科

学号	姓名	系别	补贴
----	----	----	----

房产科

学号	姓名	性别	系别	住址
----	----	----	----	----

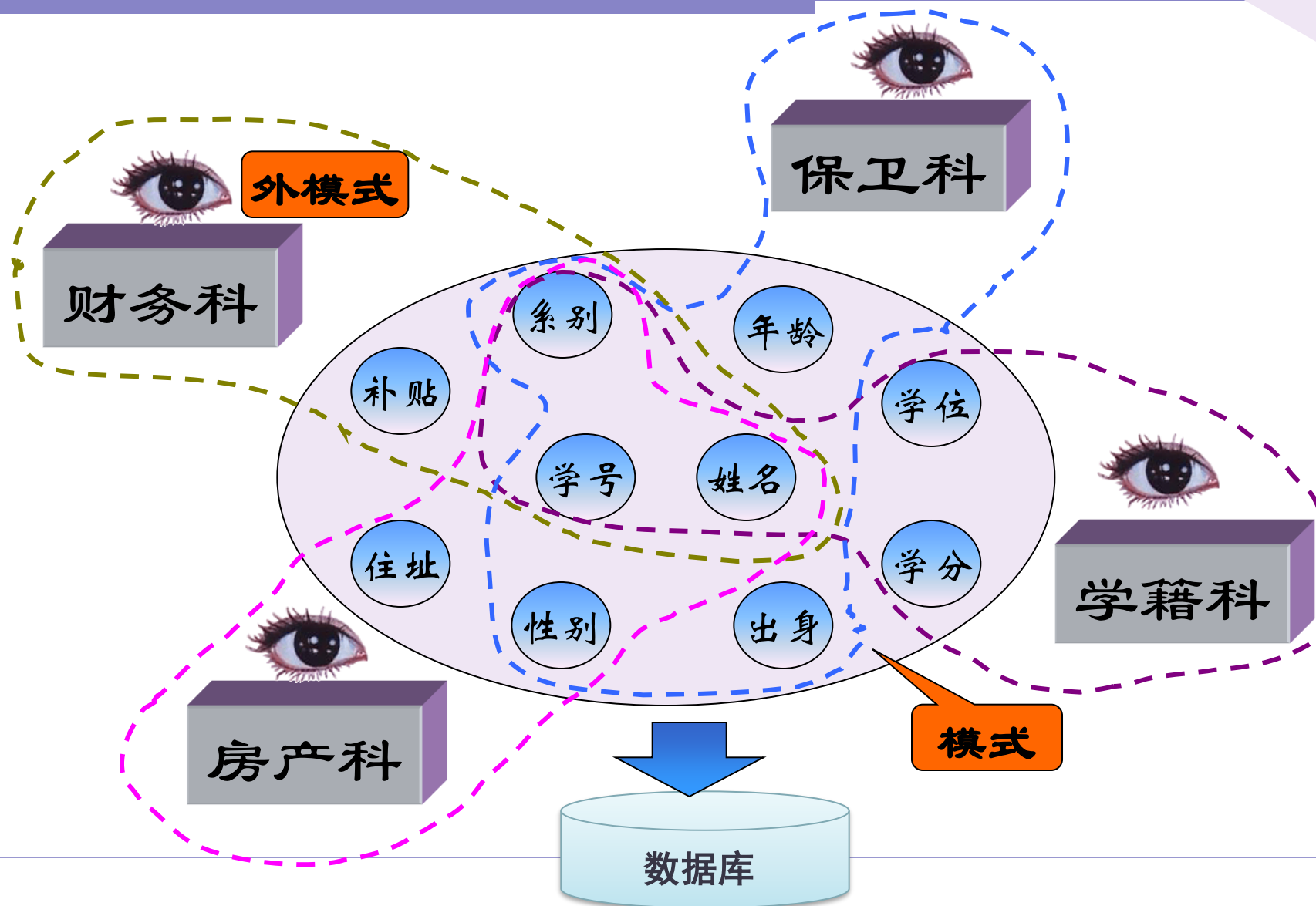
学籍科

学号	姓名	系别	学分	学位
----	----	----	----	----

人事科

学号	姓名	性别	系别	年龄	学位	出身
----	----	----	----	----	----	----

模式 vs. 外模式



模式 vs. 外模式

□ 又如，数据库模式如下：

学生 (学号, 姓名, 性别, 系别, 年龄)

课程 (课程号, 课程名)

选修 (学号, 课程号, 成绩)

■ 可定义子模式如下：

① **单关系子模式**： student_1 (学号, 姓名, 系别)

② **多关系子模式**： SC(学号, 姓名, 课程号, 课程名, 成绩)

□ 外模式的用途：

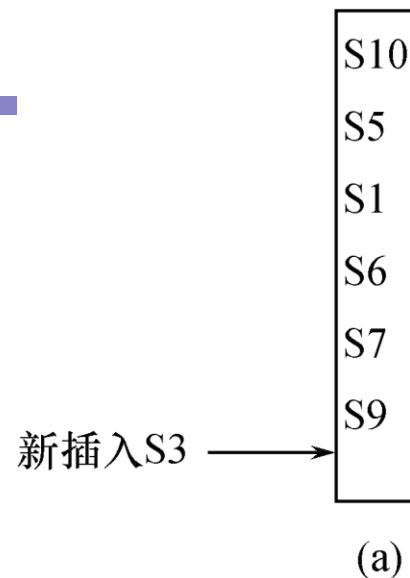
- 1) 支持不同用户建立适应局部应用特征的结构，每个用户只能看见和访问所对应的外模式中的数据；
- 2) 简化应用处理；
- 3) 提高安全性，保证数据库安全性的一个有力措施。

3. 内模式 (Internal schema)

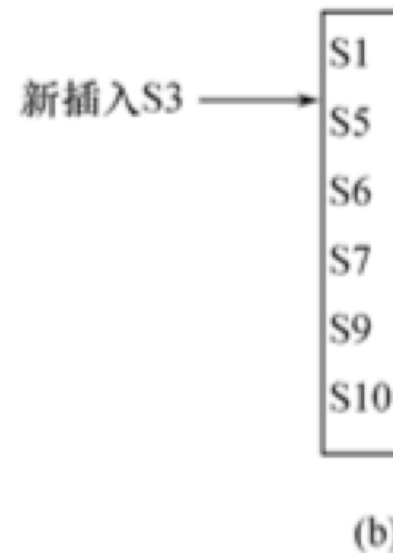
- 内模式 (也称存储模式storage schema或物理模式) :
 - 是数据物理结构和存储方式的描述;
 - 是数据在数据库内部的表示方式, 例:
 - 记录的存储方式 (顺序存储, 按照B树结构存储, 按hash方法存储)
 - 索引的组织方式
 - 数据是否压缩存储
 - 数据是否加密
 - 数据存储记录结构的规定
- 一个数据库只有一个内模式。

3. 内模式

□ 例如：学生记录，如果按**堆**存储，则插入一条新记录总是放在学生记录存储的最后，如图(a)所示。



□ 如果按学号**升序**存储，则插入一条记录就要找到它应在的位置插入，如图(b)所示。



3. 内模式

- 如果按照学生年龄聚簇存放，假如新插入的S3是16岁，则应插入的位置如图(c)所示。

内模式的作用是支持用户建立适应需求（如存取效率、空间效率、数据安全）的物理结构。

新插入S3 →

S1	16
S6	16
S10	16
S5	17
S7	17
S9	18

(c)

图 记录不同的存储方式示意图

内模式依赖于它的全局逻辑结构；
独立于数据库的用户视图，即外模式。

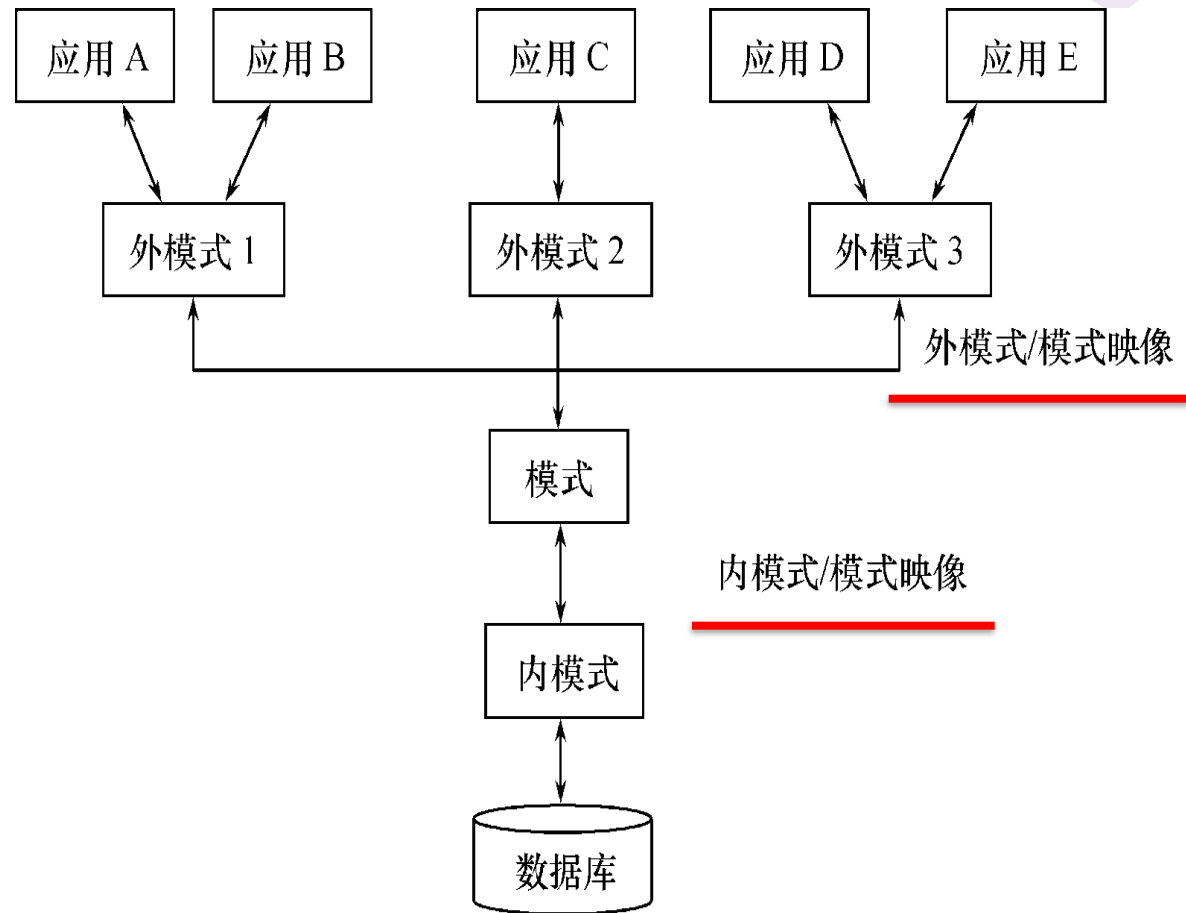
1.3.2 数据库系统的三级模式

□ 三级模式小结：

- 内模式是整个数据库实际存储的表示，反映的是数据的**存储观**；
- 概念模式是整个数据库的抽象表示，反映的是数据的**全局观**；
- 外模式是概念模式的某一部分的抽象表示，反映的是数据的**用户观**。

1.3.3 数据库的二级映像与数据独立性

- 三级模式是对数据的三个抽象级别;
- **二级映象**在DBMS内部实现这三个抽象层次的联系和转换:
 - **外模式 / 模式映像**
 - **模式 / 内模式映像**



1. 外模式 / 模式映象

- 外模式/模式映像定义某一个外模式和模式之间的对应关系。
- 作用：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。

1. 外模式 / 模式映象

student

old模式

NO	NAME	AGE	SEX
301	赵钱	20	男
302	钱	21	女



Create View Stud(学号,姓名,性别,年龄)
As Select NO, NAME, AGE, SEX,
From student

Stud (外模式)

学号	姓名	性别	年龄
301	赵钱	男	20
302	钱	女	21

**映像变化
而外模式
不变**

模式变化

student

new模式

XH	XM	XB	CSRQ
301	赵钱	男	95.07.01
302	钱	女	94.03.07



Create View Stud(学号,姓名,性别,年龄)
As Select XH,XM,XB,
datediff(year,CSRQ,getdate())
From student

2. 模式 / 内模式映像

- 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。
 - 如：说明逻辑记录 and 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。
- 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；
 - 存储结构变化的影响被限制在模式之下，这使数据的存储结构和存储方法独立于应用程序应用程序不受影响。保证了数据与程序的物理独立性，简称数据的物理独立性。

三级模式结构的优点

□ 保证数据的独立性

- 将模式和内模式分开，保证了数据的物理独立性；
- 将外模式和模式分开，保证了数据的逻辑独立性。

□ 简化了用户接口

- 按照外模式编写应用程序或敲入命令，而不需了解数据库内部的存储结构，方便用户使用系统。

□ 有利于数据共享

- 在不同的外模式下可有多个用户共享系统中数据，减少了数据冗余。

□ 利于数据的安全保密

- 在外模式下根据要求进行操作，不能对限定的数据操作，保证了其他数据的安全。

第一章 绪论

- 1.1 数据库系统概述
- 1.2 数据模型
- 1.3 数据库系统的三级模式结构
- 1.4 数据库系统的组成
- 1.5 数据库系统的体系结构

1.4 数据库系统的组成

1. 硬件平台及数据库
2. 软件
3. 人员

1.硬件平台及数据库

■数据库系统对硬件资源的要求：

(1) 足够大的内存

操作系统

DBMS的核心模块

数据缓冲区

应用程序

(2) 足够大的外存

磁盘或磁盘阵列

➤数据库

光盘、磁带

➤数据备份

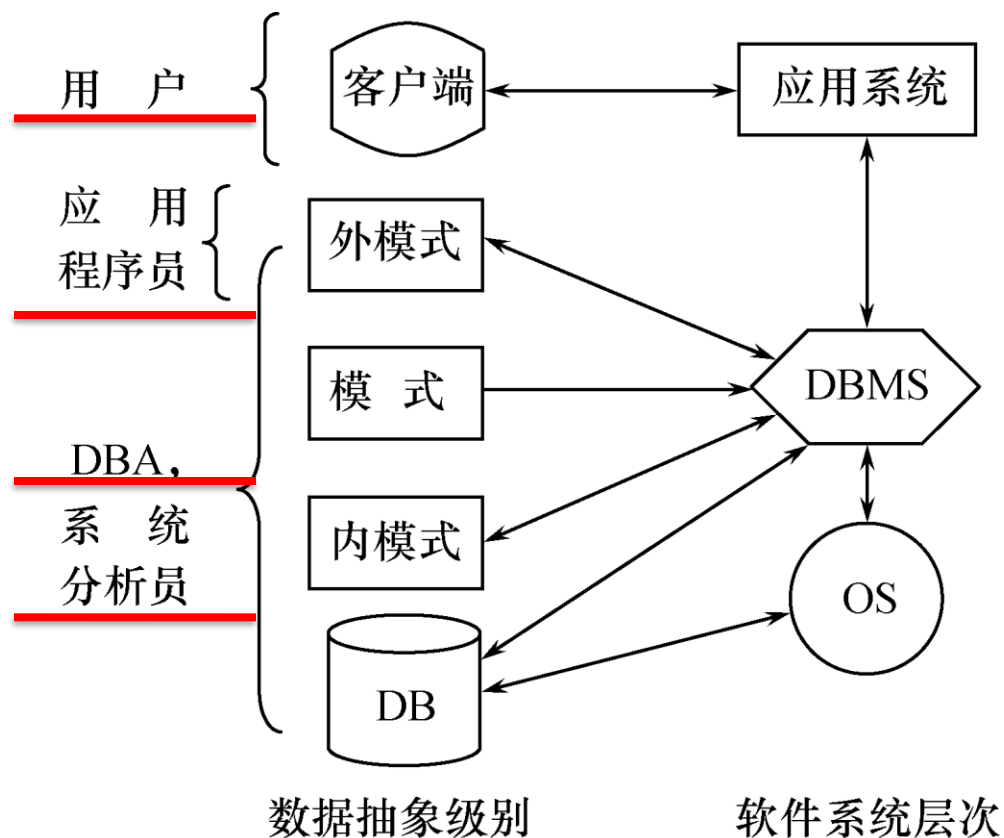
(3) 较高的通道能力，提高数据传送率

2. 软件

- 1) DBMS
- 2) 支持DBMS运行的操作系统
- 3) 与数据库接口的高级语言及其编译系统
- 4) 以DBMS为核心的应用开发工具
- 5) 为特定应用环境开发的数据库应用系统

3. 人员

□ 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，如图。



• 系统分析员

负责应用系统的需求分析和规范说明
与用户及DBA协商，确定系统的硬软件配置
参与数据库系统的概要设计

• 数据库设计人员

参加用户需求调查和系统分析
确定数据库中的数据
设计数据库各级模式

• 应用程序员

设计和编写应用系统的程序模块
进行调试和安装

- **用户**是指最终用户 (End User)
最终用户通过应用系统的用户接口使用数据库

3. 人员:数据库管理员(DBA)

- DBA: 负责全面管理和控制数据库系统。
- 具体职责:
 - (1) 决定数据库中的信息内容和结构——DB设计
 - (2) 决定数据库的存储结构和存取策略——内模式设计
 - (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
 - 周期性转储数据库: 数据文件、日志文件
 - 监视审计文件; 系统故障恢复; 介质故障恢复
 - (5) 数据库的**改进和重组**
 - 性能监控和调优; 定期对数据库进行重组织, 以提高系统的性能
 - 需求增加和改变时, 数据库须需要重构造

补：13.2.1 RDBMS层次结构

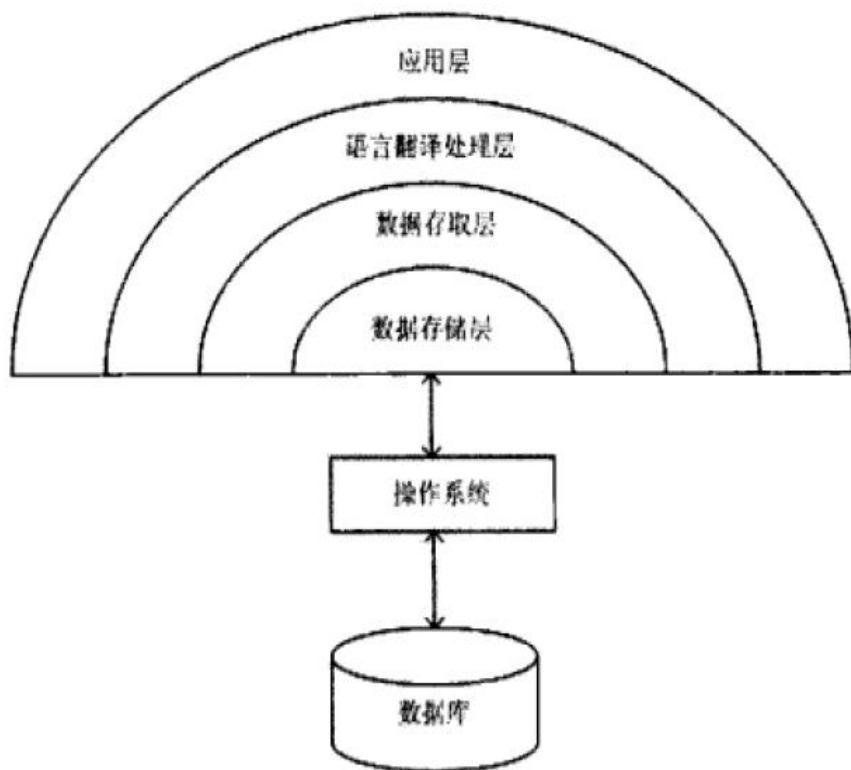


图1 RDBMS层次结构

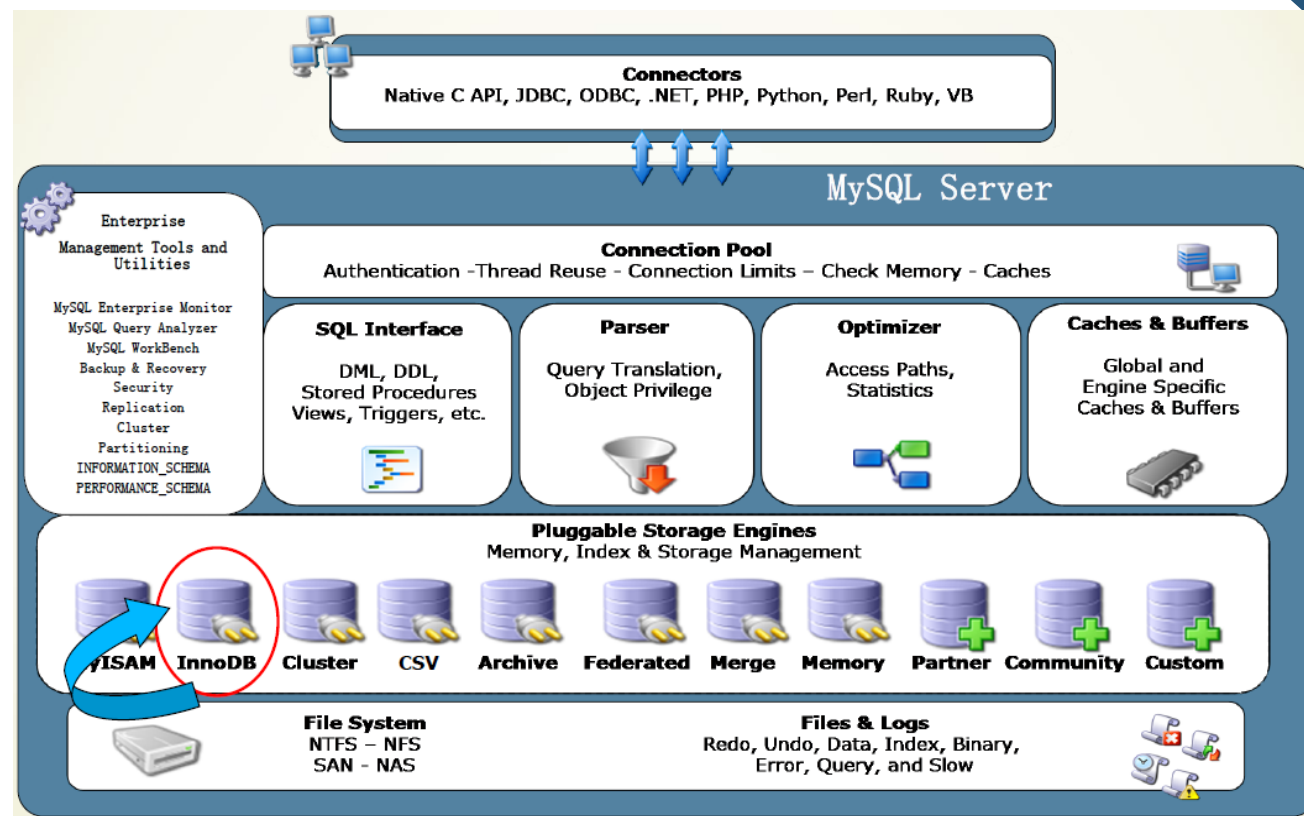
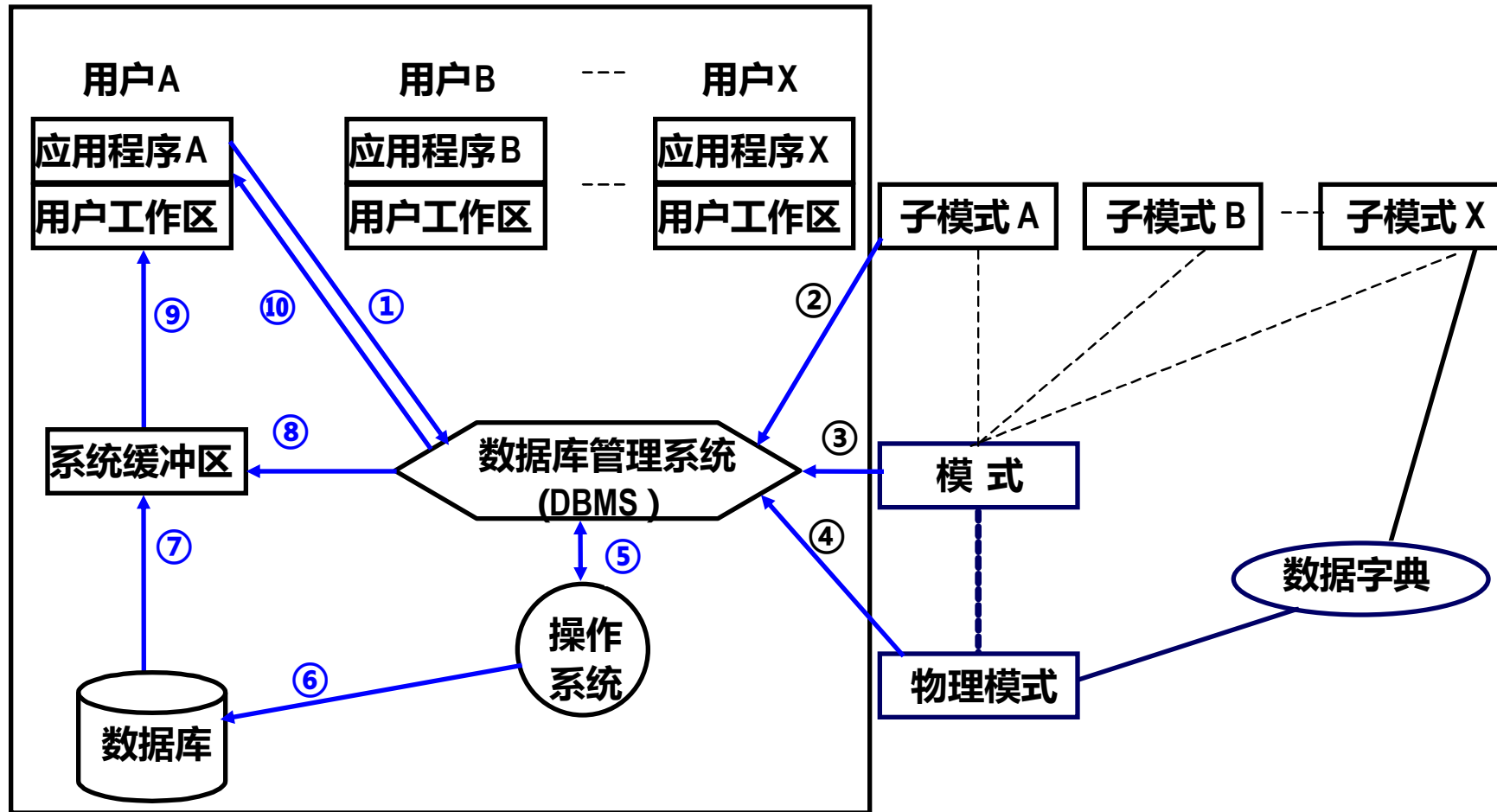


图2 Mysql服务器架构

补：13.2.2 DBMS的工作流程



DBMS工作过程

在数据库系统中，当一个应用程序或用户需要存取数据库中的数据时，应用程序、DBMS、操作系统、硬件等几个方面必须协同工作，共同完成用户的请求。

应用程序从数据库读取一个数据通常需要以下步骤：

1. 应用程序A向DBMS发出从数据库中读数据记录的命令；
2. DBMS对该命令进行语法检查、语义检查，并调用应用程序A对应的子模式，检查A的存取权限，决定是否执行命令，如果拒绝执行，则向用户返回错误信息；
3. 在决定执行该命令后，DBMS调用模式，依据子模式/模式映象的定义，确定应读入模式中的哪些记录；

DBMS工作过程（续）

4. DBMS调用物理模式，依据模式/物理模式映象的定义，决定从哪个文件、用什么存取方式、读入哪个或哪些物理记录；
5. DBMS向操作系统发出执行读取所需物理记录的命令；
6. 操作系统执行读数据的有关操作；
7. 操作系统将数据从数据库的存储区送到系统缓冲区；
8. DBMS依据子模式/模式映象的定义，导出应用程序A所要读取记录的格式；
9. DBMS将数据记录从系统缓冲区传送到应用程序A的用户工作区；
10. DBMS向应用程序返回命令执行情况的状态信息。

1.5* 数据库系统的体系结构

- 集中式数据库系统
- 客户-服务器数据库系统
- 并行数据库系统
- 分布式数据库系统
- 云数据库系统

本章小结

□ 数据库系统概述

- 数据库的基本概念
- 数据管理的发展过程

□ 数据模型

- 数据模型的三要素
- 概念模型, E-R 模型
- 三种主要数据库模型 (了解各种模型的主要特性, 数据库系统的特点)

□ 数据库系统的结构

- 数据库系统三级模式结构
- 数据库系统两层映像系统结构

□ 数据库系统的组成

作业:

P30 3, 15

(提交时间: 布置后一周)

提交方式: 学在华科大)

回顾本章主要内容

- 什么是数据库系统?
- 产生DBMS的动机是什么?
- 如何设计数据库?
- 数据库系统的组成?
- 数据库系统的结构?

数据库系统是由DB、DBMS、应用系统、DBA组成的存储、管理、处理和维护数据的系统。

见P7 表1.1, 要答出数据库系统的特点

数据模型: ER模型, 逻辑模型, 物理模型

硬件平台+DB, 软件, 人员

三级模式, 二级映像

讨论题

□ 数据库、数据库管理系统、数据库应用系统、数据库系统的区别是什么？

存在的问题：

1. “数据库应用系统是一个系统软件？”

错误！

大家学过的系统软件只有：OS，DBMS，语言处理程序（如各种编译器）。应用系统无论大小，只要是面向特定应用的，都属于应用软件。

2. “数据库应用系统属于数据库管理系统？”

错误！12306是一个应用，它能是一个DBMS（如：Mysql）的一部分吗？显然不是。



Chapter 1 ends。 。 。

追求

休息一会儿。 。 。

