

Internet of
Things

数据库基础知识



目录 Contents



相关知识

- | 数据库基础知识

- | 数据结构搭建



任务实践

- | 数据库的安装与使用

- | 数据基本操作



数据库基础知识

数据库基本概念

1

数据(Data)

- 数据是描述现实世界事物的符号记录，是用物理符号记录的可以鉴别信息。包括文字、图形、声音等，他们都是用来描述事物特性的。

2

数据库管理系统(DataBase Management System,简记为DBMS)

- 一种系统软件，用于数据库的建立、使用和维护。如Access、SQL Server、FoxPro、Oracle、Dbase、DB2、MySQL、Sybase等。数据库管理系统是位于用户与操作系统之间的一个数据管理软件，它的基本功能包括以下几个方面：

- 数据定义功能

用户通过数据定义语言（DDL,Data Definition Language）它可以方便地对数据库中的数据对象进行定义。

- 数据操纵功能

用户可以使用数据操纵语言（DML,Data Manipulation Language）操纵数据，实现对数据的基本操作。如查询、插入、删除和修改。

- 数据库的运行管理功能

数据库在建立、运行和维护时由数据库管理系统统一管理和控制，以保证数据的安全性、完整性，对并发操作的控制以及发生故障后的系统恢复等。

- 数据库的建立和维护功能

它包括数据库初始数据的输入、转换功能，数据库的转储、恢复功能，数据库的重组功能和性能监视、分析功能等。

数据库基本概念

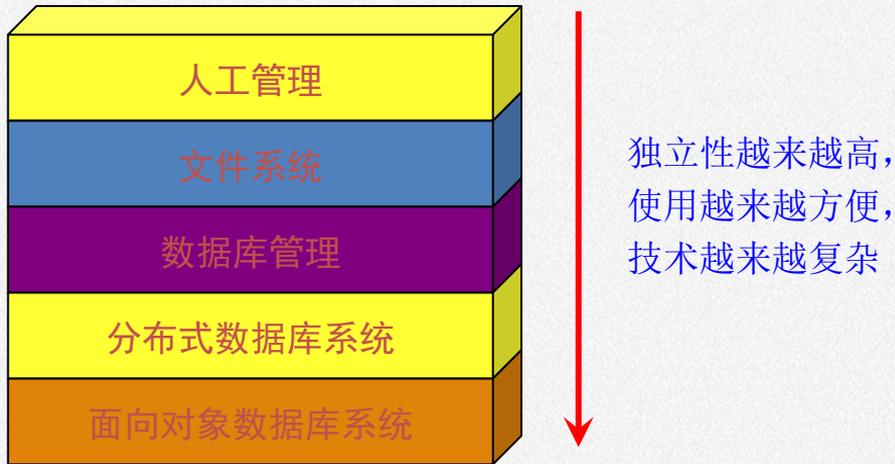
3

数据库(DataBase , 简记为DB)

- 数据库是长期存储在计算机内、有组织的、可共享的数据集合。
- 这种集合具有如下特点：
 - 最小的冗余度
 - 应用程序对数据资源共享
 - 数据独立性高
 - 统一管理和控制

计算机数据管理

- 数据处理中最重要的问题就是数据管理，包括如何对数据分类、组织、编码、存储、检索和维护。随着计算机软硬件的不断升级，数据管理经历了以下几个阶段：



数据库技术的发展

2 人工管理阶段

- ◆ 这一阶段是指20世纪50年代中期以前，计算机主要用于科学计算，当时的计算机硬件状况是：外存只有磁带、卡片、纸带，没有磁盘等直接存取的存储设备；软件状况是：没有操作系统，没有管理数据的软件，数据处理方式是批处理。人工管理阶段的特点是：数据不保存、数据无专门软件进行管理、数据不共享、数据不具有独立性、数据无结构。

2 文件系统阶段

- ◆ 这一阶段从20世纪50年代后期到60年代中期，计算机硬件和软件都有了一定的发展。计算机不仅用于科学计算，还大量用于管理。这时硬件方面已经有了磁盘、磁鼓等直接存取的存储设备。在软件方面，操作系统中已经有了数据管理软件，一般称为文件系统。处理方式上不仅有了文件批处理，而且能够联机实时处理

2

数据库系统阶段

- ◆ 20世纪60年代末数据管理进入新时代 - - 数据库系统阶段。数据库系统阶段出现了统一管理数据的专门软件系统，即数据库管理系统。数据库系统是一种较完善的高级数据管理方式，也是当今数据管理的主要方式，获得了广泛的应用。
- **数据结构化**

在描述数据时不仅描述数据本身，还描述数据之间的联系。
- **数据共享性高、冗余度小、易扩充**

数据可以被多个应用共享。这不仅大大减小了数据的冗余度、节约存储空间、减少存取时间，而且可以避免数据之间的不相容性和不一致性。
- **数据独立性高**

数据独立性包括物理独立性和逻辑独立性。数据的物理独立性是指当数据的物理存储改变时，应用程序不用改变。数据的逻辑独立性是指当数据的逻辑结构改变时，用户应用程序不用改变。
- **统一的数据管理和控制**

数据库对系统中的用户是共享资源。计算机的共享一般是并发的，即多个用户可以同时存取数据库中的数据，甚至可以同时存取数据库中同一个数据。因此，数据库管理系统必须提供以下几个方面数据控制保护功能。

2

分布式数据库系统阶段

- ◆ 分布式数据库系统是由若干个站集合而成。这些站又称为节点，它们在通讯网络中联接在一起，每个节点都是一个独立的数据库系统，它们都拥有各自的数据库、中央处理机、终端，以及各自的局部数据库管理系统。因此分布式数据库系统可以看作是一系列集中式数据库系统的联合。它们在逻辑上属于同一系统，但在物理结构上是分布式的。

2

面向对象数据库系统

- ◆ 面向对象程序语言操纵的是对象，所以面向对象数据库（简称OODB）的一个优势是面向对象语言程序员在做程序时，可直接以对象的形式存储数据

2 现实世界

- ◆ 数据库管理的对象存在于现实世界中，现实世界中的事物存在着各种各样的联系。这种联系是客观存在的，是由事物本身的性质决定的。例如，学校考务管理系统中有老师、学生、课程、教室等构成元素。

2 概念世界

- ◆ 概念世界也叫做信息世界，是现实世界在人们头脑中的反映，是对客观事物及其联系的一种抽象描述。从现实世界到概念世界是通过概念模型来表达的。如对学生的描述可分为：学号、姓名、班级、籍贯、性别等概念。

2 数据世界

- ◆ 存入计算机系统的数据是将概念世界中的事物数据化的结果。为准确地反映事物本身及事物之间的各种联系，数据库中的数据一定存在一个结构，数据模型可用来准确描述这种结构。
- ◆ 在实现数据库管理，也就是三个世界的转化过程中，概念模型和数据模型是实现现实世界数据化的桥梁，是对现实世界中的事物进行抽象的工具。

数据模型

- **模型概念**

模型（model）：一般而言，模型是现实世界某些特征的模拟和抽象，分为实物模型与抽象模型。建筑模型，汽车模型，飞机模型等都是实物模型，它们通常是客观事物的某些外观特征或者功能的模拟与刻画；数学模型 $s=\pi r^2$ 是一种抽象模型，它抽象描述了圆的面积和圆的半径之间的数量关系，揭示客观事物的某些本质的、内部的特征。

- **数据模型**

在实现数据库管理的过程中，数据模型起着关键作用。整个数据库技术的发展就是沿着数据模型的主线展开的。现有的数据库均是基于某种数据模型，了解数据模型的基本概念是学习数据库的基础。根据模型应用的不同目的，可以将这些模型划分为两类，他们分属于两个不同的层次。

第一类模型是概念模型，也称信息模型。它是按用户的观点来对数据和信息建模，主要用于数据库设计。

另一类模型是数据模型，主要包括网状模型、层次模型、关系模型等，它是按计算机系统的观点对数据建模，主要用于DBMS的实现，这种数据模型也叫结构数据模型。

1.2 数据库管理的实现

数据模型的要素:

- **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；一类是与数据之间联系有关的对象。

- **数据操作**

数据操作是指对数据库中各种对象的（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。

- **数据的约束条件**

数据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。如：入学年龄不超过30岁，成绩不得有三门以上不及格等。

1.2 数据库管理的实现

1.2.3 概念模型

概念模型的相关概念

- (1) **实体 (Entity)**：客观存在并相互区别的事物及其事物之间的联系。例如，一个学生、一门课程、学生的一次选课、一次考试等都是实体。
- (2) **属性 (Attribute)**：实体所具有的某一特性。例如，学生的学号、姓名、性别、出生年份、系、入学时间等。
- (3) **码 (Key)**：唯一标示实体的属性集。例如，学号是学生实体的码。
- (4) **域 (Domain)**：属性的取值范围。例如，年龄的域为15至35之间。
- (5) **实体型 (Entity Type)**：用实体名及其属性名集合来抽象和刻画同类实体，称为实体型。例如，学生(学号，姓名，性别，出生年份，系，入学时间)就是一个实体型。
- (6) **实体集 (Entity Set)**：同型实体的集合称为实体集。如，全体学生就是一个实体集。
- (7) **联系 (Relationship)**：实体与实体之间以及实体与组成它的各属性间的关系。

1.2 数据库管理的实现

实体间联系的三种情况

(1) 一对一联系 (1:1)

如果对于实体集A中的每一个实体，实体集B中至少有一个（也可以没有）实体与之联系，反之亦然，则称实体集A与实体集B具有一对一联系，记为1:1。例如，一个学生只能有一个学号，而一个学号只能指向一个学生，则学生与学号之间具有一对一联系。

(2) 一对多联系 (1:n)

如果对于实体集A中的每一个实体，实体集B中有n个实体 ($n \geq 0$) 与之联系，反之，对于实体集B中的每一个实体，实体集A中至多只有一个实体与之联系，则称实体集A与实体集B有一对多联系，记为1:n。例如，一个班级中有若干名学生，而每个学生只在一个班级中学习，则班级与学生之间具有一对多联系。

(3) 多对多联系 (m:n)

如果对于实体集A中的每一个实体，实体集B中有n个实体 ($n \geq 0$) 与之联系，反之，对于实体集B中的每一个实体，实体集A中也有m个实体 ($m \geq 0$) 与之联系，则称实体集A与实体集B具有多对多联系，记为m:n。例如，一门课程同时有若干个学生选修，而一个学生可以同时选修多门课程，则课程与学生之间具有多对多联系。

注意：实际上，一对一联系是一对多联系的特例，而一对多联系又是多对多联系的特例。

1.2 数据库管理的实现

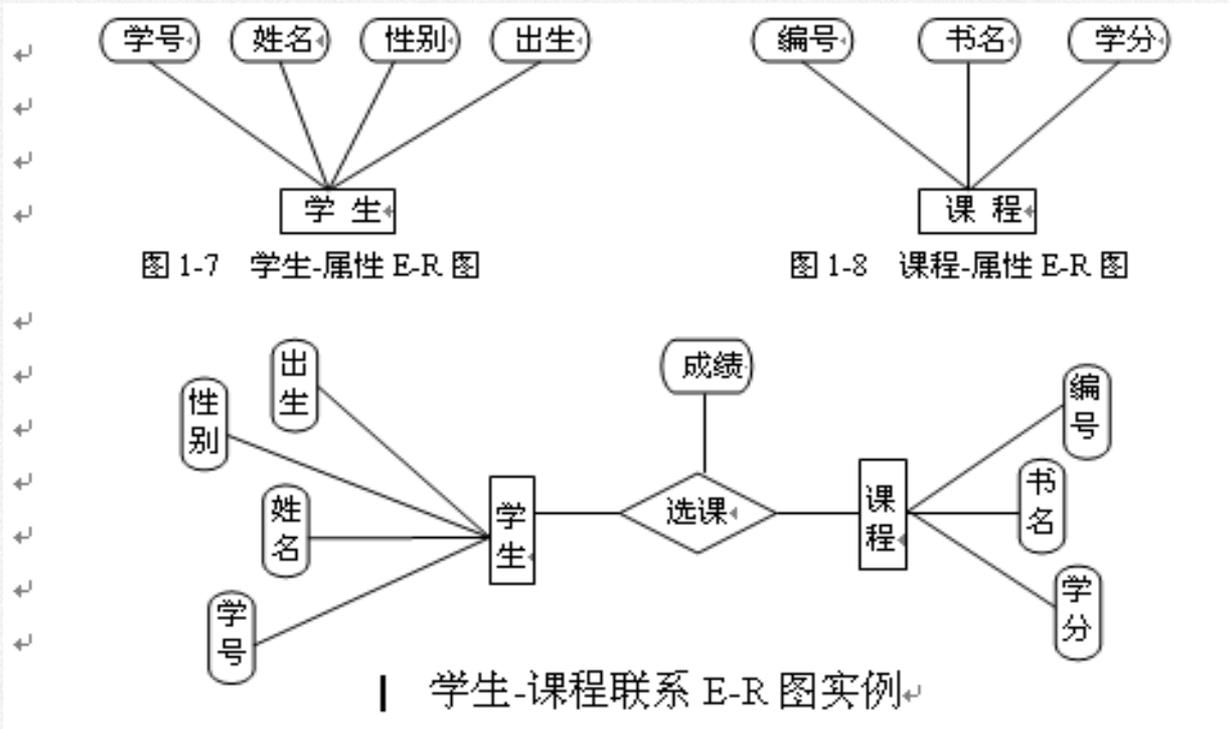
- 概念模型的表示方法

概念模型的表示方法很多，最常用的是实体—联系方法。该方法用E-R(Entity-Relationship Approach) 图来描述现实世界的概念模型。

E-R图提供了表示实体型、属性和联系的方法。E-R图有三个要素：

- (1) 实体型：用矩形表示，矩形框内写明实体名。
- (2) 属性：用椭圆形表示，并用无向边将其与相应的实体连接起来。
- (3) 联系：用菱形表示，菱形框内写明联系名，并用无向边分别与有关实体连接起来，同时在无向边旁标上联系的类型(1: 1, 1: n或m: n)。

1.2 数据库管理的实现



1.2 数据库管理的实现

1.2.4 数据模型

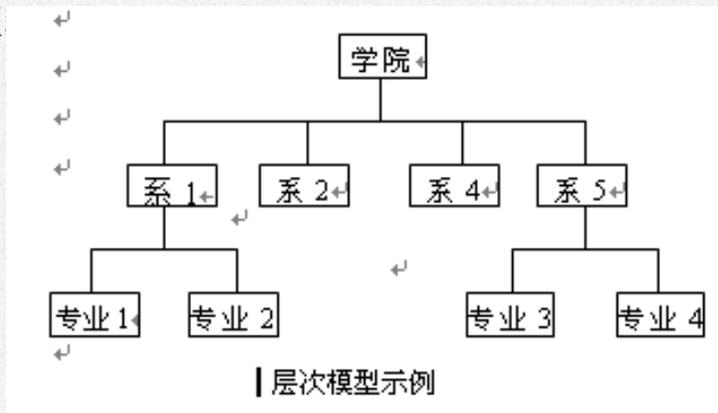
这里讲的数据模型是在概念级的模型（概念模型）的基础上形成的，为计算机上某一DBMS支持的数据模型。本书主要介绍前三种模型。

- **层次模型**

若用图来表示，层次模型是一棵倒立的树。在数据库中，满足以下条件的

数据模型
点： ②

，这个结点称为根结点



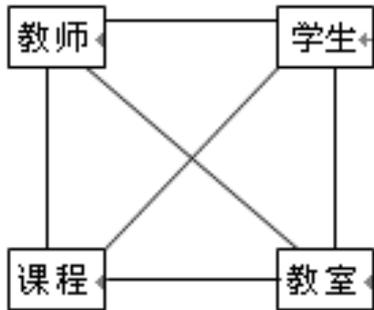
1.2 数据库管理的实现

- 网状模型

在现实世界中，事物之间的联系更多的是非层次关系的，用层次模型表示非树型结构是很不直接的，网状模型则可以克服这一弊病。网状模型是一个网络。在数据库中，满足以下两个条件的数据模型称为网状模型。① 允许一个以上的结点无父结点；② 一个结点可以有多个父结点。

从以上定义看出，网状模型构成了比层次结构复杂的网状结构，适宜表示多

上的“超链接数据”模式有类似



网状模型示例

1.2 数据库管理的实现

- **关系模型**

以二维表的形式表示实体和实体之间联系的数据模型称为关系数据模型。从模型的三

要素角度看，关系模型的内容为：（1）数据结构：一张二维表格。（2）数据操作：数据的定义，检索、维护等。（3）数据约束条件：表中列的取值范围即域值的限制条件。

关系模型的几个常见概念：

- （1）关系：一个关系就是一张二维表，每个关系都有一个关系名，即数据表名。
- （2）元组：表中的行称为元组，一行就是一个元组，对应表中一条记录。
- （3）属性：表中的列称为属性，即字段。字段名称为属性名，字段值称为属性值。
- （4）域：属性的取值范围，如，分数在0~100之间。
- （5）关键字：表中的一个属性（组），它的值可以唯一地标志一个元组。如：学号。
- （6）候选码：表中的某一个属性，它的值可以唯一地标志一个元组。一个表中可能有

多个候选码，选择一个作为主键，主键的属性称为主属性。

- （7）外关键字：如果一个关系中的属性或属性组并非该关系的关键字，但它们是另

1.2 数据库管理的实现



1.2 数据库管理的实现

- **关系模型的完整性**

关系模型对**数据**一般都具有一定的限制，这种限制称为完整性或完整性约束。

关系模型的完整性是保证关系数据表正确的关键。关系模型支持实体完整性约束、参照完整性约束和域约束3种完整性约束。

- (1) 实体完整性约束**

假设A是一个表R的主键，则A不能接收空值，即单列主键的值不能为空，复合主键的任何列也不能接收空值。例如，在学生信息表中，“学号”为该表的主键，那么在数据库的任何记录中，“学号”列的值都不能为空。这样的约束称为实体完整性约束。

1.2 数据库管理的实现

(2) 参照完整性约束

参照完整性约束关心的是逻辑相关的表中值与值之间的关系。假设X是一个表A的主键，在表B中是外键，那么若K是表B中一个外部键值，则表A中必然存在在X上的值为K的记录。例如，“系编码”是院系信息表的主键，而在学生信息表中是相对于

院系

信息表的外键（学生信息表中的主关键字是由“学号”和“系编码”组合而成），对于学生

信息表的任何记录，其所包含的“系编码”的值，在院系信息表的“系编码”列中必

一个相同

学号	系编码	姓名	性别
+ 20051201	0201	王意宝	女
+ 20051202	0201	李树财	男
+ 20051203	0101	李方浩	女
+ 20051204	0102	张敬秋	女
+ 20051205	0202	田文涛	男

系编码	系名
0101	信息工程系
0102	机械工程系
0201	外语系
0202	汽车工程系
0103	土木工程系

参照完整性约束

1.2 数据库管理的实现

(3) 域约束

域是逻辑相关的值的集合，从域中可以得出特定列的值。

例如，在学生信息表中：

- [出生日期]域的值必须按照特定的统一格式存放，而不能有时用：1986.12.23格式，有时用：12/23/1986格式，造成数据混乱；
- [学生名字]、[院系名称]等域的值必须属于字符集合；
- 对于[性别]，该域中的值必须局限于男、女等。

1.2 数据库管理的实现

1.2.5 关系运算

关系的基本运算有两类：传统的集合运算和专门的关系运算。

- **传统的集合运算**

(1) 并 (Union)：设有两个关系R和S，它们具有相同的结构。R和S的并是由属于R
或

属于S的元组组成的集合，运算符为 \cup 。例如：设每个班有一个学生关系R1, R2,
R3, ...则全校学生关系T是： $T = R1 \cup R2 \cup R3 \cup \dots$

(2) 差 (Difference)：设有两个关系R和S，它们具有相同的结构。R和S的差是由属于
R

但不属于S的元组组成的集合。运算符为 $-$ 。例如：设有参加计算机小组的学生关系
R，参加桥牌小组的学生关系S。求参加了计算机小组但没有参加桥牌小组的学生。
结果放在关系T1中，则 $T1 = R - S$ 。

1.2 数据库管理的实现

(3) 交 (Intersection)：设有两个关系R和S，它们具有相同的结构。R和S的交是由既属于R又属于S的元组组成的集合，运算符为 \cap 。交运算的结果是R和S的共同元组。

例如：设有参加计算机小组的学生关系 R，参加桥牌小组的学生关系 S。求既参加了计算机

小组又参加桥牌小组的学生。结果放在关系 T2中，则 $T2 = R \cap S$ 。关系的交可以用关系的差来表示。即 $R \cap S = R - (R - S)$ ，所以，上式也可以写成： $T2 = R - (R - S)$ 。

1.2 数据库管理的实现

- 专门的关系运算

(1) 选择 (Selection) : 按照给定条件从指定的关系中挑选出满足条件的元组构成新的关系, 其关系模式不变, 但其中元组的数目小于等于原来的关系中元组的个数, 它是原关系的一个子集。这是从行的角度进行的运算, 即水平方向抽取元组

(2) 投影 (projection) : 从指定的关系中挑选出某些属性构成新的关系, 其关系模式所包含的属性个数往往比原关系少, 或者属性的排列顺序不同。投影的结果将取消由于取消了某些列而产生的重复元组。

例如: 设有读者关系, 读者 (借书证号, 姓名, 性别, 单位, 职称, 地址)。

按照单位在前, 姓名在后的次序列出读者名单。此运算结果包括元组数目不变, 但减少了列的数目, 同时改变了列的显示顺序。

1.2 数据库管理的实现

(3) 联接 (join)：联接是将两个和多个关系模式通过公共的属性名拼接成一个更宽的关系模式，生成的新关系包含满足联接条件的元组。例如：设有三个关系：学生（学号，姓名，所在系，性别，现住址）；课程（课程号，课程名，学分）；选修（学号，课程号，成绩）。若想查询成绩90分以上的学生姓名。联接结果包括学生、课程、选修三个关系中属性的并集。自然联接：是去掉重复属性的等值连接。它属于联接运算的一个特例

1.3 关系数据库设计

1.3.1 数据库设计方法

- 手工试凑法

数据库的设计水平和与设计人员的经验有直接关系。数据库设计只是一种经验的反复实施，而不能称为是一门科学，缺乏科学分析理论基础和工程手段的支持，所以设计质量很难保证。

- 规范设计法

主要是将设计的步骤分为：需求分析、概念设计、逻辑设计和物理设计等几个步骤，并采用了许多规范化的手段和工具完成每个阶段的任务。比如基于E-R模型的数据库设计方法，基于3NF（第三范式）的设计方法，基于抽象语法规则的设计方法等，就是在数据库设计的各个过程中采用的具体的技术与方法。

注意：规范设计法仍旧是一种手工方法。对于大型的项目，规范化是必须遵循的设计思想。

1.3 关系数据库设计

1.3.2 数据库设计步骤

按照规范化的设计方法，数据库设计步骤分为六个阶段，这个过程也叫数据库的生命周期。

需求分析：这个阶段的工作是要充分调查研究，了解用户需求

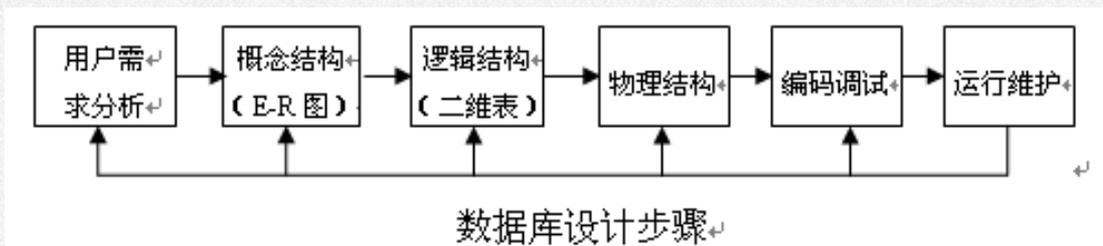
概念结构设计：概念结构是整个系统的信息结构。它是现实世界的真实反映，包括实体与实体之间的关系。

逻辑结构设计：这个阶段的任务是将概念结构转换成与所选用的DBMS所支持的数据模型相符合的过程。

物理结构设计：这个阶段的任务是为一个给定的逻辑数据模型选取一个合适的物理结构，并对物理结构进行评价。

1.3 关系数据库设计

- **物理结构设计**：这个阶段的任务是为一个给定的逻辑数据模型选取一个合适的物理结构，并对物理结构进行评价。
- **编码测试**：进入这个阶段后，就要按照逻辑设计和物理设计的结果利用DBMS的数据定义语言把数据库描述出来这个阶段
- **运行维护**：试运行的主要工作是检查应用程序的功能，测量系统的性能指标，在物理设计阶段所做的评估是否正确，此时可以得到检验。



1.3 关系数据库设计

1.3.3 数据库设计规范

- **范式的概念**：规范化是数据库设计中的一个重要过程，可以通过它来剔除数据库中冗余的数据。EF - Codd在1971年提出规范化理论，他和后来的研究人员为数据库结构定义了五种规范化模式，简称范式。规范化的优点：
 - ◆大大减少了数据冗余
 - ◆改进了数据库整体组织
 - ◆增强了数据的一致性
 - ◆增加了数据库设计的灵活性

1.3 关系数据库设计

- 第一范式

第一范式规定表的每个列的值都是不可分的简单数据项。在任何一个数据库中，第一范式都是一个最基本的要求。

不符合第一范式的表

学号	姓名	性别	联系电话	
			宿舍电话	家庭电话
20051208	刘庆	男	80687614	85987328
20051209	崔清波	男	80687615	85987329
20051210	张文强	男	80687616	85987326
20051211	李易	女	80687617	85987325
20051212	吕龙	男	80687618	85987327

修改后的表

学号	姓名	性别	宿舍电话	家庭电话
20051208	刘庆	男	80687614	85987328
20051209	崔清波	男	80687615	85987329
20051210	张文强	男	80687616	85987326
20051211	李易	女	80687617	85987325
20051212	吕龙	男	80687618	85987327

1.3 关系数据库设计

- 第二范式

第二范式有两项要求：

- (1) 所有表必须符合第一范式；
- (2) 表中每一个非主键列都必须完全函数依赖于主键（关键字）。

表的不完全函数依赖关系将导致很多问题的发生。

- ◆数据冗余：如果外语系有700个学生，就得重复700次相同系名。
- ◆更新异常：如果“外语系”升格成了“外语学院”，那么所有原外语系学生的元组都要
- ◆更新，很难保证修改不出错误。
- ◆插入异常：如果增加了一个新系，但该系还没有学生，那么这个系将不能入数据库。

不符合第二范式的表

学号	姓名	性别	系编码	所在系
20051208	刘庆	男	08	外语系
20051209	崔清波	男	08	外语系
20051210	张文强	男	01	汽车工程系
20051211	李易	女	01	土木工程系
20051212	吕龙	男	03	汽车工程系

1.3 关系数据库设计

- **第三范式**
- 第三范式有两项要求：
 - (1) 所有表必须符合第二范式。
 - (2) 表中每一个非主键列对主键都不存在传递依赖，而应是直接依赖。
- 传递依赖的存在与违背第二范式一样，也将会导致添加、删除、更新、冗余等问题。
- 表中出现上述问题的原因是对主键依赖的传递。解决的方法同样是在遵守第二范式的条件下，将主键拆成多个主键，从而消除传递依赖的情况。

不符合第三范式的表

学号	姓名	性别	系编码	所在系	系总人数
20051208	刘庆	男	08	外语系	850人
20051209	崔清波	男	08	外语系	850人
20051210	张文强	男	01	汽车工程系	2000人
20051211	李易	女	01	土木工程系	2500人

1.3 关系数据库设计

- **第四范式**
- 第四范式禁止了主键列和非主键列之间的一对多关系。
- 在表中，一个学生可以有多种喜欢的运动，也可以有多类喜欢的图书。这样在同一表中具有这样两个不相关的列，与主键存在多对一关系，使得表中存在冗余现象和不对称情况，因为不知道每个人的喜好情况，也就无法知道哪列的值多，哪列的值少，使得表格中出现了很多不规则的空白。
- 解决的方法同样是在遵守无损分解的条件下，将表分解成多个表，从而消除传递依赖的情况。

不符合第四范式的表

学号	运动特长	喜爱图书
20051208	篮球	小说
20051208	田径	
20051209	排球	散文
20051210		诗歌
20051210	摔跤	杂文
20051211	游泳	游记



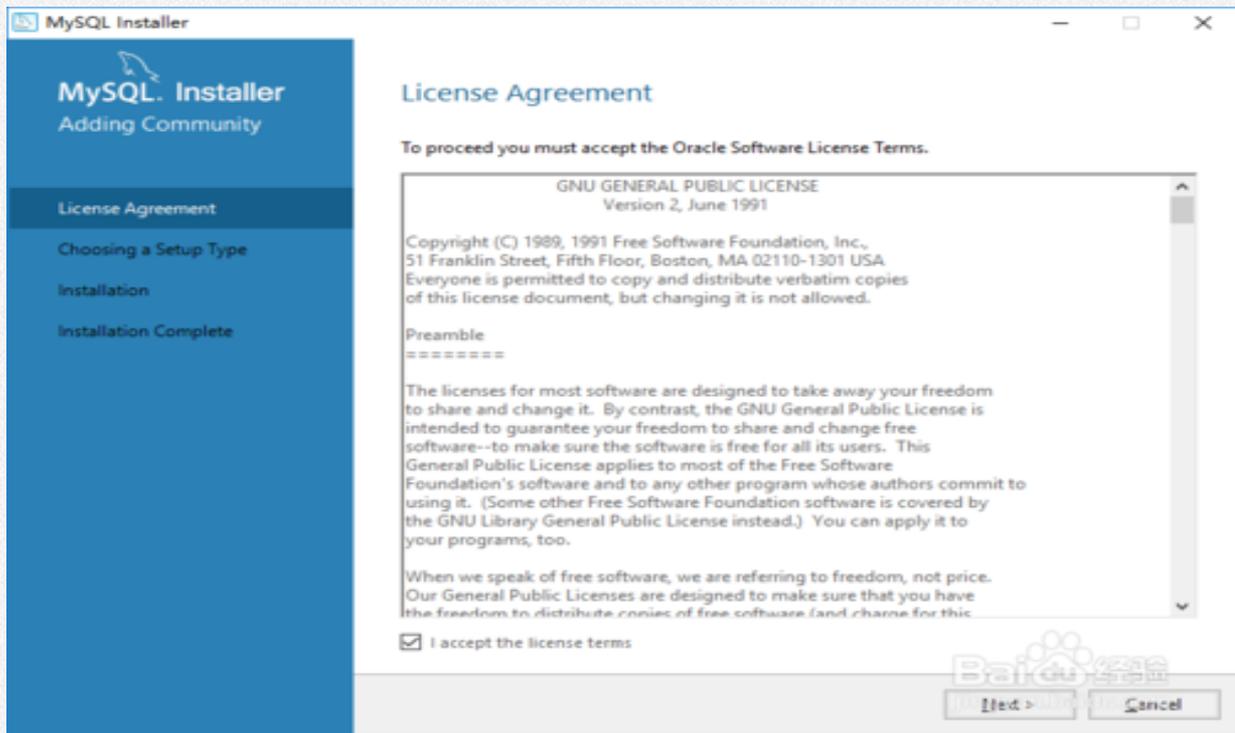
任务实践

——数据库的安装与使用

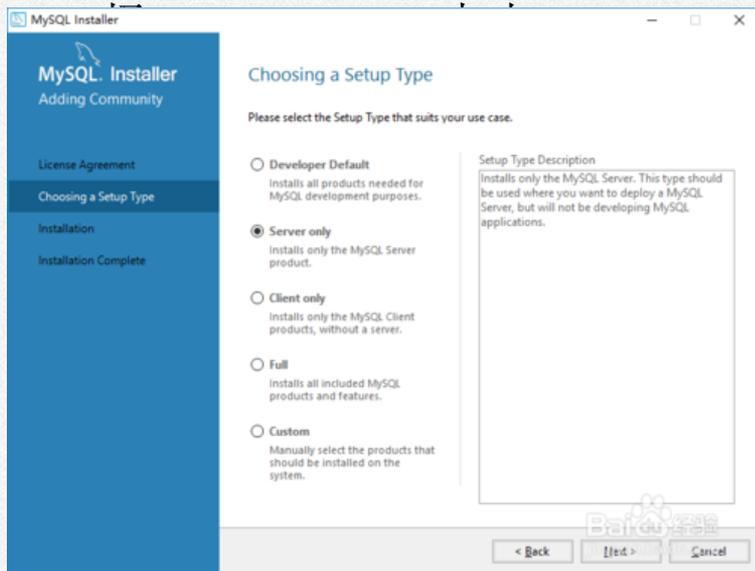
MySQL是一个关系型数据库管理系统，由瑞典**MySQL AB**公司开发，目前属于**Oracle**旗下公司。**MySQL**最流行的关系型数据库管理系统，在**WEB**应用方面**MySQL**是最好的**RDBMS (Relational Database Management System, 关系数据库管理系统)**应用软件之一。体积小、速度快、成本低、开放源代码等特点。

- 在MySQL官网下载社区版mysql-installer-community-5.6.31.0，此版本为免费的。

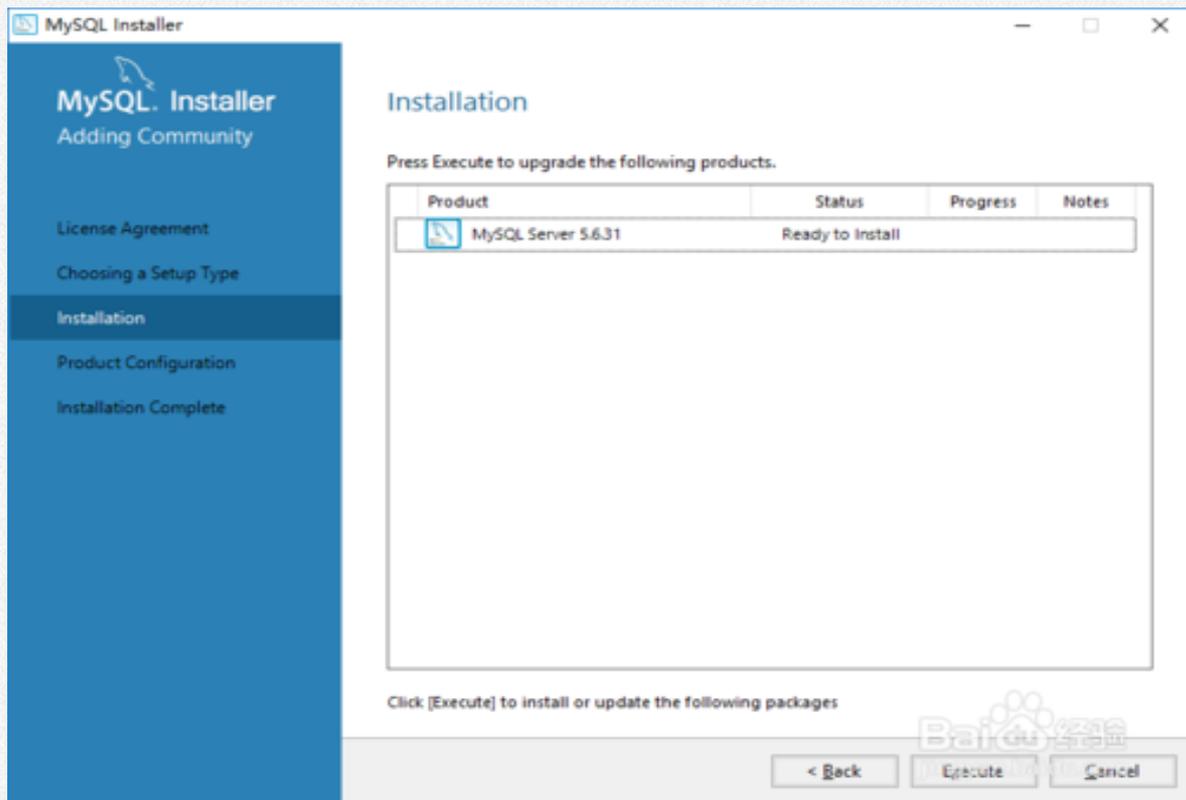
数据库的安装与使用



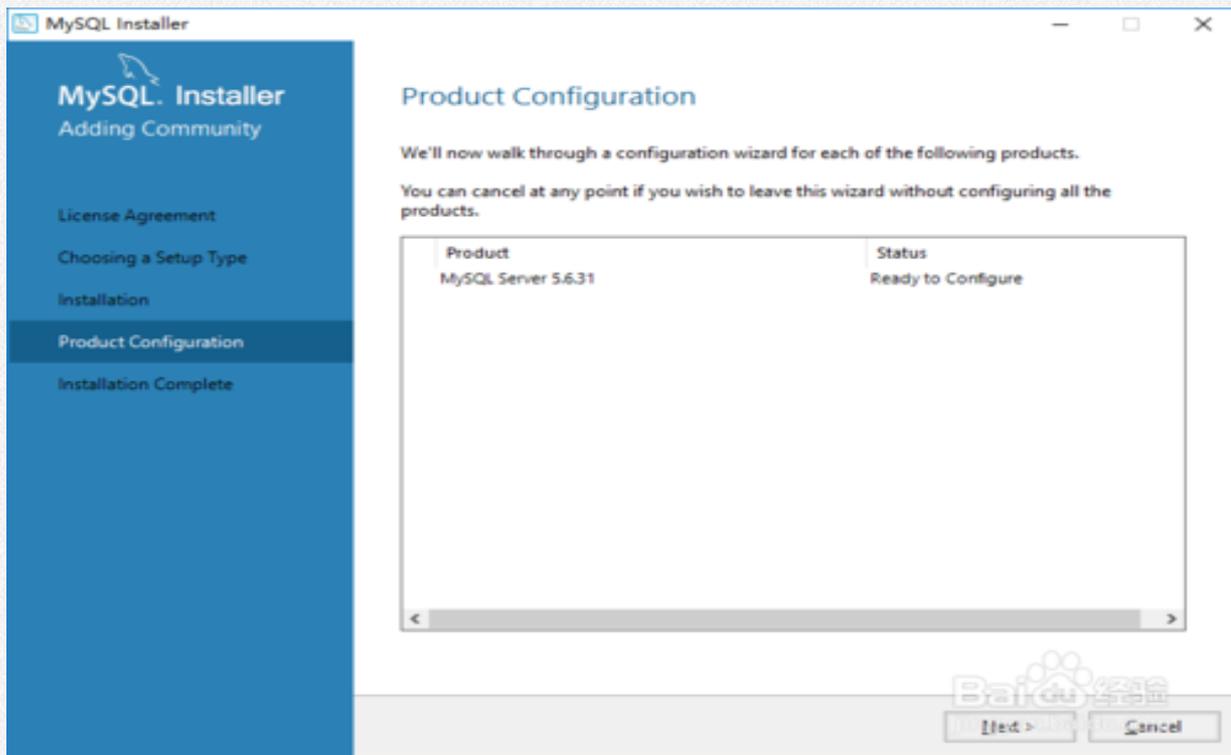
(2)如果只想安装MySQL服务，选择“Server only”，建议选



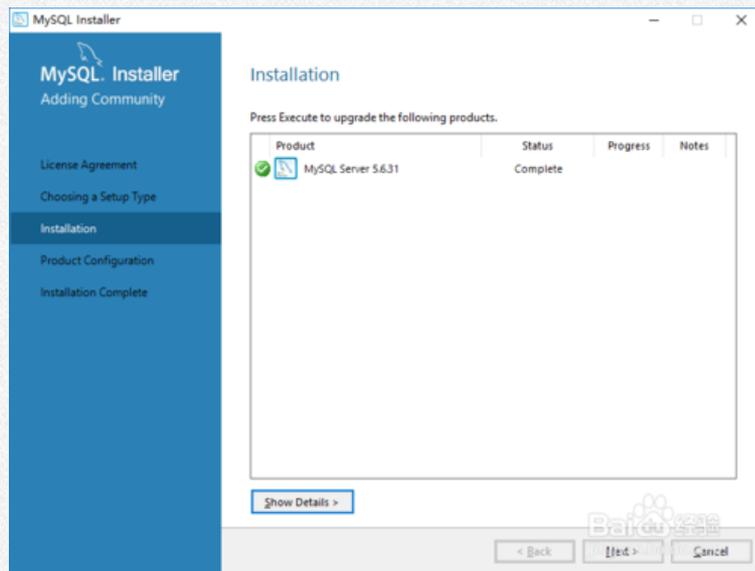
数据库基本概念



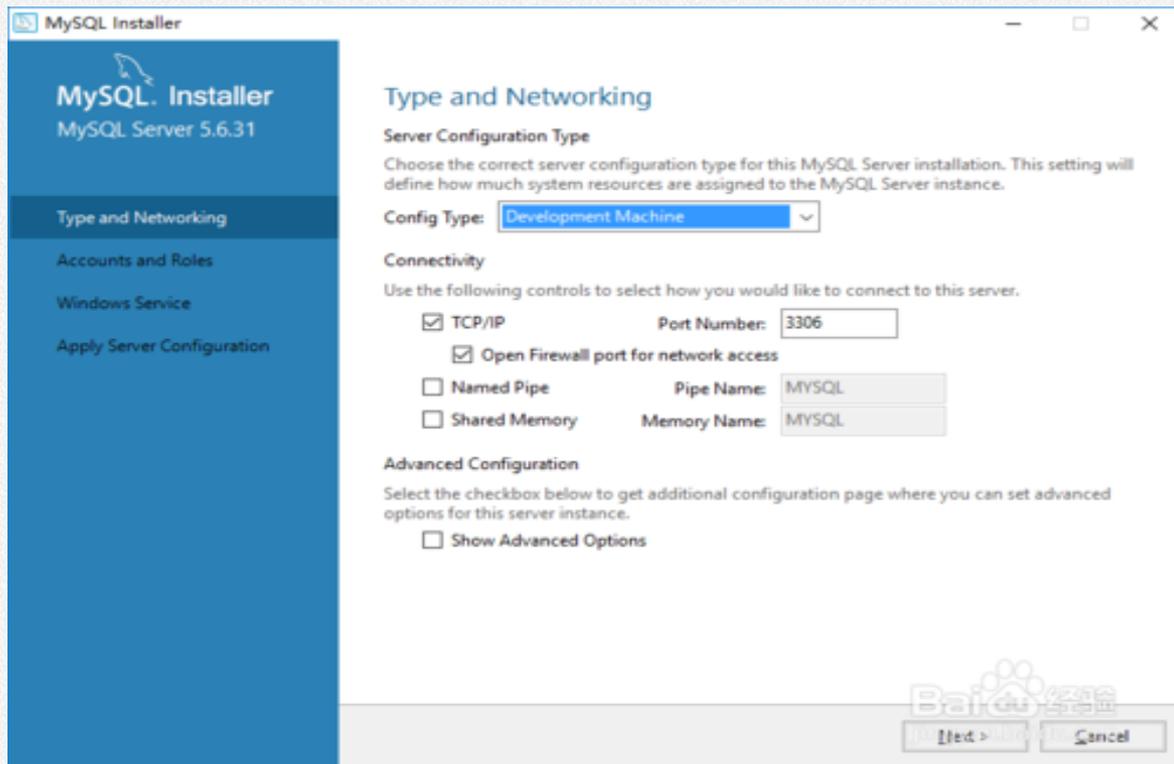
数据库的安装与使用



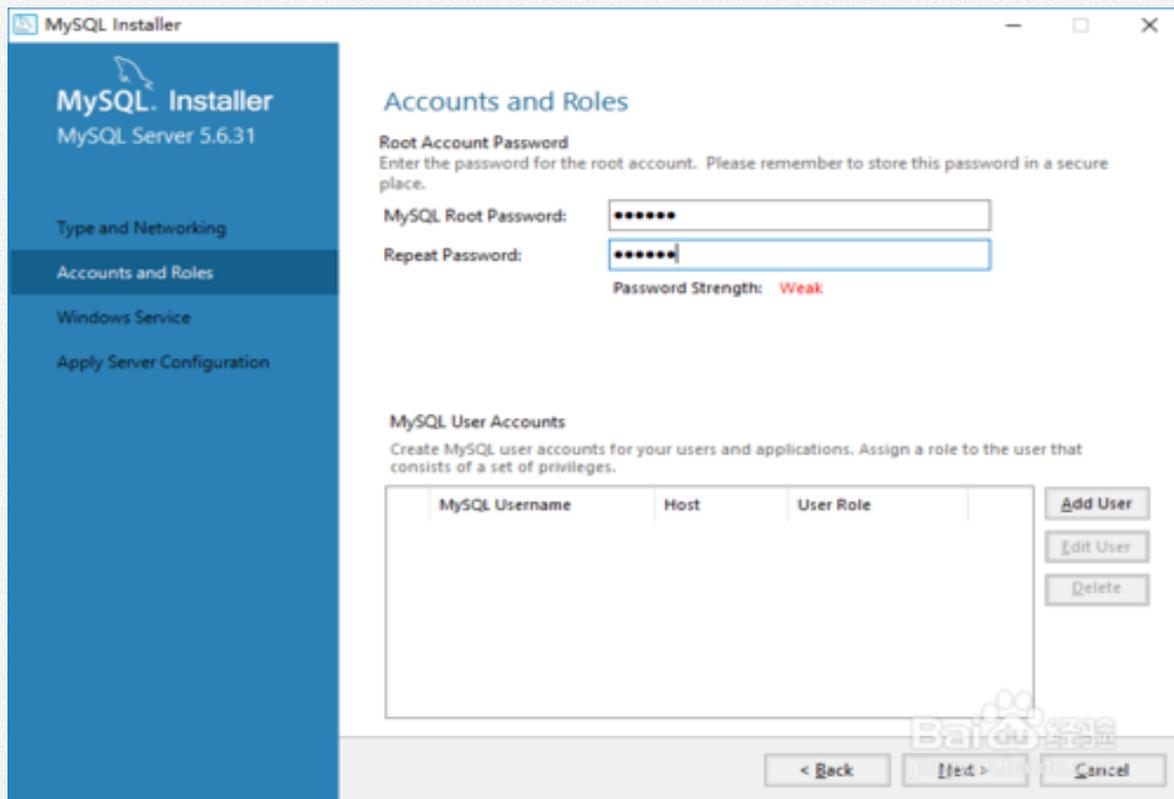
数据库基本概念



数据库的安装与使用

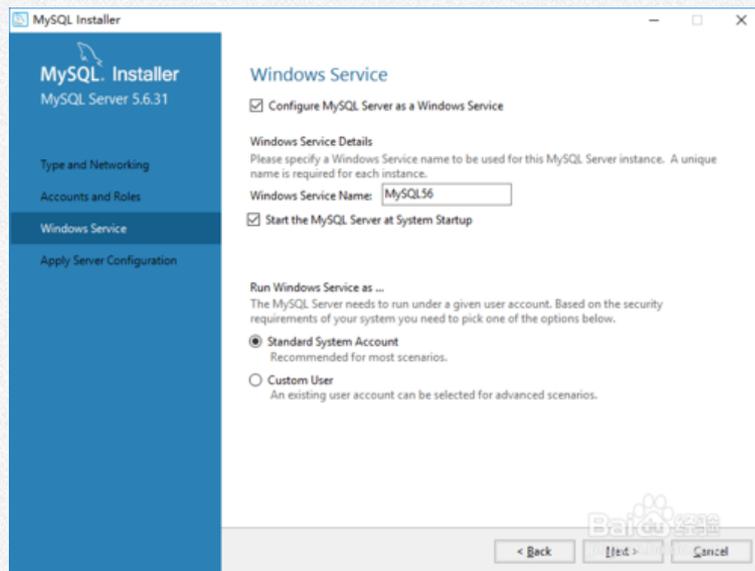


数据库基本概念

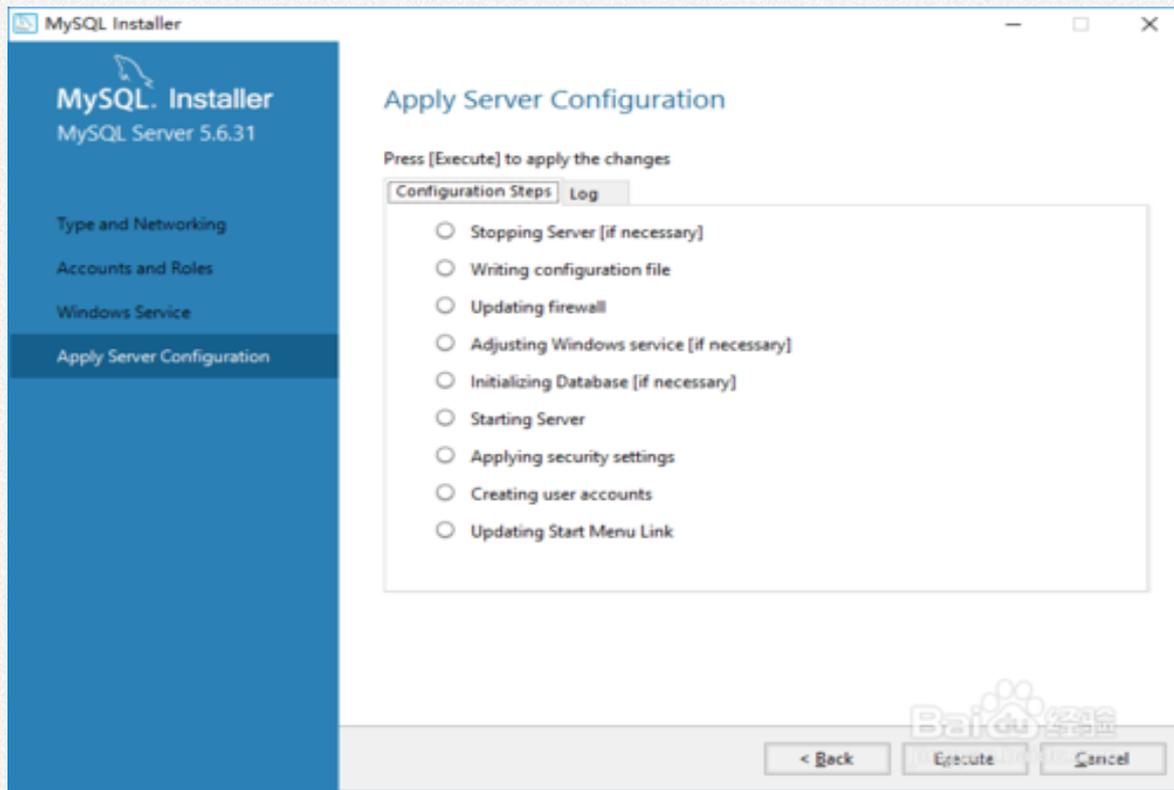


数据库的安装与使用

以系统用户运行windows服务，在Windows下MySQL服务名为:mysql56，后面步骤将讲到mysql56的用处。而且要牢牢记住此名字。



数据库基本概念



以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/135014123111011132>