数据库:

## 数据库

数据 ：能够输入计算机并能被计算机程序识别和处理的信息集合。

数据库（Database）：是按照数据结构来组织、存储和管理数据的仓库。数据库是在数据库管理系统管理和控制之下，存放在存储介质上的数据集合。

数据库管理系统：数据库管理系统是一种操纵和管理数据库的大型软件，用于建立、使用和维护数据库，简称DBMS。它对数据库进行统一的管理和控制，以保证数据库的安全性和完整性。一般支持数据库定义功能，数据库操纵功能，数据库运行控制功能，数据通信功能，支持存取海量数据等

数据库系统（DBS）：它和数据库不是一个概念，数据库系统比数据库大的多，由数据库、数据库管理系统、应用开发工具构成。常见的数据库系统：甲骨文的Oracle、IBM的DB2、微软的SQLServer、开源的MySQL等。

## 数据存储的发展

### 文件管理阶段

优点：数据可长期保存，能存储大量数据

缺点：数据冗余度（redundancy）大，数据一致性（consistency）、完整性（integrity）难以维持。数据与程序缺乏高度独立性。

### 数据库系统阶段

数据组织结构化。数据冗余度比较小，易扩充。具有较高的数据与程序之间的独立性，统一的数据控制。

## 流行数据库简介

Oracle公司是最早开发关系数据库的厂商之一，其产品支持最广泛的操作系统平台。目前Oracle关系数据库产品的市场占有率名列前茅。

IBM 的DB2是第一个具备网上功能的多媒体关系数据库管理系统，支持包括Linux在内的一系列平台。

SQLServer是微软开发的数据库产品，主要支持windows平台

mySQL是一个小型关系型数据库管理系统，开发者为瑞典MySQLAB公司，2008年被Sun公司收购。开放源码。

基于嵌入式linux的数据库主要有SQLite, Firebird, Berkeley DB, eXtremeDB   
Firebird是关系型数据库,功能强大,支持存储过程、SQL兼容等SQLite关系型数据库,体积小,支持ACID事务BerkeleyDB中并没有数据库服务器的概念，它的程序库直接链接到应用程序中eXtremeDB是内存数据库,运行效率高。

## 关系型数据库和非关系型数据库

### 关系型数据库

关系型数据库是指采用了关系模型来组织数据的数据库。

#### 关系型数据库的优点：

容易理解：二维表结构是非常贴近逻辑世界的一个概念，关系模型相对网状、层次等其他模型来说更容易理解。

使用方便：通用的SQL语言使得操作关系型数据库非常方便。

易于维护：丰富的完整性(实体完整性、参照完整性和用户定义的完整性)大大减低了数据冗余和数据不一致的概率。

#### 关系型数据库瓶颈

1. 高并发读写需求

网站的用户并发性非常高，往往达到每秒上万次读写请求，对于传统关系型数据库来说，硬盘I/O是一个很大的瓶颈

2. 海量数据的高效率读写

网站每天产生的数据量是巨大的，对于关系型数据库来说，在一张包含海量数据的表中查询，效率是非常低的

3. 高扩展性和可用性

在基于web的结构当中，数据库是最难进行横向扩展的，当一个应用系统的用户量和访问量与日俱增的时候，数据库却没有办法像web server和app server那样简单的通过添加更多的硬件和服务节点来扩展性能和负载能力。对于很多需要提供24小时不间断服务的网站来说，对数据库系统进行升级和扩展是非常痛苦的事情，往往需要停机维护和数据迁移。

对网站来说，很多关系型数据库的很多特性并不完全能够满足需求。任何大数据量的web系统，都非常忌讳多个大表的关联查询，以及复杂的数据分析类型的复杂SQL报表查询，特别是SNS类型的网站，从需求以及产品阶级角度，就避免了这种情况的产生。往往更多的只是单表的主键查询，以及单表的简单条件分页查询，SQL的功能极大的弱化了。

### 非关系型数据库

NoSQL，泛指非关系型的数据库。随着互联网web2.0网站的兴起，传统的关系数据库在应付web2.0网站，特别是超大规模和高并发的SNS类型的web2.0纯动态网站已经显得力不从心，暴露了很多难以克服的问题，而非关系型的数据库则由于其本身的特点得到了非常迅速的发展。NoSQL数据库的产生就是为了解决大规模数据集合多重数据种类带来的挑战,尤其是大数据应用难题。

#### 键值(Key-Value)存储数据库

这一类数据库主要会使用到一个哈希表，这个表中有一个特定的键和一个指针指向特定的数据。Key/value模型对于IT系统来说的优势在于简单、易部署。但是如果DBA只对部分值进行查询或更新的时候，Key/value就显得效率低下了。例如：Tokyo Cabinet/Tyrant, Redis, Voldemort, Oracle BDB.

#### 列存储数据库

这部分数据库通常是用来应对分布式存储的海量数据。键仍然存在，但是它们的特点是指向了多个列。这些列是由列家族来安排的。如：Cassandra, HBase, Riak.

#### 文档型数据库

文档型数据库的灵感是来自于Lotus Notes办公软件的，而且它同第一种键值存储相类似。该类型的数据模型是版本化的文档，半结构化的文档以特定的格式存储，比如JSON。文档型数据库可 以看作是键值数据库的升级版，允许之间嵌套键值。而且文档型数据库比键值数据库的查询效率更高。如：CouchDB,MongoDb。国内也有文档型数据库SequoiaDB，已经开源

#### 图形(Graph)数据库

图形结构的数据库同其他行列以及刚性结构的SQL数据库不同，它是使用灵活的图形模型，并且能够扩展到多个服务器上。NoSQL数据库没有标准的查询语言(SQL)，因此进行数据库查询需要制定数据模型。许多NoSQL数据库都有REST式的数据接口或者查询API。如：Neo4J, InfoGrid, Infinite Graph.

因此，我们总结NoSQL数据库在以下的这几种情况下比较适用：1、数据模型比较简单；2、需要灵活性更强的IT系统；3、对数据库性能要求较高；4、不需要高度的数据一致性；5、对于给定key，比较容易映射复杂值的环境。

在本部分课程中为同学们讲解关系型数据库，后续的课程还会讲解到非关系型数据库。

## python自身集成的数据库模块————sqlite

SQLite的源代码是C，其源代码完全开放。SQLite第一个Alpha版本诞生于2000年5月。 他是一个轻量级的数据库。在linux下进行交互操作非常方便。并且Python语言集成sqlite3数据库编程接口，也为我们在Python下使用SQLite数据库提供了方便。

### SQLite有以下特性：

### sqlite数据库的安装

要安装sqlite3，可以在终端提示符后运行下列命令:

安装后可以检查版本信息，如果安装没有问题那么通过以下命令会显示安装的sqlite3的版本信息。安装正确就可以使用了。

SQLite3 可使用 sqlite3 模块与 Python 进行集成。sqlite3 模块是由 Gerhard Haring 编写的。它提供了一个与 PEP 249 描述的 DB-API 2.0 规范兼容的 SQL 接口。你不需要单独安装该模块，因为 Python 2.5.x 以上版本默认自带了该模块。当然在这里sqlite数据库不作为重点讲解，我们重点会以mysql为例。只不过python中自带集成的sqlite3模块所以我们单独把它拿出来说。其实通过后续的学习同学们完全可以自己学习掌握sqlite数据库。

## Mysql 数据库介绍

MySQL是一个关系型数据库管理系统，由瑞典MySQL AB公司开发，由c和c++编写，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是最好的关系数据库管理系统应用软件之一。

MySQL是一种关联数据库管理系统，关联数据库将数据保存在不同的表中，而不是将所有数据放在一个大仓库内，这样就增加了速度并提高了灵活性。MySQL所使用的SQL语言是用于访问数据库的最常用标准化语言。由于其体积小、速度快、总体拥有成本低，尤其是开放源码这一特点，一般中小型网站的开发都选择 MySQL 作为网站数据库。

与其他的大型数据库例如 Oracle、DB2、SQL Server等相比，MySQL 自有它的不足之处，但是这丝毫也没有减少它受欢迎的程度。对于一般的个人使用者和中小型企业来说，MySQL提供的功能已经绰绰有余，而且由于MySQL是开放源码软件，因此可以大大降低总体拥有成本。Linux作为操作系统，Apache和 Nginx作为 Web 服务器，MySQL 作为数据库，PHP/Perl/Python作为服务器端脚本解释器。由于这四个软件都是免费或开放源码软件（FLOSS)，因此使用这种方式不用花一分钱就可以建立起一个稳定、免费的网站系统，被业界称为“LAMP“或“LNMP”组合。新浪，网易，facebook，淘宝采用的都是Mysql数据库。

## MySql数据库的安装

并不是所有的系统自来就有mysql，如果你想使用mysql数据库需要进行安装。下面以ubuntu为例进行安装，如果是windows下可以去Oracle的mysql官方页面下载安装。

1． Sudo apt-get update

2． Sudo apt-get install mysql-server

安装过程会自动弹出配置界面

安装mysql的服务端后，可以对数据库进行后端操作。还可以安装个客户端用来访问数据库

sudo apt-get install mysql-client

安装完成后通过mysql –u root -p 进入数据库,然后输入密码可以进入了。

如果想用python对mysql数据库进行操作不像sqlite那么简单，还需要安装python的MySQLdb 进行支持，实际这是一个接口程序，没安装前导入模块失败。

MySQLdb可以通过源码安装或者apt-get直接安装

sudo apt-get install python-mysqldb

安装以后可以在python中导入MySQLdb模块，如果导入成功基本就没有什么问题了。安装过程会有一些自动的路径搜索，如果你的操作系统中不同目录下安装了很多不同版本那么可能只有一个能用，需要你进行配置。当然安装方法还有很多，只要最终能够达到目的就可以了。

## SQL 语句

结构化查询语言(Structured Query Language)简称SQL，结构化查询语言是一种数据库查询和程序设计语言，用于存取数据以及查询、更新和管理关系数据库系统；sql 语句就是对数据库进行操作的一种语言。当然不是所有的数据库都使用 SQL,但所有主流的关系型数据库都使用SQL

## SQL 语句

结构化查询语言(Structured Query Language)简称SQL，结构化查询语言是一种数据库查询和程序设计语言，用于存取数据以及查询、更新和管理关系数据库系统；sql 语句就是对数据库进行操作的一种语言。当然不是所有的数据库都使用 SQL,但所有主流的关系型数据库都使用SQL

### 基本概念

数据库文件：我们之前提到过数据库的概念。在计算机中具体的来说，一个关系型数据库可以看作一个数据库文件（当然有一系列配套组件我们不用理会），我们通过对数据库文件的操作达到操作数据库中数据的目的。在使用数据库前我们都需要创建这个“数据库文件”。关系型数据库文件一般以 “.db”作为后缀。

表 ：你可以将数据库存储想像为一个表格，关系型数据库都是通过表来整理数据的。一个数据库中可以包含多个表。每个表在数据库中都有自己唯一的名字。

字段: 就像我们日常做的表格一样，表中的每一列代表着一项，每一项都有自己的名称。我们跟表中的这些项（列）称之为字段。

记录：表中的每一行都是一条记录。

操作：数据库能够被创建, 也可以被删除，表也一样。往数据库里增加一条记录称为 插入(inserting) 一条记录, 修改库中一条已有的记录则称为更新(updating),删除表中已经有的数据行称为删除(deleting)，从一个数据库中请求符合条件的数据称为查询(querying)，当你对一个数据库进行查询时, 你可以一步取回所有符合条件的数据, 也可以循环逐条取出每一行。有些数据库使用游标的概念来表示 SQL 命令, 查询,取回结果集等等。

### 语句分类

每种数据库在linux下都会有一些相应的数据库交互命令。这不算是sql语句的部分，只是数据库自身针对操作系统设计的一些使用命令。进入mysql交互模式下语句操作都是以”;“作为结束标志，

常用SQL命令 ：

help : 查看命令帮助

quit 或 exit： 退出mysql操作

show databases ： 显示数据库文件列表

use <database>：打开数据库

show tables ：显示数据库中所有表名

desc <table\_name> : 查看表的结构

在sql中关键字是不区分大小写的，操作内容区分大小写

**sql——数据定义**

**创建**一个数据库:create database <databasename>

创建后MySql在linux下默认的数据文件存储目录为/var/lib/mysql

**删除**一个数据库:

drop database <databasename>

创建一个表：create table tabname(col1 type1 [not null] [primary key],col2 type2 [not null],..)

这个命令中tabname为表名后面col为字段名称，type1为字段数据类型，后面是约束条件。

删除一个表：

drop table <table\_name>

修改数据表字段类型

alter table <tablename> modify <field> <new\_type> [first | after]

增加表字段:

alter table <tablename> add column (<field> <type>, …) [first | after]

删除表字段:

alter table <tablename> drop column <field>;

字段改名:

alter table <tablename> change <old\_field> <new\_field> <type> [first | after]

修改默认值:

ALTER TABLE <tablename> ALTER <field> SET DEFAULT 1000;

ALTER TABLE <tablename> ALTER <field> DROP DEFAULT;

更改表名:

alter table <tablename> rename [to] <new\_tablename>

## sql——数据操纵

插入记录

insert into <table\_name> (field1,field2….) values (value1, value2,…);

如果不指定类型那么默认需要按照表的字段顺序来写值

更新记录

update <table\_name> set <f1=value1>, <f2=value2> … where <expression>

删除记录

delete from <table\_name> where <expression>

查询记录

select \* from <table\_name> where <expression>

“\*”代表所有记录，也可以用都好分割的方式选定字段来代替。

条件查询

select \* from <table\_name> where <expression>

在条件中可以使用 = > < >= <= != 等比较运算符，还可以使用and，or等逻辑运算符。

## sql——数据控制

DCL语句主要是DBA用来管理系统中的对象权限时使用，一般的开发人员很少使用。

### mysql数据类型

每种数据库都有自己支持的数据类型，算然大同小异但是写法上还是有些差别。mysql中数据类型主要分为三类。

#### 数字类型IMG_256

#### 时间类型



#### 字符串类型

