

使用 Python 操作数据库

目标

1. 掌握 PyMySQL 进行数据库的连接、库表操作、事务与异常处理
2. 掌握 SQL 优化必备的锁、事务、索引设计、隔离级别等概念
3. 掌握 ORM 的使用

MySQL 数据库的安装

企业级 MySQL 部署在 Linux 操作系统上，需要注意的重点：

- 注意操作系统的平台（32 位、64 位）
- 注意安装 MySQL 的版本（MySQL 企业版、社区版、MariaDB）
- 注意安装后避免 yum 自动更新
- 注意数据库的安全性

MySQL 数据库的安装

查看字符集：

```
mysql> show variables like '%character%';
```

查看校对规则

```
mysql> show variables like 'collation_%';
```

注意： MySQL 中的 utf8 不是 UTF-8 字符集

Python 连接 MySQL 的方法

统一概念：

其他语言：连接器、绑定、binding

Python 语言：Python Database API、DB-API

注意：MySQLdb 是 Python2 的包，适用于 MySQL5.5 和 Python2.7

Python 连接 MySQL 的方法

Python3 连接 MySQL:

- Python3 安装的 MySQLdb 包叫做 mysqlclient, 加载的依然是 MySQLdb
- shell> pip install **mysqlclient**
- python> import **MySQLdb**

其他DB-API:

- shell> pip install pymysql # 流行度最高
- shell> pip install mysql-connector-python # MySQL官方

使用ORM:

- shell> pip install sqlalchemy

Python 连接 MySQL 的方法

PyMySQL 和 SQLAlchemy 连接 MySQL 数据库对比:

- `pymysql.connect("server1","testuser","testpass","testdb")`
- `engine=create_engine("mysql+pymysql://... ..",echo=True)`

那些你需要精通的SQL知识

SQL语言功能划分：

DQL: Data Query Language, 数据查询语言, 开发工程师学习的重点。

DDL: Data Definition Language, 数据定义语言, 操作库和表结构。

DML: Data Manipulation Language, 数据操作语言, 操作表中记录。

DCL: Data Control Language, 数据控制语言, 安全和访问权限控制。

那些你需要精通的 SQL 知识

创建表要注意哪些问题？

```
CREATE TABLE 'book' (  
    'book_id' int(11) NOT NULL AUTO_INCREMENT,  
    'book_name' varchar(255) ,  
    PRIMARY KEY ('book_id')  
) ENGINE = InnoDB CHARACTER SET = utf8 COLLATE = utf8_general_ci;
```

那些你需要精通的 SQL 知识

查询数据要注意哪些问题？

SELECT 查询时关键字的顺序

SELECT ... FROM ... WHERE ... GROUP BY ... HAVING ... ORDER BY ... LIMIT

注意： 1. 生产环境下因为列数相对较多，一般禁用 SELECT *
2. WHERE 字段为避免全表扫描，一般需要增加索引

那些你需要精通的 SQL 知识

SQL 函数有哪些？

算术函数、字符串函数、日期函数、转换函数、聚合函数

聚合函数

- COUNT() 行数
- MAX() 最大值
- MIN() 最小值
- SUM() 求和
- AVG() 平均值

注意：聚合函数忽略空行

那些你需要精通的 SQL 知识

什么是子查询？

需要从查询结果集中再次进行查询，才能得到想要的结果。

子查询需要关注的问题？

- 关联子查询与非关联子查询区别。
- 何时使用 IN，何时使用 EXISTS。

那些你需要精通的 SQL 知识

常见的连接（JOIN）有哪些？

- 自然连接
- ON 连接
- USING 连接
- 外连接
 - 左外连接
 - 右外连接
 - 全外连接（MySQL 不支持）

那些你需要精通的 SQL 知识

什么是事务？

要么全执行，要么不执行

事务的特性-- ACID

- A 原子性 (Atomicity)
- C 一致性 (Consistency)
- I 隔离性 (Isolation)
- D 持久性 (Durability)

那些你需要精通的 SQL 知识

事务的隔离级别

读未提交：允许读到未提交的数据

读已提交：只能读到已经提交的内容

可重复读：同一事务在相同查询条件下两次查询得到的数据结果一致

可串行化：事务进行串行化，但是牺牲了并发性能

实战 Python 操作MySQL

使用 PyMySQL 进行 MySQL 的增、删、改、查

使用 PyMySQL 取出多行数据

使用 SQLAlchemy 进行 MySQL 的增、删、改、查

使用 Django 进行 MySQL 的增、删、改、查

SQL 优化的常见手段

反范式设计

增加合理索引

MVCC 多版本并发控制技术

慢 SQL 定位工具

注意：不同的优化策略必须适应相应的场景，不能盲目背参数

实战日志数据库

使用 MySQL 保存 Nginx 日志，并采用 Python 进行查询

总结

1. 掌握 PyMySQL 进行数据库的连接、库表操作、事务与异常处理
2. 掌握 SQL 优化必备的锁、事务、索引设计、隔离级别等概念
3. 掌握 ORM 的使用

THANKS! |  极客大学