带着问题来学习：

1.如何防止数据库中的用户信息被脱库?

2.你会如何存储用户密码这么重要的数据吗？仅仅 MD5 加密一下存储就够了吗?

3.在实际开发中，我们应该如何用哈希算法解决问题?

一、什么是哈希算法?

1.定义

将任意长度的二进制值串映射成固定长度的二进制值串，这个映射的规则就是哈希算法，而通过原始数据映射之后得到的二进制值串就是哈希值。

2.如何设计一个优秀的哈希算法?

①单向哈希:

从哈希值不能反向推导出哈希值（所以哈希算法也叫单向哈希算法）。

②篡改无效:

对输入敏感，哪怕原始数据只修改一个Bit，最后得到的哈希值也大不相同。

③散列冲突:

散列冲突的概率要很小，对于不同的原始数据，哈希值相同的概率非常小。

④执行效率:

哈希算法的执行效率要尽量高效，针对较长的文本，也能快速计算哈希值。

二、哈希算法的常见应用有哪些?

7个常见应用：安全加密、唯一标识、数据校验、散列函数、负载均衡、数据分片、分布式存储。

1.安全加密

①常用于加密的哈希算法：

MD5：MD5 Message-Digest Algorithm，MD5消息摘要算法

SHA：Secure Hash Algorithm，安全散列算法

DES：Data Encryption Standard，数据加密标准

AES：Advanced Encryption Standard，高级加密标准

②对用于加密的哈希算法，有两点格外重要，第一点是很难根据哈希值反向推导出原始数据，第二点是散列冲突的概率要小。

③在实际开发中要权衡破解难度和计算时间来决定究竟使用哪种加密算法。

2.唯一标识

通过哈希算法计算出数据的唯一标识，从而用于高效检索数据。

3.数据校验

利用哈希算法对输入数据敏感的特点，可以对数据取哈希值，从而高效校验数据是否被篡改过。

4.散列函数

散列函数中用到的哈希算法更加关注散列后的值能不能平均分布，以及散列函数的执行快慢。

三、思考

1.如何防止数据库中的用户信息被脱库？你会如何存储用户密码这么重要的数据吗?

①使用MD5进行加密

②字典攻击：如果用户信息被“脱库”，黑客虽然拿到的是加密之后的密文，但可以通过“猜”的方式来破解密码，这是因为，有些用户的密码太简单。

③针对字典攻击，我们可以引入一个盐（salt），跟用户密码组合在一起，增加密码的复杂度。

2.现在，区块链是一个很火的领域，它被很多人神秘化，不过其底层的实现原理并不复杂。其中，哈希算法就是它的一个非常重要的理论基础。你能讲一讲区块链使用的是哪种哈希算法吗？是为了解决什么问题而使用的呢?

区块链是一块块区块组成的，每个区块分为两部分：区块头和区块体。

区块头保存着 自己区块体 和 上一个区块头 的哈希值。

因为这种链式关系和哈希值的唯一性，只要区块链上任意一个区块被修改过，后面所有区块保存的哈希值就不对了。

区块链使用的是 SHA256 哈希算法，计算哈希值非常耗时，如果要篡改一个区块，就必须重新计算该区块后面所有的区块的哈希值，短时间内几乎不可能做到。

极客时间文档：https://time.geekbang.org/column/article/65312