参考：<https://database.51cto.com/art/201809/583857.htm>

<https://juejin.im/entry/6844903648670007310>

<https://crossoverjie.top/2019/07/24/framework-design/sharding-db-03/>

## 为什么要分表

当一张表的数据达到几千万时，你查询一次所花的时间会变多，如果有联合查询的话，我想有可能会死在那儿了。分表的目的就在于此，减小数据库的负担，缩短查询时间。

## 方法

### 垂直拆分

所谓的垂直拆就是大表拆小表，是基于字段进行的，如果表中的字段较多，一般将数据大、长度长的拆分到扩展表

### 水平拆分

针对数据量大的单张表（比如订单表），按照某种策略规则，切分到多张表中，但这些表还是在一个库中，所以库级别的数据库操作还是有IO瓶颈。不建议采用。

#### 水平拆分策略

##### Range

按一定范围划分，

比如我们可以将某张表的创建时间按照日期划分存为月表

也可以将某张表的主键按照范围划分，比如 【1~10000】在一张表，【10001~20000】在一张表，以此类推

##### Hash

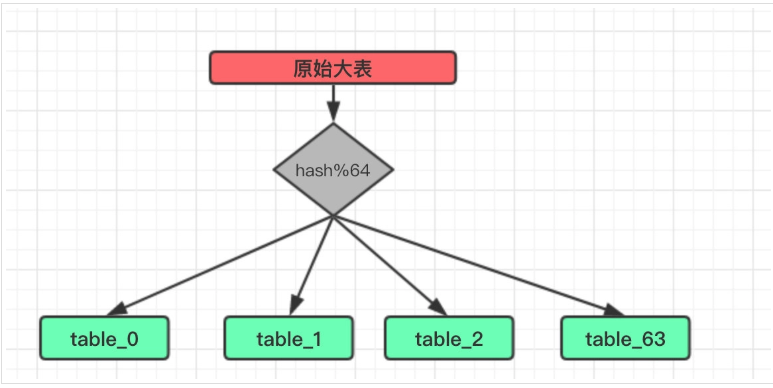
散列表，是根据键直接访问在指定储存位置数据的数据结构。通过计算一个关于键的函数也称为哈希函数

Hash 散列表，将任意长度的输入，通过散列算法变换成固定长度的输出，输出的值就是散列值。

散列函数能使对一个数据序列的访问过程更加迅速有效，是一种空间换时间的算法，通过散列函数数据元素将被更快定位。

参考

<https://baijiahao.baidu.com/s?id=1673906020696389434&wfr=spider&for=pc>

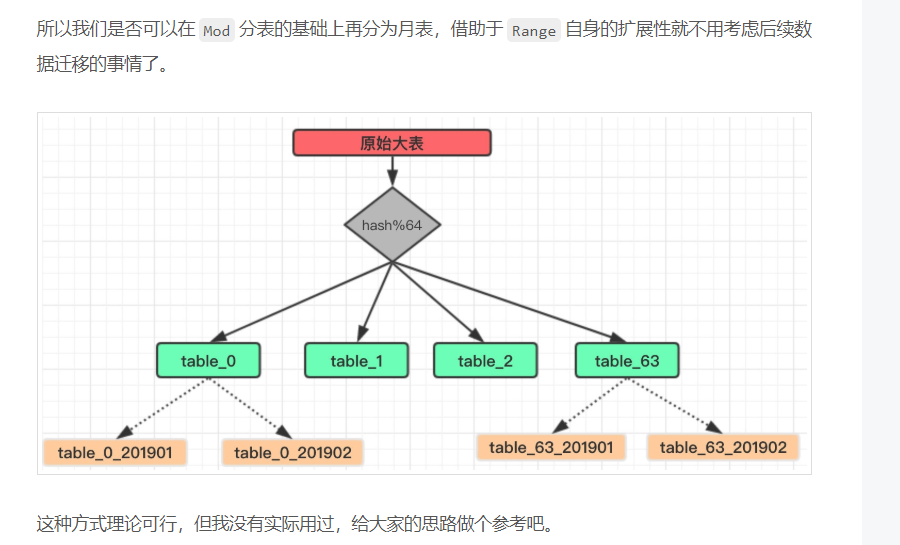


我们对主键字段直接取模也是可以的，也可以通过hash算法，获取散列值再取模 ，其本就是根据字段的值来将数据分配到不同的表中。

关于将一张数据表分成多少个表主要和你的mysql性能有关，mysqld的性能跟你的版本和使用场景有关，一般将一个分表的数据量划分到几百万即可。

如果一张单表的数据量有1亿 一般对32，64（2的n次方）取模即可

##### Range+hash



#### 局限

如果只是在一个库中做单表的分表，当数据量到达一定程度时，一个库是满足不了的。知识后就要分库

## 一致性原理与实现

参考

<https://www.jianshu.com/p/528ce5cd7e8f>

<https://baijiahao.baidu.com/s?id=1673906020696389434&wfr=spider&for=pc>

## 一致性hash分库分表