

数据库第4次作业——Chap7

软件2203班——丁海桐——202226010304

1.对于QQ数据库，其中有用户表 *user* (*userid*, *password*, *nickname*, ...)，群表 *community* (*name*, *creator_id*, ...)，群成员表 *member* (*community_id*, *member_id*, *nickname*, *act_role*, ...)。请分析群成员表 *member* 的主键应该是 (*community_id*, *member_id*)，还是 (*member_id*, *community_id*) 请替 DBMS 为群成员表 *member* 写一个创建主键索引的SQL语句。

对于群成员表最频繁的查询是，通过群号查找这个群都有哪些人，当主键的第一个字段为 *community_id* 时，主键索引会首先以 *community_id* 来排序，这样主键索引的排序特性就能充分利用。

```
CREATE INDEX idx_member_community_member ON member (community_id, member_id);
```

2.房产中介公司的业务数据库（见图2-12）中，合同表 *Lease* 是一个联系表，它的主键可以为 (*clientNo*, *propertyNo*, *fromDate*)。但是，实际上合同表的主键设为 *leaseNo*。也就是说，(*clientNo*, *propertyNo*, *fromDate*) 和 *leaseNo* 都是 *Lease* 表的候选键。

- 请基于业务特性和访问特性分析，主键到底是选 *leaseNo* 好，还是选 (*clientNo*, *propertyNo*, *fromDate*) 好？
 - 选 *leaseNo* 好
 - 使用单一的 *leaseNo* 作为主键长度更短，访问性能更好。
 - 在实际业务中，更多时候是知道订单号，而不是同时知道顾客、房产、其实日期
- 对 *leaseNo* 如何设置其构成？请从高效查询角度进行分析。
 - 可以使用 店铺编号 + 当天日期 + 自增编号
 - 也可以设计为自增整数或UUID等唯一标识符。
- 假定 *leaseNo* 为主键，那么对候选键 (*clientNo*, *propertyNo*, *fromDate*) 有必要创建索引吗？
 - 有必要，因为这些字段可能在查询中被频繁使用，特别是在按照客户、物业或起始日期进行过滤和排序时。通过为候选键创建索引，可以进一步提高查询性能。

3.对任一给定的英文文章，要统计其中出现的单词以及每个单词出现的次数。针对这样一个需求，请分析使用散列索引好还是顺序索引好。说明理由。

应该使用散列索引。原因如下：

- 散列索引适用于快速查找某个值。在单词统计的情况下，我们需要快速查找每个单词及其出现次数。
- 在处理大量不同单词的情况下，散列索引通常能够提供更好的性能，因为它能更有效地处理高基数情况。
- 散列索引适合精确查找。