

自我介绍

实习期做过的项目

实习期做爬虫的，你关于爬虫有什么心得、方法论（原问题记不清楚啦，反正答案反推，应该是类似的问题）？

Java 垃圾回收，答了引用计数、可达性分析、GC ROOT、分代收集算法、新生代老年代问为什么要搞分代收集

数据库了解多少？ 主要就用的 MySQL

架构设计题：

设计一个类似微博的东西，用户可以发微博，用户也可以刷别人发的微博

Q：你觉得微博和朋友圈的区别是什么？

A:面试官引导着答的，我开始回复的可能是关注列表、好友的数量，微信限制几 k，微博粉丝可能 100w+，

面试官说，朋友圈的好友关系是双向的、平等的，微博的关注、被关注是单向的、非对等的

Q：数据库设计

A:很多也是在面试官的引导下说出来的，首先是 user 表，其次是 content 表 -> 表切分，水平切分（大 v 与小透明）、垂直切分，分库分表（利用 hash）；

其余还有要对大 V 的发文在内存缓存，参考思路按 LRU，结合粉丝量、发文时间、点赞增长势头进行一个权重计算，高权重的缓存起来；

个人资料的离线缓存，刷微博，不能一口气所有的内容都返回，要分页返回

【这道题，差不多耗费了有 15-20 分钟，估计是】

算法题

给一个有序数组 arr，长度为 n，存在一个索引值 k，a[:k]、a[k:] 的方差最小，方差就是子数组的平均值、平方和哪一些。.. 具体公式就不写了，太麻烦啦。

给出实现思路