

牛客网算法基础提升班

第六课

牛客网最新算法课--基础提升班：针对有简单入门基础，还达不到校招水平要求的同学，左神手把手讲解高级数据结构和算法原理，结合典型例题深入讲解解题思路，并提供最优解和代码

上课时间：每周六日 09:00——11:00

上课老师：左程云，华科本科，芝加哥大学硕士，曾就职于IBM、百度、GrowingIO、亚马逊，也是牛客网的老师。

牛客网：一个提供海量校招真题及专项练习题，笔经面经，招聘信息，学习资源及交流的平台<https://www.nowcoder.com/>



笔经面经



学习交流

- 1) 哈希函数可以把数据按照种类均匀分流
- 2) 布隆过滤器用于集合的建立与查询，并可以节省大量空间
- 3) 一致性哈希解决数据服务器的负载管理问题
- 4) 利用并查集结构做岛问题的并行计算
- 5) 位图解决某一范围上数字的出现情况，并可以节省大量空间
- 6) 利用分段统计思想、并进一步节省大量空间
- 7) 利用堆、外排序来做多个处理单元的结果合并

之前的课已经介绍过前4个内容，本节内容为介绍解决大数据题目的后3个技巧

题目一

32位无符号整数的范围是0~4,294,967,295，现在有一个正好包含40亿个无符号整数的文件，所以在整个范围中必然存在没出现过的数。可以使用最多1GB的内存，怎么找到所有未出现过的数？

【进阶】

内存限制为 10MB，但是只用找到一个没出现过的数即可

题目二

有一个包含100亿个URL的大文件，假设每个URL占用64B，请找出其中所有重复的URL

【补充】

某搜索公司一天的用户搜索词汇是海量的(百亿数据量)，请设计一种求出每天热门Top100词汇的可行办法

题目三

32位无符号整数的范围是0~4294967295，现在有40亿个无符号整数，可以使用最多1GB的内存，找出所有出现了两次的数。

【补充】

可以使用最多10MB的内存，怎么找到这40亿个整数的中位数？

题目四

位运算的题目

之前介绍过一些，下面继续

给定两个有符号32位整数a和b，返回a和b中较大的。

【要求】

不用做任何比较判断。

题目五

判断一个32位正数是不是2的幂、4的幂

题目六

给定两个有符号32位整数a和b，不能使用算术运算符，分别实现a和b的加、减、乘、除运算

【要求】

如果给定a、b执行加减乘除的运算结果就会导致数据的溢出，那么你实现的函数不必对此负责，除此之外请保证计算过程不发生溢出

提升项目经验

- 课程名称：《牛客高级项目课--（牛客网）》
- 课程地址：<https://www.nowcoder.com/courses/semester/senior>
- 独家内部100元优惠券：DRMscjy



面试算法书籍

- 书名：《程序员代码面试指南—IT名企算法与数据结构题目最优解》
- 作者：左程云

THANK YOU

查看更多笔经面经

