



微信搜一搜

Java架构师进阶编程

• 百度篇

3.1.0 在函数内定义一个字符数组，用 gets 函数输入字符串的时候，如果输入越界，为什么程序会崩溃？

3.1.1 C++中引用与指针的区别

3.1.2 C/C++程序的内存分区

3.1.3 快速排序的思想、时间复杂度、实现以及优化方法

3.1.4 IO 模型——IO 多路复用机制？

3.1.5 常用的 Linux 命令

3.1.6 C 中变量的存储类型有哪些？

3.1.7 动态规划的本质

3.1.8 实践中如何优化 MySQL？

3.1.9 什么情况下设置了索引但无法使用？

3.2.0 SQL 语句的优化

3.2.1 数据库索引的底层实现原理和优化

3.2.2 HTTP 和 HTTPS 的主要区别？

3.2.3 如何设计一个高并发的系统？

3.2.4 两条相交的单向链表，如何求他们的第一个公共节点？

3.2.5 求单向局部循环链表的环入口？

3.2.6 IP 地址如何在数据库中存储？

3.2.7 new/delete 和 malloc/free 的底层实现？

3.2.8 overload、override、overwrite 的介绍？

3.2.9 小端/大端机器？

3.3.0 守护进程

3.3.1 多线程的优缺点

3.3.2 长连接与短连接

3.3.3 二分图应用于最佳匹配问题（游客对房间的满意度之和最大问题）

3.3.4 class 与 struct 的区别？

3.3.5 虚函数和纯虚函数

3.3.6 memset() 函数

3.3.7 实现一个函数，对一个正整数 n，算得到 1 需要的最少操作次数。操作规则为：如果 n 为偶数，将其除以 2；如果 n 为奇数，可以加 1 或减 1；一直处理下去。

3.3.8 找到满足条件的数组

3.3.9 一个大的含有 50M 个 URL 的记录，一个小的含有 500 个 URL 的记录，找出两个记录里相同的 URL

3.4.0 海量日志数据，提取出某日访问百度次数最多的那个 IP

3.4.1 有 10 个文件，每个文件 1G，每个文件的每一行都存放的是用户的 query，每个文件的 query 都可能重复。如何按照 query 的频度排序？

3.4.2 蚂蚁爬杆问题

3.4.3 当在浏览器中输入一个 url 后回车，后台发生了什么？比如输入 url 后，你看到了百度的首页，那么这一切是如何发生的呢？

3.4.4 判断两棵树是否相等，请实现两棵树是否相等的比较，相等返回 1，否则返回其他值，并说明算法复杂度

3.4.5 三个警察和三个囚徒的过河问题

3.4.6 从 300 万字符串中找到最热门的 10 条

3.4.7 如何找出字典中的兄弟单词。给定一个单词 **a**, 如果通过交换单词中字母的顺序可以得到另外的单词 **b**, 那么定义 **b** 是 **a** 的兄弟单词。现在给定一个字典, 用户输入一个单词, 如何根据字典找出这个单词有多少个兄弟单词?

3.4.8 找出数组中出现次数超过一半的数, 现在有一个数组, 已知一个数出现的次数超过了半, 请用 $O(n)$ 的复杂度的算法找出这个数。

3.4.9 找出被修改过的数字

3.5.0 设计 DNS 服务器中 cache 的数据结构。要求设计一个 DNS 的 Cache 结构, 要求能够满足每秒 5000 以上的查询, 满足 IP 数据的快速插入, 查询的速度要快。(题目还给出了一系列的数据, 比如: 站点数总共为 5000 万, IP 地址有 1000 万, 等等)

3.5.1 找出给定字符串对应的序号

3.5.2 找出第 k 大的数字所在的位置。写一段程序, 找出数组中第 k 大小的数, 输出数所在的位置。例如{2, 4, 3, 4, 7}中, 第一大的数是 7, 位置在 4。第二大、第三大的数都是 4, 位置在 1、3 随便输出哪一个均可。

3.5.3 给 40 亿个不重复的 `unsigned int` 的整数, 没排过序的, 然后再给几个数, 如何快速判断这几个数是否在那 40 亿个数当中?

3.5.4 在一个文件中有 10G 个整数, 乱序排列, 要求找出中位数。内存限制为 2G。

3.5.5 时分秒针在一天之类重合多少次? (24 小时)

3.5.6 将多个集合合并成没有交集的集合。

3.5.7 平面内有 11 个点, 由它们连成 48 条不同的直线, 由这些点可连成多少个三角形?

3.1.0 在函数内定义一个字符数组，用 `gets` 函数输入字符串的时候，如果输入越界，为什么程序会崩溃？

3.1.1 C++ 中引用与指针的区别

3.1.2 C/C++ 程序的内存分区

3.1.3 快速排序的思想、时间复杂度、实现以及优化方法

3.1.4 IO 模型——IO 多路复用机制？

3.1.5 常用的 Linux 命令

3.1.6 C 中变量的存储类型有哪些？

3.1.7 动态规划的本质

3.1.8 实践中如何优化 MySQL？

3.1.9 什么情况下设置了索引但无法使用？

3.2.0 SQL 语句的优化

3.2.1 数据库索引的底层实现原理和优化

3.2.2 HTTP 和 HTTPS 的主要区别？

3.2.3 如何设计一个高并发的系统？

3.2.4 两条相交的单向链表，如何求他们的第一个公共节点？

3.2.5 求单向局部循环链表的环入口？

3.2.6 IP 地址如何在数据库中存储?

3.2.7 new/delete 和 malloc/free 的底层实现?

3.2.8 overload、override、overwrite 的介绍?

3.2.9 小端/大端机器?

3.3.0 守护进程

3.3.1 多线程的优缺点

3.3.2 长连接与短连接

3.3.3 二分图应用于最佳匹配问题（游客对房间的满意度之和最大问题）

3.3.4 class 与 struct 的区别?

3.3.5 虚函数和纯虚函数

3.3.6 memset() 函数

3.3.7 实现一个函数，对一个正整数 n，算得到 1 需要的最少操作次数。操作规则为：如果 n 为偶数，将其除以 2；如果 n 为奇数，可以加 1 或减 1；一直处理下去。

3.3.8 找到满足条件的数组

3.3.9 一个大的含有 50M 个 URL 的记录，一个小的含有 500 个 URL 的记录，找出两个记录里相同的 URL

3.4.0 海量日志数据，提取出某日访问百度次数最多的那个 IP

3.4.1 有 10 个文件，每个文件 1G，每个文件的每一行都存放的是用户的 query，每个文件的 query 都可能重复。如何按照 query 的频度排序？

3.4.2 蚂蚁爬杆问题

3.4.3 当在浏览器中输入一个 url 后回车，后台发生了什么？比如输入 url 后，你看到了百度的首页，那么这一切是如何发生的呢？

3.4.4 判断两棵树是否相等，请实现两棵树是否相等的比较，相等返回 1，否则返回其他值，并说明算法复杂度

3.4.5 三个警察和三个囚徒的过河问题

3.4.6 从 300 万字符串中找到最热门的 10 条

3.4.7 如何找出字典中的兄弟单词。给定一个单词 a，如果通过交换单词中字母的顺序可以得到另外的单词 b，那么定义 b 是 a 的兄弟单词。现在给定一个字典，用户输入一个单词，如何根据字典找出这个单词有多少个兄弟单词？

3.4.8 找出数组中出现次数超过一半的数，现在有一个数组，已知一个数出现的次数超过了半，**O(n)** 的复杂度的算法找出这个数。

3.4.9 找出被修改过的数字

3.5.0 设计 DNS 服务器中 cache 的数据结构。要求设计一个 DNS 的 Cache 结构，要求能够满足每秒 5000 以上的查询，满足 IP 数据的快速插入，查询的速度要快。（题目还给出了一系列的数据，比如：站点数总共为 5000 万，IP 地址有 1000 万，等等）

3.5.1 找出给定字符串对应的序号

3.5.2 找出第 k 大的数字所在的位置。写一段程序，找出数组中第 k 大小的数，输出数所在的位置。例如{2, 4, 3, 4, 7}中，第一大的数是 7，位置在 4。第二大、第三大的数都是 4，位置在 1、3 随便输出哪一个均可。

3.5.3 给 40 亿个不重复的 **unsigned int** 的整数，没排过序的，然后再给几个数，如何快速判断这几个数是否在那 40 亿个数当中？

3.5.4 在一个文件中有 10G 个整数，乱序排列，要求找出中位数。内存限制为 2G。

3.5.5 时分秒针在一天之类重合多少次？（24 小时）

3.5.6 将多个集合合并成没有交集的集合。

3.5.7 平面内有 11 个点，由它们连成 48 条不同的直线，由这些点可连成多少个三角形？