

1. 有一个文件: c:/c.txt,写 java 程序把该文件内容复制两遍,追加到 c:/c.txt;
2. 写正则表达式 1.邮箱 2.数字
3. HashMap 改变 map 类对用户会不会有影响?
4. Linux 中需查看所有的 java 进程,用什么命令
ps -ef|grep java
5. Ajax 请求的整个流程
AJAX 在浏览器与 Web 服务器之间使用异步数据传输 (HTTP 请求),这样就可使网页从服务器请求少量的信息,而不是整个页面。

open(): 建立到服务器的新请求。

send(): 向服务器发送请求。

abort(): 退出当前请求。

readyState: 提供当前 HTML 的就绪状态。

responseText: 服务器返回的请求响应文本。

6. 写一个类实现线程同步的单例设计模式
7. 一个包含 4 块硬盘的服务器一年中至少有一块硬盘出故障的概率是 99.99%, 每块硬盘任意时刻出故障的概率服从相同的分布规律, 并且彼此独立, 问 12 块硬盘的服务器一季度内至少有一个硬盘出故障的概率是多少。
8. 有一个 size1000 的 vector<int>, 删除其中的第 5, 6, 7 号元素, 要求效率高 (C)
9. 数列 L 中有 n 个整数, 其中 K 个数字出现了两次, 1 个数字出现了一次, 所以 $n=2k+1$; 请在使用 $O(1)$ 空间的前提下, 尽快找出只出现一次的那个数字, 并说明算法的复杂度。用异或, 时间复杂度 $O(n)$
10. 有一个文件, 存在 40 亿个不重复的整数 ($0 \sim 4294967295$), 可用内存只有 256M, 32 比特的整数有 4294967295 (约 42.9 亿) 种取值可能, 如何找出不存在的 294967295 (约 2.9 亿) 个数扫描结果数据可存放到文件中, 不占用内存

分段载入内存, 排序, 输出, 一共要扫描文件 $2^{32} / (256 * 2^{20} / 32) = 512$ 遍

BITMAP 分 16 次处理

建 42.9bits 的文件, 按 200m 一段映射, 先遍历 40 亿个数, 检查 n/有 8 字节位置是否在当前映射区, 否则换映射位置, 然后标记。然后读 2.9 亿检查, 都一个道理, 建在共享内存里的 bitmap 而已。

位图算法, 用含有 1 千万个位的字符串来表示这个文件, 文件中有的数据则标识为 1, 没有则标识为 0, 最后从第一位读至最后一位, 即为有序的集合。这种算法充分利用了题目中给的条件, 但也仅仅适合本题目, (不会有重复的数字, 同时不与其余的数进行关联)

问题实例：

1).2.5 亿个整数中找出不重复的整数的个数，内存空间不足以容纳这 2.5 亿个整数。

有点像鸽巢原理，整数个数为 2^{32} ，也就是，我们可以将这 2^{32} 个数，划分为 2^8 个区域 (比如用单个文件代表一个区域)，然后将数据分离到不同的区域，然后不同的区域在利用 **bitmap** 就可以直接解决了。也就是说只要有足够的磁盘空间，就可以很方便的解决。

11. 生成递增整形数字的方法，要求高可用性（当机重启后仍递增）

直接取时间不就是递增的么？他又没说一定只能递增 1。。。

一文件记录最大使用到的数字，例如 N，内存记录当前使用最大数字例如 10，当内存记录使用到 N-20 的时候，往文件记录中写入 N+50. 确保不会出现重复数据。优点不用说，和系统无关，不需要去判定系统是怎么宕机之类的。缺点也有：宕机重启了，那么读取文件中的读数 N，然后往文件中写入 N+50，然后再开始计数。这样可导致部分自增数据没有被用上。