

阿里巴巴集团 2014 校园招聘

阿里巴巴集团

2014 秋季校园招聘

测试开发工程师 北京 笔试卷 (二)

试卷编号: 10707

姓名: _____ 手机: _____

学校 (注明校区) _____ 学历: _____

电子邮件: _____

答题说明:

- 1、答题时间为 120 分钟, 请注意把握时间;
- 2、试卷共分四部分: 单项选择题 (20 题, 50 分)、不定向选择题 (4 题, 20 分)、填空&问答题 (5 题, 30 分); JAVA 选做题 (可选做, 不计分)
- 3、测试开发工程师、算法工程师、系统工程师另有附加题, 附在卷尾, 请留意;
- 4、请将选择题答案涂在答题卡上, 其他部分答案写在答题纸上;
- 5、请务必认真清楚填写个人信息, 勿将试卷带离考场。

应聘职位 (请勾选你感兴趣的一个职位)	<input type="checkbox"/> 研发工程师	<input type="checkbox"/> 算法工程师
	<input type="checkbox"/> 测试工程师	<input type="checkbox"/> 系统工程师
擅长语言 (请勾选你最擅长的一种语言)	<input type="checkbox"/> C&C++ <input type="checkbox"/> JAVA	<input type="checkbox"/> .Net <input type="checkbox"/> PHP
	<input type="checkbox"/> Shell <input type="checkbox"/> Perl	<input type="checkbox"/> Python <input type="checkbox"/> (X)HTML\CSS
	<input type="checkbox"/> Javascript	<input type="checkbox"/> 其它

第一部分 单选题 (前 10 题, 每题 2 分; 后 10 题, 每题 3 分; 共 50 分。选对得满分, 选错倒扣 1 分, 不选得 0 分。)

1. 一次内存访问、SSD 硬盘访问和 SATA 硬盘随机访问的时间分别约是_____。
A、几微秒, 几毫秒, 几十毫秒 B、几微秒, 几十微秒, 几十毫秒
C、几十纳秒, 几十微秒, 几十毫秒 D、几十纳秒, 几十微秒, 十几毫秒
2. 8 进制数 256, 转化成 7 进制数是_____。
A、356 B、346 C、338 D、336
3. 某网络的 IP 地址空间为 192.168.5.0/24。采用定长子网划分, 子网掩码为 255.255.255.248, 则该网络的最大子网个数、每个子网内的最大可分配地址个数为各为_____。
A、32, 6 B、32, 8 C、8, 32 D、8, 30
4. 以下关于链式存储结构说法错误的是_____。
A、比顺序存储结构的存储密度小
B、每个结点是由数据域和指针域组成
C、查找结点时链式存储比顺序存储快
D、逻辑上不相邻的节点物理上可能相邻
5. 假定一个二维数组的定义语句为“int a[3][4]={{3,4},{2,8,6}};”, 则元素 a[1][2]的值为_____。
A、2 B、4 C、6 D、8
6. 下面函数的功能是_____。

```
int fun (char *s)
{
    char *p=s;
    while (*p++);
    return p-s-1;
}
```


A、求字符串的长度 B、复制一个字符串
C、计算字符串的位(bit)数 D、求字符串存放的位置
7. 判断有向图是否存在回路, 利用_____方法最佳。
A、求关键路径 B、求最短路径
C、拓扑排序 D、广度优先遍历

8. 依次读入数据元素序列{a,b,c,d,e,f,g}进栈, 元素进栈或出栈顺序是未知的, 下列序列中, 不可能成为栈空时弹出的元素构成序列的有_____。
- A、{d,e,c,f,b,g,a} B、{f,e,g,d,a,c,b} C、{e,f,d,g,c,b,a} D、{c,d,b,e,f,a,g}
9. 下列有关图的遍历说法中, 不正确的是_____。
- A、图的遍历必须用递归实现
B、基本遍历算法两种: 深度遍历和广度遍历
C、有向图和无向图都可以进行遍历操作
D、图的遍历算法可以执行在有回路的图上
10. 在 16 位机器上跑下列 foo 函数的结果是_____。
- ```
void foo()
{
 int i = 65536;
 cout << i << ", ";
 i = 65535;
 cout << i;
}
```
- A、-1,65535    B、0,65535    C、-1,-1    D、0,-1

(以下每题 3 分)

11. 有一段年代久远的 C++ 代码, 内部逻辑复杂。现在需要利用其实现一个新的需求, 假定有以下可行的方案, 应当优先选择\_\_\_\_\_。
- A、修改老代码的接口, 满足新的需求  
B、在这段代码之外写一段代码, 调用该代码的一些模块, 完成新功能需求  
C、修改老代码的内部逻辑, 满足新的需求  
D、将老代码抛弃, 自己重新实现类似的逻辑
12. 在 5 个页框上使用 LRU 页面替换算法。当页框初始为空时, 引用序列为 0、1、7、8、6、2、3、7、2、9、8、1、0、2, 系统将发生\_\_\_\_\_次缺页。
- A、11    B、12    C、13    D、8
13. 阿里巴巴有相距 1500km 的两机房 A 和 B。现有 100GB 数据需要通过一条 FTP 连接在 100s 的时间内从 A 传输到 B。已知 FTP 连接建立在 TCP 协议之上, 而 TCP 协议通过 ACK 来确认每个数据包是否正确传送。网络信号传输速度  $2 \times 10^8$  m/s, 假设机房带宽足够高, 那么 A 节点的发送缓冲区可以设置为最小\_\_\_\_\_。
- A、6MB    B、12MB    C、18MB    D、24MB

14. 有 3 个节点的二叉树可能有\_\_\_\_\_种。  
A、12                      B、13                      C、5                      D、15
15. 设某文件经内排序后得到 100 个初始归并段(初始顺串), 若使用多路归并排序算法, 且要求三趟归并完成排序, 问归并路数最少为\_\_\_\_\_。  
A、5                      B、6                      C、7                      D、8
16. 一个优化的程序可以生成  $n$  个元素集合的所有子集, 那么该程序的时间复杂度是\_\_\_\_\_。  
A、 $O(n!)$                       B、 $O(n \log n)$                       C、 $O(n^2)$                       D、 $O(2^n)$
17. 快速排序在已经有序的情况下效率最差, 复杂度为\_\_\_\_\_。  
A、 $O(n \log n)$                       B、 $O(n^2 \log n)$                       C、 $O(n^{1.5})$                       D、 $O(n^2)$
18. 从一幅牌(52 张, 不含大小怪)里抽出两张牌, 其中一红一黑的概率是\_\_\_\_\_。  
A、25/51                      B、26/51                      C、1/2                      D、1/3
19. 有一堆石子共 100 枚, 甲乙轮流从该堆中取石子, 每次可以取 2、4 或 6 枚, 取得最后的石子的玩家为赢家, 若甲先取, 则\_\_\_\_\_。  
A、甲必胜                      B、乙必胜                      C、谁都无法必胜                      D、不确定
20. 现有一完全的 P2P 共享协议。每次两个节点通讯后都能获取对方已经获取的全部信息。现在使得系统中每个节点都知道所有节点的文件信息, 共 17 个节点, 假设只能通过多次两个对等节点之间通讯的方式, 则最少需要\_\_\_\_\_次通讯。  
A、29                      B、30                      C、31                      D、32

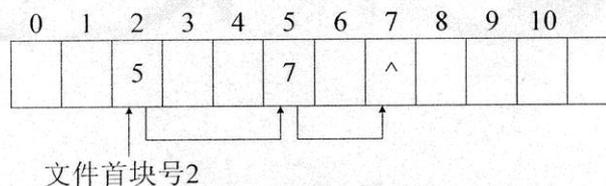
**第二部分 不定向选项 (4 题, 每题 5 分。每题有 1-5 个正确选项, 完全正确计 5 分, 漏选计 2 分, 不选计 0 分, 多选、错选计-2 分)**

21. 2-3 树是一种特殊的树, 它满足两个条件:  
(1) 每个内部结点有两个或三个子结点;  
(2) 所有的叶结点到根的路径长度相同;  
如果一棵 2-3 树有 9 个叶结点, 下列数量个非叶结点的 2-3 树可能存在的有\_\_\_\_\_。  
A、4                      B、5                      C、6                      D、7                      E、8
22. 下列有关进程的说法中, \_\_\_\_\_是错误的。  
A、进程是静态的                      B、进程是动态的过程  
C、进程与程序是一一对应的                      D、进程与作业是一一对应的

23. 下列函数定义中, 有语法错误的是\_\_\_\_\_。
- A、`void fun(int x, int *y){ x *= *y;}`  
 B、`void fun(int x, int *y){ *x *= *y;}`  
 C、`void fun(int *x, int y){ *x += y;}`  
 D、`int* fun(int *x, int y){ return x += y;}`
24. 有朋自远方来。他乘火车, 轮船, 汽车, 飞机来的概率分别是 0.3, 0.2, 0.1, 0.4。坐各交通工具迟到的概率分别是 1/4, 1/3, 1/12, 0。下列语句中正确的是\_\_\_\_\_。
- A、如果他迟到, 乘火车来的概率是 0.5。  
 B、如果他准点, 坐轮船或汽车的概率等于坐火车的概率。  
 C、如果他准点, 那么乘飞机的概率大于等于 0.5。  
 D、坐陆路(火车、汽车)交通工具准点机会比坐水路(轮船)要低。

### 第三部分 填空与问答 (5 题, 共 30 分)

25. (4 分) 文件分配表 FAT 是管理磁盘空间的一种数据结构, 用在以链接方式存储文件的系统中记录磁盘分配和跟踪空白磁盘块。整个磁盘仅设一张 FAT 表, 其结构如下图所示。如果文件块号为 2, 查找 FAT 序号为 2 的内容得知物理块 2 的后继物理块是 5; 再查 FAT 序号为 5 的内容得知物理块 5 的后继物理块是 7; 接着继续查 FAT 序号为 7 的内容为“^”, 即该文件结束标志, 所以该文件由物理块 2、5、7 组成。



假设磁盘物理块大小为 1KB, 并且 FAT 序号以 4bits 为单位向上扩充空间。请计算下列两块磁盘的 FAT 最少需要占用多大的存储空间?

- 1) 一块 540MB 的硬盘
  - 2) 一块 1.2GB 的硬盘
26. (4 分) 已知如下代码, 并在两个线程中同时执行 f1 和 f2, 待两个函数都返回后, a 的所有可能值是哪些?

```
int a = 2, b = 0, c = 0;

void f1()
{
 b = a * 2;
 a = b;
}

void f2()
{
 c = a + 11;
 a = c;
}
```

27. (6分) 设计一个最优算法来查找  $n$  个元素数组中的最大值和最小值。已知一种需要比较  $2n$  次的方法, 请给一个更优的算法。请特别注意优化时间复杂度的常数。

给出该算法最坏情况下的比较次数和该算法的步骤描述。

(不用写代码, 不给出比较次数的不得分)

28. (8分) 已知三个升序整数数组  $a[1]$ 、 $b[m]$  和  $c[n]$ 。请在三个数组中各找一个元素, 使得组成的三元组距离最小。三元组的距离定义是: 假设  $a[i]$ 、 $b[j]$  和  $c[k]$  是一个三元组, 那么距离为:

$$\text{distance} = \max(|a[i] - b[j]|, |a[i] - c[k]|, |b[j] - c[k]|)$$

请设计一求最小三元组距离的最优算法, 并分析时间复杂度。

(不用写代码, 不分析时间复杂度不得分)

29. (8分) 在黑板上写下 50 个数字: 1 至 50。在接下来的 49 轮操作中, 每次做如下动作: 选取两个黑板上的数字  $a$  和  $b$ , 擦去, 在黑板上写  $|b - a|$ 。请问最后一次动作之后剩下数字可能是什么? 为什么?

(不用写代码, 不写原因不得分)

#### 第四部分：JAVA 选做题（注：阿里有大量 JAVA 研发工程师需求；选作以下

题目有机会增加该方向面试机会）

1、请给出下面代码的输出

```
public class Test1 {
 public static int k=0;
 public static Test1 t1=new Test1("t1");
 public static Test1 t2=new Test1("t2");
 public static int i=print("i");
 public static int n=99;
 public int j=print("j");
 {
 print("构造块");
 }

 static{
 print("静态块");
 }

 public Test1(String str){
 System.out.println(++k+"."+str+" i="+i+" n="+n);
 ++i; ++n;
 }

 public static int print(String str){
 System.out.println(++k+"."+str+" i="+i+" n="+n);
 ++n;
 return ++i;
 }

 public static void main(String...strings){
 Test1 t=new Test1("init");
 }
}
```

2、已知： 为了保持请求的并行度，一台 web 应用服务器目前需要 10 个到一台 oracle 的数据库连接，目前 oracle 数据库能够稳定承载的连接数是 7000 个。而目前已经拥有的 web 应用服务器数是 400 台，并且 web 服务器以每年以 300 台的规模在增长。

问题： 目前能够采取的策略有哪些？请从原理上对这些策略为什么能够起到作用进行一下分析。

## 阿里巴巴集团 2014 校园招聘

### 【测试开发工程师】附加题 (0922)

#### 题目:

新浪 Weibo 的开发平台提供了一组 API。其提供**获取用户信息接口**, 描述如下,

users/show, 根据用户 ID 获取用户信息

URL, <https://api.weibo.com/2/users/show.json>

请求参数,

|              | 必选    | 类型及范围  | 说明                                               |
|--------------|-------|--------|--------------------------------------------------|
| source       | false | string | 采用 OAuth 授权方式不需要此参数, 其他授权方式为必填参数, 数值为应用的 AppKey。 |
| access_token | false | string | 采用 OAuth 授权方式为必填参数, 其他授权方式不需要此参数, OAuth 授权后获得。   |
| uid          | false | int64  | 需要查询的用户 ID。                                      |
| screen_name  | false | string | 需要查询的用户昵称。                                       |

注意事项, 参数 uid 与 screen\_name 二者必选其一, 且只能选其一

题目: 假设你是新浪 Weibo 开发平台的一名测试人员, 你如何测试这个 API? 思路即可,

有代码更佳。