

2019 清华大学自主命题

计算机专业基础综合（912）回忆

一、数据结构（70 分）

1. 判断题（12*2=24）

- (1) $n^{\log \log \log n} = O([\log n]!)$
- (2) 交换哈夫曼树的不同深度的节点，编码长度必然改变
- (3) 即使不使用改进的 next 表，kmp 依然可以达到线性的时间复杂度
- (4) 对于不符合局部性原理的访问，splay 的分摊复杂度不是 $\log n$
- (5) 对于二叉树，通过先序遍历和后序遍历不能确定其层次遍历
- (6) 对于叶节点为 2019 的真二叉树，其数量小于 2018 对括号所组成的合法表示式数量
- (7) 对于叶节点数量为 2018 的二叉树，层次遍历队列容量必然小于 2018
- (8) 插入排序每次插入数据，即使不增加循环节，也不至减少
- (9) 交换两个逆序对，必然会减少总逆序对数
- (10) 如果基数排序底层采用不稳定的算法，那么得到的结果可能是不正确的
- (11) 函数的调用栈中如果有相同的函数，则他们必然紧邻
- (12) 如果插入的关键词独立均匀分布，堆的插入操作平均 $O(1)$

2. 简答题，每题回答不超过 80 字（8*4=32）

- (1) 逆波兰表达式的优点？既然中缀转换为逆波兰就需要可以计算出表达式值，那逆波兰意义何在？
- (2) DFS 中何时标记前向边？何时标记后向边？
- (3) 相比锦标赛排序，败者树的优势是？
- (4) 试举出红黑树优于 AVL 树的场景，红黑树相比 AVL 树的优势
- (5) 相比开散列，闭散列的优势，试举例说明两点。
- (6) 相比选择排序，插入排序的优势，试举例说明两点。
- (7) 对于稠密图，迪杰斯特拉应该使用多叉堆替换二叉堆，为什么？多叉堆分叉数 m 怎么确定？
- (8) 在何种情形下，KMP 优于蛮力算法，为什么？

3. 算法题 (7+3+4 = 14)

返回后序遍历的第 K 个节点，时间复杂度不超过 x 的深度， $O(\text{depth}(x))$

```
struct BinNode{

    int size;    //当前节点和孩子总数

    BinNode *lchild,*rchild;

};

BinNode *rank(BinNode* t,int k){

    //有效代码行数不超过 12 行

    //不要尝试模拟后序遍历，时间复杂度会超时。

}
```

一，给出具体算法实现。

二，解释你的算法。

三，分析时间复杂度和空间复杂度。

二、操作系统（30 分）

1. 填空题（15*1=15）

(1) stride调度算法中，进程分配时间和其优先级成反比，优先级越低进程上CPU运行的时间就会更长，用八位无符号数来表示进程A的运行时间有可能会溢出，但是当步长s [1] 127时，只要做小小的技巧就仍可以正确判断时间，然后又引入进程B，a为A的stride，b为B的stride，当sign(a-b) [2] 时说明A正在运行，此时a+s [3] 256，当sign(a+s-b) [4] 0时，轮到B运行。

(2)CPL<=DPL[门]和CPL>=DPL[段]， [5] 表示请求时可以和门特权级相同， [6] 表示请求时应低于段的特权级。

(3)父进程先退出，还未退出的子进程被称为 [7] 。子进程退出时，其父进程还没wait，此时子进程被称为 [8] 。

| | | |
|-------------------|-------------------------|---------------------------|
| (4) | Semaphore::P() { | Semaphore::V() { |
| class Semaphore { | [9] ; | [11] ; |
| int sem; | if ([10]) { | if ([12]) { |
| WaitQueue q; | Add this thread t to q; | Remove a thread t from q; |
| } | block(t); | wakeup(t); |
| | } | } |
| | } | } |

(5) x86-32CPU的硬件组成，CR3寄存器用于存储页目录表起始 [13] 。

(6)A是计数为1的文件，创建A的硬链接B，再创建A的软连接C，在创建B的硬链D，B计数 [14] ,D计数 [15] 。

2. 判断题（5*1=5）

- (1) x86开启了二级页表，则一级页表可以不在内存中。
- (2)ucore时钟中断为10ms，故不能完成 10ms以下的定时。
- (3)中断向量表中存放着中断门的优先级。
- (4)不安全状态就是死锁状态
- (5)只有一个main函数的程序不能有多线程

3. UCORE （6）在以下代码中找到页表切换，堆栈切换，switch_to 函数获得两个参数的代码部分并注释：

| | |
|--|---|
| <pre>1 .text 2 .globl switch_to 3 switch_to: # switch_to(from, to) 4 5 # save from's registers 6 movl 4(%esp), %eax # eax points to from 7 popl 0(%eax) # save eip !popl 8 movl %esp, 4(%eax) 9 movl %ebx, 8(%eax) 10 movl %ecx, 12(%eax) 11 movl %edx, 16(%eax) 12 movl %esi, 20(%eax) 13 movl %edi, 24(%eax) 14 movl %ebp, 28(%eax) 15 16 # restore to's registers 17 movl 4(%esp), %eax # not 0(%esp): popped return address already 18 # eax now points to to 19 movl 28(%eax), %ebp 20 movl 24(%eax), %edi 21 movl 20(%eax), %esi 22 movl 16(%eax), %edx 23 movl 12(%eax), %ecx 24 movl 8(%eax), %ebx 25 movl 4(%eax), %esp 26 27 pushl 0(%eax) # push eip 28 29 ret</pre> | <pre>225 proc_run(struct proc_struct *proc) { 226 if (proc != current) { 227 bool intr_flag; 228 struct proc_struct *prev = current, *next = proc; 229 local_intr_save(intr_flag); 230 { 231 current = proc; 232 load_esp0(next->kstack + KSTACKSIZE); 233 lcr3(next->cr3); 234 switch_to(&(prev->context), &(next->context)); 235 } 236 local_intr_restore(intr_flag); 237 } 238 }</pre> |
|--|---|

4. 内存分析（4）给了段内存的 dump，以及一些页表项的信息，分析 A、B 进程的逻辑地址转换

三、组成原理 (30分)

1. 选择题

(1)下列哪一项没有容错能力

- A. RAID0 B. RAID1 C. RAID5 D. RAID6

(2)下列关于静态存储器和动态存储器的描述正确的是

- A.静态存储器使用触发器，需要定期刷新
B.静态存储器使用电容，不需要定期刷新
C.动态存储器使用触发器，不需要定期刷新
D.动态存储器使用电容，需要定期刷新

(3)下列哪个是对的

- A. 虚拟内存空间比实际的地址空间大
B. 虚拟内存空间比实际的地址空间小
C. 虚拟内存空间连续存放，实际内存一定连续存放
D. 虚拟内存空间不连续存放，实际内存有可能连续存放

(4)下面总线说法哪个正确 ()

- A.并行总线速度大于串行
B.异步总线速度大于同步
C.单总线速度大于双总线
D.以上说法均错误

(5)MIPS 五级流水中，有哪个数据冲突 ()

- A.RAR B.RAW C.WARD.WAW

2. 判断题

(1) C语言中， $\text{int } x > 0$,则 $x * x > 0$

(2) CPI越少的机器执行同一个程序的时间越短

(3) 对机械磁盘，读取顺序存储的文件比随机存储的文件快

(4) 冯诺依曼是数据和指令混合存储的

(5) MIPS五级流水线设计中，使用充分设置功能单元的方法可以改善结构冲突

3. 填空题

(1) +1234的补码为[](用十六进制，小端表示)

(2) 27.625的IEEE754为[](十六进制)

(3) 解决数据冲突的三种方法[] [] []

(4) cache的三种缺失为[] [] []

4. 流水线时间计算

MIPS 处理器内存延迟 10ns, ALU 延迟 6ns, 寄存器 3ns, 输入延迟 1ns, 流水线寄存器以及多周期锁存器输出延迟为 2ns, 要有计算过程, 以下是指令(可能有个别字母不对)

- 1.addn vd rs rt
- 2.suburd rs rt
- 3.ori rt rs rimm
- 4.lw rt rs imm
- 5.sw rt rs imm
- 6.beg rs rt imm
- 7.j target1.

- 1.按照单周期设计, 指令内存与数据内存分开, 计算指令延迟?
- 2.按照多周期设计, 指令内存和数据内存存在同一个内存模块, 最长和最短的指令延迟分别是指哪条指令, 分别计算对应的延迟
- 3.按照五级流水线设计, 指令内存和数据内存存在同一个内存模块, 处理器频率最高能到多少

四、计算机网络（20分）

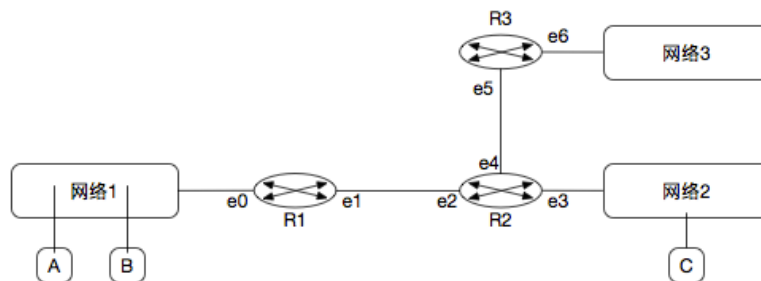
1. 选择题（1*6=6）

- (1) 电话网络和TCP网络的性质的比较，电话网应该是建立电路
- (2) 蜂窝移动网络六边形，频率840HZ，则每个点可用最大频率
- (3) 一个给了最小数据帧长度的以太网，给了信号传播速度，给了相距，问最大传输数据速度
- (4) 计算一个计算分组时间，两个转发加两个传播
- (5) 什么信道的利用率最低？相距离越远，传输速率越快
- (6) 拥塞窗口为m时发生拥塞，发送方有足够多的数据要发，问平均速度

2. 简答题（2+1+1=4）

- (1) 解释以下URL各部分的意义 `http://info.tsinghua.edu.cn:80/index.jsp`
- (2) 如域名info.tsinghua.edu.cn对应的ip为166.111.4.98，解释为何会发生如下现象：
 - ①访问`http://info.tsinghua.edu.cn/index.jsp` 正常，而访问`http://166.111.4.98/index.jsp` 异常
 - ②访问`http://166.111.4.98/index.jsp` 正常，而访问`http://info.tsinghua.edu.cn/index.jsp` 异常

3. 子网划分（6）



- (1) 如图，网络1有100台主机，网络2有50台，网络3有20台，请将166.111.4.0/24划分给网络并写出路由器接口ip
- (2) 简述AB通信时与AC通信时使用ARP协议的具体情况
- (3) 当A发送报文给C时写出各个段上报文的源IP，目的IP，源MAC，目的MAC（用MAC-A,IP-A,MAC-e0等表示）