

23-912 真题群友合力回忆版

23-清华计算机考研(912)真题-回忆版,由新威考研整理.

感谢 Ding, 雨鑫, 柯南等提供主要是回忆版的主要部分,感谢其他大佬在线直播回忆和完善其余部分.

本版本仍然可能有错误,会及时修正,最新版本加群获取:790461315

数据结构

一. 判断题(2分×8道题=16分)

1. 向量只要在即将溢出前使容量加倍,并在装填因子不足 50%时容量减半,便可保证消耗在容量调整的分摊时间复杂度为常数()
2. 插入排序可实现就地(in-place)算法
3. 按照二叉搜索树规则在 AVL 树中插入一个节点后,若造成失衡了,且在定位祖孙三代节点 g、p、v 后,只需要围绕 g 做一次旋转即可使局部恢复平衡了,则在旋转之前节点 p 的平衡因子不可能是零
4. 规模同为 2023 的 AVL 树和红黑树各做一次搜索,在最坏情况下后者的搜索长度比前者更长
5. 如果访问序列不满足局部性, Splay 树未必还能保证分摊 $O(\log n)$ 的时间性能.
6. 相对于 kruskal 算法, Prim 算法更适宜于稀疏图
7. 有有 n 个词条的跳转表(skiplist)所占用的空间了,在极端情况下会大于 $\Omega(n^2)$
8. 只要存在负权变,则 Dijkstra 算法即使能够终止,所输出的也不是最短路径.

二. 填空题(4分×6道题=24分)

1. 若某算法的运行时间满足 $T(n) = 128T\left(\frac{n}{2^{2023}}\right) + {}^{290}\sqrt{n}$, 则 $T(n) = \underline{\hspace{2cm}}$.
2. 如下逆波兰表达式的值若为 2023, 则被遮挡的那个运算符是 .

0		!	1	2	+	3	4	+	?	+	5	!	6	7	+	8	-	9	*	+	-
---	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

必须填写运算符及其对应的中文(+加法, -减法, *乘法, /除法, ^乘方, ! 阶乘)

3. 若仅使用 9 个字符,且他们出现的频率分别为{1,1,2,3,5,8,13,33,36},则按照 Huffman 算法对长度为 2023 的随机字符串做编码,所需比特位数的期望值是 .
4. 在包含 2023 个节点的左式堆中,右子堆最多可能有 个节点
5. 设文本串 T, 模式串 P 的长度分别为 2023 和 17, 且 T 并不包含 P, 借助 Boyer-Moore 算法的

坏字符表 bc（好后缀表 gs 不用）来做模式匹配，最好情况下只需做____次字符比对。

6. 采用独立链(Scparete chaining)法的某散列表长度为 119，在存在 2023 个词条时，一次成功查找的过程中所需试探次数的期望值是_____。

三. 计算（10 分）

试采用 Floyd 算法将以下向量转换为大顶堆，并按如下格式，依次给出向量在各次下滤之后的内容：

K	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
向量 A[k]	3	8	0	6	1	13	11	2	9	5	4	14	7	10	12
第 1 次下滤后															
第 2 次下滤后															
第 3 次下滤后															

相对于上一行没有变化的元素直接留白，不要重复填写

四. 算法题（7+10+3=20 分）

在以 r 为根的二叉树中，任一节点 v 与 r 之间通路上的所有节点（包括 r 及 v 本身）都成为 v 的祖先。如果节点 a 同时是节点 v 的和 u 的祖先，则称作 v 和 u 的公共祖先（当然 r 是所有节点的公共祖先）其中的深度最大值，称作最低公共祖先（Lowest Common Ancestor）

假定存在有 n 个节点的二叉树中，各节点的关键码都是取自[0,n)的正数，且彼此互异，尽管不能直接访问该树，但其先序，后序遍历序列已分别记录在数组 P[0,n)和 Q[0,n)中，试设计一个算法对[0,n)内的任意一对整数关键码 x 和 y，找出它们所对应节点的最低公共祖先 LCA(x,y)，要求算法的运行时间和辅助空间均不超过 O(n)

比如，若 $P=\{3,0,8,5,2,6,7,10,4,1,9\}$, $Q=\{5,8,2,0,10,4,7,9,1,6,3\}$ ，则有

$LCA(2,5)=0, LCA(8,4)=3, LCA(6,9)=6$

- (1)试通过文字与图示，说明算法的原理
- (2)试用伪代码描述算法的过程，并提供足够的注释
- (3)试证明算法的时间，空间复杂度均符合要求。

组成原理

一. 判断题（1 分×5 道题=5 分）

1. 用补码表示的 32 位正数加法满足交换律（）

2. IEEE754 双精浮点数加法满足结合律 ()
3. IO 端口和主存可以统一编址 ()
4. 软件管理高速缓存 Cache 与主存储器地址之间的映射关系 ()
5. 虚拟存储系统可以加快磁盘的存取速度 ()

二. 填空题 (2 分×5 个空=10 分)

1. 三位二进制数据 101 的带全局校验的海明码为_____.
2. 指令流水线冲突包括结构冲突, 数据冲突和_____冲突
3. 某 MIPS 采用标准 5 级流水线设计, 且指令存储器和数据存储器分离.假设一个时钟周期就可以完成 5 级流水线中的任意阶段, 同一周期内, 寄存器堆写入寄存器的数据即可正确读出, 执行下面的指令序列时:

addi \$2, \$3, 4; 3 号寄存器与 4 求和, 存入 2 号寄存器

lw \$1,0(\$2); 装入 2 号寄存器指向的内存数据到 1 号寄存器

sub \$4,\$1,\$5; 1 号寄存器与 5 号寄存器求差, 存入 4 号寄存器

and \$5,\$1,\$4; 1 号寄存器与 4 号寄存器求与, 存入 5 号寄存器

若处理器没有数据旁路, 则至少需要插入_____个气泡,若有数据旁路, 则至少需要插入_____个气泡.

4. 某系统总线可在一个总线周期内并行传输 16B 的数据, 一个总线周期占用 3 个时钟周期, 总线的时钟频率为 66MHz, 则总线带宽是_____.

三. 选择题 (2 分×5 道题)

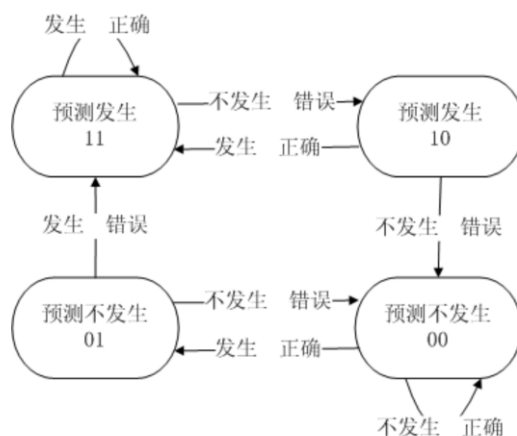
1. 应用程序执行的过程中发生了中断或者异常, 假设在指令 A 处发生了中断或者异常, 以下描述正确的是 ()
 - A. 如果返回到该应用程序, 则开始执行 A 的下一条指令
 - B. 如果返回到该应用程序, 则从指令 A 处开始执行 (A 被再一次执行)
 - C. 不会返回到该应用程序执行
 - D. 以上都有可能
2. 5 级流水线处理器实现中, 指令流经各个功能段的时间分别为 60ns, 50ns, 40ns, 60ns, 30ns 忽略流水线寄存器的延迟, 则该处理器的时钟周期至少是 ()
 - A. 60ns
 - B. 50ns

- C. 40ns
- D. 30ns
3. 下列关于闪存及 FTL(Flash Translation Layer)的说法, 错误的是 ()
- A. 闪存的写入端粒度和擦除粒度不同
 - B. 闪存每单元能编码的比特数越多, 闪存的寿命越长
 - C. FTL 提供逻辑块地址到物理块地址的映射
 - D. 闪存物理页在更新之前要进行擦除
4. 以下关于通过 DMA 方式进行数据传输的描述, 错误的是 ()
- A. DMA 用于数据的快速传输, 可用来代替中断方式
 - B. DMA 控制器与 CPU 交替使用总线
 - C. DMA 可以直接访问主存
 - D. 数据传输过程由 DMA 自行控制
5. 以下字符描述错误的是 ()
- A. ASCII 码使用一个字节进行编码
 - B. UTF-8 码使用两个字节进行编码
 - C. 字符编码不包括字符如何显示的信息
 - D. 使用矢量字体的字符放大后不会失真

四. 计算 (5 分)

分支预测包含了静态预测以及动态预测, 在下面的 MIPS 程序中, 计算不同预测方法下的预测准确率

- (1) 采用静态预测, 预测转移发生
- (2) 采用 1 位预测位, 初始预测转移不发生
- (3) 采用 2 位预测位, 初始预测转移不发生



程序 (\$0 恒为 0):

addiu \$8,\$0,1 $\$0 + 1 \rightarrow \8

addiu \$9,\$0,0 $\$0 + 1 \rightarrow \9

Loop:

add \$9,\$9,\$8 $\$8 + \$9 \rightarrow \$9$

addi \$8,\$8,1 $\$8 + 1 \rightarrow \8

addi \$1,\$8,-1 $\$8 - 1 \rightarrow \1

slti \$1,\$1,100 $\$1 = \begin{cases} 1, \$1 < 100 \\ 0, \$1 \geq 100 \end{cases}$

bne \$1,\$0,loop 若 $\$1 \neq \0 , 则 jump 到 loop

nop 空

操作系统

一. 判断题 (1 分 $\times$ 10 道题)

1. qemu-system-riscv64 可以模拟 64 位 riscv 处理机并让操作系统在其上运行 ()
2. 若应用程序执行中外设产生中断, 则操作系统接管 CPU 控制权并完成中断处理 ()
3. 进程是应用程序的一次执行过程 ()
4. GDB 调试不需要操作系统的服务来跟踪和中断被调试应用程序的执行 ()
5. 信号量可用于实现同步 (synchronization) 比如将信号量的初始值设为 0 若线程 A 对此信号量执行一个 P 操作那么该线程应当被阻塞睡眠 ()
6. 操作系统中的文件系统必须具备打开文件, 关闭文件, 读文件, 写文件的功能 ()
7. 在没有 MMU 的 CPU 上 Linux 内核可正常运行并具备基于页表机制的虚拟存储管理能力 ()
8. 应用程序每次调用 open() 函数打开文件都要通过系统调用向操作系统提出请求 ()
9. 在多核处理器中, 采用用户线程实现多线程时, 同一进程的多个线程可以并行执行 ()
10. 在多处理器场景下控制中断的打开与关闭, 可以保证只有一个线程进入临界区 ()

二. 选择题

1. 下列组件不属于操作系统内核的是

A. JavaScript 解释器

- B. 文件系统
- C. 中断处理例程
- D. 进程调度器

2. 下列关于操作系统的描述中错误的是

- A. 支持多核处理器的操作系统中允许多个执行程序并行执行
- B. 支持多核处理器的操作系统中允许多个执行程序并发执行
- C. 仅支持单核处理器的操作系统中允许多个执行程序并行执行
- D. 仅支持单核处理器的操作系统中允许多个执行程序并发执行

3. riscv 处理器在 U 态（用户态）下执行 `ecall` 指令会陷入到 S 态（内核态）中下列关于硬件会自动完成的操作中错误的是

- A. 保存进程的页面基址
- B. 保存 `ecall` 执行指令的所在地址
- C. 保存陷入原因“在 U 态执行 `ecall` 指令”
- D. 转跳到异常处理例程的入口地址

4. 下面有关进程切换错误的是

- A. 进程切换是来自两个不同进程在内核态中控制流之间的切换
- B. 进程切换的过程可以在内核态执行也可以在用户态执行
- C. 外设产生的中断会发生进程切换
- D. 进程切换的过程涉及到处理器中寄存器的保存与恢复

5. 下列关于 `fork()` 系统调用实现的说法错误的是

- A. 父子进程在返回到用户态的瞬间，都处于刚刚从一次系统调用返回的状态，但是二者的返回值不同
- B. `fork()` 系统调用会创建一个不同于父进程的子进程的内存栈
- C. `fork()` 系统调用会给予子进程创建一个进程块，其中 `TrapContext` 部分的内容和父进程相同
- D. 父进程创建的子进程后，二者的进程 ID 一定不同

三. 大题 (3+3+4)

虚拟内存管理是操作系统中非常重要的组成部分, 请回答以下问题

1. 请阐述操作系统中虚拟内存管理的基本功能
2. 请说明 `mmap()` 和 `munmap()` 系统调用的含义 (我确定这里是含义不是功能) 及基本实现机制
3. 某 32 位处理机中虚拟页面存储方式的页面大小为 4KB, 采用 32 位虚拟地址, 34 位物理地址, 二级页表, 每个页表包含 1024 项, 页表项大小为 4B, 请回答以下问题
 - A. 请画图并简要描述虚拟地址转到物理地址的地址转换过程
 - B. 请在图中标注和描述页表项各部分内容的含义和相关控制状态寄存器的作用

计算机网络

一. 判断题 (1 分×5 道题)

1. 域名解析时递归查询不一定会访问根域名服务器 ()
2. 光网络只能通过导向型介质传播 ()
3. P-坚持型 CSMA 必须在有闲槽的情况下使用 ()
4. 无线局域网面对暴露站问题可采用 RTS/CTS 解决 ()
5. 中继器是物理层中的网络互连设备 ()

二. 选择题 (1 分×5 道题)

1. 在以太网中下列那种方式是避免转发回路
 - A. 水平分割
 - B. 生成树
 - C. 指数退避
 - D. 汇聚树
2. 下列那种不是 VLAN 划分方式
 - A. 基于端口
 - B. 基于 MAC
 - C. 基于端口号
 - D. 网络地址
3. 客户端代理发送邮件所采用的协议
 - A. SMTP
 - B. POP3
 - C. IMAP
 - D. SNMP
4. DNS 缓存会造成 ()
 - A. 可以提高响应性, 没有一致性问题
 - B. 可以提高响应性, 有一致性问题
 - C. 不能提高响应性, 没有一致性问题
 - D. 不能提高响应性, 有一致性问题
5. 下列不属于信道复用技术的是 ()
 - A. 频分复用
 - B. 时分复用
 - C. 码分复用
 - D. 统计复用

三. 计算题 (10 分)

假定客户主机与 web 服务器间的瓶颈链路容量为 10 分组/秒, TCP 链接管理分组和 HTTP 请求分组的发送时间和协议栈处理时间均忽略不计, web 服务器支持的并发链接数与并发请求数均大于 5, TCP 拥塞控制算法任何时间都可以充分利用带宽客户浏览器下载封装在 10 个分组的 HTML 页面文件, 页面内嵌多个服务器本地资源文件, 可以封装在 20, 30, 40 个分组中传播.

服务器最大速率 10 个分组每秒, 往返时延 1 秒, TCP 并发请求均大于 5, 每个网页大小 10 个分组, 其中包含 3 个服务器本地资源文件, 资源文件可以分别通过第 20,30,40 个分组发送

- 1.采用 http1.0 时, 用户下载本地资源和文件(所需要的时间为多少)最短响应时间多少
- 2.采用 http1.1 时, 用户完成下载本地资源和文件(所需要时间为多少)最短响应时间多少
- 3.采用 http2.0, 平均需要多少响应时间完成页面和资源文件下载最短需要多少时间(响应时间定义为各资源文件下载完成时间的平均值), 那么这 3 个资源文件下载完成的最小响应时间是多少?