

2024 年清华大学 912 真题【知研回忆版】


数据结构

一、判断题（每题 1 分，共 12 分）

1. 排序算法的稳定性是指算法能够保持输出序列中的等值元素在原输入序列中的相对次序不变的特性。（ ）
2. 在长度为 n 的向量中删除第 r 个元素 ($0 \leq r < n$)，时间复杂度为 $O(1+r)$ 。（ ）
3. 向量的起泡排序不属于就地算法 (In-place Algorithm)。（ ）
4. 基于列表派生出的栈，无论栈顶如何设置，pop（出栈）和 push（进栈）接口的时间复杂度都是 $O(1)$ 。（ ）
5. AVL 树进行一次删除操作时，可能需要旋转的次数为 $\Omega(\log n)$ 。（ ）
6. 在一棵 B 树中，最多结点的分支数目不会超过最少结点的分支数目的 2 倍。（ ）
7. 伸展树为单链时，其势能达最小。（ ）
8. 左式堆能兼顾 merge、delmax 和 insert 接口，使三者时间复杂度同时优化到 $O(\log n)$ 。（ ）
9. 对 n 个整数排序时，胜者树可以在 $O(n \log n)$ 时间内完成，而败者树不行。（ ）
10. 采用双平方试探策略处理散列表冲突时，散列表长应取模 4、余 1 的素数。（ ）
11. 若一个有向图存在拓扑排序，则该图未必无环。（ ）
12. 在 BM 算法中，坏字符表的长度与模式串的长度无关。（ ）

二、选择题（每题 3 分，共 18 分）

1. 若以下逆波兰表达式的值为 2024，则被遮挡的运算符是（ ）。

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

- A. + B. - C. * D. / E. ^

2. 出现频率为{1,2,3,3}的一组字符，其 PFC 编码方案共有（ ）种。

- A. 16 B. 24 C. 32 D. 40

3. 在输入数据是逐个给出的应用场景中，（ ）排序算法的效率最高？

- A. 插入排序 B. 选择排序 C. 归并排序 D. 快速排序

4. 具有 n 个元素的 Skiplist（跳跃表），其期望高度为（ ）。

- A. $O(n)$ B. $O(\log n)$ C. $O(1)$ D. $O(n \log n)$

5. 某一散列表使用线性试探的方法处理冲突，其数据规模是 M ，词条数量为 N ，懒惰删除标记位为 L ，则删除失败时，期望时间复杂度是（ ）。

- A. $O(\frac{M}{M-N-L})$ B. $O(\frac{M}{M-N+L})$ C. $O(\frac{M}{M-N})$ D. $O(\frac{M-N}{N-L})$

6. 若 N 为记非负整数集，则采用以下（ ）增量序列进行希尔排序，其排序趟数为 $\Omega(\log^2 n)$ 。

- A. $\{2^p | p \in \mathbb{N}\}$ B. $\{2^{p+1}-1 | p \in \mathbb{N}\}$
C. $\{2^p \times 3^q | p, q \in \mathbb{N}\}$ D. $\{9 \times 4^p - 9 \times 2^p + 1 | p \in \mathbb{N}\} \cup \{4^p - 3 \times 2^p + 1 | p \in \mathbb{N}\}$

三、填空题（每题 5 分，共 20 分）

1. 将以 x 为根的子树的结点个数记为 $x.size$ ，如果将二叉树中满足 $x.size \times 2 \leq x.parent.size$ 的非根节点 x 称作“轻节点”，则结点总数为 2024 的二叉树中，一个节点最多可能有（ ）个轻的真祖先。

2. 由（ ）个元素构成栈混洗序列，与由 2024 个括号构成合法表达式一样多。

3. 一棵二叉搜索树有 8 个叶结点，且内部结点、平衡因子都是+1，则该二叉树共有（ ）个结点。

4. 在 KMP 算法中，模式串“BABABBBBABBBABBA”的改进版 $next[]$ 表最大项与最小项的差为（ ）。

5. 以下函数时间复杂度为（ ）。

```
void F(intn){    //假定字宽足够，运算不会溢出，且每次只需常数时间

    for (int i=1,j=1; i < n; j <= 1)

        i <= j;

}
```

四、算法题（1-6 小题分值分别为 2、2、8、4、2、2，共 20 分）

1. 所谓的有根有序树 T ，都有唯一的根节点 $r = \text{root}(T)$ ，且子树可编号为 $\{T_1, T_2, \dots, T_d\}$ 。子树总数 d 亦称作 r 的度数。可递归定义 T 的先序遍历序列，其中“+”为序列串接运算，如下：

$$\text{pre}(T) = \{r\} + \text{pre}(T_1) + \text{pre}(T_2) + \dots + \text{pre}(T_d)$$

用类似的方法定义后序遍历序列，如下：

$$\text{post}(T) = \text{post}(T_1) + \text{post}(T_2) + \dots + \text{post}(T_d) + \{r\}$$

请回答以下问题：

- （1）已知某有根有序树 T 的结点用整数编号，其先序遍历序列为 $\text{pre}(T) = \{9, 2, 3, 4, 1, 5, 8, 7, 6, 0\}$ 对应的结点度数序列，按上述次序给出了各结点的度数： $\text{deg}\{T\} = \{3, 2, 0, 0, 1, 3, 0, 0, 0, 0\}$ 试画出该树的结构（子树必须自左而右“有序”）。
- （2）试给出该树的后序遍历序列 $\text{post}(T)$ 。
- （3）试用伪代码结合足够的注释，描述如下算法：

```
pre2post(stack<int>&pre, stack<int>&deg, stack<int> &post);
```

给出任意一棵有根有序树的先序遍历序列，以及对应的节点度数序列，计算出对应的后序遍历序列。初始时先序遍历序列，节点度数序列分别自顶而下地存放在栈 pre 、 deg ， post 栈为空；算法终止时， pre 、 deg 栈都应为空，后序遍历序列则应自底而上存放于 post 栈中。算法运行时间应线性正比于树的规模。算法过程中任何时刻的空间占用了（含调用栈），相对于初始时的空间占用量不超过

$O(1)$ 。

- (4) 试证明算法的正确性。
- (5) 试证明算法的时间复杂度符合要求。
- (6) 试证明算法的空间复杂度符合要求。

计算机组成原理

一、判断题（每题 1 分，共 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、选择题（每题 2 分，共 10 分）

- 1. 指令系统操作数的长度为 6 位，有 14 条一地址，253 条二地址，20 条三地址，则指令的长度至少为（ ）。
- A. 23 B. 24 C. 26 D. 32
- 2. 若某一程序的功能是将 1 到 100 进行求和，则当程序执行时，下列关于 MIPS 流水线的描述中正确的是（ ）。
- A. beq 作为跳转指令，可能会有控制冲突；
- B. 相邻的两轮循环可能因为读写相同的寄存器造成结构冲突；
- C. 分支缓冲优化对这个程序不起作用；
- D. 数据旁路无法完全消除该程序中数据冲突带来的性能损失
- 3. 指令在（ ）阶段不会发生异常。

A.取指 B.执行 C.访存 D.写回

4.若 x 和 y 是补码, 则 ()

A. 若 $-x < 0$ 则 $x > 0$ B. 若 $-x > 0$ 则 $x < 0$

C. 若 $-x < -y$ 则 $x > y$ D. 若 $-x > -y$ 则 $x < y$

5.某磁盘的转速为 10000 转/min, 传输速率为 20MB/s, 平均寻道时间为 7ms, 控制器延迟为 0.3ms, 则该磁盘读取 4KB 扇区数据时, 所需要的时间为 ()。

A.10.1ms B.10.5ms C.11.5ms D.13.5ms

三、填空 (每题 1 分, 共 5 分)

1.在 IEEE754 单精浮点数表示中, 最小非规格化数的十六进制表示为 ()。

2.在经典五级流水线 MIPS 处理器中, 无法通过数据旁路方式消除数据冲突指令序列是_____、_____ (请填写两条指令)。

3.RAID5 同一条带中数据块对应位置的二进制数据分别为 1011,0110,1001,1110, 那么校验块对应位置的数据为_____。

4.已知 Cache 命中的访问时间为 2ns, Cache 缺失访问时间为 50ns, CPU 执行一段程序时候, CPU 访问内存系统共 10000 次, 其中缺失次数为 500 次, 那么 CPU 平均访存时间是_____。

5.在经典五级流水线 MIPS 处理器中, 各个流水段所需要延迟如下 (不考虑寄存器的延迟):

IF 阶段 130ps、ID 阶段 100ps;

无转发的 EXE 阶段 120ps、带全部转发的 EXE 阶段 150ps;

仅从 EXE/MEM 转发的 EXE 阶段 140ps、仅从 MEM/WB 转发的 EXE 阶段 130ps;

MEM 阶段 120ps、WB 阶段 100ps

(1) 如果实现了全部的转发, 则该处理器的时钟周期不短于_____ps。

(2) 如果实现中不使用转发,则该处理器的时钟周期不短于_____ps。

(3) 执行如下的指令序列:

```
addr5, r2, r1  
lwr3, 4(r5)  
lwr2, 0(r2)  
orr3, r5, r3  
swr3, 0(r5)
```

若要实现全部转发,则需要_____个时钟周期;若仅实现从 EXE/MEM 转发,则需要_____个时钟周期;若仅实现从 MEM/WB 转发,则需要_____个时钟周期。
(请写出计算过程)

操作系统

一、判断题 (每题 1 分, 共 12 分)

1. 进程切换可以发生在用户态, 也可以发生在内核态。 ()
2. 在支持多线程的操作系统中, 可以通过进程控制块访问该进程的线程执行状态。 ()
3. 最短作业优先调度算法 (SJF) 可以保证作业的平均等待时间最短。 ()
4. 银行家算法可以有效的避免死锁。 ()
5. 在一个进程的页表中, 页目录项和页表项不会同时存在。 ()
6. 请求分页存储管理需要 CPU 硬件的支持。 ()
7. 对于单处理器系统, 关闭中断可以保证临界区的互斥访问。 ()
8. 文件控制块除了包含元数据信息, 也可以包含文件数据信息。 ()
9. 若父进程先于子进程退出, 则致子进程会变成僵尸进程。 ()
10. 应用程序执行系统调用可能会引起进程上下文切换。 ()

- 11.微内核操作系统的设备驱动程序大多运行在内核态。（ ）
- 12.虚拟存储系统允许进程的地址空间大于实际物理内存大小。（ ）

二、计算题（每题 1 分，共 12 分）

有一台假想计算机，虚拟页和物理页帧大小均为 16Bytes，支持 256Bytes 的虚拟地址空间，有 256Bytes 的物理内存；采用二级页表，一个页目录项大小为 1Byte，一个页表项大小为 1Byte，1 个页目录大小为 4Bytes，页目录基址寄存器保存了页目录表的物理页帧号，占 1Byte。PTE 和 PDE 格式（8Bytes）如下所示，其中 I,R,W,X 都是标志位，分别是存在位、可读位、可写位、执行位，Frame Number 是物理页帧号。

Frame Number	I	R	W	X
4bits	1	1	1	1

页目录基址寄存器的内容为 0x00，页目录基址寄存器格式为：

Frame Number	Reserved
4bits	4bits

256Bytes 的整个物理内存内容如下：（数值为十六进制）

```

0x00  0C 1C 2C 3C 00 00 00 00 00 00 00 00 00 00 00 00
0x10  4C 5C 6C 7C 00 00 00 00 00 00 00 00 00 00 00 00
0x20  8C 9C AC BC 00 00 00 00 00 00 00 00 00 00 00 00
0x30  CC DC EC FC 00 00 00 00 00 00 00 00 00 00 00 00
0x40  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0x50  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0x60  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0x70  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0x80  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0x90  00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00

```

0xa0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
0xb0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
0xc0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
0xd0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
0xe0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
0xf0	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00

(1) 在这台假象计算机的虚拟地址中，页目录号、页表号和页内偏移各占多少位？请简明解释求解过程。

(2) 在这台假想的计算机上执行访存指令时，虚拟地址 0x32 经地址转换后得到的物理地址为多少？内容是什么？请简明解释过程。

(3) 已知这台假想的计算机上执行访存指令时，虚拟地址 0x12 经过地址转换后得到的物理地址为 0x12，请回答对应页目录号和内容，对应页表号和内容分别是什么？请简明解释过程。

计算机网络

一、选择题（每题 2 分，共 12 分）

1. 发送 ARP 消息时，采用（ ）

A. 单播 B. 多播 C. 组播 D. 广播

2. 是以太网交换机收到一数据帧，经转发表对比该帧目的 Mac 地址不在表项中，则（ ）

A. 将该帧转发给除接受该帧端口外所有端口；

B. 将该帧传送给特定端口进行 ARP 查询；

C. 将该帧发送给 HDCP 服务器；

D. 丢弃该帧；

3.关于无线局域网，下列说法正确的是（ ）

- A.为了减少冲突发生，无线局域网在发送信号时需要进行冲突检测；
- B.无线局域网可采用 CSMA/CD 来进行冲突检测；
- C.采用 RTS/CTS 方式，如果某一站点监听到 CTS 信号，但没有监听到对应 RTS，那么该站点不需要为了避免冲突而静默。
- D.SIFS、DIFS、PIFS 的帧间间隔，主要是用于控制站点消息的优先级。

4.局域网路由表如下：

地址/子网掩码	端口
192.168.160.0 /19	0
192.168.192.0 /18	1
192.168.128.0 /17	2
0.0.0.0	3

当目的地址分别是 192.168.64.18 和 192.168.172.17 的两个数据包到达此路由设备时，该设备将使用的转发接口分别是（ ）

- A.1,0 B.1,2 C.3,0 D.3,2

5.主机 A 和主机 B 之间采用 TCP 传输，A 向 B 传输 3 个 TCP 段，其有效载荷长度分别为 200 字节、300 字节、500 字节，第三个 TCP 段起始序号为 800，A 向 B 传输第一个段成功，第二个段丢失，则主机 B 在收到第三个段后，返回的确认序号是（ ）

- A. 200 B. 500 C. 800 D. 1000

6.HTTP、QUIC、SMTP 在传输层采用的协议分别为（ ）

- A.TCP, UDP, TCP B.TCP, TCP, TCP
C.UDP, TCP, UDP D.UDP, UDP, UDP

二、简答题（每题 4 分，共 8 分）

7.为了增强互联网规模可扩大性，请分别从四个方面为 OSPF 和 BGP 提出优化建

议。

8.相比于 TCP 的 reno 算法，请列举 cubic 和 bbr 算法的改进性。