

# 哈工大 2020 试题

哈尔滨工业大学

2020年

硕士研究生招生考试试题

考试科目代码：[854] 考试科目：计算机基础

考生注意：答案务必写在答题纸上，并标明题号。答在试题上无效。

题号	一	二	三	四	五	六	七	八	九	十	总分
分数	10	10	20	30	20	20	10	5	10	15	150

## 第 I 部分 计算机系统 (70 分)

### 一、单项选择题 (每题 1 分, 共 10 分)

- 计算机系统层次结构中, (A) 是软硬件系统的交界面, 系统程序员与硬件设计者都能够看到。  
A. 操作系统 B. ISA 指令集体系结构 C. 系统函数库 D. 微程序
- 执行程序的生成, 对于 C 标准函数库中的函数, 不能采用 (D) 连接方式。  
A. 操作系统加载时 B. 静态库连接 C. 运行过程中连接 D. 编译时
- Linux 操作系统的内核采用 (A) 编码。  
A. ASCII B. UTF-8 C. UTF-16 D. UTF-32
- C 语言中 int 与 float 所表示的数据个数, 哪一个多 ( )。  
A. int 多 B. float 多 C. 一样多 D. 与编译时采用 m32/m64 模式有关
- Linux 下 C 程序中出现整数除法错, 如除以 0 等, 会输出 (A), 并退出程序。  
A. Divide Overflow B. Floating point exception C. NULL D. 空白行
- 软件工具 objdump 反编译的汇编语言中某个操作数 1024(RIP)最可能是 ( )。  
A. 参数 B. 局部变量 C. 静态局部变量 D. 全局变量
- 进程的状态不包括 (B)。  
A. 运行 B. 睡眠 C. 停止 D. 终止
- 计算机系统中下列存储器最快的是 (A)。  
A. L1 Cache B. 动态内存 DRAM C. SSD 固态硬盘 D. TLB
- 父进程与其创建的子进程的堆栈段 (C)。  
A. 物理存储空间上完全独立 B. 虚拟存储空间共享  
C. 虚拟存储空间独立 D. 物理存储空间上完全共享
- 程序 hello world! 执行第一条机器指令时, 正确的是 ( )。  
A. 所有的代码与数据都已经由操作系统加载到物理内存 B. 此指令是 call main  
C. 会产生缺页中断 D. 此指令已经在物理内存中

二、 填空题 (每题 2 分, 共 10 分)

11. C 语言 `int x=-2`; 则 `&x` 地址处依次存放的 4 个字节为 2-11
12. 连接重定位时, 指令 `400800: e8` 2-12 `callq 400810`
13. TLB(翻译后备缓冲器) 俗称快表, 是 2-13 的缓存
14. Linux 下程序运行时, 在键盘上输入 Ctrl-C, 会产生 2-14 信号, 终止程序。
15. 十进制的有符号常数, 由 2-15 转换成二进制补码。

三、 分析题 (每题 5 分, 共 20 分)

16. 请按步骤算出 float 数 0.1 的二进制表示、规格化表示、阶码、尾数、其在内存从低地址到高地址的 4 个字节内容 (16 进制)。

17. Intel i7 CPU 的虚拟地址 48 位, 物理地址 52 位, 每一页面 4KB。其内部结构如下图所示:



分析填写如下数据:

虚拟地址中的 VPN 占 ① 位; 其一级页表为 ② 项; L1 数据 TLB 的组索引位数 TLBI 为 ③ 位。L1 数据 Cache 共 ④ 组。用物理地址访问 L1 数据 Cache 时, Cache 标记 CT 占 ⑤ 位。

18. 分析如下子程序, 写出其对应的 C 语言源程序(函数名、参数名可自定义)。

```
0x40114a  mov    (%rsi),%eax
0x40114c  neg    %eax
0x40114e  mov    (%rdi),%edx
0x401150  neg    %edx
0x401152  mov    %edx,(%rsi)
0x401154  mov    %eax,(%rdi)
0x401156  retq
```

19. 下列子程序存在着缓冲器溢出漏洞, 请分析漏洞的产生原因, 说明如何攻击?

```
0x401152    push    %rbp
0x401153    mov     %rsp,%rbp
0x401156    sub     $0x28,%rsp
0x40115a    mov     %rdi,%rsi
0x40115d    lea     -0x20(%rbp),%rdi
0x401161    callq   0x401030 <strcpy@plt>
0x401166    lea     -0x20(%rbp),%rdi
0x40116a    callq   0x401040 <puts@plt>
0x40116f    leaveq
0x401170    retq
```

#### 四、 综合设计题 (每题 10 分, 共 30 分)

20. 在 Y86-64 的 SEQ 顺序结构 CPU 设计中, 采用取指、译码、执行、访存、写回、更新 PC 共 6 个阶段, 请写出 ret 每一阶段的微操作。在 PIPE 五段流水线 CPU 设计中, 怎么避免 ret 指令的控制冒险?

21. 计算两个向量的内积。向量类型定义如下, 函数 long vec\_length(vec v) 函数返回向量 v 的元素个数; 函数 double get\_vec\_element(vec u, long i) 函数返回向量 u 的第 i 个元素; 一个初步的内积计算函数 inner0 如下:

```
typedef struct {
    long    len;    //向量的元素个数
    double  *data;
} vec;
```

```
void inner0(vec *u, vec *v, double *dest) {
    for(long i=0; *dest=0; i<vec_length(u); i++)
        *dest = *dest + get_vec_element(u,i) * get_vec_element(v,i);
}
```

请面向 C 编译器, 进行初步优化, 以提高程序的性能, 编写优化后的函数 inner1, 并说明优化依据。

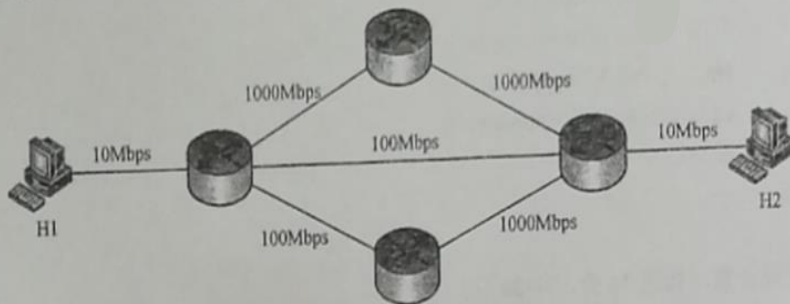
22. 针对上题的函数 inner0, 采用 Intel I7 CPU, 其有 2 个带地址计算的数据加载器、2 个浮点乘法器, 一个浮点加法器。u、v 的元素个数是 4 的倍数, 采用 2\*2 循环展开方式, 编写面向 CPU 友好的子程序 inner2。这个 inner2 程序是性能最优的吗? 为什么? 怎么获得面向本 cpu 最优的程序?

五、单项选择题（每小题 2 分，共 20 分）

1. 在 TCP/IP 参考模型中，实现端到端可靠数据传输功能的层是

- A. 应用层      B. 传输层      C. 网络层      D. 数据链路层

2. 如下图所示网络。若 H1 欲向 H2 发送 1 个大小为 10 MB 的文件，则从 H1 开始发送时刻起，到 H2 收到文件为止，采用报文交换和分组交换（分组长度为 1000B）所用时间至少分别约为

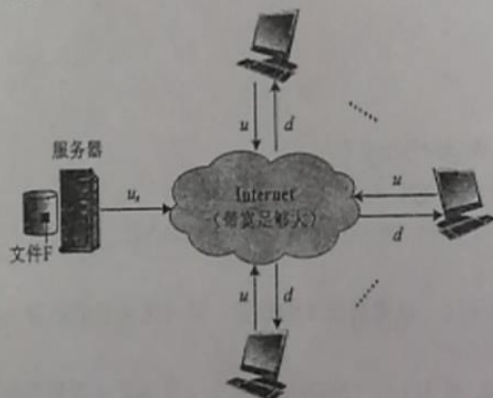


- A. 16.16s, 8s      B. 16.8s, 8s      C. 2.1s, 1s      D. 2.02s, 0.1s

3. 如果浏览器分别采用非流水的 HTTP/1.1 和支持并行 TCP 连接的 HTTP/1.0 请求一个引用 4 个 JPEG 小图片的 Web 页，则需要的时间至少分别是

- A. 3RTT, 4RTT      B. 5RTT, 3RTT      C. 6RTT, 4RTT      D. 6RTT, 10RTT

4. 如下图所示，为向 1000 个客户进行文件分发的应用示意图，其中  $F=1\text{MB}$ ,  $u_s=1000\text{Mbps}$ ,  $d=10\text{Mbps}$ ,  $u_c=1\text{Mbps}$ （注： $M=10^6$ ）。若分别采取 C/S 分发方式和 P2P 分发方式，则完成文件 F 分发所需时间至少分别为



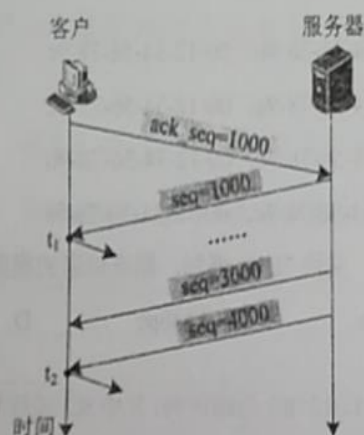
- A. 8s, 4s      B. 8s, 0.8s      C. 1s, 0.5s      D. 0.8s, 0.008s



5. 主机甲通过一条链路连接主机乙，链路带宽为 10 Mbps ( $M=10^6$ )，单向传播延时为 7 ms。若甲采用 GBN 协议向乙发送帧长为 1000B 的数据帧，乙采用帧长为 250B 的确认帧进行确认，数据帧序号为 4 比特编号，则甲的最大信道利用率约为

A. 10%      B. 80%      C. 85%      D. 100%

6. 服务器通过一个 TCP 连接向客户发送数据的部分过程如下图所示。若服务器均以最大段长 MSS=1000B 向客户发送 TCP 段，则客户在  $t_1$  时刻和  $t_2$  时刻发送给服务器的 TCP 段中正确的确认序号分别是

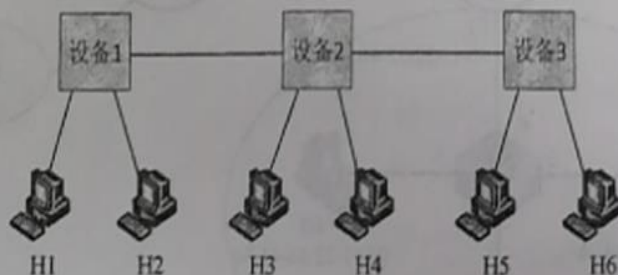


A. 1000, 4000      B. 2000, 2000      C. 3000, 4000      D. 2000, 5000

7. 若客户与服务器建立 TCP 连接时，第三次握手 TCP 段的序号为 1020，当客户向服务器发送了 1000B 的数据后，首先请求与服务器断开连接，则客户收到的第二次挥手 TCP 段的确认序号是

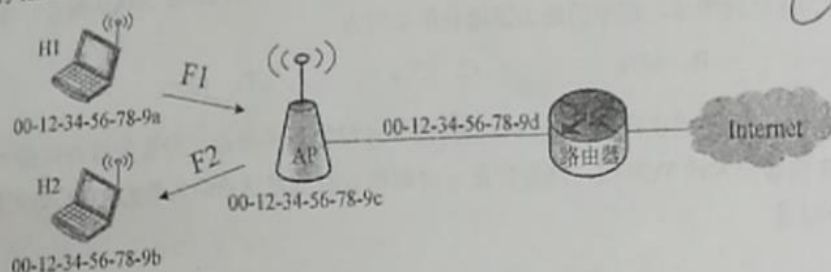
A. 2019      B. 2020      C. 2021      D. 2022

8. 如下图所示以太网局域网中，若 H1 与 H2 属于同一个广播域，但不属于同一个冲突域；H2 与 H3 不属于同一个广播域；H5 和 H6 属于同一个冲突域，则设备 1、设备 2 和设备 3 分别是



A. 路由器、交换机、集线器      B. 交换机、路由器、交换机  
C. 路由器、集线器、交换机      D. 交换机、路由器、集线器

9. 在下图所示的网络中, 主机 H1 向主机 H2 发送一个 IP 分组。若主机 H2 接收到的封装该 IP 分组的 IEEE 802.11 数据帧为 F2, 则帧 F2 的地址 1、地址 2 和地址 3 分别是

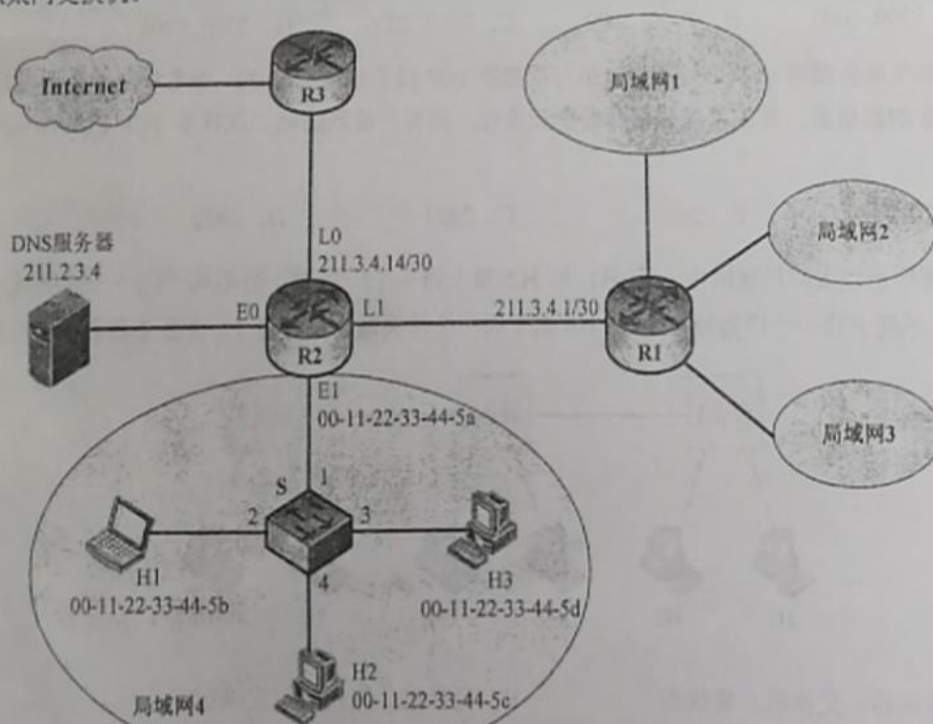


- A. 00-12-34-56-78-9a, 00-12-34-56-78-9b, 00-12-34-56-78-9c  
 B. 00-12-34-56-78-9b, 00-12-34-56-78-9a, 00-12-34-56-78-9c  
 C. 00-12-34-56-78-9a, 00-12-34-56-78-9c, 00-12-34-56-78-9b  
 D. 00-12-34-56-78-9b, 00-12-34-56-78-9c, 00-12-34-56-78-9a
10. 若某无噪声信道带宽为 2 MHz, 采用 2PSK 调制, 则该信道的极限数据速率是

- A. 1 Mbps B. 2 Mbps C. 4 Mbps D. 8 Mbps

#### 六、综合应用题 (共 20 分)

某网络拓扑如下图所示, 图中 R1、R2、R3 为路由器, 其中 R2 运行 DHCP 服务; S 为 100Base-T 以太网交换机。



R2 的路由表结构为:

目的网络	子网掩码	下一跳	接口
------	------	-----	----

交换机 S 的交换表结构为:

MAC 地址	端口
--------	----

请回答下列问题。

- (1) 请将 IP 地址空间 201.2.3.0/24 分配给局域网 1~4, 要求局域网 1 可分配 IP 地址数不少于 120 个, 局域网 2 和局域网 3 可分配 IP 地址数分别不少于 30 个, 局域网 4 可分配 IP 地址数不少于 60 个。给出地址分配方案。
- (2) 基于 (1) 的 IP 地址分配方案, 请给出 R2 的路由表, 要求路由表项尽可能少。
- (3) 基于 (1) 的 IP 地址分配方案, 为路由器 R2 的 E1 接口分配 IP 地址, 并在下表中为主机 H2 配置 IP 地址信息:

IP 地址	
子网掩码	
默认网关 IP 地址	

- (4) 若 S 交换表为空时, H1 接入局域网 4, 并通过 DHCP 请求动态分配 IP 地址, 则 H1 发送的第一个 DHCP 报文是什么报文? 封装该报文的 IP 数据报的源 IP 地址和目的 IP 地址分别是什么? 请给出 H1 完成 IP 地址配置后的交换机 S 的交换表。

III. 数据结构部分 (共 40 分)

七、单项选择题: (第 1-5 题, 每题 2 分, 共 10 分。在每题给出的四个选项中, 请选出一项最符合题目要求的)

1. 利用直接插入排序法的思想建立一个有序线性表的时间复杂度为 (①)。C  
A.  $O(n)$     B.  $O(n \log_2 n)$     C.  $O(n^2)$     D.  $O(\log_2 n)$
2. 将 10 阶对称矩阵压缩存储到一维数组 A 中, 则数组 A 的长度最少为 (②)。C  
A. 100    B. 40    C. 55    D. 80
3. 给定 n 个字符的字符表与字符频率, 则哈夫曼编码算法的时间复杂度为 (③)。B  
A.  $O(n^2)$     B.  $O(n \log_2 n)$     C.  $O(\log_2 n)$     D.  $O(n)$
4. 用 Prim 算法和 Kruskal 算法构造的最小生成树, 所得到的最小生成树 (④)。C  
A. 相同    B. 不同    C. 可能相同, 可能不同    D. 无法比较
5. 设散列表有 m 个存储单元, 散列函数为  $H(\text{key}) = \text{key} \% p$ , 则 p 最好选择 (⑤)。B  
A. 小于等于 m 的最大奇数    B. 小于等于 m 的最大素数  
C. 小于等于 m 的最大偶数    D. 小于等于 m 的最大合数

八、填空题: (第 6-10 题, 每空 1 分, 共 5 分)

6. 算术表达式  $a + (b * (c - d) - e) / f$  的后缀表达式为 (⑥)。
7. 已知一棵完全二叉树中共有 2019 个结点, 则该树中共有 (⑦) 个叶子结点。
8. 在含 n 个顶点和 e 条边的无向图的邻接矩阵中, 零元素的个数为 (⑧)。
9. 设有向图 G 中有向边的集合  $E = \{ \langle 1, 2 \rangle, \langle 2, 3 \rangle, \langle 1, 4 \rangle, \langle 4, 2 \rangle, \langle 4, 3 \rangle \}$ , 则该图的一个拓扑序列为 (⑨)。
10. 在 4 阶 B 树中, 每个非失败结点的关键字个数范围是 (⑩)。

九、简答题: (第 11 题, 共 10 分)

11. (10 分) 已知二叉查找树的后序遍历序列 (比如 5, 7, 6, 9, 11, 10, 8), 请问能否唯一的确定一棵二叉查找树。如果能, 给出构造二叉查找树的方法; 如果不能, 请举出反例。



十、算法设计题：（第 12 题，共 15 分）

按以下要求设计算法：

- （1）给出算法的基本设计思想。
- （2）使用 C 或 C++ 或 Java 语言，给出相关数据类型定义。
- （3）根据设计思想，采用 C 或 C++ 或 Java 语言描述算法，栈和队列的操作可以直接使用。

12. (15 分) 假设两个有序数组 A 和 B，长度分别是  $m$  和  $n$ ，设计一个尽可能高效算法查找第  $k$  小元素，并给出所设计算法的时间复杂度。