

2012 年全国硕士研究生入学统一考试

计算机科学与技术学科联考计算机学科专业基础综合试题

一、单项选择题：第 1~40 小题，每小题 2 分，共 80 分。下列每题给出的四个选项中，只有一个选项最符合试题要求。

1. 求整数 $n(n \geq 0)$ 阶乘的算法如下，其时间复杂度是_____。

```
int fact(int n){
    if (n<=1) return 1;
    return n*fact(n-1);
}
```

- A. $O(\log_2 n)$ B. $O(n)$ C. $O(n \log_2 n)$ D. $O(n^2)$

2. 已知操作符包括 ‘+’、‘-’、‘*’、‘/’、‘(’ 和 ‘)’。将中缀表达式 $a+b-a*((c+d)/e-f)+g$ 转换为等价的后缀表达式 $ab+acd+e/f-*g+$ 时，用栈来存放暂时还不能确定运算次序的操作符，若栈初始时空，则转换过程中同时保存在栈中的操作符的最大个数是_____。

- A. 5 B. 7 C. 8 D. 11

3. 若一棵二叉树的前序遍历序列为 a, e, b, d, c，后序遍历序列为 b, c, d, e, a，则根结点的孩子结点_____。

- A. 只有 e B. 有 e、b C. 有 e、c D. 无法确定

4. 若平衡二叉树的高度为 6，且所有非叶结点的平衡因子均为 1，则该平衡二叉树的结点总数为_____。

- A. 10 B. 20 C. 32 D. 33

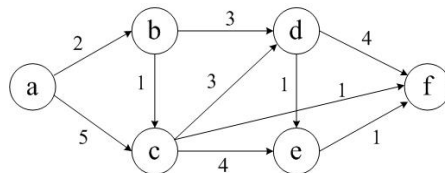
5. 对有 n 个结点、 e 条边且使用邻接表存储的有向图进行广度优先遍历，其算法时间复杂度是_____。

- A. $O(n)$ B. $O(e)$ C. $O(n+e)$ D. $O(n*e)$

6. 若用邻接矩阵存储有向图，矩阵中主对角线以下的元素均为零，则关于该图拓扑序列的结论是_____。

- A. 存在，且唯一 B. 存在，且不唯一
C. 存在，可能不唯一 D. 无法确定是否存在

7. 如右图所示的有向带权图，若采用迪杰斯特拉 (Dijkstra) 算法求从源点 a 到其他各顶点的最短路径，则得到的第一条最短路径的目标顶点是 b，第二条最短路径的目标顶点是 c，后续得到的其余各最短路径的目标顶点依次是_____。



- A. d, e, f B. e, d, f C. f, d, e D. f, e, d

8. 下列关于最小生成树的叙述中，正确的是_____。

I. 最小生成树的代价唯一

II. 所有权值最小的边一定会出现在所有的最小生成树中

III. 使用普里姆 (Prim) 算法从不同顶点开始得到的最小生成树一定相同

IV. 使用普里姆算法和克鲁斯卡尔 (Kruskal) 算法得到的最小生成树总不相同

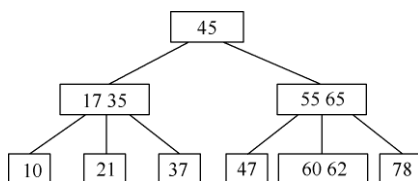
A. 仅 I

B. 仅 II

C. 仅 I、III

D. 仅 II、IV

9. 已知一棵 3 阶 B-树，如下图所示。删除关键字 78 得到一棵新 B-树，其最右叶结点中的关键字是_____。



A. 60

B. 60, 62

C. 62, 65

D. 65

10. 在内部排序过程中，对尚未确定最终位置的所有元素进行一遍处理称为一趟排序。下列排序方法中，每一趟排序结束都至少能够确定一个元素最终位置的方法是_____。

I. 简单选择排序

II. 希尔排序

III. 快速排序

IV. 堆排序

V. 二路归并排序

A. 仅 I、III、IV

B. 仅 I、III、V

C. 仅 II、III、IV

D. 仅 III、IV、V

11. 对一待排序序列分别进行折半插入排序和直接插入排序，两者之间可能的不同之处是_____。

A. 排序的总趟数

B. 元素的移动次数

C. 使用辅助空间的数量

D. 元素之间的比较次数

12. 假定基准程序 A 在某计算机上的运行时间为 100 秒，其中 90 秒为 CPU 时间，其余为 I/O 时间。若 CPU 速度提高 50%，I/O 速度不变，则运行基准程序 A 所耗费的时间是_____。

A. 55 秒

B. 60 秒

C. 65 秒

D. 70 秒

13. 假定编译器规定 int 和 short 型长度分别为 32 位和 16 位，执行下列 C 语言语句：

```
unsigned short x=65530;
unsigned int y=x;
```

得到 y 的机器数为_____。

A. 0000 7FFAH

B. 0000 FFFAH

C. FFFF 7FFAH

D. FFFF FFFAH

14. float 类型（即 IEEE754 单精度浮点数格式）能表示的最大正整数是_____。

A. $2^{126}-2^{103}$

B. $2^{127}-2^{104}$

C. $2^{127}-2^{103}$

D. $2^{128}-2^{104}$

15. 某计算机存储器按字节编址，采用小端方式存放数据。假定编译器规定 int 型和 short 型长度分别为 32 位和 16 位，并且数据按边界对齐存储。某 C 语言程序段如下：

```
struct{
    int a;
    char b;
    short c;
} record;
record.a=273;
```

若 record 变量的首地址为 0xC008，则地址 0xC008 中内容及 record.c 的地址分别为_____。

- A. 0x00、0xC00D
- B. 0x00、0xC00E
- C. 0x11、0xC00D
- D. 0x11、0xC00E

16. 下列关于闪存（Flash Memory）的叙述中，错误的是_____。

- A. 信息可读可写，并且读、写速度一样快
- B. 存储元由 MOS 管组成，是一种半导体存储器
- C. 掉电后信息不丢失，是一种非易失性存储器
- D. 采用随机访问方式，可替代计算机外部存储器

17. 假设某计算机按字编址，Cache 有 4 个行，Cache 和主存之间交换的块大小为 1 个字。若 Cache 的内容初始为空，采用 2 路组相联映射方式和 LRU 替换策略。访问的主存地址依次为 0,4,8,2,0,6,8,6,4,8 时，命中 Cache 的次数是_____。

- A. 1
- B. 2
- C. 3
- D. 4

18. 某计算机的控制器采用微程序控制方式，微指令中的操作控制字段采用字段直接编码法，共有 33 个微命令，构成 5 个互斥类，分别包含 7、3、12、5 和 6 个微命令，则操作控制字段至少有_____。

- A. 5 位
- B. 6 位
- C. 15 位
- D. 33 位

19. 某同步总线的时钟频率为 100MHz，宽度为 32 位，地址/数据线复用，每传输一个地址或数据占用一个时钟周期。若该总线支持突发（猝发）传输方式，则一次“主存写”总线事务传输 128 位数据所需要的时间至少是_____。

- A. 20ns
- B. 40ns
- C. 50ns
- D. 80ns

20. 下列关于 USB 总线特性的描述中，错误的是_____。

- A. 可实现外设的即插即用和热拔插
- B. 可通过级联方式连接多台外设
- C. 是一种通信总线，连接不同外设
- D. 同时可传输 2 位数据，数据传输率高

21. 下列选项中，在 I/O 总线的数据线上传输的信息包括_____。

- I. I/O 接口中的命令字
- II. I/O 接口中的状态字
- III. 中断类型号
- A. 仅 I、II
- B. 仅 I、III
- C. 仅 II、III
- D. I、II、III

22. 响应外部中断的过程中，中断隐指令完成的操作，除保护断点外，还包括_____。

- I. 关中断
- II. 保存通用寄存器的内容
- III. 形成中断服务程序入口地址并送 PC
- A. 仅 I、II
- B. 仅 I、III
- C. 仅 II、III
- D. I、II、III

23. 下列选项中，不可能在用户态发生的事件是_____。

- A. 系统调用
- B. 外部中断
- C. 进程切换
- D. 缺页

24. 中断处理和子程序调用都需要压栈以保护现场，中断处理一定会保存而子程序调用不需要保存其内容的是_____。

- A. 程序计数器
- B. 程序状态字寄存器
- C. 通用数据寄存器
- D. 通用地址寄存器

25. 下列关于虚拟存储器的叙述中，正确的是_____。

- A. 虚拟存储只能基于连续分配技术
- B. 虚拟存储只能基于非连续分配技术
- C. 虚拟存储容量只受外存容量的限制

D. 虚拟存储容量只受内存容量的限制

26. 操作系统的 I/O 子系统通常由四个层次组成，每一层明确定义了与邻近层次的接口。其合理的层次组织排列顺序是_____。

- A. 用户级 I/O 软件、设备无关软件、设备驱动程序、中断处理程序
- B. 用户级 I/O 软件、设备无关软件、中断处理程序、设备驱动程序
- C. 用户级 I/O 软件、设备驱动程序、设备无关软件、中断处理程序
- D. 用户级 I/O 软件、中断处理程序、设备无关软件、设备驱动程序

27. 假设 5 个进程 P0、P1、P2、P3、P4 共享三类资源 R1、R2、R3，这些资源总数分别为 18、6、22。T0 时刻的资源分配情况如下表所示，此时存在的一个安全序列是_____。

进程	已分配资源			资源最大需求		
	R1	R2	R3	R1	R2	R3
P0	3	2	3	5	5	10
P1	4	0	3	5	3	6
P2	4	0	5	4	0	11
P3	2	0	4	4	2	5
P4	3	1	4	4	2	4

- A. P0, P2, P4, P1, P3
- B. P1, P0, P3, P4, P2
- C. P2, P1, P0, P3, P4
- D. P3, P4, P2, P1, P0

28. 若一个用户进程通过 read 系统调用读取一个磁盘文件中的数据，则下列关于此过程的叙述中，正确的是_____。

- I. 若该文件的数据不在内存，则该进程进入睡眠等待状态
- II. 请求 read 系统调用会导致 CPU 从用户态切换到核心态
- III. read 系统调用的参数应包含文件的名称

- A. 仅 I、II
- B. 仅 I、III
- C. 仅 II、III
- D. I、II 和 III

29. 一个多道批处理系统中仅有 P1 和 P2 两个作业，P2 比 P1 晚 5ms 到达，它们的计算和 I/O 操作顺序如下：

P1：计算 60ms，I/O 80ms，计算 20ms

P2：计算 120ms，I/O 40ms，计算 40ms

若不考虑调度和切换时间，则完成两个作业需要的时间最少是_____。

- A. 240ms
- B. 260ms
- C. 340ms
- D. 360ms

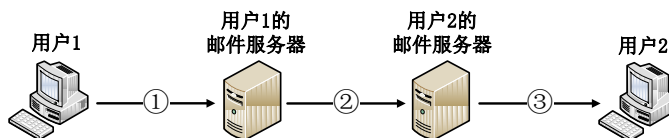
30. 若某单处理器多进程系统中有多个就绪态进程，则下列关于处理机调度的叙述中，错误的是_____。

- A. 在进程结束时能进行处理机调度
- B. 创建新进程后能进行处理机调度
- C. 在进程处于临界区时不能进行处理机调度
- D. 在系统调用完成并返回用户态时能进行处理机调度

31. 下列关于进程和线程的叙述中，正确的是_____。

- A. 不管系统是否支持线程，进程都是资源分配的基本单位
- B. 线程是资源分配的基本单位，进程是调度的基本单位
- C. 系统级线程和用户级线程的切换都需要内核的支持
- D. 同一进程中的各个线程拥有各自不同的地址空间

32. 下列选项中，不能改善磁盘设备 I/O 性能的是_____。
- A. 重排 I/O 请求次序 B. 在一个磁盘上设置多个分区
- C. 预读和滞后写 D. 优化文件物理块的分布
33. 在 TCP/IP 体系结构中，直接为 ICMP 提供服务的协议是_____。
- A. PPP B. IP C. UDP D. TCP
34. 在物理层接口特性中，用于描述完成每种功能的事件发生顺序的是_____。
- A. 机械特性 B. 功能特性 C. 过程特性 D. 电气特性
35. 以太网的 MAC 协议提供的是_____。
- A. 无连接不可靠服务 B. 无连接可靠服务
- C. 有连接不可靠服务 D. 有连接可靠服务
36. 两台主机之间的数据链路层采用后退 N 帧协议 (GBN) 传输数据，数据传输速率为 16kbps，单向传播时延为 270ms，数据帧长度范围是 128~512 字节，接收方总是以与数据帧等长的帧进行确认。为使信道利用率达到最高，帧序号的比特数至少为_____。
- A. 5 B. 4 C. 3 D. 2
37. 下列关于 IP 路由器功能的描述中，正确的是_____。
- I. 运行路由协议，设置路由表
- II. 监测到拥塞时，合理丢弃 IP 分组
- III. 对收到的 IP 分组头进行差错校验，确保传输的 IP 分组不丢失
- IV. 根据收到的 IP 分组的目的 IP 地址，将其转发到合适的输出线路上
- A. 仅 III、IV B. 仅 I、II、III
- C. 仅 I、II、IV D. I、II、III、IV
38. ARP 协议的功能是_____。
- A. 根据 IP 地址查询 MAC 地址 B. 根据 MAC 地址查询 IP 地址
- C. 根据域名查询 IP 地址 D. 根据 IP 地址查询域名
39. 某主机的 IP 地址为 180.80.77.55，子网掩码为 255.255.252.0。若该主机向其所在子网发送广播分组，则目的地址可以是_____。
- A. 180.80.76.0 B. 180.80.76.255 C. 180.80.77.255 D. 180.80.79.255
40. 若用户 1 与用户 2 之间发送和接收电子邮件的过程如下图所示，则图中①、②、③阶段分别使用的应用层协议可以是_____。



- A. SMTP、SMTP、SMTP B. POP3、SMTP、POP3
- C. POP3、SMTP、SMTP D. SMTP、SMTP、POP3

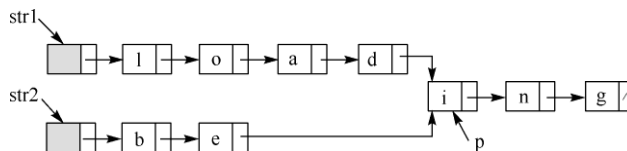
二、综合应用题：第 41~47 题，共 70 分。

41. 设有 6 个有序表 A、B、C、D、E、F，分别含有 10、35、40、50、60 和 200 个数据元素，各表中元素按升序排列。要求通过 5 次两两合并，将 6 个表最终合并成 1 个升序表，并在最坏情况下比较的总次数达到最小。请回答下列问题。

- 1) 给出完整的合并过程，并求出最坏情况下比较的总次数。

2) 根据你的合并过程, 描述 N ($N \geq 2$) 个不等长升序表的合并策略, 并说明理由。

42. 假定采用带头结点的单链表保存单词, 当两个单词有相同的后缀时, 则可共享相同的后缀存储空间, 例如, “loading” 和 “being” 的存储映像如下图所示。



设 $str1$ 和 $str2$ 分别指向两个单词所在单链表的头结点, 链表结点结构为

data	next
------	------

, 请设计一个时间上尽可能高效的算法, 找出由 $str1$ 和 $str2$ 所指向两个链表共同后缀的起始位置 (如图中字符 i 所在结点的位置 p)。要求:

1) 给出算法的基本设计思想。

2) 根据设计思想, 采用 C 或 C++ 或 Java 语言描述算法, 关键之处给出注释。

3) 说明你所设计算法的时间复杂度。

43. 假定某计算机的 CPU 主频为 80MHz, CPI 为 4, 平均每条指令访存 1.5 次, 主存与 Cache 之间交换的块大小为 16B, Cache 的命中率为 99%, 存储器总线宽度为 32 位。请回答下列问题。

1) 该计算机的 MIPS 数是多少? 平均每秒 Cache 缺失的次数是多少? 在不考虑 DMA 传送的情况下, 主存带宽至少达到多少才能满足 CPU 的访存要求?

2) 假定在 Cache 缺失的情况下访问主存时, 存在 0.0005% 的缺页率, 则 CPU 平均每秒产生多少次缺页异常? 若页面大小为 4KB, 每次缺页都需要访问磁盘, 访问磁盘时 DMA 传送采用周期挪用方式, 磁盘 I/O 接口的数据缓冲寄存器为 32 位, 则磁盘 I/O 接口平均每秒发出的 DMA 请求次数至少是多少?

3) CPU 和 DMA 控制器同时要求使用存储器总线时, 哪个优先级更高? 为什么?

4) 为了提高性能, 主存采用 4 体低位交叉存储模式, 工作时每 1/4 个存储周期启动一个体。若每个体的存储周期为 50ns, 则该主存能提供的最大带宽是多少?

44. 某 16 位计算机中, 带符号整数用补码表示, 数据 Cache 和指令 Cache 分离。题 44 表给出了指令系统中部分指令格式, 其中 Rs 和 Rd 表示寄存器, mem 表示存储单元地址, (x) 表示寄存器 x 或存储单元 x 的内容。

指令系统中部分指令格式

名称	指令的汇编格式	指令功能
加法指令	ADD Rs, Rd	$(Rs) + (Rd) \rightarrow Rd$
算术/逻辑左移	SHL Rd	$2 * (Rd) \rightarrow Rd$
算术右移	SHR Rd	$(Rd) / 2 \rightarrow Rd$
取数指令	LOAD Rd, mem	$(mem) \rightarrow Rd$
存数指令	STORE Rs, mem	$(Rs) \rightarrow mem$

该计算机采用 5 段流水方式执行指令, 各流水段分别是取指 (IF)、译码/读寄存器 (ID)、执行/计算有效地址 (EX)、访问存储器 (M) 和结果写回寄存器 (WB), 流水线采用 “按序发射, 按序完成” 方式, 没有采用转发技术处理数据相关, 并且同一个寄存器的读和写操作不能在同一个时钟周期内进行。请回答下列问题:

1) 若 int 型变量 x 的值为 -513, 存放在寄存器 $R1$ 中, 则执行指令 “SHR $R1$ ” 后, $R1$ 的内容是多少? (用十六进制表示)

2) 若某个时间段中, 有连续的 4 条指令进入流水线, 在其执行过程中没有发生任何阻塞, 则执行这 4 条指令所需的时钟周期数为多少?

3) 若高级语言程序中某赋值语句为 $x=a+b$, x 、 a 和 b 均为 `int` 型变量, 它们的存储单元地址分别表示为 $[x]$ 、 $[a]$ 和 $[b]$ 。该语句对应的指令序列及其在指令流水线中的执行过程如下图所示。

	时间单元													
指令	1	2	3	4	5	6	7	8	9	10	11	12	13	14
I_1	IF	ID	EX	M	WB									
I_2		IF	ID	EX	M	WB								
I_3			IF				ID	EX	M	WB				
I_4							IF				ID	EX	M	WB

图 指令序列及其执行过程示意图

则这 4 条指令执行过程中, I_3 的 ID 段和 I_4 的 IF 段被阻塞的原因各是什么?

4) 若高级语言程序中某赋值语句为 $x=x*2+a$, x 和 a 均为 `unsigned int` 类型变量, 它们的存储单元地址分别表示为 $[x]$ 、 $[a]$, 则执行这条语句至少需要多少个时钟周期? 要求模仿题 44 图画出这条语句对应的指令序列及其在流水线中的执行过程示意图。

45. 某请求分页系统的局部页面置换策略如下:

系统从 0 时刻开始扫描, 每隔 5 个时间单位扫描一轮驻留集 (扫描时间忽略不计), 本轮没有被访问过的页框将被系统回收, 并放入到空闲页框链尾, 其中内容在下一次分配之前不被清空。当发生缺页时, 如果该页曾被使用过且还在空闲页链表中, 则重新放回进程的驻留集中; 否则, 从空闲页框链表头部取出一个页框。

假设不考虑其他进程的影响和系统开销。初始时进程驻留集为空。目前系统空闲页框链表中页框号依次为 32、15、21、41。进程 P 依次访问的<虚拟页号, 访问时刻>是: <1,1>、<3,2>、<0,4>、<0,6>、<1,11>、<0,13>、<2,14>。请回答下列问题。

1) 访问<0,4>时, 对应的页框号是什么? 说明理由。

2) 访问<1,11>时, 对应的页框号是什么? 说明理由。

3) 访问<2,14>时, 对应的页框号是什么? 说明理由。

4) 该策略是否适合于时间局部性好的程序? 说明理由。

46. 某文件系统空间的最大容量为 4TB (1TB= 2^{40} B), 以磁盘块为基本分配单位。磁盘块大小为 1KB。文件控制块 (FCB) 包含一个 512B 的索引表区。请回答下列问题。

1) 假设索引表区仅采用直接索引结构, 索引表区存放文件占用的磁盘块号, 索引表项中块号最少占多少字节? 可支持的单个文件最大长度是多少字节?

2) 假设索引表区采用如下结构: 第 0~7 字节采用<起始块号, 块数>格式表示文件创建时预分配的连续存储空间, 其中起始块号占 6B, 块数占 2B; 剩余 504 字节采用直接索引结构, 一个索引项占 6B, 则可支持的单个文件最大长度是多少字节? 为了使单个文件的长度达到最大, 请指出起始块号和块数分别所占字节数的合理值并说明理由。

47. 主机 H 通过快速以太网连接 Internet, IP 地址为 192.168.0.8, 服务器 S 的 IP 地址为 211.68.71.80。H 与 S 使用 TCP 通信时, 在 H 上捕获的其中 5 个 IP 分组如题 47-a 表所示。

题 47-a 表

编号	IP 分组的前 40 字节内容（十六进制）				
1	45 00 00 30	01 9b 40 00	80 06 1d e8	c0 a8 00 08	d3 44 47 50
	0b d9 13 88	84 6b 41 c5	00 00 00 00	70 02 43 80	5d b0 00 00
2	43 00 00 30	00 00 40 00	31 06 6e 83	d3 44 47 50	c0 a8 00 08
	13 88 0b d9	e0 59 9f ef	84 6b 41 c6	70 12 16 d0	37 e1 00 00
3	45 00 00 28	01 9c 40 00	80 06 1d ef	c0 a8 00 08	d3 44 47 50
	0b d9 13 88	84 6b 41 c6	e0 59 9f f0	50 f0 43 80	2b 32 00 00
4	45 00 00 38	01 9d 40 00	80 06 1d de	c0 a8 00 08	d3 44 47 50
	0b d9 13 88	84 6b 41 c6	e0 59 9f f0	50 18 43 80	e6 55 00 00
5	45 00 00 28	68 11 40 00	31 06 06 7a	d3 44 47 50	c0 a8 00 08
	13 88 0b d9	e0 59 9f f0	84 6b 41 d6	50 10 16 d0	57 d2 00 00

回答下列问题。

1) 题 47-a 表中的 IP 分组中, 哪几个是由 H 发送的? 哪几个完成了 TCP 连接建立过程? 哪几个在通过快速以太网传输时进行了填充?

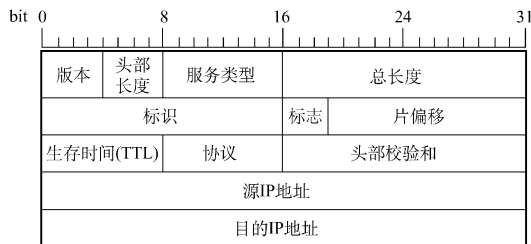
2) 根据题 47-a 表中的 IP 分组, 分析 S 已经收到的应用层数据字节数是多少?

3) 若题 47-a 表中的某个 IP 分组在 S 发出时的前 40 字节如题 47-b 表所示, 则该 IP 分组到达 H 时经过了多少个路由器?

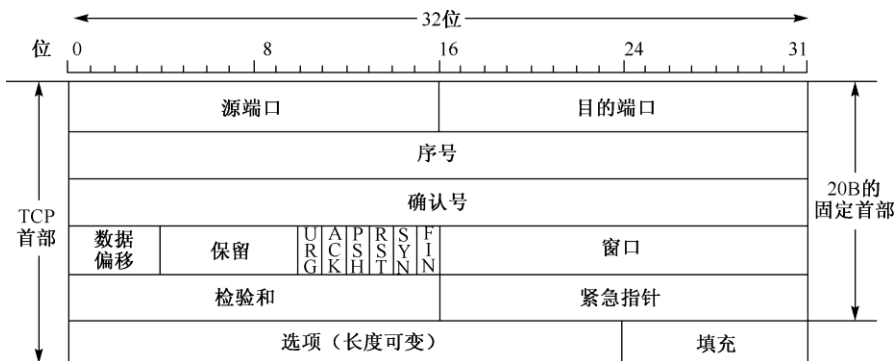
题 47-b 表

来自 S 的分组	45 00 00 28	68 11 40 00	40 06 ec ad	d3 44 47 50	ca 76 01 06
	13 88 a1 08	e0 59 9f f0	84 6b 41 d6	50 10 16 d0	b7 d6 00 00

注: IP 分组头和 TCP 段头结构分别如题 47-a 图, 题 47-b 图所示。



题 47-a 图 IP 分组头结构



题 47-b 图 TCP 段头结构

2012 年计算机学科专业基础综合试题参考答案

一、单项选择题

- | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 1. B | 2. A | 3. A | 4. B | 5. C | 6. C | 7. C | 8. A |
| 9. D | 10. A | 11. D | 12. D | 13. B | 14. D | 15. D | 16. A |
| 17. C | 18. C | 19. C | 20. D | 21. D | 22. B | 23. C | 24. B |
| 25. B | 26. A | 27. D | 28. A | 29. B | 30. C | 31. A | 32. B |
| 33. B | 34. C | 35. A | 36. B | 37. C | 38. A | 39. D | 40. D |

二、综合应用题

41. 解答：

本题同时对多个知识点进行了综合考查。对有序表进行两两合并考查了归并排序中的 Merge() 函数；对合并过程的设计考查了哈夫曼树和最佳归并树。外部排序属于大纲新增考点。

1) 对于长度分别为 m, n 的两个有序表的合并，最坏情况下是一直比较到两个表尾元素，比较次数为 $m+n-1$ 次。故最坏情况的比较次数依赖于表长，为了缩短总的比较次数，根据哈夫曼树（最佳归并树）思想的启发，可采用如图所示的合并顺序。

根据上图中的哈夫曼树，6 个序列的合并过程为：

第 1 次合并：表 A 与表 B 合并，生成含有 45 个元素的表 AB；

第 2 次合并：表 AB 与表 C 合并，生成含有 85 个元素的表 ABC；

第 3 次合并：表 D 与表 E 合并，生成含有 110 个元素的表 DE；

第 4 次合并：表 ABC 与表 DE 合并，生成含有 195 个元素的表 ABCDE；

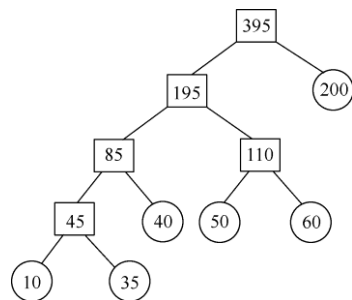
第 5 次合并：表 ABCDE 与表 F 合并，生成含有 395 个元素的最终表。

由上述分析可知，最坏情况下的比较次数为：第 1 次合并，最多比较次数 $=10+35-1=44$ ；第 2 次合并，最多比较次数 $=45+40-1=84$ ；第 3 次合并，最多比较次数 $=50+60-1=109$ ；第 4 次合并，最多比较次数 $=85+110-1=194$ ；第 5 次合并，最多比较次数 $=195+200-1=394$ 。

故比较的总次数最多为： $44+84+109+194+394=825$ 。

2) 各表的合并策略是：在对多个有序表进行两两合并时，若表长不同，则最坏情况下总的比较次数依赖于表的合并次序。可以借用哈夫曼树的构造思想，依次选择最短的两个表进行合并，可以获得最坏情况下最佳的合并效率。

【1) 2) 评分说明】①对于用类似哈夫曼树（或最佳归并树）思想进行合并，过程描述正确，给 5 分。按其他策略进行合并，过程描述正确，给 3 分。

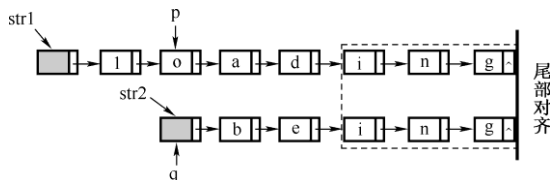


② 正确算出与合并过程一致的总比较次数，给 2 分。若计算过程正确，但结果错误，可给 1 分。

③ 考生只要说明采用的是类似哈夫曼树（或最佳归并树）的构造方法作为合并策略，即可给 3 分。如果采用其他策略，只要能够完成合并，给 2 分。

42. 解答：

顺序遍历两个链表到尾结点时，并不能保证两个链表同时到达尾结点。这是因为两个链表的长度不同。假设一个链表比另一个链表长 k 个结点，我们先在长链表上遍历 k 个结点，之后同步遍历两个链表，这样就能够保证它们同时到达最后一个结点。由于两个链表从第一个公共结点到链表的尾结点都是重合的，所以它们肯定同时到达第一个公共结点。



算法的基本设计思想：

① 分别求出 $str1$ 和 $str2$ 所指向的两个链表的长度 m 和 n 。

② 将两个链表以表尾对齐：令指针 p 、 q 分别指向 $str1$ 和 $str2$ 的头结点，若 $m \geq n$ ，则使 p 指向链表中的第 $m-n+1$ 个结点；若 $m < n$ ，则使 q 指向链表中的第 $n-m+1$ 个结点，即使指针 p 和 q 所指的结点到表尾的长度相等。

③ 反复将指针 p 和 q 同步向后移动，并判断它们是否指向同一结点。若 p 和 q 指向同一结点，则该点即为所求的共同后缀的起始位置。

2) 算法的 C 语言代码描述：

```

LinkNode *Find_1st_Common(LinkList str1, LinkList str2) {
    int len1 = Length(str1), len2 = Length(str2);
    LinkNode *p, *q;
    for (p = str1; len1 > len2; len1--) // 使 p 指向的链表与 q 指向的链表等长
        p = p->next;
    for (q = str2; len1 < len2; len2--) // 使 q 指向的链表与 p 指向的链表等长
        q = q->next;
    while (p->next != NULL && p->next != q->next) { // 查找共同后缀起始点
        p = p->next; // 两个指针同步向后移动
        q = q->next;
    }
    return p->next; // 返回共同后缀的起始点
}

```

【1)、2) 的评分说明】 ① 若考生所给算法实现正确，且时间复杂度为 $O(m+n)$ ，可给 12 分；若算法正确，但时间复杂度超过 $O(m+n)$ ，则最高可给 9 分。

② 若在算法的基本设计思想描述中因文字表达没有非常清晰反映出算法思路，但在算法实现中能够清晰看出算法思想且正确的，可参照①的标准给分。

③ 若算法的基本设计思想描述或算法实现中部分正确，可参照①中各种情况的相应给分标准酌情给分。

3) 时间复杂度为： $O(len1+len2)$ 或 $O(\max(len1, len2))$ ，其中 $len1$ 、 $len2$ 分别为两个链表的长度。

【3）的评分说明】 若考生所估计的时间复杂度与考生所实现的算法一致，可给 1 分。

43. 解答：

本题综合涉及到多个考点：计算机的性能指标、存储器的性能指标、DMA 的性能分析，DMA 方式的特点，多体交叉存储器的性能分析。

1) 平均每秒 CPU 执行的指令数为： $80\text{M}/4=20\text{M}$ ，故 MIPS 数为 20；（1 分）

平均每条指令访存 1.5 次，故平均每秒 Cache 缺失的次数= $20\text{M} \times 1.5 \times (1-99\%)=300\text{k}$ ；（1 分）

当 Cache 缺失时，CPU 访问主存，主存与 Cache 之间以块为传送单位，此时，主存带宽为 $16\text{B} \times 300\text{k/s} = 4.8\text{MB/s}$ 。在不考虑 DMA 传送的情况下，主存带宽至少达到 4.8MB/s 才能满足 CPU 的访存要求。（2 分）

2) 题中假定在 Cache 缺失的情况下访问主存，平均每秒产生缺页中断 $300000 \times 0.0005\% = 1.5$ 次。因为存储器总线宽度为 32 位，所以每传送 32 位数据，磁盘控制器发出一次 DMA 请求，故平均每秒磁盘 DMA 请求的次数至少为 $1.5 \times 4\text{KB}/4\text{B} = 1.5\text{K} = 1536$ 。（2 分）

3) CPU 和 DMA 控制器同时要求使用存储器总线时，DMA 请求优先级更高；（1 分）

因为 DMA 请求得不到及时响应，I/O 传输数据可能会丢失。（1 分）

4) 4 体交叉存储模式能提供的最大带宽为 $4 \times 4\text{B}/50\text{ns} = 320\text{MB/s}$ 。（2 分）

44. 解答：

1) x 的机器码为 $[x]_{\text{机}} = 1111\ 1101\ 1111\ 1111\text{B}$ ，即指令执行前 $(R1) = \text{FDFFH}$ ，右移 1 位后位 $1111\ 1110\ 1111\ 1111\text{B}$ ，即指令执行后 $(R1) = \text{FEFFH}$ 。（2 分）

2) 每个时钟周期只能有一条指令进入流水线，从第 5 个时钟周期开始，每个时钟周期都会有一条指令执行完毕，故至少需要 $4 + (5 - 1) = 8$ 个时钟周期。（2 分）

3) I_3 的 ID 段被阻塞的原因：因为 I_3 与 I_1 和 I_2 都存在数据相关，需等到 I_1 和 I_2 将结果写回寄存器后， I_3 才能读寄存器内容，所以 I_3 的 ID 段被阻塞（1 分）。 I_4 的 IF 段被阻塞的原因：因为 I_4 的前一条指令 I_3 在 ID 段被阻塞，所以 I_4 的 IF 段被阻塞（1 分）。

注意：要求“按序发射，按序完成”，故第 2 小问中下一条指令的 IF 必须和上一条指令的 ID 并行，以免因上一条指令发生冲突而导致下一条指令先执行完。

4) 因 $2 \times x$ 操作有左移和加法两种实现方法，故 $x = x * 2 + a$ 对应的指令序列为

```

I1  LOAD      R1, [x]
I2  LOAD      R2, [a]
I3  SHL       R1          //或者   ADD    R1, R1
I4  ADD       R1, R2
I5  STORE     R2, [x]

```

这 5 条指令在流水线中执行过程如下图所示。

指令	时间单元																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
I1	IF	ID	EX	M	WB												
I2		IF	ID	EX	M	WB											
I3			IF			ID	EX	M	WB								
I4						IF				ID	EX	M	WB				
I5										IF				ID	EX	M	WB

故执行 $x = x * 2 + a$ 语句最少需要 17 个时钟周期。

45. 解答：

1) 页框号为 21。理由：因为起始驻留集为空，而 0 页对应的页框为空闲链表中的第三个空闲页框 21，其对应的页框号为 21。

2) 页框号为 32。理由：因 $11 > 10$ 故发生第三轮扫描，页号为 1 的页框在第二轮已处于空闲页框链表中，此刻该页又被重新访问，因此应被重新放回驻留集中，其页框号为 32。

3) 页框号为 41。理由：因为第 2 页从来没有被访问过，它不在驻留集中，因此从空闲页框链表中取出链表头的页框 41，页框号为 41。

4) 合适。理由：如果程序的时间局部性越好，从空闲页框链表中重新取回的机会越大，该策略的优势越明显。

46. 解答：

1) 文件系统中所能容纳的磁盘块总数为 $4\text{TB}/1\text{KB}=2^{32}$ 。要完全表示所有磁盘块，索引项中的块号最少要占 $32/8=4\text{B}$ 。而索引表区仅采用直接索引结构，故 512B 的索引表区能容纳 $512\text{B}/4\text{B}=128$ 个索引项。每个索引项对应一个磁盘块，所以该系统可支持的单个文件最大长度是 $128 \times 1\text{KB}=128\text{KB}$ 。

2) 这里的考查的分配方式不同于我们所熟悉的三种经典分配方式，但是题目中给出了详细的解释。所求的单个文件最大长度一共包含两部分：预分配的连续空间和直接索引区。

连续区块数占 2B ，共可以表示 2^{16} 个磁盘块，即 2^{26}B 。直接索引区共 $504\text{B}/6\text{B}=84$ 个索引项。所以该系统可支持的单个文件最大长度是 $2^{26}\text{B}+84\text{KB}$ 。

为了使单个文件的长度达到最大，应使连续区的块数字段表示的空间大小尽可能接近系统最大容量 4TB 。分别设起始块号和块数分别占 4B ，这样起始块号可以寻址的范围是 2^{32} 个磁盘块，共 4TB ，即整个系统空间。同样的，块数字段可以表示最多 2^{32} 个磁盘块，共 4TB 。

47. 解答：

【解析】1) 由题 47-a 表看出，源 IP 地址为 IP 分组头的第 13~16 字节。在表 5-1 中 1、3、4 号分组的源 IP 地址均为 $192.168.0.8$ (c0a8 0008H)，所以 1、3、4 号分组是由 H 发送的。

题 47-a 表中，1 号分组封装的 TCP 段的 $\text{SYN}=1$ ， $\text{ACK}=0$ ， $\text{seq}=846\text{b } 41\text{c5H}$ ；2 号分组封装的 TCP 段的 $\text{SYN}=1$ ， $\text{ACK}=1$ ， $\text{seq}=\text{e059 } 9\text{fefH}$ ， $\text{ack}=846\text{b } 41\text{c6H}$ ；3 号分组封装的 TCP 段的 $\text{ACK}=1$ ， $\text{seq}=846\text{b } 41\text{c6H}$ ， $\text{ack}=\text{e059 } 9\text{ff0H}$ ，所以 1、2、3 号分组完成了 TCP 连接的建立过程。

由于快速以太网数据帧有效载荷的最小长度为 46 字节，表 5-1 中 3、5 号分组的总长度为 40 (28H) 字节，小于 46 字节，其余分组总长度均大于 46 字节。所以 3、5 号分组通过快速以太网传输时需要填充。

2) 由 3 号分组封装的 TCP 段可知，发送应用层数据初始序号为 $\text{seq}=846\text{b } 41\text{c6H}$ ，由 5 号分组封装的 TCP 段可知， ack 为 $\text{seq}=846\text{b } 41\text{d6H}$ ，所以 S 已经收到的应用层数据的字节数为 $846\text{b } 41\text{d6H} - 846\text{b } 41\text{c6H}=10\text{H}=16$ 字节。

【评分说明】其他正确解答，亦给 2 分；若解答结果不正确，但分析过程正确给 1 分；其他情况酌情给分。

3) 由于 S 发出的 IP 分组的标识= 6811H ，所以该分组所对应的是表 5-1 中的 5 号分组。S 发出的 IP 分组的 $\text{TTL}=40\text{H}=64$ ，5 号分组的 $\text{TTL}=31\text{H}=49$ ， $64-49=15$ ，所以，可以推断该 IP 分组到达 H 时经过了 15 个路由器。

【评分说明】若解答结果不正确，但分析过程正确给 1 分；其他情况酌情给分。