

中国科学院大学
2019年招收攻读硕士研究生入学统一考试试题
科目名称：计算机学科综合（专业）

考生须知：

1. 本试卷满分150分，全部考试总计时间180分钟。
 2. 所有答案必须写在答题纸上，写在试题纸上或草稿纸上一律无效。
-

1. 单项选择题：第1~40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. 通常操作系统区分用户态和内核态的主要目的是（ ）。
A. 提高执行性能 B. 实现数据保护
C. 代码结构清楚 D. 方便数据通信
2. 进程控制块是操作系统管理进程的重要数据结构，通常进程控制块不包含（ ）。
A. 进程标识符 B. 处理机状态 C. I/O缓冲区 D. 进程资源清单
3. 关于多进程临界区同步机制的准则，以下表述错误的是（ ）。
A. 当无进程处于临界区时，应允许一个请求进入临界区的进程立即进入自己的临界区
B. 当已有进程进入临界区时，其它试图进入临界区的进程必须等待
C. 对要求访问临界资源的进程，应保证在有限时间内能进入自己的临界区
D. 当进程不能进入自己的临界区时，应保持已有资源并持续探测临界区是否空闲
4. 相比于进程，线程的优点包括（ ）。
A. 创建开销小 B. 并发性高 C. 切换的开销小 D. 以上全部
5. 下列作业调度的CPU调度算法中，公平性最好的调度算法是（ ）。
A. 先入先出调度 B. 优先级抢占式调度
C. 时间片轮转调度 D. 最短作业优先调度
6. 某处理器具有36位虚拟地址空间，页面大小为4KB。该处理器采用了3级页表，每级页表包含12位页号，那么第一级页表的一个表项，对应了（ ）的虚拟内存空间。
A. 64GB B. 32GB C. 4GB D. 256MB
7. 某存储系统采用5块磁盘组成RAID-5模式（4个数据，1个校验），那么应用写一个完整扇区大小的数据块时，通常会导致（ ）块磁盘参与写操作。
A. 1 B. 2 C. 4 D. 5
8. 设备管理中，（ ）不是引入缓冲区的主要原因。
A. 缓和CPU与I/O设备速度不匹配的矛盾
B. 减少对CPU的中断频率
C. 隔离CPU和I/O设备提高系统安全
D. 提高CPU和I/O设备之间的并行性
9. 文件系统中，在定位一个文件某偏移量所在磁盘块位置时，下列哪个信息是不需要的？（ ）
A. 页表 B. 索引块 C. 磁盘分区表 D. 根目录

10. 系统安全管理中, 通常公开密钥(非对称密钥)算法不用于()。

A. 数字签名 B. 身份认证 C. 大量数据加密 D. 密钥交换

11. 以下关于冯·诺依曼计算机结构及工作原理的表述, 正确的是()。

- A. 须由运算器、存储器、加速器、控制器和输入/输出(I/O)设备组成
- B. 指令使用二进制代码进行存储, 数据使用十六进制代码进行存储
- C. 作为冯·诺依曼计算机的基础, 存储程序是指被程序处理的数据与程序算法模型
的参数一并存放于存储器中
- D. 数据和指令以同等地位存储, CPU依赖指令周期的不同阶段加以区分

12. 对于传统的冯诺依曼机器来讲, 哪个位于主存储器中

- A. 地址寄存器和数据寄存器
- B. 程序计数器和指令寄存器
- C. 指令寄存器和数据寄存器
- D. 数据寄存器和累加寄存器

13. 某一型号智能手机使用ARMv8指令集体系结构, 其CPU主频为1.80 GHz; 另一型号便携式计算机使用x86-64指令集体系结构, 其CPU主频为2.90 GHz。使用同一个程序算法的C语言源代码, 分别在这两台计算设备上成功编译为各自的可执行二进制代码, 运行后得到平均CPI分别为2.90和1.80。据此, 可推算出此程序在两台计算设备上运行时间的比值是()。注: 计算结果精确到小数点后两位。

A. 0.39 B. 1.00 C. 2.60 D. 无法确定

14. 对于n路组相连接映的Cache, 在保存n及缓存内部数据字存储总量不变的前提下, 将Cache块大小增加一倍, 以下表述正确的是()

- A. 字块内地址(offset)位宽加一, 主存字块标记字段(tag)位宽减一, cache中总块数减少一半
- B. 字块内地址(offset)位宽加一, 主存字块标记字段(tag)位宽不变, cache中总块数减少一半
- C. 字块内地址(offset)位宽减一, 主存字块标记字段(tag)位宽加一, cache中总块数不变
- D. 字块内地址(offset)位宽增加一倍, 主存字块标记字段(tag)位宽减少一半, cache中总块数不变

15. 利用了程序访存局部性原理的cache替换算法或策略是()。

- I. 随机算法(Random) II. 近期最少使用算法(Least Recently Used, LRU)
 - III. 先进先出算法(First-In-First-Out, FIFO) IV. 直接映射cache中的替换策略
- A. 只有II B. II和III C. I、III和IV D. II、III和IV

解析: 近期少使用算法考虑了程序局部访问性原理, 其他三项均为考虑程序局部性原理, 故选A

16. 假设sizeof(int)=4 bytes, 考虑以下C语言代码:

```
int cod=2;
printf("%d\n", ~cod);
```

上述程序段的输出结果为()。注: “~”为按位取反操作。

A. -3 B. -2 C. 1 D. 4294967293

16. 假设sizeof(int)=4 bytes, 考虑以下C语言代码:

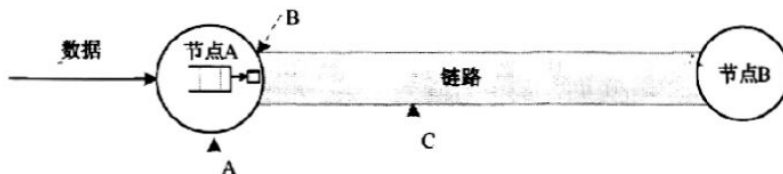
```
int cod=2;
printf("%d\n", ~cod);
```

上述程序段的输出结果为()。注: “~”为按位取反操作。

A. -3 B. -2 C. 1 D. 4294967293

17. 某团队设计了一款面向深度神经网络模型训练算法加速的专用处理芯片, 此专用加速芯片通过总线访问内存, 总线地址位宽18位, 总线数据位宽64位, 总线时钟频率800MHz, 每个时钟周期在上升沿和下降沿各传输一次数据, 则此加速芯片与内存之间的最大数据传输率为()。
- A. 16400 MB/s B. 12800 MB/s C. 6400 MB/s D. 3600 MB/s
18. 为了通过软硬件协同来实现上述专用处理芯片的灵活可编程特性, 团队拟设计一套专用指令系统。已确定的指令条数为30条。考虑到未来算法更新将可能产生新的操作类别, 决定再预留40条指令定义。假设二进制编码的操作码长度固定, 那么此指令系统至少需要()位的操作码。
- A. 5 B. 6 C. 7 D. 8
19. 超标量流水线技术的核心是()。
- A. 每个时钟周期内同时并发多条指令, 且指令之间独立
B. 将多条可并行操作的指令通过编译技术组合成一条具有多个操作码字段的指令
C. 将一个时钟周期再分段, 复用功能部件
D. 同一时间内, 某些段正在实现某种运算操作, 而另外一些段却在实现另一种运算
20. 某台计算机有4级中断, 硬件排队优先级从高到低为1→2→3→4。修改设置1级中断的屏蔽字为1001, 2级中断的屏蔽字为1111, 3级中断的屏蔽字为1011, 4级中断的屏蔽字为0001, 则修改后, 中断响应优先级为()。
- A. 1→2→3→4 B. 2→3→1→4 C. 3→1→4→2 D. 4→1→3→2
21. 以下关于CSMA/CD的表述, 正确的是()。
- A. 站点在发送完帧之后再对冲突进行检测
B. 站点在发送帧期间, 同时对冲突进行检测
C. 发送帧和检测冲突并不是在同一个站上进行
D. 在一个站上发送的帧, 只有当另一个站没有收到时, 才对冲突进行检测
22. IP数据包需要经由源主机和中途路由器到达目的主机, 以下表述正确的是()。
- A. 源主机和中途路由器都知道数据包到达目的主机需要经过的完整路径
B. 源主机知道数据包到达目的主机需要经过的完整路径, 而中途路由器不知道
C. 源主机不知道数据包到达目的主机需要经过的完整路径, 而中途路由器知道
D. 源主机和中途路由器都不知道数据包到达目的主机需要经过的完整路径
23. 差错检测技术用于检测传输过程中出现的错误, 以下关于差错检测技术的表述, 错误的是()。
- A. 差错检测的功能并不一定由网络体系中的某一确定层次完成, 可能有多层次同时采用不同的差错检测技术对网络所传输的信息进行差错检测
B. 绝大部分的数据链路层技术都提供差错检测功能, 循环冗余校验是该层常用的技术之一
C. IP报头采用校验和技术, 使得最终的目的节点在接收到数据包之后, 能够判断该数据包报头各字段是否在传输过程中被篡改过
D. 传输层协议采用校验和技术, 对伪首部及传输层报文进行校验, 以验证消息是否在两个确定的端点之间被正确传输
24. 数据链路层采用选择重传协议(SR)传输数据, 发送方已发送了0-3号数据帧, 现已收到0号帧和2号帧的确认, 而1、3号帧依次超时, 则此时需要重传的帧数是()。
- A. 1 B. 2 C. 3 D. 4
25. 下面可用于流量控制的是()。
- A. 滑动窗口 B. SNMP C. UDP D. RARP

26. 时延是评价计算机网络的重要指标, 下图标示了时延可能发生的位置, 以下关于时延的表述, 错误的是 ()。



- A. 数据从源节点到目的节点传输过程中需经历的时延并不只是由链路带宽决定
- B. 处理时延发生在图中位置A, 指节点在收到数据时进行一些必要的处理所花费的时间
- C. 发送时延发生在图中位置B, 指节点发送数据所需要的时间
- D. 传播时延发生在图中位置C, 由链路带宽等链路特性所决定

27. 一个长度为4000字节的IP分组(含报头与数据部分), 需经过MTU为1500字节的网络, 被其传输路径上的路由器分片为三个小的IP分组。以下关于这一分片事件的表述, 正确的是 ()。

- A. 分片后的三个分组, 其头部偏移字段的值分别为0、1500、3000
- B. 分片后的三个分组, 其头部校验和字段的值与原IP分组校验和字段的值相同
- C. 分片后形成的IP分组可以在后续传输路径上再被分片
- D. 当新形成的三个IP分组都到达下一跳路由器时, 将被重组为原IP分组后进行转发

28. 设有2条路由222.15.221.0/24和222.15.218.0/24, 如果进行路由汇聚, 能覆盖这2条路由的是 ()。

- A. 222.15.208.0/20
- B. 222.15.208.0/21
- C. 222.15.216.0/22
- D. 222.15.220.0/22

29. 主机A向主机B连续发送了两个TCP报文段, 其序号分别为80和120; 主机B收到这两个报文段后向主机A回复确认消息, 其中的确认号为191。这两个报文段携带的数据大小分别为 ()。

- A. 第一个报文段携带40字节, 第二个报文段携带70字节
- B. 第一个报文段携带40字节, 第二个报文段携带71字节
- C. 第一个报文段携带41字节, 第二个报文段携带72字节
- D. 第一个报文段携带41字节, 第二个报文段携带71字节

30. 某自治系统采用RIP协议, 若路由器R1收到其邻居路由器R2的距离矢量中包含信息<net1, 16>, 可能得出的结论是 ()。

- A. R2可以经过R1到达net1, 跳数为17
- B. R1和R2都可以到达net1, 跳数为16
- C. R1可以经过R2到达net1, 跳数为17
- D. R1不能经过R2到达net1

31. 若某个线性表最常用的操作是在值为x的结点之前和之后插入元素, 那选择存储方式 () 可使得最常用操作最快。

- A. 顺序表
- B. 单链表
- C. 双向链表
- D. 单循环链表

32. 循环队列A[0...m-1]存放数据元素, 队列长度为n, 指针r指向队尾, 则当前队头元素的实际位置为 ()。

- A. $(*r-n+1) \bmod m$
- B. $(*r-n+m) \bmod m$
- C. $(*r-n+1+m) \bmod m$
- D. $(*r-n) \bmod m$

33. 按照KMP算法, 模式串p="ACABBACA"的next函数为 ()

- A. 01121123
- B. 01011123
- C. 01111123
- D. 01121113

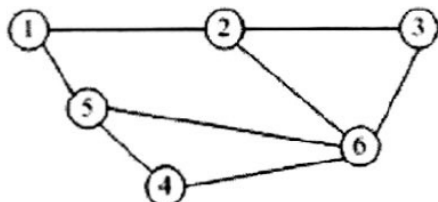
34. 设有广义表L=((x, (a, b)), (x, (a,b), y)), 则运算head(head(tail(A)))的结果是 ()。

- A. (x)
- B. x
- C. ()
- D. a

35. 有5个字符, 根据其使用频率对其进行哈夫曼编码, 以下不可能是哈夫曼编码的是 ()。

- A. 000, 001, 010, 011, 1 B. 0000, 0001, 001, 01, 1
C. 000, 001, 01, 10, 11 D. 00, 100, 101, 110, 111

36. 对下图从顶点2出发, 进行深度优先遍历, 正确的顶点序列为 ()



- A. 2, 3, 6, 5, 1, 4 B. 2, 3, 1, 6, 5, 4
C. 2, 1, 5, 3, 6, 4 D. 2, 3, 5, 6, 4, 1

37. 已知有向图 $G=(V, E)$, 其中 $V=\{1, 2, 3, 4, 5, 6, 7\}$, $E=\{<1, 2>, <1, 3>, <1, 4>, <2, 5>, <3, 5>, <3, 6>, <4, 6>, <5, 7>, <6, 7>\}$, G 的拓扑排序序列是 ()。

- A. 1, 3, 4, 6, 2, 5, 7 B. 1, 3, 2, 6, 4, 5, 7
C. 1, 3, 4, 5, 2, 6, 7 D. 1, 2, 5, 3, 4, 6, 7

38. 已知AOE网 $G=(V, E)$, 其中 $V=\{1, 2, 3, 4, 5, 6\}$, $E=\{<1, 2, 3>, <1, 3, 8>, <3, 2, 4>, <2, 4, 9>, <2, 5, 6>, <3, 5, 10>, <4, 6, 6>, <5, 6, 9>\}$ (注: 边用三元组表示, 分别表示边的起点、终点和边上的权重), 那么, 加快 () 边的进度就可以缩短工期。

- A. $<1, 2>$ B. $<3, 2>$ C. $<2, 5>$ D. $<3, 5>$

39. 在顺序表(2, 6, 9, 17, 19, 24, 28, 30)中, 用折半查找查找关键字29需要与表中的元素进行 () 次比较, 才能得到查找失败的结果。

- A. 2 B. 3 C. 4 D. 5

40. 某地中学生生物竞赛有2000人参加, 只录取成绩优异的前10名, 对落选的其他考生, 无需排出名次。这种情况下, 用 () 排序方法能最快输出结果。

- A. 希尔(Shell)排序 B. 堆排序 C. 快速排序 D. 归并排序

2. 综合应用题: 41~48小题, 共70分。

41. `int sleep(int seconds)` 系统调用的作用是将进程暂时挂起一段时间。试简要描述一个进程A调用`sleep(5)`睡眠后恢复执行的过程中, 操作系统内部的处理流程(包括时钟中断的作用)。(8分)

42. 某计算机系统采用请求分页虚拟存储系统，页面大小为4KB，进程数据可用的物理内存页面为2MB（内核和进程的代码、栈不参与换页），页面替换采用LRU算法。已知每次访问需要100ns，每次进行一个页面磁盘交换需要10ms。如下所示程序，请回答以下问题：（7分）

```
#define PAGE_SIZE 4096
#define PAGE_NUM 1048
char data[PAGE_SIZE*PAGE_NUM];
1: for(i=0;i<PAGE_NUM;i++) data[i*PAGE_SIZE] = 1;
2: data[1024] = 2;
3: data[512*PAGE_SIZE+2048] = 3;
4: data[3072] = 4;
5: data[768*PAGE_SIZE] = 5;
```

（1）假设语句1执行前，data数组都不在cache和内存中，那么语句1执行完大约需要多少时间？

（2）假设条件同（1），语句1执行后，语句2、3、4、5执行时间分别是多少？

（3）假设系统采用的是随机页面替换策略，且执行语句3时发生了换页，那么语句4执行时间大于1ms的概率是多少？

43. 为了实现面向深度神经网络模型训练算法加速的专用处理芯片与CPU处理器芯片之间的数据交互，系统架构师拟采用程序查询或程序中断两种方式。假设CPU处理器时钟主频为1GHz，需要从专用加速芯片读取1000字节的处理结果数据。专用芯片每0.02毫秒（ms）可产生1字节输出结果数据供处理器读取。忽略结果数据的处理时间以及数据在接口电路中的缓冲延时和传输等时间。请回答下列问题：（10分）

（1）如果CPU采用程序查询方式读取专用芯片输出结果数据，并且每次查询操作耗时60个时钟周期。那么此专用芯片每提供1字节数据需要花费多少个处理器时钟周期？

(2) 结合问题 (1) 的场景, CPU 处理器芯片每读取专用芯片 1 字节数据, 需要在这之前查询多少次? 注: 如果计算结果不为整数, 请进行向上取整操作。

(3) 结合问题 (1) 和 (2) 的场景, 读取 1000 字节的数据需要花费多少个处理器时钟周期?

44. 一位计算机高级软件架构师希望通过在某台特定计算设备上运行精巧设计的测试代码, 未获得其内部处理器高速数据缓存的硬件组织结构, 从而编写更好的应用软件以充分利用cache并发挥此计算设备性能。此架构师通过一个测试代码在这台计算设备上分别产生了四个不同的访存地址序列, 并测得相对应的缓存命中率, 如下表所示。

访存序列编号	访存地址序列 (地址使用十进制表示)	缓存命中率
A	0, 2, 4, 8, 16, 32	33.33%
B	0, 512, 1024, 1536, 2048, 1536, 1024, 512, 0	33.33%
C	0, 64, 128, 256, 512, 256, 128, 64, 0	33.33%
D	0, 512, 1024, 0, 1536, 0, 2048, 512	25%

此外, 有以下前提假设:

- I. 在每个访存地址序列发出之前, 测试代码均将缓存清空。
- II. 主存储器按字节编址。
- III. 处理器每次访问主存储器均为单字节访问。

现在, 架构师希望推理出此高速缓存的硬件组织结构及相关参数, 具体包括:

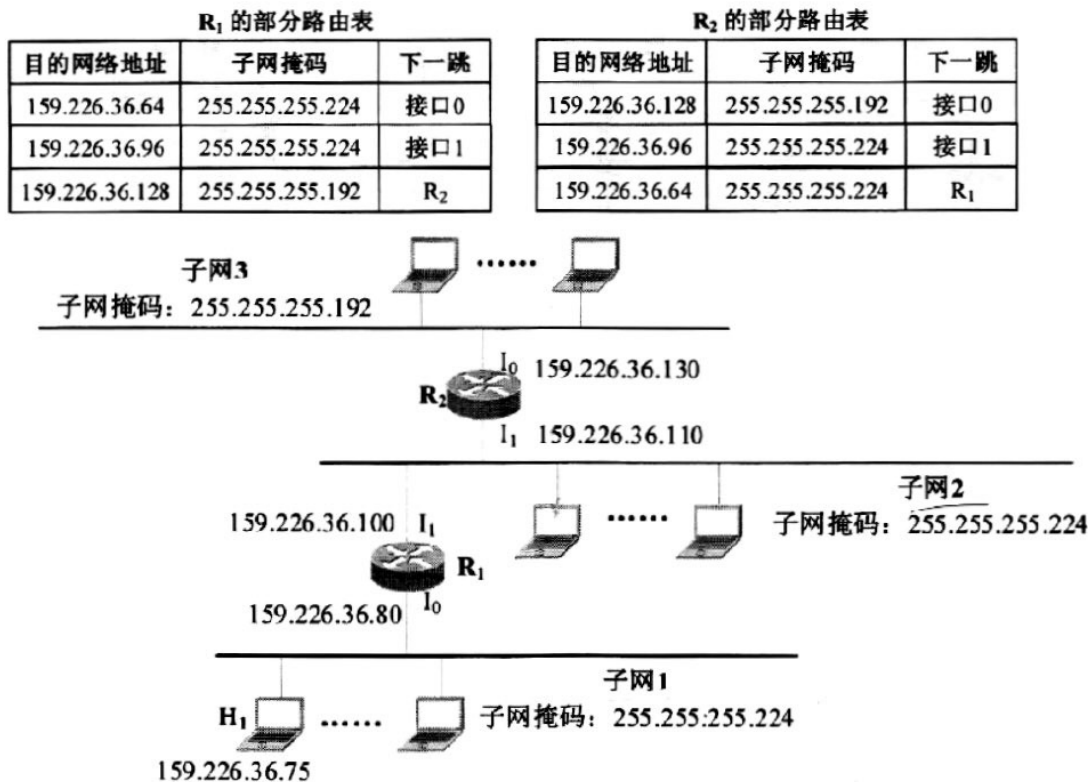
- (1) Cache块大小是多少字节? (可选值范围: 1、2、4、8、16或32字节)
- (2) 此cache是1路、2路还是4路组相联?
- (3) 此cache的容量是256字节还是512字节?
- (4) 此cache采用的是FIFO还是LRU替换算法?

请根据上述信息协助架构师进行推理, 在上述可选值范围内选择你认为最为合理的数值或算法名称, 给出你的结果并简要阐述推理过程。(10分)

45. 已知某个C类网络: 211.81.192.0; 现要将该网络划分为14个子网, 每个子网至少可容纳12台主机 (子网号和主机号均不能为全0和全1)。请回答下列问题: (7分)

- (1) 求出子网的掩码。
- (2) 每个子网最多能容纳的主机数。
- (3) 分别写出第一个子网和最后一个子网的网络地址空间。

46. 在下图所示的网络中，子网1、子网2的子网掩码均为255.255.255.224，子网3的子网掩码为255.255.255.192。子网1与子网2通过路由器R1相连接，R1的I0接口地址配置为159.226.36.80，I1接口地址配置为159.226.36.100。子网2与子网3通过路由器R2相连接，R2的I1接口地址配置为159.226.36.130，I0接口地址配置为159.226.36.110。图中给出了路由器R1和R2的部分路由表信息，请回答以下问题：（8分）



(1) 请分析，当子网1中的主机H1（IP地址为159.226.36.75）发送目的地址为159.226.36.91的IP分组时，相关节点（包括源主机及分组传输路径上的每一个路由节点）在网络层的查表及转发过程。

(2) 请分析，当子网1中的主机H1（IP地址为159.226.36.75）发送目的地址为159.226.36.179的IP分组时，相关节点（包括源主机及分组传输路径上的每一个路由节点）在网络层的查表及转发过程。

47. 将关键字序列{116, 100, 101, 115, 117, 103}依次插入到初始为空的平衡二叉树 (AVL树), 给出每插入一个关键字后的平衡树, 并说明其中可能包含的平衡调整步骤 (即, 先说明是哪个结点失去平衡, 然后说明做了什么平衡处理); 然后分别给出前序、中序和后序遍历该二叉树的输出结果。
(10分)

48. 待排序列采用带头结点的单链表, 请写出在其上进行快速排序的递归算法。需要给出相关的数据结构、算法和算法思想。(10分)