

计算机操作系统-题库

第一章 操作系统引论

一. 选择题

1. 操作系统是一种_____。

- A. 通用软件
- B. 系统软件
- C. 应用软件
- D. 软件包

答: B

2. 操作系统的_____管理部分负责对进程进行调度。

- A. 主存储器
- B. 控制器
- C. 运算器
- D. 处理机

答: D

3. 操作系统是对_____进行管理的软件。

- A. 软件
- B. 硬件
- C. 计算机资源
- D. 应用程序

答: C

4. 从用户的观点看, 操作系统是_____。

- A. 用户与计算机之间的接口
- B. 控制和管理计算机资源的软件
- C. 合理地组织计算机工作流程的软件
- D. 由若干层次的程序按一定的结构组成

答: A

5. 操作系统的功能是进行处理机管理、_____管理、设备管理及信息管理。

- A. 进程
- B. 存储器
- C. 硬件
- D. 软件

答: B

6. 操作系统中采用多道程序设计技术提高 CPU 和外部设备的_____。

- A. 利用率
- B. 可靠性
- C. 稳定性
- D. 兼容性

答: A

7. 操作系统是现代计算机系统不可缺少的组成部分, 是为了提高计算机的_____和方便用户使用计算机而配备的一种系统软件。

- A. 速度
- B. 利用率
- C. 灵活性
- D. 兼容性

答: B

8. 操作系统的基本类型主要有_____。

- A. 批处理系统、分时系统及多任务系统
- B. 实时操作系统、批处理操作系统及分时操作系统

C. 单用户系统、多用户系统及批处理系统

D. 实时系统、分时系统和多用户系统

答: B

9. 所谓_____是指将一个以上的作业放入主存, 并且同时处于运行状态, 这些作业共享处理机的时间和外围设备等其他资源。

A. 多重处理

B. 多道程序设计

C. 实时处理

D. 并行执行

答: B

10. 下面关于操作系统的叙述中正确的是_____。

A. 批处理作业必须具有作业控制信息。

B. 分时系统不一定都具有人机交互功能。

C. 从响应时间的角度看, 实时系统与分时系统差不多。

D. 由于采用了分时技术, 用户可以独占计算机的资源。

答: A

11. 如果分时操作系统的时间片一定, 那么_____, 则响应时间越长。

A. 用户数少

B. 用户数越多

C. 内存越少

D. 内存越多

答: B

12. 实时操作系统必须在_____内完成来自外部的事件。

A. 响应时间

B. 周转时间

C. 规定时间

D. 调度时间

答: C

13. 分时系统中为了使多个用户能够同时与系统交互, 最关键的问题是_____。

A. 计算机具有足够快的运算速度

B. 能快速进行内外存之间的信息交换

C. 系统能够及时接收多个用户的输入

D. 短时间内所有用户程序都能运行

答: C

14. 分时操作系统通常采用_____策略为用户服务。

A. 可靠性和灵活性

B. 时间片轮转

C. 时间片加权分配

D. 短作业优先

答: B

15. 在 _____操作系统控制下计算机系统能及时处理由过程控制反馈的数据并作出响应。

A. 实时

B. 分时

C. 分布式

D. 单用户

答: A

16. 下面 6 个系统中, 必须是实时操作系统的有 _____ 个。

计算机辅助设计系统

航空订票系统

过程控制系统

机器翻译系统

办公自动化系统

计算机激光照排系统

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

答：C

17. 若把操作系统看作计算机系统资源的管理者，下列的_____不属于操作系统所管理的资源。

A. 程序 B. 内存 C. CPU D. 中断

答：D

18. 在下列操作系统的各个功能组成部分中，_____不需要硬件的支持。

A. 进程调度 B. 时钟管理
C. 地址映射 D. 中断系统

答：A

19. 设计实时操作系统时，首先应考虑系统的_____。

A. 可靠性和灵活性 B. 实时性和可靠性
C. 多路性和可靠性 D. 优良性和分配性

答：B

20. 分时系统需要使用下面哪些成分？

(1)多道程序设计技术 (2)作业说明书
(3)终端命令解释程序 (4)中断处理
(5)优先级调度 (6)系统调用

答：(1)、(3)、(4)、(6)

21. 时间片概念一般用于_____。

A. 批处理操作系统 B. 分时操作系统
C. 实时操作系统 D. 都不是

答：B

22. 批处理系统的主要缺点是_____。

A. CPU 使用效率低 B. 无并行性
C. 无交互性 D. 都不是

答：C

23. 下面_____不属于操作系统功能。

A. 用户管理 B. CPU 和存储管理
C. 设备管理 D. 文件和作业管理

答：A

24. 当 CPU 执行系统程序时，CPU 处于_____。

A. 管态 B. 目态
C. 系统态 D. A 和 C

答：D

25. 能直接对系统中各类资源进行动态分配和管理，控制、协调各任务的并行执行且系统中主机无主次之分，并向用户提供统一的、有效的软件接口的系统是_____。

A. 分布式操作系统 B. 实时操作系统

C. 网络操作系统 D. 批处理操作系统

答: A

26. 下列选择中, _____不是操作系统关心的主要问题。

- A. 管理计算机裸机 B. 设计、提供用户程序与计算机硬件系统的界面
C. 管理计算机系统资源 D. 高级程序设计语言的编译器

答: D

27. 实时操作系统对可靠件和安全性的要求极高, 它_____。

- A. 十分注意系统资源的利用率 B. 不强调响应速度
C. 不强求系统资源的利用率 D. 不必向用户反馈信息

答: C

28. 火车站的售票系统属于_____系统。

- A. 单道批处理 B. 多道批处理
C. 实时 D. 分时

答: C

29. 在分时系统中, 通常的时间片是_____。

- A. 几分钟 B. 几十秒
C. 几十毫秒 D. 几十微秒

答: C

30. 用于控制生产流水线、进行工业处理控制的操作系统是_____。

- A. 分时系统 B. 网络操作系统
C. 实时系统 D. 批处理系统

答: C

31. 分时操作系统适用于_____。

- A. 控制生产流水线 B. 调试运行程序
C. 大量的数据处理 D. 多个计算机资源共享

答: B

32. 下列系统中, 属于实时系统的是_____。

- A. 办公自动化系统 B. 计算机辅助设计系统
C. 飞机订票系统 D. 计算机激光照排系统

答: C

33. 目前个人计算机中的操作系统主要是_____。

- A. 网络操作系统 B. 批处理操作系统
C. 单用户操作系统 D. 单道单用户操作系统

答: C

34. 下列操作系统中强调并行计算的操作系统是_____。

- A. 分时系统 B. 实时系统
C. 网络操作系统 D. 分布式操作系统

答: D

35. 能够实现通信及资源共享的操作系统是_____。

- A. 批处理操作系统 B. 分时操作系统
C. 实时操作系统 D. 网络操作系统

答：D

36. UNIX 操作系统是一种_____。

- A. 分时操作系统 B. 批处理操作系统
C. 实时操作系统 D. 分布式操作系统

答：A

37. 以下操作系统中属于网络操作系统的是_____。

- A. MS-DOS B. Windows98 C. UNIX D. WindowsNT

答：D

二. 填空题

1. 操作系统的基本功能包括 ①_____管理、②_____管理、③_____管理、④_____管理。除此之外还为用户使用操作系统提供了用户接口。

答：①处理机 ②存储器 ③ 设备 ④信息(或文件)

2. 如果一个操作系统兼有批处理、分时处理和实时处理操作系统三者或其中两者的功能，这样的操作系统称为_____。

答：通用操作系统

3. 在分时和批处理系统结合的操作系统中引入了“前台”和“后台”作业的概念，其目的是_____。

答：为了提高 CPU 的利用率

4. 实时操作系统与分时操作系统的主要区别是_____。

答：及时性和高可靠性

5. 在主机控制下进行的输入 / 输出操作称为_____操作。

答：联机输入输出

6. 按内存中同时运行程序的数目可以将批处理系统分为两类：①_____和 ②_____。

答：①单道批处理系统 ②多道批处理系统

7. _____系统不允许用户随时干预自己程序的运行。

答：批处理

8. 多道程序设计的特点是____道、_____和_____。

答：宏观上的并行、微观上的串行

9. 单道批处理系统是在解决_____和_____的矛盾中发展起来的。

答：人机矛盾、CPU 和 I/O 设备之间速度不匹配

10. 对操作系统的总体设计目标来说，批处理操作系统应注重提高计算机的效率，尽量增加系统的_____，分时操作系统应保证用户_____；而实时操作系统则应在保证及时响应和处理有关事件的前提下，再考虑_____。

答：平均吞吐量、所能忍受的响应时间、系统资源的利用率

11. 分时系统中的_____是衡量一个分时系统性能的重要指标。

答：响应时间

12. 导弹飞行控制系统属于_____系统。

答：实时

13. 在操作系统的发展过程中, 多道程序设计和_____的出现, 标志了操作系统的正式形成。

答: 分时系统

14. 实时系统按应用的不同分为_____和_____两种。

答: 过程控制系统、数据处理系统

15. 分时操作系统的用户通过与计算机相连的_____来使用计算机系统, 允许多个用户同时上机。

答: 终端

16. 个人计算机操作系统每次只允许一个用户使用计算机, 称为_____。

答: 单用户操作系统

17. 为计算机网络配置的操作系统称_____。

答: 网络操作系统

18. 网络操作系统的主要功能是实现_____以及_____。

答: 各台计算机之间的通信、网络中各种资源的共享

19. 近年来, 适合开放系统的操作系统成为一个重要的研究课题, IEEE 把它命名为_____, 即计算机环境可移植操作系统。

答: POSIX

20. 在_____控制下, 用户在终端设备上可以直接输入、调试和运行自己的程序, 能直接修改程序中的错误, 并直接获得结果。

答: 分时系统

21. 根据服务对象不同, 常用的单处理机 OS 可以分为如下三种类型:

允许多个用户在其终端上同时交互地使用计算机的 OS 称为_____系统, 它通常采用_____策略为用户服务;

允许用户把若干个作业提交计算机系统集中处理的 OS 称为_____系统, 衡量这种系统性能的一个主要指标是系统的_____;

在_____系统控制下, 计算机系统能及时处理由过程控制反馈的数据并做出响应。设计这种系统时应首先考虑系统的_____。

答: 分时、时间片轮转、批处理、吞吐量、实时、及时性和可靠性

三. 简答题

1. 叙述操作系统在计算机系统的位置。

答: 操作系统是运行在计算机硬件系统上的最基本的系统软件。它控制和管理着所有的系统硬件(CPU、主存、各种硬件部件和外部设备等), 也控制和管理着所有的系统软件(系统程序和用户进程等), 操作系统对计算机使用者提供了一种良好的操作环境, 也为其他各种应用系统提供了最基本的支撑环境。

现代操作系统是一个复杂的软件系统, 它与计算机硬件系统有着千丝万缕的联系, 也与用户有着密不可分的关系, 它在计算机系统中位于计算机裸机和计算机用户之间, 如下图所示(图 1.1), 紧挨着硬件的就是操作系统, 它通过系统核心程序对计算机系统中的几类资源进行管理, 如处理机、存储器、输入输出设备、数据与文档资源、用户作业等, 并向用户提供若干服务, 通过这些服务将所有对硬件的复杂操作隐藏起来, 为用户提供一个透明的操作环境。

在操作系统的外层是其他系统软件。操作系统是最基本的系统软件。用户可以直接通过系统软件层与计算机打交道, 也可以建立各类应用软件和系统, 通过它们来解决用户的问题。

由此可见, 操作系统是介于计算机硬件和用户之间的一个接口。

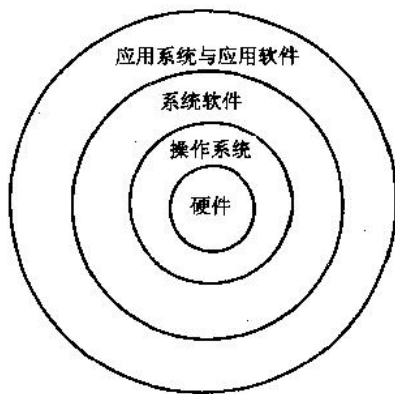


图 1.1 操作系统的位置

2. 怎样理解“由于计算机上装有操作系统，从而扩展了原计算机的功能”？

答：计算机系统的硬件结构和机器一级的操作包含了诸如指令集、存储器组织、总线结构和输入输出部件等的操作与控制，这些最基本的操作恰恰又是最复杂和最难以由用户直接进行的操作。例如：用户要进行文件读写，而文件是以二进制代码的方式存放在磁盘、磁带等存储装置中，需要有一种途径把用户的要求转换成对具体的硬件部件、电路信号、选择开关等的细微操作，用户自己不可能完成这些操作，但操作系统把用户的高级操作转换成一系列的低级操作，最终完成文件的读写。所有的低级操作对用户来讲都是透明的，即无需用户关心的、看不见的，操作系统把硬件全部隐藏起来，给用户提供了一个友好的、易于操作的界面。此外，操作系统还要进行大量的系统事务处理。如响应中断的发生、处理定时操作、管理存储器及其他低级操作。所以，可以说操作系统是硬件系统的扩展，从而扩展了原计算机的功能，它比直接对计算机硬件系统进行操作要容易得多。

3. 试对分时操作系统和实时操作系统进行比较。

答：我们可以从以下几个方面对这两种操作系统进行比较：

(1) 实时信息处理系统与分时操作系统一样都能为多个用户服务，系统按分时原则为多个终端用户服务；而对实时控制系统，则表现为经常对多路现场信息进行采集以及对多个对象或多个执行机构进行控制。

(2) 实时信息处理系统与分时操作系统一样，每个用户各占一个终端，彼此独立操作，互不干扰。因此用户感觉就像他一人独占计算机；而在实时控制系统中信息的采集和对对象的控制也都是彼此互不干扰的。

(3) 实时信息系统对响应时间的要求与分时操作系统类似，都是以人所能接受的等待时间来确定的；而实时控制系统的响应时间则是以控制对象所能接受的延时来确定的。

(4) 分时操作系统是一种通用系统，主要用于运行终端用户程序，因此它具有较强的交互能力。而实时操作系统虽然也有交互能力，但其交互能力不及前者。

(5) 分时操作系统要求系统可靠，相比之下，实时操作系统则要求系统高度可靠。

4. 设计实时操作系统的主要困难是什么？

答：设计适用于实时环境的操作系统的主要困难是：在实时环境规定的时间限额内对用户作出相应的反应。如果系统不能在规定的时间内完成指定的任务，那么就可能导致整个实时系统的崩溃。因此，在设计这类操作系统时，设计者必须保证所采用的调度策略及相关技术不会使响应时间超过实时环境所规定的时间限额。

5. 采用多道程序设计的主要优点是什么？

答：多道程序设计考虑到作业的运行规律是交替使用 CPU 和 I/O，故将多道程序同时保存于系统中，使各作业对 CPU 与 I/O 的使用在时间上重叠，提高了 CPU 和 I/O 设备的利用率。

6. 一个分层结构操作系统由裸机，用户，CPU 管理，文件管理，作业管理，内存管理，设备管理，命令管理等部分组成。试按层次结构的原则从内到外将各部分重新排列。

答：按层次结构原则从内到外依次为：裸机，CPU 管理，内存管理，设备管理，文件管理，作业管理，命令管理，用户。

7. 什么是操作系统？它有什么基本特征？

答：操作系统是指控制和管理计算机的软、硬件资源，合理组织计算机的工作流程、方便用户使用的程序集合。操作系统具有 4 个基本特征：

- (1) 并发性：宏观上在一段时间内有多道程序在同时运行，而微观上这些程序是在交替执行。
- (2) 共享性：因程序的并发执行而使系统中的软、硬件资源不再为某个程序独占，而是由多个程序共同使用。
- (3) 虚拟性：多道程序设计技术把一台物理计算机虚拟为多台逻辑上的计算机，使得每个用户都感觉自己是“独占”计算机。
- (4) 异步性(不确定性)：多道程序系统中，各程序之间存在着直接或间接的联系，程序的推进速度受到其他程序的影响，这样，程序运行的顺序、程序完成的时间以及程序运行的结果都是不确定的。

8. 操作系统是随着多道程序设计技术的出现逐步发展起来的，要保证多道程序的正常运行，在技术上需要解决哪些基本问题？

- 答：(1) 在多道程序之间应如何分配被它们共享的处理机，使得 CPU 既能满足各程序运行的需要，又有较高的利用率；此外，一旦将 CPU 分配给某程序后，应何时收回。
- (2) 如何为每道程序分配必要的内存空间，使它们各得其所但又不会因相互重叠而丢失信息；此外，应如何防止因某道程序出现异常情况而破坏其他程序。
- (3) 系统中可能有多种类型的 I/O 设备供多道程序共享，应如何分配这些 I/O 设备，如何做到既方便用户的使用，又能提高设备的利用率。
- (4) 在现代计算机系统中通常都存在着大量的程序和数据，应如何组织它们才便于用户使用；此外，还有信息保存的安全性和一致性问题。
- (5) 对于系统中的各种应用程序，它们有的属于计算型，有的属于 I/O 型，有些既重要又紧迫，有些又要求系统能及时响应，这时系统应如何组织和安排这些作业(程序)的工作流程。

实际上，这些问题的全体就是操作系统的核心内容。

9. 如何理解操作系统的不确定性？

答：在单道程序环境下，由于程序独占全机软、硬件资源且完全是在一个封闭环境下运行，所以，如果使用相同的输入，则多次运行的结果必然相同。

引入多道程序技术后，由于多道程序共享全机软、硬件资源，即各程序之间存在着直接或间接的联系。此时程序的运行就不是在一个封闭的环境下，程序的运行要受到其他程序和系统资源的影响，这会引引起程序运行的结果、程序运行的顺序和完成的时间不确定。此外，在多道程序环境下，外部设备的中断、输入输出的请求等出现的时间都是不确定的。这些不确定因素给操作系统设计带来了很大的复杂性。对于程序执行结果的不确定是绝不允许的，这是操作系统必须解决的主要问题；而程序运行顺序及完成时间的不确定是多道程序环境所决定的，因此，操作系统应能对运行顺序进行协调和管理。至于外部设备的中断、输入输出请求甚至程序运行故障等不确定性，操作系统应在出现的时刻及时响应并予以解决，如果操作系统的设计考虑不周，也可能造成系统的混乱和崩溃。

第二章 进程管理

一. 选择题

1. 在进程管理中, 当_____时进程从阻塞状态变为就绪状态.

- A. 进程被进程调度程序选中 B. 等待某一事件
C. 等待的事件发生 D. 时间片用完

答: C

2. 分配到必要的资源并获得处理机时的进程状态是_____。

- A. 就绪状态 B. 执行状态
C. 阻塞状态 D. 撤消状态

答: B

3. P、V 操作是_____。

- A. 两条低级进程通信原语 B. 两组不同的机器指令
C. 两条系统调用命令 D. 两条高级进程通信原语

答: A

4. 对进程的管理和控制使用_____。

- A. 指令 B. 原语
C. 信号量 D. 信箱通信

答: B

5. 进程的并发执行是指若干个进程_____。

- A. 同时执行 B. 在执行的时间上是重叠的
C. 在执行的时间上是不可重叠的 D. 共享系统资源

答: B

6. 若 P、V 操作的信号量 S 初值为 2, 当前值为-1, 则表示有_____等待进程。

- A. 0 个 B. 1 个 C. 2 个 D. 3 个

答: B

7. 进程的三个基本状态在一定条件下可以相互转化, 进程由就绪状态变为运行状态的条件是__①__；由运行状态变为阻塞状态的条件是__②__。

- A. 时间片用完 B. 等待某事件发生
C. 等待的某事件已发生 D. 被进程调度程序选中

答, ①D ②B

8. 下列的进程状态变化中, _____变化是不可能发生的。

- A. 运行—就绪 B. 运行—等待
C. 等待—运行 D. 等待—就绪

答: C

9. 一个运行的进程用完了分配给它的时间片后, 它的状态变为_____。

- A. 就绪 B. 等待
C. 运行 D. 由用户自己确定

答: A

10. 用 P、V 操作管理临界区时, 信号量的初值应定义为_____。

- A. -1 B. 0 C. 1 D. 任意值

答: C

11. 用 V 操作唤醒一个等待进程时, 被唤醒进程的状态变为_____.

- A. 等待 B. 就绪 C. 运行 D. 完成

答: B

12. 进程间的同步是指进程间在逻辑上的相互_____关系。

- A. 联接 B. 制约 C. 继续 D. 调用

答: B

13. _____是一种只能进行 P 操作和 V 操作的特殊变量。

- A. 调度 B. 进程 C. 同步 D. 信号量

答: D

14. 下面对进程的描述中, 错误的是 _____。

- A. 进程是动态的概念 B. 进程执行需要处理机
C. 进程是有生命期的 D. 进程是指令的集合

答: D

15. 下面的叙述中正确的是_____。

- A. 操作系统的一个重要概念是进程, 因此不同进程所执行的代码也一定不同。
B. 为了避免发生进程死锁, 各进程只能逐个申请资源。
C. 操作系统用 PCB 管理进程, 用户进程可以从此 PCB 中读出与本身运行状况有关的信息
D. 进程同步是指某些进程之间在逻辑上的相互制约关系

答: D

16. 进程控制就是对系统中的进程实施有效的管理, 通过使用_____、进程撤销、进程阻塞、进程唤醒等进程控制原语实现。

- A. 进程运行 B. 进程管理
C. 进程创建 D. 进程同步

答: C

17. 操作系统通过_____对进程进行管理。

- A. JCB B. PCB C. DCT D. CHCT

答: B

18. 用 P、V 操作可以解决_____互斥问题。

- A. 一切 B. 某些 C. 正确 D. 错误

答: A

19. 通常, 用户进程被建立后, _____。

- A. 便一直存在于系统中, 直到被操作人员撤消
B. 随着作业运行正常或不正常结束而撤消
C. 随着时间片轮转而撤消与建立
D. 随着进程的阻塞或唤醒而撤消与建立

答: B

20. 一个进程被唤醒意味着_____。

- A. 该进程重新占有了 CPU B. 它的优先权变为最大
C. 其 PCB 移至等待队列队首 D. 进程变为就绪状态

答：D

21. 下面所述步骤中，_____不是创建进程所必需的。

- A. 由调度程序为进程分配 CPU
- B. 建立一个进程控制块
- C. 为进程分配内存
- D. 将进程控制块链入就绪队列

答：A

22. 多道程序环境下，操作系统分配资源以_____为基本单位。

- A. 程序
- B. 指令
- C. 进程
- D. 作业

答：C

23. 对于两个并发进程，设互斥信号量为 mutex，若 mutex=0，则_____。

- A. 表示没有进程进入临界区
- B. 表示有一个进程进入临界区
- C. 表示有一个进程进入临界区，另一个进程等待进入
- D. 表示有两个进程进入临界区

答：B

24. 两个进程合作完成一个任务，在并发执行中，一个进程要等待其合作伙伴发来消息，或者建立某个条件后再向前执行，这种制约性合作关系被称为进程的_____。

- A. 同步
- B. 互斥
- C. 调度
- D. 执行

答：A

25. 为了进行进程协调，进程之间应当具有一定的联系，这种联系通常采用进程间交换数据的方式进行，这种方式称为_____。

- A. 进程互斥
- B. 进程同步
- C. 进程制约
- D. 进程通信

答：D

26. 设有 5 个进程共享一个互斥段，如果最多允许有 3 个进程同时进入互斥段，则所采用的互斥信号量的初值应是_____。

- A. 5
- B. 3
- C. 1
- D. 0

答：B

27. 支持多道程序设计的操作系统在运行过程中，不断地选择新进程运行来实现 CPU 的共享，但其中_____不是引起操作系统选择新进程的直接原因。

- A. 运行进程的时间片用完
- B. 运行进程出错
- C. 运行进程要等待某一事件的发生
- D. 有新进程进入就绪状态

答：D

28. 在多道程序设计系统中，多个计算问题同时装入计算机系统的主存储器_____。

- A. 并发执行
- B. 顺序执行
- C. 并行执行
- D. 同时执行

答：A

29. 引入多道程序设计技术后，处理器的利用率_____。

- A. 有所改善
- B. 极大地提高
- C. 降低了
- D. 无变化，仅使程序执行方便

答：B

30. 从总体上说, 采用多道程序设计技术可_____单位时间的算题量, 但对每一个算题, 从算题开始到全部完成所需的时间比单道执行所需的时间可能要_____。

- A. 增加, 减少 B. 增加, 延长
C. 减少, 延长 D. 减少, 减少

答: B

31. 同一程序运行在不同的数据集上, 创建了_____的进程。

- A. 同步 B. 互斥
C. 相同 D. 不同

答: D

32. 关于进程, 下列叙述不正确的是_____。

- A. 进程包含了数据和运行其上的程序
B. 同一程序运行在不同的数据集合上时, 构成不同的进程
C. 一个被创建了的进程, 在它消亡之前, 总是处于三种基本状态之一
D. 若干个进程在单 CPU 系统中必须依次执行, 即一个进程完成后, 另一个进程才能开始工作

答: D

33. 对于一个单 CPU 系统, 允许若干进程同时执行, 轮流占用 CPU 称它们为_____的。

- A. 顺序执行 B. 同时执行
C. 并行执行 D. 并发执行

答: D

34. 一个程序与和它有关的进程的对应关系是_____。

- A. 一对一 B. 多对一
C. 一对多 D. 多对多

答: C

35. 在操作系统中, 使计算机系统能实现进程并发执行, 保证系统正常工作的机制是_____。

- A. 中断 B. 查询
C. 同步 D. 互斥

答: A

二. 填空题

1. 进程的基本特征有 ①、②、独立性、异步性及结构特征。

答: ①动态性 ②并发性

2. 信号量的物理意义是当信号量值大于零时表示 ①; 当信号量值小于零时其绝对值表示 ②。

答: ①可用资源的数目 ②因请求该资源而被阻塞的进程数目

3. 临界资源的概念是 ①, 而临界区是指 ②。

答: ①一次仅允许一个进程访问的资源

②进程中访问临界资源的那段程序代码

4. 进程在运行过程中有三种基本状态, 它们是 ①、②、③。

答: ①运行 ②就绪 ③等待(阻塞)

5. 进程主要由①、②、③三部分内容组成, 其中④是进程存在的唯一标志, 而⑤部分也可以为其他进程共享。

答: ①程序段 ②数据段 ③PCB ④PCB ⑤程序段

6. 系统中各进程之间逻辑上的相互制约关系称为_____。

答: 进程同步

7: 若一个进程已进入临界区, 其他欲进入临界区的进程必须_____。

答: 等待

8. 将进程的_____链接在一起就形成了进程队列。

答: PCB

9. 用P、V操作管理临界区时, 任何一个进程在进入临界区之前应调用①操作, 在退出临界区时应调用②操作。

答: ①P ②V

10. 用信箱实现通信时, 应有①和②两条基本原语。

答: ①发送 ②接收

11. 在多道程序系统中, 进程之间存在着的不同制约关系可以划分为两类: ①与②。③指进程间具有的一定逻辑关系; ④是指进程间在使用共享资源方面的约束关系。

答: ①同步 ②互斥 ③同步 ④互斥

12. 有m个进程共享同一临界资源, 若使用信号量机制实现对临界资源的互斥访问, 则信号量值的变化范围是_____。

答: 1至-(m-1)

13. 设系统中有 $n(n>2)$ 个进程, 且当前不再执行进程调度程序, 试考虑下述4种情况:

①没有运行进程, 有2个就绪进程, n个进程处于等待状态。

②有1个运行进程, 没有就绪进程, $n-1$ 进程处于等待状态。

③有1个运行进程, 有1个就绪进程, $n-2$ 进程处于等待状态。

④有1个运行进程, $n-1$ 个就绪进程, 没有进程处于等待状态。

上述情况中, 不可能发生的情况是_____。

答: ①

14. 在一个单处理机系统中, 若有5个用户进程, 且假设当前时刻为用户态, 则处于就绪状态的用户进程最多有①个, 最少有②个。

答: ①4 ②0

15. 操作系统中, 对信号量S的P原语操作定义中, 使进程进入相应等待队列等待的条件是_____。

答: $S<0$

16. 下面关于进程的叙述不正确的是_____。

①进程申请CPU得不到满足时, 其状态变为等待状态。

②在单CPU系统中, 任一时刻至多有一个进程处于运行状态。

③优先级是进行进程调度的重要依据, 一旦确定不能改变。

④进程获得处理机而运行是通过调度而实现的。

答: ③

17. 信箱逻辑上分成①和②两部分。①中存放有关信箱的描述。②由若干格子组成,

每格存放一信件，格子的数目和大小在创建信箱时确定。

答：①信箱头 ②信箱体

18. 当多个进程等待分配处理机时，系统按一种规定的策略从多个处于_____状态的进程选择一个进程，让它占有处理机，被选中的进程就进入了_____状态。

答：就绪、执行

19. 若使当前运行的进程总是优先级最高的进程，应选择_____进程调度算法。

答：抢占式（剥夺式）

20. 操作系统中用于完成一些特定功能的、不可中断的过程称为_____。

答：原语

三. 简答题

1. 在操作系统中为什么要引入进程概念?它与程序的关系是怎样的?

答：在操作系统中，由于多道程序并发执行时共享系统资源，共同决定这些资源的状态，因此系统中各程序在执行过程中就出现了相互制约的新关系，程序的执行出现“走走停停”的新状态。这些都是在程序执行的动态过程中发生的。用程序这个静态概念已不能如实反映程序并发执行过程中的这些特征。为此，人们引入“进程”这一概念来描述程序动态执行过程的性质。

进程和程序是既有联系又有区别的两个概念，进程不能脱离具体程序而虚设，程序规定了相应进程所要完成的动作。它们的主要区别如下：

- (1)程序是指令的有序集合，其本身没有任何运行的含义，它是一个静态的概念；进程是程序在处理机上的一次执行过程，它是一个动态概念。
- (2)程序的存在是永久的，而进程则是有生命期的，它因创建而产生，因调度而执行，因得不到资源而暂停，因撤销而消亡。
- (3)程序仅是指令的有序集合。而进程则由程序、数据和进程控制块组成。
- (4)在传统的操作系统中，进程是资源分配和调度运行的基本单位，而程序不是。
- (5)进程与程序之间不是一一对应的，即同一程序同时运行于若干不同的数据集合上，它将属于若干个不同的进程；而一个进程至少对应执行一个程序。

2. 为了实现并发进程间的合作和协调工作，以及保证系统的安全，操作系统在进程管理方面应做哪些工作?

- 答：(1)进程控制：系统必须设置一套控制机构来实现进程创建、进程撤消以及进程在运行过程中的状态转换。
- (2)进程同步：系统必须设置同步机制来对所有进程的运行进行协调，协调的方式包括进程的互斥和进程的同步。
- (3)进程通信：多道程序环境下可能需要诸进程合作完成一个任务，这些进程相互间需要通过交换信息来协调各自工作的进度。因此系统必须具有进程之间通信(交换信息)的能力。
- (4)进程调度：系统必须能够在处理机空闲时，按一定算法从就绪进程队列中选择一个就绪进程，把处理机分配给它，并为之设置运行的现场使其投入运行。

3. 在操作系统中引入线程概念的主要目的是什么?

答：减少程序并发执行时所需付出的时空开销，提高程序执行的并发度。特别有利于共享存储器的多处理机系统，和更加适合 c / s 模型。

4. 在一个单 CPU 的多道程序设计系统中，若在某一时刻有 N 个进程同时存在，那么处于运行态、等待态和就绪态进程的最小和最大值分别可能是多少?

答：若多道程序设计系统中，在某一时刻有 N 个进程同时存在，则处于运行态的进程最少 0 个，最大 1 个；处于就绪态的进程最少 0 个，最大 N—1 个；处于等待态的进程最少 0 个，最大 N 个。

5. 进程之间存在哪几种相互制约关系?各是什么原因引起的?下列活动分别属于哪种制约关系?

- (1)若干同学去图书馆借书;
- (2)两队举行篮球比赛;
- (3)流水线生产各道工序;
- (4)商品生产和社会消费

答:进程之间的制约关系分为直接制约关系(即同步)和间接制约关系(即互斥)。同步是因合作进程之间协调彼此的工作而控制自己的执行速度,即因相互合作、相互等待而产生的制约关系;而互斥是进程之间竞争临界资源而禁止两个以上的进程同时进入临界区所发生的制约关系。

- (1)属于互斥关系,因为一本书只能借给一个同学。
- (2)属于互斥关系,篮球只有一个,两队都要争夺。
- (3)居于同步关系,各道工序的开始都依赖前一道工序的完成。
- (4)属于同步关系,商品没生产出来则消费无法进行,商品没有消费完则无须再生产。

6. 为了让用户进程互斥地进入临界区,可以把整个临界区实现成不可中断的过程,即让用户具有屏蔽所有中断的能力。每当用户程序进入临界区的时候,屏蔽所有中断;当出了临界区的时候,再开放所有中断。你认为这种方法有什么缺点。

答:用户进程进入临界区时屏蔽所有中断,应当也包括系统程序。若屏蔽的是用户进程,的确可以保护临界资源,但若系统所发出的中断也被屏蔽,则会引起错误,因为系统的中断往往与当前运行的程序无关,却可能是一些重要的硬件中断,如电源故障等,故不可盲目屏蔽所有中断;又或者当时发出故障中断的中断源恰好是该临界资源,则更应该及时响应处理。

7. 设有 n 个进程共享一个互斥段,对于如下两种情况:

- (1)如果每次只允许一个进程进入互斥段;
- (2)如果每次最多允许 m 个进程 ($m < n$) 同时进入互斥段。

试问:所采用的互斥信号量初值是否相同?信号量的变化范围如何?

- 答: (1)信号量的初值为 1。信号量的变化范围是 $1, 0, -1, \dots, -(n-1)$ 。
- (2)信号量的初值为 m 。信号量的变化范围是: $m, m-1, \dots, 1, 0, \dots, -(n-m)$ 。

四. 应用题

1. 若程序 A 和 B 单独执行时分别需要 1 小时和 1.5 小时,其中 CPU 工作时间分别为 18 分钟和 27 分钟。若采用多道程序设计方法,让 A 和 B 并行工作,假定 CPU 利用率达到 50%,另加 15 分钟系统开销,请问系统效率能提高多少?

解:在多道系统中,程序 A 和 B 共用的 CPU 时间为: $(18 + 27) / 50\% = 90$ 分钟

系统效率提高 = $(A \text{ 和 } B \text{ 单独执行的时间总和} - \text{多道方式下总时间}) / A \text{ 和 } B \text{ 单独执行的时间总和}$,即 $((60 + 90) - (90 + 15)) / (60 + 90) = 45 / 150 = 30\%$

2. 在测量控制系统中,数据采集任务把所采集的数据送入一单缓冲区;计算任务从该单缓冲区中取出数据进行计算。试写出利用信号量机制实现两者共享单缓冲区的同步算法。

解:在本题中,应设置两个信号量 sf , se , 信号量 sf 表示缓冲区中是否有可供打印的计算结果,其初值为 0;信号量 se 用于表示缓冲区有无空位置存放新的信息,其初值为 1。

本题的同步描述如下:

```
int se=1;
int sf=0;
```

```

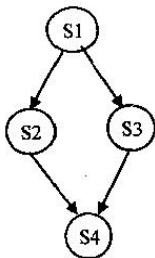
main()
{
    cobegin
        get ();
        compute ();
    coend
}

get ()
{
    while (采集工作未完成)
    {
        采集一个数据;
        p(se);
        将数据送入缓冲区;
        v(sf);
    }
}

compute ()
{
    while (计算工作未完成)
    {
        p(sf);
        从缓冲区中取出数据;
        v(se);
        进行数据计算;
    }
}

```

3. 下图给出了四个进程合作完成某一任务的前趋图，试说明这四个进程间的同步关系并用 P、V 操作描述它。



解：上图说明任务启动后 S1 先执行。当 S1 结束后，S2、S3 可以开始执行。S2、S3 完成后，S4 才能开始执行。为了确保这一执行顺序，设三个同步信号量 b2、b3、b4 分别表示进程 S2、S3、S4 是否可以开始执行，其初值均为 0。这四个进程的同步描述如下：

```

int b2=0;
int b3=0;

```

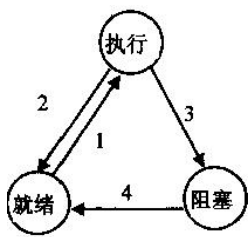


```

int b4=0;
main ( )
{
    cobegin
        s1 ( );
        s2 ( );
        s3 ( );
        s4 ( );
    coend
}
s1 ( )
{
    ...
    v(b2);
    v(b3);
}
s2 ( )
{
    p(b2);
    ...
    v(b4);
}
s3 ( )
{
    p(b3);
    ...
    v(b4);
}
s4 ( )
{
    p(b4);
    p(b4);    /*因在 s2 和 s3 完成后均对 b4 作了 v 操作，因此这里要用两个 p 操作*/
    ...
}

```

4. 某系统的进程状态转换图如下，请说明：



(1) 引起各种状态转换的典型事件有哪些？

(2) 当我们观察系统中某些进程时，能够看到某一进程产生的一次状态转换能引起另一进程作一次状态转换。在什么情况下，当一个进程发生转换 3 时能立即引起另一个进程发生转换 1？

(3) 试说明是否会发生下述因果转换：

2 → 1
3 → 2
4 → 1

答：(1) 在本题所给的进程状态转换图中，存在四种状态转换。当进程调度程序从就绪队列中选取一个进程投入运行时引起转换 1；正在执行的进程如因时间片用完而被暂停执行就会引起转换 2；正在执行的进程因等待的事件尚未发生而无法执行(如进程请求完成 I/O)则会引起转换 3；当进程等待的事件发生时(如 I/O 完成)则会引起转换 4。

(2) 如果就绪队列非空，则一个进程的转换 3 会立即引起另一个进程的转换 1。这是因为一个进程发生转换 3 意味着正在执行的进程由执行状态变为阻塞状态，这时处理机空闲，进程调度程序必然会从就绪队列中选取一个进程并将它投入运行，因此只要就绪队列非空，一个进程的转换 3 能立即引起另一个进程的转换 1。

(3) 所谓因果转换指的是有两个转换，一个转换的发生会引起另一个转换的发生，前一个转换称为因，后一个转换称为果，这两个转换称为因果转换。当然这种因果关系并不是什么时候都能发生，而是在一定条件下才会发生。

2 → 1：发生转换 2 时，就必然引起另一进程的转换 1。因为当发生转换 2 时，正在执行的进程从执行状态变为就绪状态，进程调度程序必然会从就绪队列中选取一个进程投入运行，即发生转换 1。

3 → 2：某个进程的转换 3 决不可能引起另一进程发生转换 2。这是因为当前执行进程从执行状态变为阻塞状态，不可能又从执行状态变为就绪状态。

4 → 1：当处理机空闲且就绪队列为空时，某一进程的转换 4 就会引起该进程的转换 1。因为此时处理机空闲，一旦某个进程发生转换 4，就意味着有一个进程从阻塞状态变为就绪状态，因而调度程序就会将就绪队列中的此进程投入运行。

5. 某车站售票厅，任何时刻最多可容纳 20 名购票者进入，当售票厅中少于 20 名购票者时，厅外的购票者可立即进入，否则需在外面等待。若把一个购票者看作一个进程，请回答下列问题：

(1) 用 PV 操作管理这些并发进程时，应怎样定义信号量，写出信号量的初值以及信号量各种取值的含义。

(2) 根据所定义的信号量，把应执行的 P、V 操作填入下面横线上，以保证进程能够正确地并发执行。

(3) 若欲购票者最多为 n 个人，写出信号量可能的变化范围(最大值和最小值)。

答：(1) 定义一信号量 S，初始值为 20，其意义如下：

$S > 0$ S 的值表示可继续进入售票厅的人数

$S = 0$ 表示售票厅中已有 20 名顾客(购票者)

$S < 0$ |S| 的值为等待进入售票厅的人数

(2) 根据所定义的信号量，把应执行的 P、V 操作填入下面横线上，以保证进程能够正确地并发执行。

COBEGIN PROCESS $P_i (i=1, 2, \dots)$

```

begin;
P(S)
进入售票厅;
购票;
退出;
V(S)
end;
COEND

```

(3) S 的最大值为 20; S 的最小值为 $20 - n$

6. 理发店里有一位理发师, 一把理发椅和 N 把供等候理发的顾客坐的椅子. 如果没有顾客, 则理发师便在理发椅上睡觉. 当一个顾客到来时, 他必须先唤醒理发师. 如果顾客到来时理发师正在理发, 则如果有空椅子, 可坐下来等; 否则离开。

解: 定义信号量如下:

```

Var Sn: semaphore; {位子数目, 初值为 n}
S: semaphore; {理发师睡觉, 初值为 1}
mutex: semaphore; {初值为 1}

```

用 P、V 操作实现如下:

顾客进程 i:

```

P(Sn); {门外观望}
P(mutex);
进门;
V(mutex);
V(S); {if (sn==n-1) v(s); }
等候;
理发;
V(Sn)
P(mutex);
出门;
V(mutex);

```

理发师进程 :

```

Repeat
P(S);
P(mutex);
叫人理发;
V(mutex);
理发;
Until false;

```

7. 试写出用加锁原语和开锁原语实现两个进程关于临界资源的操作的描述。

答: Program test

begin

s:=0 (表示该资源可用)

cobegin (1)

A: begin

⋮

加锁原语;

临界区 A;

开锁原语;

⋮

end

B: begin

⋮

加锁原语;

临界区 B;

开锁原语;

⋮

end

conend

end

8. 桌子上有一只盘子, 每次只能放入一只水果。爸爸专向盘中放苹果, 妈妈专向盘中放桔子, 一个儿子专等吃盘中的桔子, 一个女儿专等吃盘中的苹果。请利用 P、V 操作实现他们之间的同步。

解: 在本题中, 应设置三个信号量 s、so、sa, 信号量 s 表示盘子是否为空, 其初值为 1; 信号量 so 表示盘中是否有桔子, 其初值为 0; 信号量 sa 表示盘中是否有苹果, 其初值为 0。同步描述如下:

```
int s=1;
```

```
int sa=0;
```

```
int so=0;
```

```
main ( )
```

```
{
```

```
    cobegin
```

```
        father ( );
```

```
        son ( );
```

```
        daughter ( );
```

```
    coend
```

```
}
```

```
father ( )
```

```
{
```

```
    p(s);
```

```
    将水果放入盘中;
```

```

        if(放入的是桔子) v(so);
        else v(sa);
    }
    son ( )
    {
        p(so);
        从盘中取出桔子;
        v(s);
        吃桔子;
    }
    daughter ( )
    {
        p(sa);
        从盘中取出苹果;
        v(s);
        吃苹果;
    }
}

```

9. 桌子上有一只盘子，最多可容纳两个水果，每次只能放入或取出一个水果。爸爸专向盘子中放苹果(apple)，妈妈专向盘子中放桔子(orange)，两个儿子专等吃盘子中的桔子，两个女儿专等吃盘子中的苹果。请用 Pv 操作来实现爸爸、妈妈、儿子、女儿之间的同步与互斥关系。

解：盘子为互斥资源，因可以放两个水果，empty 初值为 2；再设信号量 mutex 初值为 1，控制对盘子的互斥访问；apple 表示盘中苹果个数，表示盘中桔子个数，初值均为 0。

parbegin

Father: begin

```

    L1: p(empty);
        P(mutex);
        放苹果;
        V(mutex);
        V(apple);
    Goto L1;
End;

```

Mother: begin

```

    L2: P(empty);
        P(mutex);
        放桔子;
        V(mutex);
        V(orange);
    Goto L2;
End;

```

Daughter: begin

```

L3: p(apple);
    P(mutex);
    取苹果;
    V(mutex);
    V(empty);
    Goto L3;
End;

```

Son: begin

```

L4: P(orange);
    P(mutex);
    取桔子;
    V(mutex);
    V(empty);
    Goto L4;
End;

```

Parend

10. 现为某临界资源设一把锁 w ，当 $w=1$ 时，表示关锁， $w=0$ 时，表示锁已打开，试写出开锁和关锁的原语，并说明如何利用它们去控制对该临界资源的互斥访问？

解：① 开锁原语 $unlock(w)$ 如下：

```
unlock(w):  w: =0
```

关锁原语 $lock(w)$ 如下：

```
Lock(w):  while w=1 do skip;    w: =1;
```

② 可设临界段 cs 放在两者之间来实现互斥，即

```
Lock(w);
```

```
cs;
```

```
unlock(w)
```

11. 用 PV 操作解决读者写者问题的正确程序如下：

```

begin S, Sr: Semaphore; rc: integer;
S:=1; Sr:=1; rc:=0;
cobegin PROCESS Reader i ( i=1,2...)
begin P(Sr)
rc:=rc+1;
if rc=1 then P(S);
V(Sr);
read file;
P(Sr);
rc:=rc-1
if rc=0 then V(S);
V(Sr);
end ;

```

PROCESS Writer j (j=1,2,...)

begin P(S);

Write file;

V(S)

end;

coend ;

end;

请回答：(1) 信号量 Sr 的作用；

(1) 程序中什么语句用于读写互斥，写写互斥；

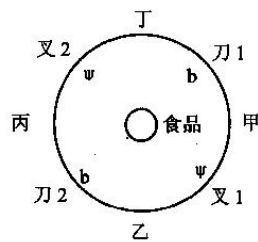
(2) (3) 若规定仅允许 5 个进程同时读怎样修改程序？

答：(1) Sr 用于读者计数 rc 的互斥信号量；

(2) if rc=1 then P (S) 中的 P (S) 用于读写互斥，写者进程中的 P (S) 用于写写互斥，读写互斥。

(3) 程序中增加一个信号量 S5，初值为 5，P (S5) 语句加在读者进程 P (Sr) 之前，V (S5) 语句加在读者进程第 2 个 V (Sr) 之后。

12. 哲学家甲请哲学家乙、丙、丁到某处讨论问题，约定全体到齐后开始讨论；在讨论的间隙四位哲学家进餐，每人进餐时都需使用刀、叉各一把，餐桌上的布置如下图所示。请用信号量及 P、v 操作说明这四位哲学家的同步、互斥过程。



(其中 b 表示刀，w 表示叉)

解：在本题中，应设置四个信号量 fork1、fork2、knife1、knife2，其初值均为 1，分别表示资源叉 1、叉 2、刀 1、刀 2 是否可用。同步描述如下：

```
int fork1=1;
```

```
int fork2=1;
```

```
int knife1=1;
```

```
int knife2=1;
```

```
main ( )
```

```
{
```

```
    cobegin
```

```
        pa ( );
```

```
        pb ( );
```

```
        pc ( );
```

```
        pd ( );
```

```
    coend;
```

```
}
```

```

pa ()
{
    while (1);
    {
        p (knife1);
        p(fork1);
        进餐;
        v(knife1);
        v(fork1);
        讨论问题;
    }
}

```

```

pb ()
{
    while (1);
    {
        p (knife2);
        p(fork1);
        进餐;
        v(knife2);
        v(fork1);
        讨论问题;
    }
}

```

```

pc ()
{
    while (1);
    {
        p (knife2);
        p(fork2);
        进餐;
        v(knife2);
        v(fork2);
        讨论问题;
    }
}

```

```

pd ()

```



```

{
    while (1);
    {
        p (knife1);
        p(fork2);
        进餐;
        v(knife1);
        v(fork1);
        讨论问题;
    }
}

```

13. 某数据库有一个写进程，多个读进程，它们之间读、写操作的互斥要求是：写进程正在写该数据库时不能有其他进程读该数据库，也不能有其他进程写该数据库；读进程之间不互斥，可以同时读该数据库。请用信号量及 P、V 操作描述这一组进程的工作过程。

解：在本题中，允许读进程同时读数据库，但写进程正在写数据库时不允许其他进程读数据库，也不允许其他进程写该数据库。为了解决读、写进程之间的同步，应设置两个信号量和一个共享变量：读互斥信号量 `rmutex`，用于使读进程互斥地访问共享变量 `count`，其初值为 1；写互斥信号量 `wmutex`，用于实现写进程与读进程的互斥及写进程与写进程的互斥，其初值为 1，共享变量 `count`，用于记录当前正在读数据库的读进程数目，初值为 0。

其工作过程如下：

```

int rmutex=1;
int wmutex=1;
int count=0;
main ( )
{
    cobegin
        reader ( );
        write ( );
    coend
}
reader ( )
{
    while(1)
    {
        p(rmutex);
        if(count=0) p(wmutex);
        count ++;
        v(rmutex);
        读数据库;
        p(rmutex);
    }
}

```

```

        count --
        if(count=0) v(wmutex);
        v(rmutex);
    }
}

write( )
{
    while(1)
    {
        p(wmutex);
        写数据库;
        v(wmutex);
    }
}

```

14. 设公共汽车上，司机和售票员的活动分别是：

司机的活动： 启动车辆；
 正常行车；
 到站停车；

售票员的活动：关车门；
 售票；
 开车门；

在汽车不断的到站、停车、行驶过程中，这两个活动有什么同步关系？用信号量和 P、V 操作实现他们的同步。

解：在汽车行驶过程中，司机活动与售票员活动之间的同步关系为：售票员关车门后，向司机发开车信号，司机接到开车信号后启动车辆，在汽车正常行驶过程中售票员售票，到站时司机停车，售票员在车停后开车门让乘客上下车。因此司机启动车辆的动作必须与售票员关车门的动作取得同步；售票员开车门的动作也必须与司机停车取得同步，在本题中，应设置两个信号量：s1、s2，s1 表示是否允许司机启动汽车，其初值为 0；

s2 表示是否允许售票员开门，其初值为 0。用 P、v 原语描述如下：

```

int s1=0;
int s2=0;
main ( )
{
    cobegin
        driver ( );
        busman ( );
    coend
}

driver ( )
{

```

```

while(1)
{
    p(s1);
    启动车辆;
    正常行车;
    到站停车;
    v(s2);
}
}
busman ( )
{
    while(1)
    {
        关车门;
        v(s1);
        售票;
        p(s2);
        开车门;
        上下乘客;
    }
}

```

15. 有一个仓库，可以存放 A 和 B 两种产品，仓库的存储空间足够大，但要求：

(1) 一次只能存入一种产品 (A 或 B)；

(2) $-N < A \text{ 产品数量} - B \text{ 产品数量} < M$ 。

其中，N 和 M 是正整数。试用“存放 A”和“存放 B”以及 P、v 操作描述产品 A 与产品 B 的入库过程。

解：由于不能直接在程序中使用表达式的值来控制产品 A 与产品 B 入库的同步与互斥过程，因此应将表达式转换成制约条件，即首先将表达式分解为：

$A \text{ 产品数量} - B \text{ 产品数量} < M$

$B \text{ 产品数量} - A \text{ 产品数量} < N$

进一步分析如下：

(1) 若只放入 A 产品而不放入 B 产品，则 A 产品最多可放 M-1 次。因此，设信号量 Sa 初值为 M-1。此外，每当放入一个 B 产品时，则使信号量 Sa 增 1，即 A 产品随之增加一次放入的机会。

(2) 若只放入 B 产品而不放入 A 产品，则 B 产品最多可放 N-1 次。因此，设信号量 Sb 初值为从 N-1。此外，每当放入一个 A 产品时，则使信号量 Sb 增 1，即 B 产品随之增加一次放入的机会。

因为每次只允许一种产品入库，所以设置信号量 mutex 控制对仓库的互斥访问。并发程序如下：

```

begin
    mutex:=1;
    Sa:=M-1;
    Sb:=N-1;

```

```

Cobegin
    PA: begin
        Repeat
            P(Sa);
            P(mutex);
            A 产品入库;
            V(mutex);
            V(Sb);
        Until false
    End
    PB: begin
        Repeat
            P(Sb);
            P(mutex);
            B 产品入库;
            V(mutex);
            V(Sa);
        Until false
    End
Coend
end

```

第三章 处理机调度与死锁

一. 选择题

1. 在为多道程序所提供的可共享的系统资源不足时，可能出现死锁。但是，不适当的_____也可能产生死锁。

- A. 进程优先权 B. 资源的线性分配
- C. 进程推进顺序 D. 分配队列优先权

答：C

2. 采用资源剥夺法可解除死锁，还可以采用_____方法解除死锁。

- A. 执行并行操作 B. 撤消进程
- C. 拒绝分配新资源 D. 修改信号量

答：B

4. 产生死锁的必要条件有四个，要防止死锁的发生，可以破坏这四个必要条件之一，但破坏_____条件是不大实际的。

- A. 互斥 B. 不可抢占
- C. 部分分配 D. 循环等待

答：A

5. 在分时操作系统中, 进程调度经常采用_____算法.

- A. 先来先服务 B. 最高优先权
- C. 时间片轮转 D. 随机

答: C

6. 资源的按序分配策略可以破坏_____条件。

- A. 互斥使用资源 B. 占有且等待资源
- C. 非抢夺资源 D. 循环等待资源

答: D

7. 在_____情况下, 系统出现死锁。

- A. 计算机系统发生了重大故障
- B. 有多个封锁的进程同时存在
- C. 若干进程因竞争资源而无休止地相互等待他方释放已占有的资源
- D. 资源数大大小于进程数或进程同时申请的资源数大大超过资源总数

答: C

8. 银行家算法是一种_____算法。

- A. 死锁解除 B. 死锁避免
- C. 死锁预防 D. 死锁检测

答: B

9. _____优先权是在创建进程时确定的, 确定之后在整个进程运行期间不再改变。

- A. 先来先服务 B. 静态
- C. 动态 D. 短作业

答: B

10. 某系统中有 3 个并发进程, 都需要同类资源 4 个, 试问该系统不会发生死锁的最少资源数是_____。

- A. 9 B. 10 C. 11 D. 12

答: B

11. 在下列解决死锁的方法中, 属于死锁预防策略的是_____。

- A. 银行家算法 B. 资源有序分配法
- C. 死锁检测法 D. 资源分配图化简法

答: B

12. 以下叙述中正确的是_____。

- A. 调度原语主要是按照一定的算法, 从阻塞队列中选择一个进程, 将处理机分配给它。
- B. 预防死锁的发生可以通过破坏产生死锁的四个必要条件之一来实现, 但破坏互斥条件的可能性不大。
- C. 进程进入临界区时要执行开锁原语。
- D. 既考虑作业等待时间, 又考虑作业执行时间的调度算法是先来先服务算法

答: B

13. _____是作业存在的惟一标志。

- A. 作业名 B. 进程控制块
C. 作业控制块 D. 程序名

答: C

14. 作业调度算法的选择常考虑因素之一是使系统有最高的吞吐量, 为此应_____。

- A. 不让处理机空闲 B. 能够处理尽可能多的作业
C. 使各类用户都满意 D. 不使系统过于复杂

答: B

15. 用户使用操作系统通常有三种手段, 它们是终端命令、系统调用命令和 _____。

- A. 计算机高级指令 B. 宏命令
C. 作业控制语言 D. 汇编语言

答: C

16. 在分时操作系统环境下运行的作业通常称为_____。

- A. 后台作业 B. 长作业
C. 终端型作业 D. 批量型作业

答: C

17. 当作业进入完成状态, 操作系统_____。

- A. 将删除该作业并收回其所占资源, 同时输出结果
B. 将该作业的控制块从当前作业队列中剔除, 收回其所占资源
C. 将收回该作业所占资源并输出结果
D. 将输出结果并删除内存中的作业

答: B

18. 在各种作业调度算法中, 若所有作业同时到达, 则平均等待时间最短的算法是_____。

- A. 先来先服务 B. 优先权调度
C. 最高响应比优先 D. 短作业优先

答: D

19. 既考虑作业等待时间, 又考虑作业执行时间的调度算法是_____。

- A. 响应比高者优先 B. 短作业优先
C. 优先级调度 D. 先来先服务

答: A

20. 作业调度程序从处于_____状态的队列中选取适当的作业投入运行。

- A. 运行 B. 提交 C. 完成 D. 后备

答: D

21. _____是指从作业提交给系统到作业完成的时间间隔。

- A. 周转时间 B. 响应时间
C. 等待时间 D. 运行时间

答: A

22. 作业从进入后备队列到被调度程序选中的时间间隔称为_____。

- A. 周转时间 B. 响应时间
C. 等待时间 D. 触发时间

答：C

23. 假设下述四个作业同时到达，当使用最高优先数优先调度算法时，作业的平均周转时间为_____小时。

作业	所需运行时间	优先数
1	2	4
2	5	9
3	8	1
4	3	8

- A. 4.5 B. 10.5 C. 4.75 D. 10.25

答：D

24. 下述作业调度算法中，_____调度算法与作业的估计运行时间有关。

- A. 先来先服务 B. 短作业优先
C. 优先权 D. 时间片轮转

答：B

25. 用户通过终端使用计算机系统控制作业执行的方式称为_____方式。

- A. 自动 B. 联机 C. 脱机 D. 假脱机

答：B

26. 作业生存期共经历四个状态，它们是提交、后备、_____和完成。

- A. 就绪 B. 执行 C. 等待 D. 开始

答：B

27. 系统在_____, 发生从目态到管态的转换。

- A. 发出 P 操作时 B. 发出 v 操作时
C. 执行系统调用时 D. 执行置程序状态字时

答：C

28. 以下叙述中正确的是_____。

- A. 操作系统的作业管理是一种微观的低级管理。
B. 作业的提交方式有两种，但对应的作业控制方式只有一种。
C. 一个作业从进入系统到运行结束，一般要经历的状态是：后备状态、就绪状态和完成状态。
D. 多道批处理与单道批处理的主要区别在于它必须有作业调度功能和进程调度功能，内存中可以存放多道作业。

答：D

29. 设有 4 个作业同时到达，每个作业的执行时间均为 2 小时，它们在一台处理机上按单道方式运行，则平均周转时间为_____。

- A. 1 小时 B. 5 小时 C. 2.5 小时 D. 8 小时

答：B

30. 现有 3 个同时到达的作业 J1, J2 和 J3, 它们的执行时间分别是 T1、T2 和 T3, 且 $T1 < T2 < T3$ 。

系统按单道方式运行且采用短作业优先算法，则平均周转时间是_____。

- A. $T_1 + T_2 + T_3$ B. $(T_1 + T_2 + T_3)/3$
C. $(3T_1 + 2T_2 + T_3)/3$ D. $(T_1 + 2T_2 + 3T_3)/3$

答：C

31. 一作业 8:00 到达系统，估计运行时间为 1 小时。若 10:00 开始执行该作业，其响应比是_____。

- A. 2 B. 1 C. 3 D. 0.5

答：C

32. 设有 3 个作业，它们的到达时间和运行时间如下表所示，并在一台处理机上按单道方式运行。如按响应比高者优先算法，则作业执行的次序是_____。

作业	到达时间	运行时间
1	8:00	2 时间
2	8:30	1 小时
3	9:30	0.25 小时

- A. J1,J2,J3 B. J1,J3,J2 C. J2,J3,J1 D. J3,J2,J1

答：B

33. 响应比是指_____。

- A. 作业计算时间与作业等待时间之比 B. 作业周转时间与作业计算时间之比
C. 系统调度时间与作业等待时间之比 D. 系统调度时间与作业计算时间之比

答：B

35. 设有 3 个作业 J1、J2、J3，其运行的时间分别为 1、2、3 小时；假定这些作业同时到达，并在一台处理机上按单道运行，则平均用转时间最小的执行序列是_____。

- A. J1,J2,J3 B. J1,J3,J2 C. J2,J1,J3 D. J2,J3,J1

答：

36. 当作业进入完成状态，操作系统_____。

- A. 将删除该作业并回收其所占资源，同时输出结果
B. 将该作业的作业控制块从当前作业队列中删除，回收其所占资源并输出结果
C. 将回收该作业所占资源并输出结果
D. 将删除内存中的作业并输出结果

答：A

37. 一个作业的完成要经过若干加工步骤，这每个步骤称为_____。

- A. 作业流 B. 子程序 C. 子进程 D. 作业步

答：D

38. 在批处理方式下，操作员把一批作业组织成_____向系统成批输入。

- A. 作业步 B. 作业流 C. 子程序 D. 程序组

答：B

39. 根据进程的紧迫性程度进行进程调度，应采用_____。

- A. 先来先服务调度算法 B. 最高优先级调度算法
C. 时间片轮转调度算法 D. 分级调度算法

答：B

40. 采用最高优先级调度算法时，对那些具有相同优先级的进程分配 CPU 的次序是_____。

- A. 时间片轮转 B. 运行时间长短
C. 先来先服务 D. 使用外围设备多少

答：C

二. 填空题

1. 死锁是指在系统中的多个_____无限期地等待永远不会发生的条件。

答：进程

2. 进程调度负责_____的分配工作。

答：处理机

3. 在_____调度算法中，按照进程进入就绪队列的先后次序来分配处理机。

答：先来先服务

4. 死锁产生的必要条件有四个，即_____①_____、_____②_____、_____③_____、_____④_____。

答：①互斥条件 ②不剥夺条件 ③请求和保持条件 ④环路等待条件

5. 解除死锁常用的方法有两种，_____①_____是从其他进程那里剥夺足够数量的资源给_____②_____进程以解除死锁状态。

答：①资源剥夺法 ②死锁

6. 银行家算法中，当一个进程提出的资源请求将导致系统从_____①_____进入_____②_____时，系统就拒绝它的资源请求。

答：①安全状态 ②不安全状态

7. 如果要求所有进程一次性申请它所需要的全部资源，若系统有足够的资源分配给进程，便一次把所有的资源分配给该进程。但在分配时只要有一种资源要求不能满足，则资源全不分配，进程等待，这种死锁预防方法破坏了死锁产生必要条件中的_____条件。

答：请求和保持

8. 对待死锁，一般应考虑死锁的预防、避免、检测和解除四个问题。典型的银行家算法是属于_____①_____，破坏环路等待条件是_____②_____，而剥夺资源是_____③_____的基本方法。

答：①死锁的避免 ②死锁的预防 ③死锁的解除

9. 在一个具有分时兼批处理的计算机操作系统中，如果有终端作业和批处理作业混合同时执行，_____作业应优先占用处理器。

答：终端型作业

10. 一个作业可以分成若干顺序处理的加工步骤，每个加工步骤称为一个_____。

答：作业步

11. 作业调度又称_____①_____，其主要功能是_____②_____，并为作业做好运行前的准备工作和作业完成后的善后处理工作。

答：①高级调度 ②按照某种原则从后备作业队列中选取作业

12. 确定作业调度算法时应注意系统资源的均衡使用，使_____①_____作业和_____②_____作业搭配运行。

答：①I/O ②CPU 繁忙

13. 设有一组作业，它们的提交时间及运行时间如下：

作业号	提交时间	运行时间（分钟）
1	9:00	70
2	9:40	30
3	9:50	10
4	10:10	5

在单道方式下，采用短作业优先调度算法，作业的执行顺序是_____。

答：1、4、3、2

14. _____调度是处理机的高级调度，_____调度是处理机的低级调度。

答：作业、进程

15. 如果系统中所有作业是同时到达的，则使作业平均周转时间最短的作业调度算法是_____。

答：短作业优先（SJF）调度算法

16. 一个理想的作业调度算法应该是既能_____，又能使进入系统的作业_____得到计算结果。

答：提高系统效率、及时

17. 响应比高者优先算法综合考虑了作业的_____和_____。

答：等待时间、计算时间

18. 时间片是指允许进程一次占用处理器的_____。时间片轮转调度算法常用于_____操作系统中。

答：最长时间、分时

19. 时间片取值的大小关系到_____和_____，所以时间片的值应根据进程要求系统给出应答的时间和进入系统的进程数来决定。

答：计算机系统的效率、用户的满意程度

20. 进程的切换是由_____引起的，总是发生在_____发生之后。

答：进程状态的变化、中断事件

三. 简答题

1. 在单处理机的分时系统中，分配给进程 P 的时间片用完后，系统进行切换，结果调度到的仍然是进程 P。有可能出现上述情形吗？如果可能请说明理由。

答：有可能出现上述情况。例如，若在进程 P 时间片用完后，被迫回到就绪队列时，就绪队列为空，这样进程 P 就是就绪队列中惟一的一个进程，于是调度程序选中的进程必然是进程 P；又如在按优先数调度的系统中，就绪队列按进程优先级排列，在进程 P 时间片用完之后回到就绪队列时，若其优先级高于当前就绪队列中的其他进程，则它将排在就绪队列之首，从而再次被调度程序选中并投入运行。

2. 进程调度中"可抢占"和"非抢占"两种方式，哪一种系统的开销更大？为什么？

答：可抢占式会引起系统的开销更大。可抢占式调度是严格保证任何时刻，让具有最高优先权的进程占有处理机运行，因此增加了处理机调度的时机，引起为退出处理机的进程保留现场，为占有处理机的进程恢复现场等时间

(和空间)开销增大。

3. 简单叙述作业调度和进程调度的区别。

答: (1) 作业调度是选符合条件的收容态作业装入内存。进程调度是从就绪态进程中选一个占用处理机。

(2) 作业调度是高级调度, 它位于操作系统的作业管理层次。进程调度是低级调度, 它位于操作系统分层结构的最内层。

(3) 作业调度是宏观调度, 它所选择的作业只是具有获得处理机的资格, 但尚未占有处理机, 不能立即在其上实际运行。进程调度是微观调度, 动态地把处理机实际地分配给所选择的进程, 使之真正活动起来。另外, 进程调度相当频繁, 而作业调度执行的次数一般很少。

4. 某进程被唤醒后立即投入运行, 我们就说这个系统采用的是剥夺式调度方式, 对吗?为什么?

答: 不对。剥夺式调度方式是指进程在申请资源时系统无法满足, 则强行剥夺其他进程所占有的资源, 分配给该进程。而进程进入睡眠状态则表明当前并不是剥夺调度方式, 被唤醒后立即投入运行, 只能说资源可以分配给该进程使用, 但并不一定是从其他进程那里剥夺过来的, 也可以是进程自动释放的。

5. 在某一时刻, 系统中既无执行态进程又无就绪态进程, 是否可能? 若可能, 在什么情况下会产生?

答: 有可能。在系统死锁的状态下, 进程组处于占有等待资源的状态, 应当既不属于执行态, 也不属于就绪态。

6. 关于处理机调度, 试问:

(1)什么是处理机三级调度?

(2)处理机三级调度分别在什么情况下发生?

(3)各级调度分别完成什么工作?

答: (1)处理机三级调度是: 高级调度(作业调度)、中级调度(交换调度)和低级调度(进程调度), 它们构成了操作系统内的多级调度, 不同类型的操作系统不一定都有这三种调度。

(2)高级调度是在需要从后备作业队列调度作业进入内存运行时发生; 低级调度是在处理机空闲时需要调度一个就绪进程投入运行时发生; 中级调度是在内存紧张不能满足进程运行需要时发生。

(3)高级调度决定把外存中处于后备队列的哪些作业调入内存, 并为它们创建进程和分配必要的资源, 然后将新创建的进程接入就绪队列准备执行。低级调度则决定就绪队列中的哪个进程将获得处理机, 并将处理机分配给该进程使用。中级调度是在内存资源紧张的情况下暂时将不运行的进程调至外存, 待内存空闲时再将外存上具有运行条件的就绪进程重新调入内存。

7. 何谓 JCB?其作用是什么?它由谁在什么时候创建?它由谁在什么时候撤消?JCB 至少包括哪些内容?

答: 为了管理和调度外存中的作业, 就必须记录已进入系统中的各作业情况, 因此, 系统为每个作业设置了一个作业控制块 JCB(Job control block)来记录作业的有关信息。JCB 是作业存在的唯一标志, 系统通过 JCB 而感知作业的存在, 并通过 JCB 对作业进行控制和管理。

系统在作业进入后备状态时由作业注册程序为作业建立 JCB, 从而使该作业可以被作业调度程序感知而实施管理。当作业执行完毕进入完成状态时由“终止作业”程序撤消其 JCB, 释放有关资源, 该作业就被撤消而不再存在了。不同的系统 JCB 所包含的信息有所不同, 这取决于系统对作业的调度要求。通常 JCB 包含如下内容:

(1)资源要求: 要求运行时间、最迟完成时间、需要的内存容量、外设的种类及数量等。

(2)资源使用情况: 作业进入系统的时间、开始运行时间、已运行时间、内存地址、外部设备号等。

(3)作业控制方式: 联机作业控制还是脱机作业控制。

(4)作业类型：是终端型作业还是批量型作业，是 I/O 繁忙型作业还是 CPU 繁忙型作业。

8. 现有两道作业同时执行，一道以计算为主，另一道以输入输出为主，你将怎么赋予作业进程占有处理器的优先级?为什么?

答：优先级调度算法的依据是：

- (1)对某些时间要求紧迫的作业给予优先照顾；
- (2)为充分发挥外部设备的作用，优先照顾 I/O 繁忙的作业；
- (3)为保证终端会话型作业有合理的响应时间，应优先照顾终端会话型作业。

因此，应赋予输入输出为主的作业比以计算为主的作业更高的优先级。

9. 什么是作业? 作业步? 作业流?

答：作业是用户一次请求计算机系统为其完成任务所做工作的总和。作业步是处理作业各个独立的子任务，即一个作业处理过程中相对独立的加工步骤。在批处理系统中把若干作业依次放在输入设备上，并在操作系统控制下将其送入辅存，形成一个作业流，也称为输入流。作业流是由若干作业组成的。

10. 在一个分时系统中，若为每个进程分配时间片可以随进程的执行特点和执行情况而变化。现有两类进程，一类进程经常产生中断，另一类进程中断次数很少，请问如何为这两类进程分配时间片?为什么?更进一步，它们的优先级如何确定?

答：系统分配时间片，经常产生中断的进程应分配较短的时间片，很少产生中断的进程应分配较长的时间片。这是因为经常中断的进程连续运行的时间相对较短，即使给它分配较长的时间片，也可能在时间片未用完之前产生中断，这样不利于 CPU 效率提高；而中断次数很少的进程，它连续运行的时间较长，为减少调度次数应分配较长的时间片，尽可能让它较长时间占用 CPU，可减少系统在调度上的开销。

经常产生中断的进程的优先级应高于中断次数少的进程的优先级。因为这类进程经常产生中断，当它占用 CPU 后因某事件主动让出 CPU 时，就可以让其他进程使用 CPU，使 CPU 与其他部件并行工作。

11. 考虑一个有 3 类资源的系统，系统所有可用资源为(4, 2, 2)，进程 A 申请(2, 2, 1)，可满足；进程 B 请求(1, 0, 1)，可以满足，若 A 再请求(0, 0, 1)，则被阻塞。此时，若 C 请求(2, 0, 0)，它可以分到剩余资源(1, 0, 0)，并从 A 已分到的资源中获得一个资源，于是进程 A 的分配向量变成(1, 2, 1)，而需求向量变成(1, 0, 1)。

(1)这种分配方式会导致死锁吗?如果会，请举一个例子；如果不会，请说明产生死锁的哪一个必要条件不成立? (2)这种分配方式会导致某些进程的无限等待吗?为什么?

答：(1)这种分配方式不会导致死锁。因为死锁的必要条件：请求并保持条件不成立。

(2)有可能导致进程 A 的无限等待。因为进程 A 因处于等待状态而被剥夺了资源，今后若进程 B、C 释放所占资源，但后续进程源源不断到来，相继申请并占有进程 A 所需的资源，则进程 A 有可能进入无限等待状态。

12. 试说明资源的静态分配策略能防止死锁的原因。

答：资源静态分配策略要求每个进程在开始执行前申请所需的全部资源，仅在系统为之分配了所需的全部资源后，该进程才开始执行。这样，进程在执行过程中不再申请资源，从而破坏了死锁的四个必要条件之一“占有并等待条件”，从而防止死锁的发生。

13. 采用资源顺序分配法可破坏死锁的必要条件之一：“环路等待条件”，简述该方法的思想，并说明这种方法为什么能预防死锁。

答：首先给系统中的资源编号，即作一个函数 $F: R \rightarrow N$ ，其中 R 为资源类型集合， N 为自然数集合，进程只能按编号从小到大的顺序申请资源，否则拒绝分配。若存在循环等待的一组进程，不妨记为 $\{P_0, P_1 \dots P_a\}$ 且 P_i 拥有资源 r_i ，则根据顺序分配原则有： $F(r_0) < F(r_1) < \dots < F(r_a) < F(r_0)$ ，即 $F(r_0) < F(r_0)$ 显然与资源编号的唯一性相矛盾。故不会存在“环路等待”。

14. 为什么说多级反馈队列调度算法能较好地满足各类用户的需要？

答：多级反馈队列调度算法能较好地满足各种类型用户的需要。对终端型用户而言，由于终端型作业用户所提交的作业大部分属于交互型作业，作业通常比较短小，系统只要能使这些作业在第一队列所规定的时间片内完成，便可使终端型作业用户感到满意；对于短批处理作业用户而言，他们的作业开始时像终端型作业一样，如果仅在第一队列中执行一个时间片即可完成，便可获得与终端型作业一样的响应时间，对于稍长的作业，通常也只需要在第二队列和第三队列中各执行一个时间片即可完成，其周转时间仍然较短；对于长批处理作业用户而言，他们的长作业将依次在第 1, 2, ..., 直到第 n 个队列中运行，然后再按轮转方式运行，用户不必担心其作业长期得不到处理。

15. 为什么说采用有序资源分配法不会产生死锁？

答：为了便于说明，不妨设系统中有 m 类资源， n 个进程，分别用 R_1, R_2, \dots, R_m ($1, 2, \dots, m$ 可看作资源编号)和 P_1, P_2, \dots, P_n 表示。根据有序资源分配法可知，进程申请资源时必须按照资源编号的升序进行，即任何进程在占有了 R_1 类资源后，再申请的资源 R_j 的编号 j 一定大于 1。因此在任一时刻，系统中至少存在一个进程 P_k ，它占有了较高编号的资源 R_h ，且它继续请求的资源必然是空闲的，因而 P_k 可以一直向前推进直至完成，当 P_k 运行完成后即会释放它占有的所有资源；在 P_k 完成之后，剩下的进程集合中同样会存在一个进程，它占有了较高编号的资源，且它继续请求的资源必然是空闲的，因而它可以一直向前推进直至完成，以此类推，所有进程均可运行完成，故不会发生死锁。

16. 有相同类型的 5 个资源被 4 个进程所共享，且每个进程最多需要 2 个这样的资源就可以运行完毕。试问该系统是否会由于对这种资源的竞争而产生死锁。

答：该系统不会由于对这种资源的竞争而产生死锁。因为在最坏情况下，每个进程都需要 2 个这样的资源，且每个进程都已申请到了 1 个资源，那么系统中还剩下 1 个可用资源。无论系统为了满足哪个进程的资源申请而将资源分配给该进程，都会因为该进程已获得了它所需要的全部资源而确保它运行完毕，从而可将它占有的 2 个资源归还给系统，这就保证了其余三个进程能顺利运行。由此可知，该系统不会由于对这种资源的竞争而产生死锁。

四. 应用题

1. 假定在单 CPU 条件下有下列要执行的作业：

作业	运行时间	优先级
1	10	2
2	4	3
3	3	0

作业到来的时间是按作业编号顺序进行的(即后面作业依次比前一个作业迟到一个时间单位)。

(1) 用一个执行时间图描述在采用非抢占式优先级算法时执行这些作业的情况。

(2) 对于上述算法，各个作业的周转时间是多少？平均周转时间是多少？

(3) 对于上述算法，各个作业的带权周转时间是多少？平均带权周转时间是多少？

解： (1)非抢占式优先级算法作业的执行情况如下：

作业	到达时间	运行时间	完成时间	周转时间	带权周转时间
1	0	10	10	10	1.0
2	1	4	17	16	4.0
3	2	3	13	11	3.7
平均周转时间		12.3			
平均带权周转时间		2.9			

2. 若在后备作业队列中等待运行的同时有三个作业 1、2、3，已知它们各自的运行时间为 a、b、c，且满足关系 $a < b < c$ ，试证明采用短作业优先调度算法能获得最小平均周转时间。

证明：由于短作业优先调度算法总是在后备作业队列中选择运行时间最短的作业作为调度对象，因此对短作业优先调度算法而言，这三个作业的总周转时间为

$$T1 = a + (a+b) + (a+b+c) = 3a + 2b + c \quad \dots\dots (1)$$

若不按短作业优先调度算法来调度这三个作业，不失一般性，假定调度顺序为 2、1、3，则其周转时间为

$$T2 = b + (b+a) + (b+a+c) = 3b + 2a + c \quad \dots\dots (2)$$

由(1)、(2)两式可得： $T2 - T1 = b - a > 0$

由此可见，短作业优先调度算法能获得最小平均周转时间。

3. 设有 4 道作业，它们的提交时间及执行时间如下：

作业号	提交时间	执行时间
1	10.0	2.0
2	10.2	1.0
3	10.4	0.5
4	10.5	0.3

试计算在单道程序环境

下，采用先来先服务调度算法和最短作业优先调度算法时的平均周转时间和平均带权周转时间，并指出它们的调度顺序。(时间单位：小时，以十进制进行计算。)

解：若采用先来先服务调度算法，则其调度顺序为 1、2、3、4。

作业号	提交时间	执行时间	开始时间	完成时间	周转时间	带权周转时间
1	10.0	2.0	10.0	12.0	2.0	1.0
2	10.2	1.0	12.0	13.0	2.8	2.8
3	10.4	0.5	13.0	13.5	3.1	6.2
4	10.5	0.3	13.5	13.8	3.3	11.0

平均周转时间 $T = (2.0 + 2.8 + 3.1 + 3.3) / 4 = 2.8$

平均带权周转时间 $W = (1 + 2.8 + 6.2 + 11) / 4 = 5.25$

若采用短作业优先调度算法，则其调度顺序为 1、4、3、2

作业号	提交时间	执行时间	开始时间	完成时间	周转时间	带权周转时间
1	10.0	2.0	10.0	12.0	2.0	1.0
4	10.5	0.3	12.0	12.3	1.8	6.0
3	10.4	0.5	12.3	12.8	2.4	4.8
2	10.2	1.0	12.8	13.8	3.6	3.6

平均周转时间为 $T = (2.0 + 1.8 + 2.4 + 3.6) / 4 = 2.45$

平均带权周转时间 $W = (1 + 6 + 4.8 + 3.6) / 4 = 3.85$

4. 假设有四个作业，它们的提交、运行时间如下表所示。若采用高响应比优先调度算法，试问平均周转时间和平均带权周转时间为多少？（时间单位小时，以十进制进行计算。）

作业号	到达时间	运行时间
1	8.0	2.0
2	8.3	0.5
3	8.5	0.1
4	9.0	0.4

解：根据响应比的定义

每次调度前计算出各作业的响应比，得到四个作业的调度次序为：作业 1、作业 3、作业 2、作业 4。

作业号	到达时间	运行时间	开始时间	完成时间	周转时间	带权周转时间
1	8.0	2.0	8.0	10.0	2.0	1.0
2	8.3	0.5	10.1	10.6	2.3	4.6
3	8.5	0.1	10.0	10.1	1.6	16.0
4	9.0	0.4	10.6	11.0	2.0	5.0

平均周转时间为 $T = (2.0 + 2.3 + 1.6 + 2.0) / 4 = 1.975$

平均带权周转时间 $W = (1 + 4.6 + 16 + 5) / 4 = 6.65$

5. 有一个具有两道作业的批处理系统，作业调度采用短作业优先的调度算法，进程调度采用以优先数为基础的抢占式调度算法。在下表所示的作业序列，作业优先数即为进程优先数，且优先数越小优先级越高。

作业名	到达时间	估计运行时间	优先数
A	10:00	40 分	5
B	10:20	30 分	3
C	10:30	50 分	4
D	10:50	20 分	6

(1) 列出所有作业进入内存时间及结束时间

(2) 计算平均周转时间。

分析：在本题中，每个作业的运行将经历两级调度：作业调度和进程调度。作业调度采用短作业优先调度算法，进程调度采用基于优先数的抢占式调度算法，高优先级的进程可以抢占系统处理机。只有当作业调度程序将作业装入内存后，方能参与进程调度。本题中的批处理系统是两道作业系统，因此每次只能有两道作业进入系统内存。本题中的作业和进程推进顺序如下：

10:00 时, A 作业到达。因系统的后备作业队列中没有其他作业, 进程就绪队列中也没有进程, 故作业调度程序将作业 A 调入内存并将它排在就绪队列上, 进程调度程序调度它运行。

10:20 时, B 作业到达。因系统的后备作业队列中没有其他作业, 故作业调度程序将作业 B 调入内存并将它排在就绪队列上。而作业 B 的优先级高于作业 A 的优先级, 进程调度程序停止作业 A 的运行, 将作业 A 放入就绪队列, 调度作业 B 运行。此时, 系统中已有两道作业在内存中运行, 作业 A 已运行 20 分钟, 还需运行 20 分钟才能完成。

10:30 时, C 作业到达。因系统中已有两道作业在内存中运行, 故作业 C 只能在后备作业队列中等待作业调度。此时, 作业 B 已运行了 10 分钟并将继续运行, 还需运行 20 分钟才能完成, 作业 A 已等待 10 分钟并将继续等待、还需运行 20 分钟才能完成。

10:50 时, B 作业运行 30 分钟结束运行, D 作业到达。因系统中只有作业 A 在内存中运行, 作业后备队列中有 C、D 两道作业, 按短作业优先的作业调度策略, 作业 D 被作业调度程序选中, 装入内存运行, 作业 C 仍在后备作业队列中等待作业调度。在内存中, 作业 A 的优先级高于作业 D, 进程调度程序调度作业 A 运行, 作业 D 在就绪队列中等待进程调度。此时, 作业 A 已运行了 20 分钟, 在就绪队列中等待了 30 分钟, 还需运行 20 分钟才能完成; 作业 C 已在后备队列中等待了 20 分钟并将继续等待。

11:10 时, A 作业运行 40 分钟结束运行。因系统中只有作业 D 在内存中运行, 作业后备队列中只有作业 C 在等待, 作业调度程序将作业 C 装入内存运行。因作业 C 的优先级高于作业 D, 进程调度程序调度作业 C 运行, 作业 D 仍在就绪队列中等待进程调度。此时作业 D 已在就绪队列中等待了 20 分钟并将继续等待。

12:00 时, C 作业运行 50 分钟结束运行。因系统中只有作业 D 在内存, 进程调度程序调度作业 D 运行。

12:20 时, D 作业运行 20 分钟结束运行。

解: (1) 由上述分析可得出所有作业的进入内存时间和结束时间:

作业名	进入内存时间	结束时间
A	10:00	11:10
B	10:20	10:50
C	11:10	12:00
D	10:50	12:20

(2) 各作业执行时的周转时间为:

作业 A: 70 分钟

作业 B: 30 分钟

作业 c: 90 分钟

作业 D: 90 分钟

作业的平均周转时间为:

$(70 + 30 + 90 + 90) / 4 = 70$ 分钟。

6. 已知 3 个批处理作业中: 第一个作业 10.00 时到达, 需要执行 2 小时; 第二个作业在 10.1 时到达, 需要执行 1 小时; 第三个作业在 10.5 小时到达, 需要执行 0.5 小时。如果分别采用如下的表 1 和表 2 所示的两种作业调度算法。

(1) 计算各调度算法下的作业平均周转时间:

(2) 这两种调度算法各可能是什么作业调度算法?

表 1 调度算法 1

作业	提交时间	开始时刻	完成时刻
1	10.0	10.0	12.0
2	10.1	12.0	13.0
3	10.5	13.0	13.5

表 2 调度算法 2

作业	提交时间	开始时刻	完成时刻
1	10.0	10.0	12.0
3	10.5	12.0	12.5
2	10.1	12.5	13.5

解：(1)采用调度算法 1 的作业运行情况如下表 3 所示：

表 3 采用调度算法 1 的作业运行情况表

作业执行次序	提交时间	运行时间	等待时间	开始时刻	完成时刻	周转时间
1	10.0	2	0	10.0	12.0	2
2	10.1	1	1.9	12.0	13.0	2.9
3	10.5	0.5	2.5	13.0	13.5	3
作业平均周转时间		$T=(2+2.9+3)/3=2.63$				

采用调度算法 2 的作业运行情况如下表 4 所示：

表 4 采用调度算法 2 的作业运行情况表

作业执行次序	提交时间	运行时间	等待时间	开始时刻	完成时刻	周转时间
1	10.0	2	0	10.0	12.0	2
3	10.5	0.5	1.5	12.0	12.5	2
2	10.1	1	2.4	12.5	13.5	2.5
作业平均周转时间		$T=(2+2+2.5)/3=2.17$				

(2) 调度算法 1 是按照作业到达的先后次序执行的，所以它是先来先服务调度算法。调度算法 2 满足短作业优先的调度原则，所以它属于短作业优先调度算法。此外，从响应比高者优先调度算法来看，当作业 1 在 12.0 完成时，作业 2 和作业 3 的响应比如下：作业 2 的响应比 $=1+1.9/1=2.9$ ，作业 3 的响应比 $=1+1.5/0.5=4$ 也即，调度算法 2 也可能是响应比高者优先调度算法。

7. 有三个进程 P1, P2 和 P3 并发工作。进程 P1 需用资源 S3 和 S1；进程 P2 需用资源 S1 和 S2；进程 P3 需用资源 S2 和 S3。回答：

- (1) 若对资源分配不加限制，会发生什么情况？为什么？
- (2) 为保证进程正确工作，应采用怎样的资源分配策略？为什么？

答：(1) 可能会发生死锁

例如：进程 P1, P2 和 P3 分别获得资源 S3, S1 和 S2 后再继续申请资源时都要等待，这是循环等待。(或进程在等待新源时均不释放已占资源)

(2) 可有几种答案：

A. 采用静态分配 由于执行前已获得所需的全部资源，故不会出现占有资源又等待别的资源的现象(或不会出现循环等待资源现象)。

或 B. 采用按序分配 不会出现循环等待资源现象。

或 C. 采用银行家算法 因为在分配时，保证了系统处于安全状态。

8. 设系统有三种类型的资源，数量为(4, 2, 2)，系统中有进程 A, B, C 按如下顺序请求资源：

进程 A 申请 (3, 2, 1)

进程 B 申请 (1, 0, 1)

进程 A 申请 (0, 1, 0)

进程 C 申请 (2, 0, 0)

请你给出一种防止死锁的资源剥夺分配策略，完成上述请求序列，并列出资源分配过程，指明哪些进程需要等待，哪些资源被剥夺。

解：① 分配策略为：当进程 P_i 申请 r_i 类资源时，检查 r_i 中有无可分配的资源，有则分配给 P_i ；否则将 P_i 占有的资源全部释放而进入等待状态。（ P_i 等待原占有的所有资源和新申请的资源）

② 资源分配过程： 剩余资源

进程 A: (3, 2, 1) (1, 0, 1)

进程 B: (1, 0, 1) (0, 0, 0)

进程 C: (2, 0, 0) (1, 2, 1)

进程 A: (0, 1, 0) (不满足) (3, 2, 1)，A 的所有资源被剥夺，A 处于等待，C，B 完成之后，A 可完成。

9. 某系统中有 10 台打印机，有三个进程 P_1 ， P_2 ， P_3 分别需要 8 台，7 台和 4 台。若 P_1 ， P_2 ， P_3 已申请到 4 台，2 台和 2 台。试问：按银行家算法能安全分配吗？请说明分配过程。

答：系统能为进程 P_3 分配二台打印机。因为尽管此时 10 台打印机已分配给进程 P_1 4 台， P_2 2 台和 P_3 4 台，全部分配完，但 P_3 已分配到所需要全部 4 台打印机，它不会对打印机再提出申请，所以它能顺利运行下去，能释放占用的 4 台打印机，使进程 P_1 ， P_2 均可能获得乘余的要求 4 台和 5 台，按银行家算法是安全的。

10. 在生产者—消费者问题中，如果对调生产者进程中的两个 P 操作和两个 V 操作，则可能发生什么情况？

解：如果对调生产者进程中的两个 P 操作和两个 v 操作，则生产者—消费者问题的同步描述为：

```
int full=0;
```

```
int empty =n;
```

```
int mutex=1;
```

```
main ( )
```

```
{
```

```
  cobegin
```

```
    producer ( );
```

```
    consumer ( );
```

```
  coend
```

```
}
```

```
producer ( )
```

```
{
```

```
  while (生产未完成)
```

```
  {
```

```
    生产一个产品;
```

```

    p(mutex);
    p(empty);
    送一个产品到有界缓冲区;
    v(full);
    v(mutex);
}
}
consumer ( )
{
    while(还要继续消费)
    {
        p(full);
        p(mutex);
        从有界缓冲区中取产品;
        v(mutex);
        v(empty);
        消费一个产品;
    }
}

```

由于 V 操作是释放资源，因此对调 V 操作的次序无关紧要。而对调 P 操作的次序则可能导致死锁。这是因为对调 P 操作后，有可能出现这样一种特殊情况：在某一时刻缓冲区中已装满了产品且缓冲区中无进程工作（这时信号量 full 的值为 n，信号量 empty 的值为 0，信号量 mutex 的值为 1），若系统此时调度生产者进程运行，生产者进程又生产了一个产品，

它执行 P(mutex) 并顺利进入临界区（这时 mutex 值为 0），随后它执行 p(empty) 时因没有空闲缓冲单元而受阻等待，等待消费者进程进入缓冲区取走产品以释放出缓冲单元；消费者进程执行 p(full) 后再执行 p(mutex) 时，因缓冲区被生产者进程占据而无法进入。这样就形成了生产者进程在占有临界资源的情况下，等待消费者进程取走产品，而消费者进程又无法进入临界区取走产品的僵局，此时两进程陷入死锁。

11. n 个进程共享某种资源 R，该资源共有 m 个可分配单位，每个进程一次一个地申请或释放资源单位。假设每个进程对该资源的最大需求量均小于 m，且各进程最大需求量之和小于 m + n，试证明在这个系统中不可能发生死锁。

解：设 $\max(i)$ 表示第 i 个进程的最大资源需求量， $\text{need}(i)$ 表示第 i 个进程还需要的资源量， $\text{alloc}(i)$ 表示第 i 个进程已分配的资源量。由题中所给条件可知：

$$\max(1) + \max(2) + \dots + \max(n) = (\text{need}(1) + \dots + \text{need}(n)) + (\text{alloc}(1) + \dots + \text{alloc}(n)) < m + n$$

如果在这个系统中发生了死锁，那么一方面 m 个资源应该全部分配出去，即

$$\text{alloc}(1) + \dots + \text{alloc}(n) = m$$

另一方面所有进程将陷入无限等待状态。

由上述两式可得： $\text{need}(1) + \dots + \text{need}(n) < n$

上式表示死锁发生后，n 个进程还需要的资源量之和小于 n，这意味着此刻至少存在一个进程 i， $need(i)=0$ ，即它已获得了所需要的全部资源。既然该进程已获得了它所需要的全部资源，那么它就能执行完成并释放它占有的资源，这与前面的假设矛盾，从而证明在这个系统中不可能发生死锁。

12. 在银行家算法中，若出现下述资源分配情况：

	Allocation	Need	Available
P0	0 0 3 2	0 0 1 2	1 6 2 2
P1	1 0 0 0	1 7 5 0	
P2	1 3 5 4	2 3 5 6	
P3	0 3 3 2	0 6 5 2	
P4	0 0 1 4	0 6 5 6	

试问：(1) 该状态是否安全？

(2) 如果进程 P2 提出请求 $Request_2(1, 2, 2, 2)$ 后，系统能否将资源分配给它？

解：(1) 利用银行家算法对此时刻的资源分配情况进行分析，可得此刻的安全性分析情况：

	Work	Need	Allocation	Work+Allocation	Finish
P0	1 6 2 2	0 0 1 2	0 0 3 2	1 6 5 4	true
P3	1 6 5 4	0 6 5 2	0 3 3 2	1 9 8 6	true
P4	1 9 8 6	0 6 5 6	0 0 1 4	1 9 9 10	true
P1	1 9 9 10	1 7 5 0	1 0 0 0	2 9 9 10	true
P2	2 9 9 10	2 3 5 6	1 3 5 4	3 12 14 14	true

从上述分析中可以看出，此时存在一个安全序列 {P0, P3, P4, P1, P2}，故该状态是安全的。

(2) P2 提出请求 $Request_2(1, 2, 2, 2)$ ，按银行家算法进行检查：

$$Request_2(1, 2, 2, 2) \leq Need_2(2, 3, 5, 6)$$

$$Request_2(1, 2, 2, 2) \leq Available(1, 6, 2, 2)$$

试分配并修改相应的数据结构，资源分配情况如下：

	Allocation	Need	Available
P0	0 0 3 2	0 0 1 2	0 4 0 0
P1	1 0 0 0	1 7 5 0	
P2	2 5 7 6	1 1 3 4	
P3	0 3 3 2	0 6 5 2	
P4	0 0 1 4	0 6 5 6	

再利用安全性算法检查系统是否安全，可用资源 Available (0, 4, 0, 0) 已不能满足任何进程的需要，故系统进入不安全状态，此时系统不能将资源分配给 P2。

13. 有相同类型的 5 个资源被 4 个进程所共享，且每个进程最多需要 2 个这样的资源就可以运行完毕。试问该系统是否会由于对这种资源的竞争而产生死锁。

答：该系统不会由于对这种资源的竞争而产生死锁。因为在最坏情况下，每个进程都需要 2 个这样的资源，且每个进程都已申请到了 1 个资源，那么系统中还剩下 1 个可用资源。无论系统为了满足哪个进程的资源申请而将资源分配给该进程，都会因为该进程已获

得了它所需要的全部资源而确保它运行完毕，从而可将它占有的 2 个资源归还给系统，这就保证了其余三个进程能顺利运行。由此可知，该系统不会由于对这种资源的竞争而产生死锁。

14. 考虑下列资源分配策略：对资源的申请和释放可以在任何时候进行。如果一个进程提出资源请求时得不到满足，若此时无由于等待资源而被阻塞的进程，则自己就被阻塞；若此时已有等待资源而被阻塞的进程，则检查所有由于等待资源而被阻塞的进程。如果它们有申请进程所需要的资源，则将这些资源取出分配给申请进程。

例如，考虑一个有 3 类资源的系统，系统所有可用资源为 (4, 2, 2)，进程 A 申请 (2, 2, 1)，可满足；进程 B 申请 (1, 0, 1)，可满足；若 A 再申请 (0, 0, 1)，则被阻塞。此时，若 C 请求 (2, 0, 0)，它可以分到剩余资源 (1, 0, 0)，并从 A 已分到的资源中获得一个资源，于是进程 A 的分配向量变成 (1, 2, 1) 而需求向量变成 (1, 0, 1)。

①这种分配策略会导致死锁吗？如果会，请举一个例子；如果不会，请说明产生死锁的哪一个必要条件不成立。

②这种分配方式会导致某些进程的无限等待吗？为什么？

解：①本题所给的资源分配策略不会产生死锁。因为本题给出的分配策略规定若一进程的资源得不到满足，则检查所有由于等待资源而被阻塞的进程，如果它们有申请进程所需要的资源，则将这些资源取出分配给申请进程。从而破坏了产生死锁必要条件中的不剥夺条件，这样系统就不会产生死锁。

②这种方法会导致某些进程无限期的等待。因为被阻塞进程的资源可以被剥夺，所以被阻塞进程所拥有的资源数量在其被唤醒之前只可能减少。若系统中不断出现其他进程申请资源，这些进程申请的资源与被阻塞进程申请或拥有的资源类型相同且不被阻塞，则系统无法保证被阻塞进程一定能获得所需要的全部资源。例如，本题中的进程 A 申请 (2, 2, 1) 后再申请 (0, 0, 1) 被阻塞。此后，进程 C 又剥夺了进程 A 的一个资源，使得进程 A 拥有的资源变为 (1, 2, 1)，其需求向量为 (1, 0, 1)。之后，若再创建的进程总是只申请第 1 和第 3 类资源，总是占有系统所剩下的第 1 和第 3 类资源的全部且不阻塞，那么进程 A 将会无限期地等待。

15. 一台计算机有 8 台磁带机。它们由 N 个进程竞争使用，每个进程可能需要 3 台磁带机。请问 N 为多少时，系统没有死锁危险，并说明原因。

解：当 N 为 1, 2, 3 时，系统没有产生死锁的危险。因为，当系统中只有 1 个进程时，它最多需要 3 台磁带机，而系统有 8 台磁带机，其资源数目足够系统内的 1 个进程使用，因此绝不可能发生死锁；当系统中有 2 个进程时，最多需要 6 台磁带机，而系统有 8 台磁带机，其资源数目也足够系统内的 2 个进程使用，因此也不可能发生死锁；当系统中有 3 个进程时，在最坏情况下，每个进程都需要 3 个这样的资源，且假定每个进程都已申请到了 2 个资源，那么系统中还剩下 2 个可用资源，无论系统为了满足哪个进程的资源申请而将资源分配给该进程，都会因为该进程已获得了它所需要的全部资源而确保它运行完毕，从而可将它占有的 3 个资源归还给系统，这就保证了其余进程能顺利运行完毕。由此可知，当 N 为 1, 2, 3 时，该系统不会由于对这种资源的竞争而产生死锁。

16. 假设就绪队列中有 10 个进程，系统将时间片设为 200ms，CPU 进行进程切换要花费 10ms，试问系统开销所占的比率约为多少？

解：因就绪队列中有 10 个进程，它们以时间片轮转的方式使用 CPU，时间片长度为 200ms。当一个时间片用完时，调度进程将当前运行进程设置为就绪状态并放入就绪队列尾，再从就绪队列队首选择进程投入运行，这一过程（进程切换）要花费时间 10ms。因此系统开销所占比率为： $10 / (200 + 10) = 4.8\%$

17. 如果 P、V 操作设计不当，则有可能产生死锁。假如系统中有输入机和打印机两类资源各一台，有两个进程 P1 和 P2 都要求使用输入机和打印机。我们可以利用 P、V 操作来实现互斥：定义 s1、s2 分别为代表输入机和打印机能否被使用的信号量，初值均为 1，并且按如下方法请求使用和归还资源：

Process P1

begin

P(s1);

使用输入机;

P(s2);

使用打印机;

V(s2);

V(s1);

End;

Process P2

Begin

P(s2);

使用打印机;

P(s1);

使用输入机;

V(s2);

V(s1);

End;

结合死锁产生的必要条件，分析此种方法是否会造成死锁，若不会，给出理由；若会产生死锁，则修改上面程序，使 P1、P2 两进程能够互斥的使用资源，并且能够顺利完成。

解：此种方法可能会出现 P1 得到输入机而 P2 得到打印机，双方在不释放已有资源的情况下，又去申请新的资源，从而造成死锁。可以采用为资源编号的方法，让进程按序申请资源，来避免死锁。程序可修改如下：

Process P2

Begin

P(s2);

使用输入机;

P(s1);

使用打印机;

V(s2);

V(s1);

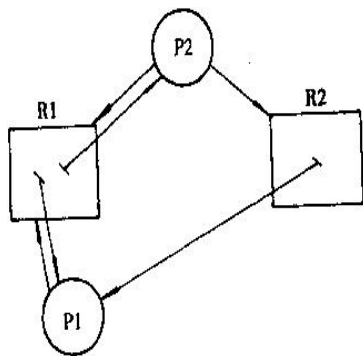
End;

18. 假定某计算机系统有 R1 和 R2 两类可再使用资源(其中 R1 有两个单位，R2 有一个单位)。它们被进程 P1 和 P2 所共享，且已知两个进程均以

→申请 R1→ 申请 R2→ 申请 R1→ 释放 R1→ 释放 R2→ 释放 R→

所示的顺序使用两类资源。试求出系统运行过程中可能到达的死锁点，并画出死锁点的资源分配图(或称进程—资源图)。

解：该题答案不惟一。从已知条件可知，P1、P2 两进程都是各自按顺序申请系统中所有资源，并在全部得到满足之后才会依次释放；由此可得，只要让 P1、P2 分别占有其中某个资源，即不把全部资源都交给一个进程，则会发生死锁。进程—资源图如下：



19. 某系统有 R1、R2 和 R3 共 3 种资源。在 T0 时刻 P1、P2、P3 和 P4 这 4 个进程对资源的占用和需求情况见下表，此刻系统的可用资源向量为 (2, 1, 2)，问题：

- (1) 将系统中各种资源总数和此刻各进程对各资源的需求数目用向量或矩阵表示出来；
- (2) 如果此时 P1 和 P2 均发出资源请求向量 Request (1, 0, 1)，为了保持系统安全性，应该如何分配资源给这两个进程？说明你所采用策略的原因；
- (3) 如果 (2) 中两个请求立刻得到满足后，系统此刻是否处于死锁状态？

进程	最大需求			占有量		
	R1	R2	R3	R1	R2	R3
P1	3	2	2	1	0	0
P2	6	1	3	4	1	1
P3	3	1	4	2	1	1
P4	4	2	2	0	0	2

解：(1) 系统资源总数为 (9, 3, 6)。

各进程对资源需求矩阵为：

2	2	2
2	0	2
1	0	3
4	2	0

(2) 采用银行家算法进行计算得：系统不可以将资源分配给进程 P1，虽然剩余资源还可以满足进程 P1 现在的需求，但是一旦分配给进程 P1 后，就找不到一个安全执行的序列保证各个进程能够正常运行下去。因此进程 P1 进入等待状态。系统可以满足 P2 的请求，因为分配完成后，至少还可以找到一个安全序列，如 (P2P1P3P4)，使各进程可以运行至结束。

(3) 系统满足进程 P1 和 P2 的请求后，没有立即进入死锁状态，因为此时所有进程还处于运行状态，没有被阻塞；只有等到进程继续申请资源并因得不到满足而全部进入阻塞状态，死锁才真正发生了。

第四章 存储器管理

一. 选择题

1. 在存储管理中, _____ 可与覆盖技术配合。

- A. 页式管理
- B. 段式管理
- C. 段页式管理
- D. 可变分区管理

答: D

2. 采用覆盖与交换技术的目的是_____。

- A. 节省主存空间
- B. 物理上扩充主存容量
- C. 提高 CPU 效率
- D. 实现主存共享

答: A

3. 动态重定位技术依赖于_____。

- A. 重定位装入程序
- B. 重定位寄存器
- C. 地址机构
- D. 目标程序

答: B

4. 虚拟存储器的最大容量_____。

- A. 为内外存容量之和
- B. 由计算机的地址结构决定
- C. 是任意的
- D. 由作业地址空间决定

答: B

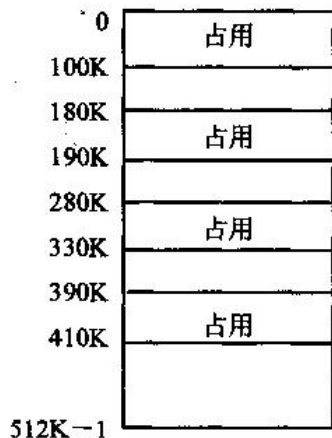
5. 在虚拟存储系统中, 若进程在内存中占 3 块(开始时为空), 采用先进先出页面淘汰算法, 当执行访问页号序列为 1、2、3、4、1、2、5、1、2、3、4、5、6 时, 将产生_____次缺页中断。

- A. 7
- B. 8
- C. 9
- D. 10

答: D

6. 设内存的分配情况如下图所示。若要申请一块 40K 字节的内存空间, 若采用最佳适应算法, 则所得到的分区首址为_____。

- A. 100K
- B. 190K
- C. 330K
- D. 410K



答: C

7. 很好地解决了“零头”问题的存储管理方法是_____。

- A. 页式存储管理
- B. 段式存储管理
- C. 多重分区管理
- D. 可变式分区管理

答: A

8. 系统“抖动”现象的发生是由_____引起的。

- A. 置换算法选择不当
- B. 交换的信息量过大
- C. 内存容量不足
- D. 请求页式管理方案

答: A

9. 在可变式分区存储管理中的拼接技术可以_____。

- A. 集中空闲区
- B. 增加主存容量
- C. 缩短访问周期
- D. 加速地址转换

答: A

10. 分区管理中采用“最佳适应”分配算法时, 宜把空闲区按_____次序登记在空闲区表中。

- A. 长度递增
- B. 长度递减
- C. 地址递增
- D. 地址递减

答: A

11. 在固定分区分配中, 每个分区的大小是_____。

- A. 相同
- B. 随作业长度变化
- C. 可以不同但预先固定
- D. 可以不同但根据作业长度固定

答: C

12. 实现虚拟存储器的目的是_____。

- A. 实现存储保护
- B. 实现程序浮动
- C. 扩充辅存容量
- D. 扩充主存容量

答: D

13. 采用段式存储管理的系统中, 若地址用 24 位表示, 其中 8 位表示段号, 则允许每段的最大长度是 _____。

- A. 2^{24}
- B. 2^{16}
- C. 2^8
- D. 2^{32}

答: B

14. 作业在执行中发生了缺页中断, 经操作系统处理后, 应让其执行_____指令。

- A. 被中断的前一条
- B. 被中断的
- C. 被中断的后一条
- D. 启动时的第一条

答: B

15. 把作业地址空间中使用的逻辑地址变成内存中物理地址的过程称为 _____。

- A. 重定位
- B. 物理化
- C. 逻辑化
- D. 加载

答: A

16. 在分页系统环境下, 程序员编制的程序, 其地址空间是连续的, 分页是由_____完成的。

- A. 程序员
- B. 编译地址
- C. 用户
- D. 系统

答: D

17. 在请求分页存储管理中, 若采用 FIFO 页面淘汰算法, 则当分配的页面数增加时, 缺页中断的次数_____。

- A. 减少
- B. 增加
- C. 无影响
- D. 可能增加也可能减少

答: D

18. 在段页式存储管理系统中, 内存等分成_____, 程序按逻辑模块划分成若干_____。

- A. 块, 页
- B. 块, 段
- C. 分区, 段
- D. 段, 页

答: B

19. 虚拟存储管理系统的基础是程序的_____理论。

- A. 局部性
- B. 全局性
- C. 动态性
- D. 虚拟性

答: A

20. 下述_____页面淘汰算法会产生 Belady 现象。

- A. 先进先出
- B. 最近最少使用
- C. 最不经常使用
- D. 最佳

答: A

21. 某段表的内容如下:

段号	段首址	段长度
0	120K	40K
1	760K	30K
2	480K	20K
3	370K	20K

一逻辑地址为(2154), 它对应的物理地址为_____。

- A. 120K + 2 B. 480K + 154
C. 30K + 154 D. 2 + 480K

答: B

22. 在一个页式存储管理系统中, 页表内容如下:

页号	块号
0	2
1	1
2	6
3	3
4	7

若页的大小为 4K, 则地址转换机构将逻辑地址 0 转换成的物理地址为_____。

- A. 8192 B. 4096 C. 2048 D. 1024

答: A

23. 如果一个程序为多个进程所共享, 那么该程序的代码在执行的过程中不能被修改, 即程序应该是_____。

- A. 可执行码 B. 可重入码
C. 可改变码 D. 可再现码

答: B

24. 在分时系统中, 可将作业不需要或暂时不需要的部分移到辅存, 让出主存空间以调入其他所需数据, 称为_____。

- A. 覆盖技术 B. 对换技术
C. 虚拟技术 D. 物理扩充

答: B

25. 支持多道程序设计, 算法简单, 但存储碎片多的存储管理方式是_____。

- A. 段式 B. 页式
C. 固定分区 D. 段页式

答: C

26. 碎片是指_____。

- A. 存储分配完后所剩的空闲区
B. 没有被使用的存储区
C. 不能被使用的存储区
D. 未被使用, 而又暂时不能使用的空闲区

答: D

27. 碎片现象的存在使得_____。

- A. 主存空间利用率降低 B. 主存空间利用率提高
C. 主存空间利用率得以改善 D. 主存空间利用率不受影响

答: A

28. 可变分区管理方式按作业需求量分配主存分区, 所以_____。

- A. 分区的长度是固定的
- B. 分区的个数是确定的
- C. 分区的长度和个数都是确定的
- D. 分区的长度不是预先固定的、分区的个数是不确定的

答: D

29. 在可变分区方式管理下收回主存空间时, 若已判定“空闲区表第 j 栏开始地址=归还的分区开始地址+长度”, 则表示_____。

- A. 归还区有上邻空闲区
- B. 归还区有下邻空闲区间
- C. 归还区有上、下邻空闲区
- D. 归还区无相邻空闲区间

答: B

30. 当可变分区方式管理主存空间回收时, 要检查有无相邻的空闲区。若归还区开始地址为 S , 长度为 L , 下列表示归还区有上邻空闲区的是_____。

- A. 第 j 栏开始地址 $= S + L$
- B. 第 j 栏开始地址+长度 $= S$
- C. 第 j 栏开始地址+长度 $= S$ 且第 k 栏开始地址 $= S + L$
- D. 以上都不对

答: B

31. 在可变分区存储管理中, 回收一个分区时, 反而使空闲区的个数减 1 的情况是该分区_____。

- A. 只有上邻
- B. 只有下邻
- C. 既有上邻又有下邻
- D. 无上、下邻

答: B

32. 操作系统对已在主存储器中的作业根据需要改变存放位置, 称为_____。

- A. 覆盖技术
- B. 交换技术
- C. 移动技术
- D. 虚拟技术

答: C

33. 在可变分区方式管理主存时、采用移动技术能提高主存的利用率, 下列不能移动作业的情况是_____。

- A. 正在取主存中的数据准备计算
- B. 正在计算一个表达式的值
- C. 正在把计算结果写入主存
- D. 正在等待外围设备传输消息

答: C

34. 最容易形成很多小碎片的可变分区算法是_____。

- A. 最先适应算法
- B. 最优适应算法
- C. 最坏适应算法
- D. 以上都不是

答: B

35. 段式存储管理中分段是用户决定的, 因此_____。

- A. 段内的地址和段间的地址都是连续的
- B. 段内的地址是连续的, 段间的地址是不连续的
- C. 段内的地址是不连续的, 段间的地址是连续的
- D. 段内的地址和段间的地址都是不连续的

答: B

36. 段页式存储管理汲取了页式存储管理和段式存储管理的长处, 其实现原理结合了页式和段式管理的基本思想, 即用_____。

- A. 分段方法来分配和管理物理存储空间, 用分页方法来管理逻辑地址空间
- B. 分段方法来分配和管理逻辑地址空间, 用分页方法来管理物理存储空间
- C. 分段方法来分配和管理主存空间, 用分页方法来管理辅存空间
- D. 分段方法来分配和管理辅存空间, 用分页方法来管理主存空间

答: B

37. 下列存储管理方案中, 不采用动态重定位的是_____。

- A. 页式管理 B. 可变分区
C. 固定分区 D. 段式管理

答: C

38. 采用虚拟存储器的前提是程序执行时某些部分的互斥性和_____。

- A. 顺序性 B. 局部性
C. 并发性 D. 并行性

答: B

39. 下列存储管理方案中, 不要求将作业全部调入并且也不要求连续存储空间的是_____。

- A. 固定分区 B. 可变分区
C. 页式存储管理 D. 页式虚拟存储管理

答: D

40. 下列不适合多道程序系统的存储管理是_____存储管理。

- A. 单用户连续 B. 固定分区
C. 可变分区 D. 段页式

答: A

41. 与虚拟存储技术不能配合使用的是_____。

- A. 分区存储管理 B. 页式存储管理
C. 段式存储管理 D. 段页式存储管理

答: A

42. 不可能产生系统抖动现象的存储管理是_____。

- A. 固定分区管理 B. 分页式虚拟存储管理
C. 段式虚拟存储管理 D. 以上都不对

答: A

二. 填空题

1. 主存的空间划分为_____和用户占用区, 其中用户区用来存放用户的_____。

答: 系统占用区、数据和程序

2. 操作系统的存储管理负责管理_____。

答: 计算机系统的主存储器

3. 区分不同的存储管理方式, 主要是指采用的主存空间的_____的不同。

答: 分配策略

4. 将程序中的逻辑地址转换成绝对地址, 这种地址转换工作称为_____。

答: 重定位

5. 主存空间的共享是为了_____, 主存空间的存储保护是为了_____。

答: 提高主存空间的利用率、防止各存储区域中的程序相互干扰

6. 主存空间的共享包含两方面含义: 共享_____和共享_____。

答: 主存储器, 主存储器的某些区域

7. 由绝对地址对应的主存空间称_____, 由逻辑地址对应的主存空间称_____。

答: 物理地址空间, 逻辑地址空间

8. 重定位分为两种: _____和_____. 前者是在作业装入时进行的, 后者是在作业_____过程中进行的。

答: 静态重定位, 动态重定位, 执行

9. 能方便进行作业在主存中移动的重定位是_____。

答: 动态重定位

10. 固定分区采用_____重定位方式进行地址转换, 可变分区采用_____重定位方式进行地址转换。

答: 静态, 动态

11. 使用最先适应分配算法, 可以把空闲区按地址_____的顺序登记在空闲区表中, 这样做有利于大作业的装入。

答: 从小到大

12. 最优适应分配算法可按作业要求从所有空闲区中挑选一个能满足作业要求的_____空闲区, 以保证不去分割更大的空闲区。

答: 最小

13. 采用最坏适应分配算法, 空闲区表中的第一个登记项所对应的空闲区总是_____的。

答: 最大

14. 移动技术可集中分散的_____, 提高主存空间的利用率。

答: 空闲区、利用率

15. 某个作业在执行过程中正在等待_____, 则该作业不能移动。

答: 外围设备传输信息

16. 在可变分区方式下采用移动技术可以把主存中的“_____”汇集成一个大的空闲区。

答: 碎片

17. 我们把操作系统所占用的系统资源和所需的处理器时间称为_____。

答: 系统开销

18. 页式存储管理中的页表指出了_____与_____之间的对应关系。

答: 页、块

19. 页式存储管理按给定的逻辑地址读写时, 要访问两次主存: 第一次_____, 第二次_____。

答: 按页号读出页表中所对应的块号、按计算出来的绝对地址进行读写

20. 段式存储管理以段为单位进行存储空间的管理, _____的地址是连续的, _____的地址是不连续的。

答: 段内、段与段之间

21. 分页是由_____自动完成的, 而分段是由_____决定的。

答: 系统、用户

22. 虚拟存储器实际上是为_____而采用的一种设计技巧, 并非真正的存储器。

答: 扩大主存容量

23. 虚拟存储器不能无限大, 它的容量由计算机的_____和_____决定, 而与实际的主存容量无关。

答: 地址结构、辅存的容量

24. 分页式虚拟存储管理的页表增加标志位, 指出对应页是否已经装入_____, 若未装入, 则产生_____中断。

答: 主存、缺页

25. 一个好的页面调度算法应该避免和减少_____现象。

答: 抖动

26. 常用的页面调度算法中, 总是淘汰最先进入主存的那一页的, 称为_____调度算法; 最近最少使用调度算法选择最近一段时间里_____的页调出。

答: 先进先出(或 FIFO)、最久没有被使用过

27. 最近最少使用调度算法是基于程序执行的_____理论, 即程序一旦访问到某些位置的数据或指令, 可能在一段时间里经常会访问它们。

答: 局部性

三. 简答题

1. 操作系统中存储器管理的主要功能是什么? 什么叫虚拟存储器, 并举例说明操作系统是如何实现虚拟内存的?

答: 存储器管理的主要功能是: 内存分配, 地址映射, 内存保护, 内存扩充。

虚拟存储器是用户能作为可编址内存对待的存储空间, 在这种计算机系统中虚地址被映象成实地址。简单地说, 虚拟存储器是由操作系统提供的一个假想的特大存储器, 它通过把主、辅存统一起来管理, 给用户造成一种仿佛系统内有巨大主存供用户使用的假象。例如页式虚存管理, 一道作业被划分成若干页, 其中较活跃的几页放在内存, 而其余不活跃的页被放在辅存, 当需要访问辅存内的页时, 就可通过页面调度将其调入内存运行; 但用户感觉不到这种变化, 他会以为作业的所有部分都存在于主存。这样可以让更多的作业进入主存, 提高系统的效率。

2. 存储管理的基本任务是什么?

答: (1)管理内存空间;

(2)进行虚拟地址(或: 逻辑地址)到物理地址的转换;

(3)实现内存的逻辑扩充;

(4)完成内存信息的共享和保护。

3. 覆盖技术与虚拟存储技术有何本质不同?交换技术与虚存中使用的调入 / 调出技术有何相同与不同之处?

答:覆盖技术中,覆盖段由用户设计,用户自身对内存的划分要参与操作;虚拟存储技术是由系统提供逻辑空间给用户使用,而用户并不真正了解内存的情况,物理空间的划分和管理由系统完成。

交换技术是将内存中处于就绪队列或等待队列的进程暂时调出内存,放入磁盘空间,以便让更多的作业被选择进入内存,提高系统效率。虚存中使用的调入 / 调出技术是利用磁盘空间对内存进行扩充,提供一个大于实际内存的逻辑空间给用户使用。它们的相同之处是:都将本应处于实际内存的内容调至辅存,提高系统效率;不同之处是:交换技术并未提供大于实际内存空间的逻辑空间以供用户使用,该技术并不是直接面向用户的;而虚存技术则是提供更大的逻辑空间以供用户使用,是直接面向用户的。

4. 简述固定分区和可变分区在管理方式上的区别。

答:固定分区的管理方式:分区的个数、大小均固;一个分区只放一个作业。可变分区的管理方式:分区大小和个数依作业情况而定;作业进入主存时才建分区。

5. 试述缺页中断与一般中断的主要区别。

答:缺页中断作为中断,同样需要经历保护 CPU 现场、分析中断原因、转缺页中断处理程序进行处理、恢复 CPU 现场等步骤。但缺页中断又是一种特殊的中断,它与一般中断的主要区别是:

(1)在指令执行期间产生和处理中断信号。通常, CPU 都是在一条指令执行完后去检查是否有中断请求到达。若有便去响应中断;否则继续执行下一条指令。而缺页中断是在指令执行期间,发现所要访问的指令或数据不在内存时产生和处理的。

(2)一条指令在执行期间可能产生多次缺页中断。例如,对于一条读取数据的多字节指令,指令本身跨越两个页面,假定指令后一部分所在页面和数据所在页面均不在内存,则该指令的执行至少产生两次缺页中断。

6. 试述分页系统和分段系统的主要区别。

答:分页和分段有许多相似之处,比如两者都不要求作业连续存放。但在概念上两者完全不同,主要表现在以下几个方面:

(1)页是信息的物理单位,分页是为了实现非连续分配,以便解决内存碎片问题,或者说分页是由于系统管理的需要。段是信息的逻辑单位,它含有一组意义相对完整的信息,分段的目的是为了更好地实现共享,满足用户的需要。

(2)页的大小固定且由系统确定,将逻辑地址划分为页号和页内地址是由机器硬件实现的,而段的长度却不固定,决定于用户所编写的程序,通常由编译程序在对源程序进行编译时根据信息的性质来划分。

(3)分页的作业地址空间是一维的;分段的地址空间是二维的。

7. 假定占有 m 块(初始为空)的进程有一个页访问串,这个页访问串的长度为 p ,其中涉及到 n 个不同的页号。对于任何页面替换算法,求出;

(1)缺页中断次数的下界是多少?

(2)缺页中断次数的上界是多少?

答:(1)缺页中断次数的下界是 m

(2)缺页中断次数的上界是 p 。

8. 引起系统抖动的原因是什么?系统如何检测抖动?一旦检测出抖动后,系统怎样消除它?

答:由于分配给进程的页面数少于进程所需的最低页面数,导致出现接连不断的缺页中断,从而引起系统抖动。系统可以利用将 CPU 的利用率与多道程序的度数进行比较的方法来检测系统抖动,一旦发生抖动,可通过减少多道程序度数的办法来消除它。

9. 试说明多级存储方法与虚拟存储方法的主要区别。

答:所谓多级存储方法是指在主存以外再配上后援存储器来补充主存,先把用户作业放在后援存储器上,然后等到要真正执行它时再设法把这部分作业调入主存。它们之间的调度由用户自己负责,这造成用户使用上的困难。由于系统不同,其主存大小也会随之不同,用户往往不得不修改程序。在批量分时系统中多个用户同时共享主存、因此,用户就不能直接参与主存和后援存储器之间的调度的了。这就需要由操作系统来对它进行统一的、自动的管理,并采用虚拟存储技术;有了虚拟存储器之后,用户就不感到主存和后援存储器之间的区别,而在统一的逻辑存储器上安排程序和数据。由于虚拟存储器比实际存储器大得多,因而可以同时装入许多作业。

10. 在页式虚存管理的页面替换算法中,对于任何给定的物理内存块数,在什么样的引用

串情况下，FIFO 与 LRU 替换算法一样(即缺页中断和中断时被替换的页完全一样)?举例说明。

答：条件是分配的物理内存块数任意，若要 FIFO 与 LRU 算法完全一样，则访问串必须具有其特点。考虑到 FIFO 是替换最早进入主存的页；LRU 是替换上次访问距当前最远的页。二者要相当，则首先进入主存的页在该次缺页中断之前不应被再次访问，也即某页被替换出主存前，仅被访问过一次(非连续的)。由此可得，访问串的规律为：

当发生缺页中断时，最先进入主存的页仅被访问一次。

例如：分配的物理块数为 4，有如下访问串：

1 2 3 4 1 2 3 6

当 6 号页被调入主存时，应在 1, 2, 3, 4 页中选择一个淘汰，FIFO 中应选择的是 1 号页，而 LRU 会选择 4 号页，原因是 1 号页被再次访问，不是上次访问距当前最远的页。但若把对 1 号页的第二次请求去掉，则 LRU 也会选择 1 号页，因为此时 1 号页成为上次访问距当前最远的页。据此类推，整个访问串中每一个页号都应有这种特点，所以最后的答案为：访问串中所有页号均不同，若相同则必须排列在一起。

11. 什么叫地址重定位？动态地址重定位的特点是什么？

答：重定位是指作业装入与其逻辑地址空间不同的物理空间所引起的地址变换过程。

动态地址重定位的特点是：(1)由硬件实现；(2)在程序运行过程中进行地址变换。

12. 在内存管理中“内零头”和“外零头”各指的是什么？在固定式分区分配、可变式分区分配、页式虚拟存储系统、段式虚拟系统中，各会存在何种零头？为什么？

答：内零头(又称内部碎片)：若存储单元长度为 n ，该块存储的作业长度为 m ，则剩下的长度为 $(n-m)$ 的空间称为该单元的内部碎片；若存储单元长度为 n ，在该系统所采用的调度算法下较长时间内无法选出一道长度不超过该块的作业，则称该块为外零头(外部碎片)。

在固定式分区分配中两种零头均会存在，因为空间划分是固定的，无论作业长短，存储单元均不会随之变化，若作业短而存储块长则产生内零头，若作业长而存储块短则产生外零头。

在可变式分区分配中只有外零头而无内零头，因为空间划分是依作业长度进行的，要多少给多少，但剩下的部分太短而无法再分则成为外零头。

页式虚存中会存在内零头而无外零头，因存储空间与作业均分为等长单元，所以不存在无法分配的单元，但作业长度并不刚好为页面大小的整数倍，因此在最后一页会有剩余空间，即为内零头。

段式虚存中会存在外零头而无内零头，因段式的空间划分类似于可变分区分配，根据段长分配，要多少给多少，但会剩余小空间无法分配，则为外零头。

13. 考虑一个请求分页系统，测得如下的时间利用率：CPU:20%，分页磁盘:97.7%，其它外设:5%。下列措施中哪个(些)可改善 CPU 的利用率？说明理由。

更换速度更快的 CPU；

更换更大容量的分页磁盘；

增加内存中的用户进程数；

挂起内存中某个(些)用户进程

采用更快的 I/O 设备。

答：CPU 的利用率指的是系统整个运行时间里，CPU 有多少时间是真正处理用户的数据。由于磁盘利用率已达 97.70%，说明磁盘空间紧张而影响 CPU 利用率，更换大容量的磁盘可以提高 CPU 利用率。慢速外设往往同快速 CPU 不匹配，因此提高 I/O 设备的速度，应当可以改善 CPU 的效率。另外，内存当中进程数增加，可令 CPU 空闲时间减少，也可以在一定程度上提高 CPU 的利用率。

14. 在各种存储管理方案中，哪些方案适合于虚拟存储管理？

答：页式、段式和段页式存储管理都适合进行虚拟存储管理，因为它们都不需要作业在主存中连续存放，这样就可以实现作业的部分装入，进而通过多次对换的方式使作业得以运行。而单用户连续方式、固定分区和可变分区方式都不能实现虚拟存储管理。

15. 简要比较各种存储管理方法的功能和实现特点。

答：列表比较如下：

	单用户 连续方式	固定分区	可变分区	页式	段式	段页式
主存 分配方式	连续区	连续区	连续区	主存块 可以不连续	主存块 可以不连续	主存块 可以不连续
主存 分配表	——	分区分配表	已分配区表 空闲区表	位示图	已分配区表 空闲区表	页式结合 段式
调度算法	——	顺序分配	最先适应 最优适应 最坏适应	页面调度： FIFO、LRU、 LFU		
重定位	静态	静态	动态	动态	动态	动态
地址转换	——	绝对地址 = 下限寄存器的 值 + 逻辑地址	绝对地址 = 基址寄存器的 值 + 逻辑地址	绝对地址 = 块号 × 块长 + 页内地址	绝对地址 = 主存中该段的 起始地址 + 段内地址	段式与 页式结合
存储保护	界限地址 ≤ 绝对地址 ≤ 主存最大地址	下限地址 ≤ 绝对地址 < 上限地址	基址寄存器的 值 ≤ 绝对地址 ≤ 限长寄存器 的值	页表	段表	段表与 页表结合
虚拟存储器	——	——	——	可用	可用	可用
使用系统	单道	多道	多道	多道	多道	多道

四. 应用题

1. 在一个请求分页存储管理系统中, 一个作业的页面走向为 4、3、2、1、4、3、5、4、3、2、1、5, 当分配给该作业的物理块数分别为 3、4 时, 试计算采用下述页面淘汰算法时的缺页次数(假设开始执行时主存中没有页面), 并比较所得结果。

(1)最佳置换法(OPT)

(2)先进先出法(FIFO)

解: (1)根据所给页面走向, 使用最佳页面置换算法时, 页面置换情况如下: (略)

物理块为 3 时, 缺页次数为 7;

物理块为 4 时, 缺页次数为 6。

由上述结果可以看出, 增加分配给作业的内存块数可以降低缺页次数。

(2)根据所给页面走向, 使用先进先出页面置换算法时, 页面置换情况如下: (略)

物理块为 3 时, 缺页次数为 9;

物理块为 4 时, 缺页次数为 10。

由上述结果可以看出, 对先进先出算法而言, 增加分配给作业的内存块数反而出现缺页次数增加的异常现象。

2. 某采用页式存储管理的系统, 接收了一个共 7 页的作业, 作业执行时依次访问的页为: 1、2、3、4、2、1、5、6、2、1、2、3、7。当内存块数量为 4 时, 请分别用先进先出(FIFO)调度算法和最近最少使用(LRU)调度算法, 计算作业执行过程中会产生多少次缺页中断? 写出依次产生缺页中断后应淘汰的页。(所有内存开始时都是空的, 凡第一次用到的页面都产生一次缺页中断。要求写出计算过程)

答: 采用先进先出(FIFO)调度算法, 共产生 10 次缺页中断, 依次淘汰的页是 1、2、3、4、5、6, (页面调度过程略);

采用最近最少使用(LRU)调度算法, 共产生 8 次缺页中断, 依次淘汰的页是 3、4、5、6, (页面调度过程略)。

3. 在一个多道程序系统中，设用户空间为 200K，主存空间管理采用最先适应分配算法，并采用先来先服务算法管理作业。今有如下所示的作业序列，请列出各个作业开始执行时间、完成时间和周转时间。（注意：忽略系统开销，时间用 10 进制。）

作业名	进入输入井时间	需计算时间	主存需求量
JOB1	8.0 时	1 小时	20K
JOB2	8.2 时	0.6 小时	60K
JOB3	8.4 时	0.5 小时	25K
JOB4	8.6 时	1 小时	20K

答：

作业	开始时间	完成时间	周转时间
JOB1	8	9	1
JOB2	9	9.6	1.4
JOB3	9.6	10.1	1.7
JOB4	10.1	11.1	2.5

4. 设某作业的程序部分占一页，A 是该作业的一个 100×100 的数组，在虚空间中按行主秩序存放（即按如下顺序存放：A (1,1), A (1,2), ..., A (1,100), A (2,1), ..., A (2,100) ... A (100,1) ... A (100,100)，页面大小为 100 个字，驻留集为 2 个页帧。若采用 LRU 替换算法，则下列两种对 A 进行初始化的程序段引起的缺页中断次数各是多少？并说明原因。

(a) for j:=1 to 100 do

for i:=1 to 100 do A (i,j) :=0

(b) for i:=1 to 100 do

for j:=1 to 100 do A (i,j) :=0

答：由于程序占 1 页，故放入一个页帧中。只发生 1 次缺页中断。

(a)种情况是按 A (1,1)、A (2,1) ... A (100,1) A (1,2), A (2,2) ..., A (100, 2), A (1,100)、A (2, 100) ... A (100, 100) 次序初始化，而数组的存放为

A (1,1), A (1,2) ... A (1,100)1 页

A (2,1), A (2,2) ... A (2,100)2 页

.....

A (100,1) A (100,2) ... A (100,100)100 页

故对数组中每个元素初始化时，均要发生缺页中断，共发生 100×100 次再加上程序加载所发生的一次，共计 10001 次缺页中断。

(b) 种情况是按：A (1,1)、A (1,2) ... A (1,100), A (2,1), A (2,2) ..., A (2,100) A (100,1) A (100,2) ... A (100, 100) 次序进行初始化，则每初始化一页发一次缺页中断，所以共 101 次。

5. 在一个采用页式虚拟存储管理的系统中，有一用户作业，它依次要访问的字地址序列是：115，228，120，88，446，102，321，432，260，167，若该作业的第0页已经装入主存，现分配给该作业的主存共300字，页的大小为100字，请回答下列问题：按（1）FIFO调度算法（2）LRU调度算法将产生多少次缺页中断，缺页中断率为多少，依次淘汰的页号是什么。

答：（1）按FIFO调度算法将产生5次缺页中断；

依次淘汰的页号为：0，1，2；

缺页中断率为： $5/10=50\%$ 。

（2）按LRU调度算法将产生6次缺页中断；

依次淘汰的页号为：2，0，1，3；

缺页中断率为： $6/10=60\%$ 。

6. 已知页面走向为1、2、1、3、1、2、4、2、1、3、4，且开始执行时主存中没有页面。若只给该作业分配2个物理块，当采用FIFO页面淘汰算法时缺页率为多少？假定现有一种淘汰算法，该算法淘汰页面的策略为当需要淘汰页面时，就把刚使用过的页面作为淘汰对象，试问就相同的页面走向，其缺页率又为多少？

解：根据所给的页面走向，采用FIFO置换算法的页面置换情况如下：

页面走向	1	2	1	3	1	2	4	2	1	3	4
物理块1	1	1		3	3	2	2		1	1	4
物理块2		2		2	1	1	4		4	3	3
缺页	缺	缺		缺	缺	缺	缺		缺	缺	缺

从上述页面置换中可以看出：页面引用次数为11次，缺页次数为9次，所以缺页率为9/11。

若采用后一种置换算法：其页面置换情况如下：

页面走向	1	2	1	3	1	2	4	2	1	3	4
物理块1	1	1		3	1		1	1		3	4
物理块2		2		2	2		4	2		2	2
缺页	缺	缺		缺	缺		缺	缺		缺	缺

从上述页面置换图可以看出：页面引用次数为11次，缺页次数为8次，所以缺页率为8/11。

7. 在一个请求分页系统中，假定系统分配给一个作业的物理块数为3，并且此作业的页面走向为2、3、2、1、5、2、4、5、3、2、5、2。试用FIFO和LRU两种算法分别计算出程序访问过程中所发生的缺页次数。

解：在本题中，分配给作业的物理块数为3。

（1）根据所给页面走向，使用FIFO算法时，页面置换情况如下：

走向	2	3	2	1	5	2	4	5	3	2	5	2
块1	2	2		2	5	5	5		3		3	3
块2		3		3	3	2	2		2		5	5
块3				1	1	1	4		4		4	2
缺页	缺	缺		缺	缺	缺	缺		缺		缺	缺

缺页次数为9。

（2）根据所给页面走向，使用LRU算法时，页面置换情况如下：

走向	2	3	2	1	5	2	4	5	3	2	5	2
块1	2	2		2	2		5		5	5		
块2		3		3	5		2		3	3		
块3				1	1		4		4	2		
缺页	缺	缺		缺	缺		缺		缺	缺		

缺页次数为7。

8. 有一页式系统，其页表存放在主存中，(1)如果对主存的一次存取需要 1.5 微秒，试问实现一次页面访问的存取时间是多少？

(2)如果系统加有快表，平均命中率为 85%，当页表项在快表中时，其查找时间忽略为 0，试问此时的存取时间为多少？

解：若页表存放在主存中，则要实现一次页面访问需两次访问主存，一次是访问页表，确定所存取页面的物理地址，第二次才根据该地址存取页面数据。

(1)由于页表存放在主存，因此 CPU 必须两次访问主存才能获得所需数据，所以实现一次页面访问的存取时间是 $1.5 \times 2 = 3$ 微秒

(2)在系统增加了快表后，在快表中找到页表项的概率为 85%，所以实现一次页面访问的存取时间为 $0.85 \times 1.5 + (1 - 0.85) \times 2 \times 1.5 = 1.725$ 微秒

9. 在一个段式存储管理系统中，段表内容如下：

段号	内存起始地址	段长
0	210	500
1	2350	20
2	100	90
3	1350	590
4	1938	95

试求下述逻辑地址对应的物理地址是什么？

段号	段内位移
0	430
1	10
2	500
3	400
4	112
5	32

解：(1)由于第 0 段的内存始址为 210，段长为 500，故逻辑地址[0, 430]是合法地址。逻辑地址[0, 430]对应的物理地址为 $210 + 430 = 640$ 。

(2)由于第 1 段的内存始址为 2350，段长为 20，故逻辑地址[1, 10]是合法地址。逻辑地址[1, 10]对应的物理地址为 $2350 + 10 = 2360$ 。

(3)由于第 2 段起始地址为 100，段长为 90，所给逻辑地址[2, 500]非法。

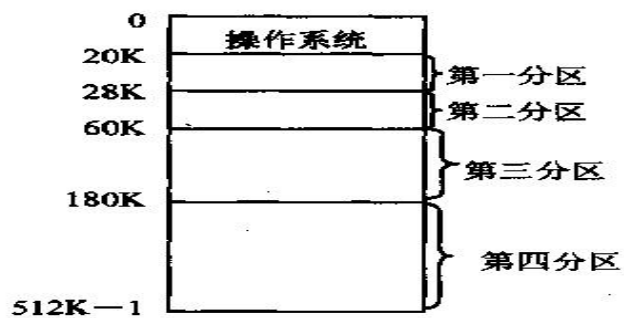
(4)由于第 3 段的内存始址为 1350，段长为 590，故逻辑地址[3, 400]是合法地址。逻辑地址[3, 400]对应的物理地址为 $1350 + 400 = 1750$ 。

(5)由于第 4 段的内存始址为 1938，段长为 95，所给逻辑地址[4, 112]非法。

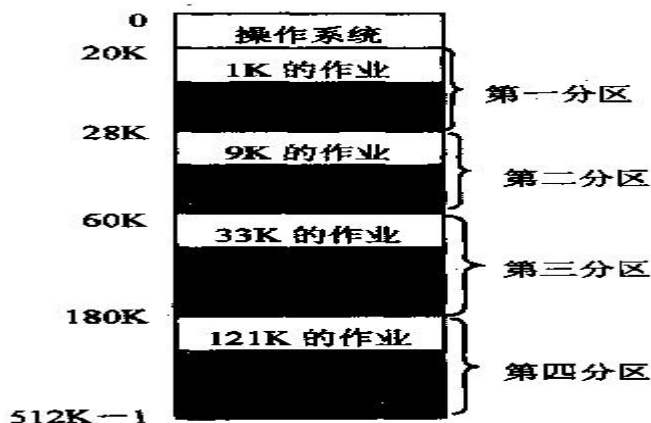
(6)由于系统中不存在第 5 段，所给逻辑地址[5, 32]非法。

10. 在某系统中，采用固定分区分配管理方式，内存分区(单位字节)情况如图 a 所示。现有大小为 1K、9K、33K、121K 的多个作业要求进入内存，试画出它们进入内存后的空间分配情况，并说明主存浪费有多大？

解：从图 a 可以看出，该系统中共有四个分区，第一分区的大小为 8k，第二分区的大小为 32K,第三分区的大小为 120K，第四分区的大小为 332K。作业进入系统后的内存分配情况，如图 b 所示(每个分区中未被利用的那部分空间用阴影表示)：



(图 a 某系统内存分配情况)



(图 b 作业进入系统后的分配情况)

从图 b 可以看出，作业进入系统后，第一分区剩余空间为 7K，第二分区剩余空间为 23K，第三分区剩余空间为 87K，第四分区剩余空间为 211K。主存空间浪费 328K。

11. 下表给出了某系统中的空闲分区表，系统采用可变式分区存储管理策略。现有以下作业序列：96k、20K、200k。若用首次适应算法和最佳适应算法来处理这些作业序列，试问哪一种算法可以满足该作业序列的请求，为什么？

空闲分区表：

分区号	大小	起始地址
1	32K	100K
2	10K	150K
3	5K	200K
4	218K	220K
5	96K	530K

解：若采用最佳适应算法，在申请 96K 存储区时，选中的是 5 号分区，5 号分区大小与申请空间大小一致，应从空闲分区表中删去该表项；接着申请 20K 时，选中 1 号分区，分配后 1 号分区还剩下 12K；最后申请 200K，选中 4 号分区，分配后剩下 18K。显然采用最佳适应算法进行内存分配，可以满足该作业序列的需求。为作业序列分配了内存空间后，空闲分区表如表 a 所示。

若采用首次适应算法，在申请 96K 存储区时，选中的是 4 号分区，进行分配后 4 号分区还剩下 122K；接着申请 20K，选中 1 号分区，分配后剩下 12K；最后申请 200K，现有的五个分区都无法满足要求，该作业等待。显然采用首次适应算法进行内存分配，无法满足该作业序列的需求。这时的空闲分区表如表 b 所示。

分区号	大小	起始地址
1	12K	100K
2	10K	150K
3	5K	200K
4	18K	220K

(分配后的空闲分区表 a)

分区号	大小	起始地址
1	12K	100K
2	10K	150K
3	5K	200K
4	122K	220K
5	96K	530K

(分配后的空闲分区表 b)

12. 在某个采用页式存储管理的系统中, 现有 J1, J2, J3 共 3 个作业同驻主存。其中 J2 有 4 个页面, 被分别装入到主存的第 3, 4, 6, 8 块中。假定页面和存储块的大小均为 1024 字节, 主存容量为 10K 字节。

(1) 写出 J2 的页面映像表;

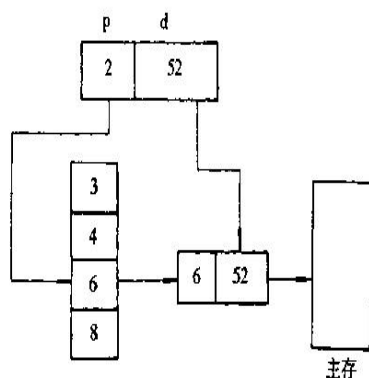
(2) 当 J2 在 CPU 上运行时, 执行到其地址空间第 500 号处遇到一条传指令 MOV 2100, 3100

请用地址变换图计算出 MOV 指令中两个操作数的物理地址。

解: 该题已知条件很多, 但实质还是给出逻辑地址, 要求转换成物理地址。首先得出页表项如图 1 所示, 页面大小为 1024 字节, 可得页内位移为 10 位。逻辑地址 2100 页号为 2, 页内位移 52; 3100 页号为 3。页内位移 28。转换过程如图 2 所示。

0	3
1	4
2	6
3	8

(图 1)



(图 2)

13. 假设有一台计算机, 它有 1M 内存, 操作系统占用 200K, 每个用户进程也占用 200K, 用户进程等待 I/O 的时间为 80%, 若增加 1M 内存, 则 CPU 的利用率将提高多少?

答: 由题目所给条件可知, 当前 1M 内存可支持 4 个用户进程, 由于每个用户进程等待 I/O 的时间为 80%, 故 CPU 的利用率为:

$$1 - (80\%)^4 = 1 - (0.8)^4 = 59\%$$

若增加 1M 内存, 则系统中可同时运行 9 个进程, 则 CPU 的利用率为

$$1 - (0.8)^9 = 87\%$$

故增加 1M 内存使 CPU 的利用率提高了

$$87\% \div 59\% = 147\%$$

$$147\% - 100\% = 47\%$$

所以增加 1M 内存使 CPU 的利用率提高了 47% 的利用率。

第五章 设备管理

一. 选择题

1. 假脱机技术是指_____。

- A. 联机同时外围设备操作技术
- B. 对换技术和覆盖技术
- C. SPOOLing 技术
- D. A 和 C

答: D

2. 缓冲技术中的缓冲池在_____中。

- A. 主存
- B. 外存
- C. ROM
- D. 寄存器

答: A

3. 引入缓冲的主要目的是_____。

- A. 改善 CPU 和 I/O 设备之间速度不匹配的情况
- B. 节省内存
- C. 提高的 CPU 利用率
- D. 高 I/O 设备的效率

答: A

4. CPU 输出数据的速度远远高于打印机的打印速度, 为了解决这一矛盾, 可采用_____。

- A. 并行技术
- B. 通道技术
- C. 缓冲技术
- D. 虚存技术

答: C

5. 为了使多个进程能有效地同时处理输入和输出, 最好使用_____结构的缓冲技术。

- A. 缓冲池
- B. 闭缓冲区环
- C. 单缓冲区
- D. 双缓冲区

答: A

6. 通过硬件和软件的功能扩充, 把原来独立的设备改造成能为若干用户共享的设备, 这种设备称为_____。

- A. 存储设备
- B. 系统设备
- C. 用户设备
- D. 虚拟设备

答: D

7. 如果 I/O 设备与存储设备进行数据交换不经过 CPU 来完成, 这种数据交换方式是_____。

- A. 程序查询
- B. 中断方式
- C. DMA 方式
- D. 无条件存取方式

答: C

8. 中断发生后, 应保留_____。

- A. 缓冲区指针
- B. 关键寄存器内容
- C. 被中断的程序
- D. 页表

答: B

9. 设备管理程序对设备的管理是借助一些数据结构来进行的, 下面的_____不属于设备管理数据结构。

- A. JCB B. DCT C. COCT D. CHCT

答: A

10. 大多数低速设备都属于_____设备。

- A. 独享 B. 共享 C. 虚拟 D. Spool

答: A

11. _____是直接存取的存储设备。

- A. 磁盘 B. 磁带
C. 打印机 D. 键盘显示终端

答: A

12. 以下叙述中正确的为_____。

- A. 在现代计算机中, 只有 I/O 设备才是有效的中断源
B. 在中断处理过程中必须屏蔽中断
C. 同一用户所使用的 I/O 设备也可能并行工作
D. Spooling 是脱机 I/O 系统

答: C

13. _____是操作系统中采用的以空间换取时间的技术。

- A. Spooling 技术 B. 虚拟存储技术
C. 覆盖与交换技术 D. 通道技术

答: A

14. 操作系统中 Spooling 技术, 实质是将_____转化为共享设备的技术。

- A. 虚拟设备 B. 独占设备 、
C. 脱机设备 D. 块设备

答: B

15. Spooling 系统提高了_____的利用率。

- A. 独占设备 B. 共享设备
C. 文件 D. 主存储器

答: A

16. 在操作系统中, _____指的是一种硬件机制。

- A. 通道技术 B. 缓冲池
C. Spooling 技术 D. 内存覆盖技术

答: A

17. 在操作系统中, 用户在使用 I/O 设备时, 通常采用_____。

- A. 物理设备名 B. 逻辑设备名
C. 虚拟设备名 D. 设备牌号

答: B

18. 采用假脱机技术，将磁盘的一部分作为公共缓冲区以代替打印机，用户对打印机的操作实际上是对磁盘的存储操作，用以代替打印机的部分是_____。

- A. 独占设备 B. 共享设备
- C. 虚拟设备 D. 一般物理设备

答：C

19. 按_____分类可将设备分为块设备和字符设备。

- A. 从属关系 B. 操作特性
- C. 共享属性 D. 信息交换单位

答：D

20. _____算法是设备分配常用的一种算法。

- A. 短作业优先 B. 最佳适应
- C. 先来先服务 D. 首次适应

答：C

21. 利用虚拟设备达到 I/O 要求的技术是指_____。

- A. 利用外存作缓冲，将作业与外存交换信息和外存与物理设备交换信息两者独立起来，并使它们并行工作的过程
- B. 把 I/O 要求交给多个物理设备分散完成的过程
- C. 把 I/O 信息先存放在外存，然后由一台物理设备分批完成 I/O 要求的过程
- D. 把共享设备改为某个作业的独享设备，集中完成 I/O 要求的过程

答：A

22. 将系统中的每一台设备按某种原则进行统一的编号，这些编号作为区分硬件和识别设备的代号，该编号称为设备的_____。

- A. 绝对号 B. 相对号 C. 类型号 D. 符号名

答：A

23. 通道是一种_____。

- A. I/O 端口 B. 数据通道
- C. I/O 专用处理器 D. 软件工具

答：C

24. SPOOLing 系统在工作过程中会和下面_____操作系统组成部分发生联系。

- A. 内存管理 B. 处理机管理
- C. 文件管理 D. 设备管理

答：D

二．简答题

1. 简述 Spooling 系统是如何模拟脱机外围设备操作的。

答：预输入程序模拟控制输入的外围机，缓输出程序模拟控制输出的外围机，输入井和输出井模拟脱机外围设备操作的两个磁盘。

2. 在 Spooling 系统中设计了一张"缓输出表"，请问哪些程序执行时要访问缓输出表，简单说明之。

答：井管理程序把作业执行结果文件登记在缓输出表中；缓输出程序从缓输出表中查找结果文件并打印输出。

3. 什么是缓冲?为什么要引入缓冲?

答：缓冲是用来在两种不同速度的设备之间传输信息时平滑传输过程的常用手段。除了在关键的地方采用少量硬件缓冲器之外，大都采用软件缓冲。软件缓冲区是指在 I/O 操作期间用来临时存放输入输出数据的一块存储区域。在操作系统中，引入缓冲的主要原因可归结为以下几点：

(1) 缓和 CPU 与 I/O 设备间速度不匹配的矛盾。一般情况下，程序的运行过程是时而进行计算，时而进行输入或输出。以输出为例，如果没有缓冲，则程序在输出时，必然由于打印机的速度跟不上而使 CPU 停下来等待；然而在计算阶段，打印机又无事可做。如果设置一个缓冲区，程序可以将待输出的数据先输出到缓冲区中，然后继续执行；而打印机则可以从缓冲区取出数据慢慢打印。

(2) 减少中断 CPU 的频率，放宽对 CPU 中断响应时间的限制。例如，假定从系统外传来的数据只用一位二进制位来接收，则每收到一位二进制数就要中断 CPU 一次，如果数据通信速率为 9.6kbs，则中断 CPU 的频率也为 9.6kHz，即每 100us 就要中断 CPU 一次，若设置一个具有 8 位的缓冲寄存器，则可使 CPU 被中断的次数降低为前者的 1/8。

(3) 提高 CPU 和 I/O 设备之间的并行性。

4. 什么是虚拟设备?为什么在操作系统中引入虚拟设备?

答：虚拟设备是指通过虚拟技术将一台独占设备变换为若干台逻辑设备，供若干个用户进程同时使用，通常把这种经过虚拟技术处理后的设备称为虚拟设备。

在操作系统设备管理中，引入虚拟设备是为了克服独占设备速度较慢、降低设备资源利用率的缺点，从而提高设备的利用率。

5. 假脱机系统的基本工作原理是什么?

答：假脱机技术主要由输入程序模块和输出程序模块所组成，系统分别为之创建输入进程和输出进程，它们的优先级高于一般用户进程。输入进程负责通过通道将信息从输入设备送到盘区的输入井中，输出进程负责通过通道将信息从盘区的输出井送到输出设备。主机仅和快速存储设备磁盘中的输入井和输出井交换信息，大大提高了信息处理的速率。

6. 简述设备分配的过程。

答：设备分配程序要用到系统设备表、设备控制表、控制器控制表和通道控制表。设备分配时要考虑到设备的固有属性、分配的算法、防止死锁以及用户程序与实际使用的物理设备无关等特性。设备分配的过程主要是：

(1) 从系统设备表 SDT 中找到需要的物理设备的设备控制表 DCT；

(2) 若设备闲，则分配，然后从设备控制表 DCT 中找到控制器控制表指针所指出的控制器控制表 COCT；

(3) 若控制器闲，则分配，然后从控制器控制表 COCT 中找到通道控制表指针所指出的通道控制表 CHCT；

(4) 根据通道控制表 CHCT 中的状态信息来判断是否可以启动 I/O 设备传送信息，若闲则可以，若忙则把该进程插入到等待通道的队列中去。

7. 在某系统中，从磁盘将一块数据输入到缓冲区需要花费的时间为 T，CPU 对一块数据进行处理的时间为 C，将缓冲区的数据传送到用户区所花时间为 M，那么在单缓冲和双缓冲情况下，系统处理大量数据时，一块数据的处

理时间为多少?

答: (1) 在无缓冲的情况下, 先从磁盘把一块数据输入到用户数据区, 所花费的时间为 T ; 然后便由 CPU 对这一块数据进行计算, 计算时间为 C , 所以每一块数据的处理时间为 $T+C$ 。

(2) 在单缓冲的情况下, 先从磁盘把一块数据输入到缓冲区, 所花费的时间为 T ; 然后由操作系统将缓冲区的数据传送到用户区, 其所花的时间为 M ; 接下来便由 CPU 对这一块数据进行计算, 计算时间为 C 。由于 CPU 的计算操作与磁盘的数据输入操作可以并行, 因此一块数据的处理时间为 $\max(C, T)+M$ 。

(3) 在双缓冲的情况下, 先从磁盘把一块数据输入到第一个缓冲区, 装满后便将数据输入第二个缓冲区, 与此同时操作系统可将第一个缓冲区的数据传送到用户区; 当 CPU 对第一块数据进行计算时, 又可将磁盘输入数据送入第一个缓冲区, 当计算完成后又可从第二个缓冲区中取数据, 如此反复交替使用两个缓冲区。当 $C>T$ 时, 计算操作比输入操作慢, 在此情况下, 上一块数据计算完成后, 仍需将一个缓冲区中的数据传送到用户区, 花费时间为 M , 再对这块数据进行计算, 花费时间 C , 所以一块数据的处理时间为 $C+M$, 即 $\max(C, T)+M$; 当 $C<T$ 时, 输入操作比计算操作慢, 在此情况下, 由于 M 远小于 T , 故在将磁盘上的一块数据传送到一个缓冲区期间(花费时间为 T), 计算机已完成了将另一个缓冲区中的数据传送到用户区并对这块数据进行计算的工作, 所以一块数据的处理时间为 T , 即

$\max(C, T)$ 。

8. 在某计算机系统中, 时钟中断处理程序每次执行的时间为 2ms(包括进程切换开销)。若时钟中断频率为 60Hz, 试问 CPU 用于时钟中断处理的时间比率为多少?

答: 在计算机系统中, 时钟以固定的频率中断 CPU, 以增加日历计数或控制系统中的一些定时操作。由题目所给条件可知, 时钟中断频率为 60Hz, 所以时钟周期为:

$1/60s=50/3ms$: 在每个时钟周期中, CPU 要用 2ms 时间执行中断程序, 所以 CPU 用于时钟中断处理的时间比率为:
 $2/(50/3)=6/50=12\%$

9. 当平均队列长度较小时, 所有的磁盘调度算法都退化为 FCFS 调度算法, 请说明原因。

答: 当队列长度 L 为 1 时, 算法 FCFS、SSTF、LOOK 和 C—LOOK 功能都是等同的, 当 L 为 2 时, 通过把 FCFS 和 SSTF、LOOK 和 C—LOOK 作比较, 就会发现它们功能等同的情况大约是 50%, 而且访问的局部性也可以使调度算法之间差异变得极小。

10. 除 FCFS 外的所有磁盘调度算法都不是真正公平的(例如,会出现饥饿现象)。

(1)说明为什么?

(2)提出一个确保公平性的方案。

(3)为什么公平性在分时系统中是一个很重要的指标?

答: (1)对位于当前磁头所在的磁道上的新请求, 从理论上讲, 只要它们一到达就可得到服务, 对位于其他磁道上的请求则不然。

(2)预定一时间限额, 把所有在这期间内尚未服务的请求“强行”移到队列的顶部, 并置其相关的位以指明任何新的请求都不得移到这些请求之前。对 SSTF 而言, 必须相对于这些“老”请求的最后一个, 重新组织队列的剩余部分。

(3)将公平性作为分时系统中的重要指标的目的在于, 为了避免造成响应时间过长而使某些进程等待太久。

11. SSTF 算法比较有利于中间柱面的磁道, 说明这是为什么?

答：磁盘的中心是到盘中所有其他磁道的平均距离最短的位置，这样在服务了第一个请求后，磁头将很可能更靠近中心磁道。因此，使用 SSTF 算法将有更多的机会首先到达中心磁道。一旦到达某个特定的磁道上，SSTF 往往把磁头保持在该磁道附近，这样就可以调整最初的趋势而使磁头向中间移动。

12. 实现 SPOOLing 技术系统需付出哪些代价？

- 答：(1) 占用大量内存作为外设间传送用的缓冲区，系统所用的表格页占用不少内存空间；
(2) 占用大量磁盘空间用作输入和输出；
(3) 增加了系统的复杂性。

13. 如何将独占设备改造成可共享使用的虚设备？

答：独占型设备的使用效率低，在一段时间内只能由一个用户所占用，从而常常影响整个系统的效率。如果采用 SPOOLing 技术实现将程序运行过程中所需的输入信息从独占型设备上输至输入井，从而使程序运行时可以直接从输入井获得所需的信息。对用户来说，只要保证信息的正确存取就行，至于信息是从磁盘获得还是从输入设备上获得都无关紧要。系统借助于硬件的中断装置和通道技术，使得中央处理器与各种外围设备以及各外围设备之间均可并行工作。作业执行是从磁盘上读信息来代替从输入机的输入操作，不仅使多个作业可以同时执行，而且加快了作业的执行速度，提高了单位时间内处理作业的能力。在作业执行的同时，还可利用输入设备继续预输入作业信息，这种联机同时操作极大地提高了独占设备的利用率。

14. 在设备管理中，何谓设备独立性？如何实现设备独立性？

答：设备独立性，又称设备无关性，指用户程序独立于所使用的具体物理设备，即用户只使用逻辑设备名。其体现在两个方面。一方面是从程序设计的角度看待设备。从这个角度看，各种设备所体现的接口应该都是一致的。程序中可使用相同的命令读出不同设备上的数据，也可以用相同的命令将输出数据送到各种不同的设备上，不同设备之间的差异由操作系统来处理，对程序加以屏蔽。设备无关性的另一方面是指，在操作系统管理设备和相应操作时，对所有的设备都采用统一的方式进行。由于各类设备之间的差异，软件实现时，很难达到真正的一致。一般采用层次式、模块化的思想来实现设备管理子系统。低层程序用来屏蔽设备的具体细节，高层软件将各类设备的操作都以一致的界面对用户提供。与设备无关性紧密相关的是统一命名法。一个文件或设备名将简单地只是一个字符串或一个整数，而完全不依赖于设备。

为实现设备独立性，系统应为每个用户进程配置 1 张用于联系逻辑设备名和物理设备名的映射表，表中一般应包含：逻辑设备名、物理设备名和驱动程序入口地址。

15. 何谓虚拟设备？请说明 Spooling 系统是如何实现虚拟设备的。

答：操作系统利用共享设备来模拟独占设备的工作，当系统只有一台输入设备或一台输出设备的情况下，可允许两个以上的作业并行执行，并且每个作业都感觉到获得了供自己独占使用的输入设备和输出设备，我们说，操作系统采用的这种技术为用户提供了虚拟设备。

Spooling 技术借助磁盘和通道实现了输入 / 输出过程的共享。当用户提出输入 / 输出请求时，系统及时响应，此时用户会认为已独占输入输出设备；但事实上，有多道作业同时进入该过程，并分别占用各个阶段。可假设如下情况：第一道作业提出打印申请，得到响应后正在打印机上输出；此时第二道作业提出输出请求，系统响应后将其送入磁盘输出井，一旦第一道作业打印结束，第二道作业可马上开始打印；接着第三道、第四道作业也源源不断地提出输出请求并得到响应，先后进入输出井及占用打印机。若系统控制得好，可令整个过程被数道作业共享，而每一个进入输出过程的作业都会认为自己在独占打印机。可以说，该系统向用户提供了多台打印机。

16. 磁盘是用来存放文件和数据，因此可以说磁盘仅仅是用作文件系统的设备。这句话对吗？为什么？

答：不正确，磁盘不仅用于存放文件，还可作为主存的延伸，即提供虚存管理。在虚存的设计思想中，主存作为实际的物理空间，仅存放目前较为活跃的程序部分，其它不活跃部分暂存于辅存(即磁盘空间)，等待调度程序在主、辅存间进行交换调度。

17. 什么是通道？通道有哪几种类型？

答：通道又称 I/O 处理机，它主要用于主存与外设之间的信息传输，从属于中央处理机，能与中央处理机并行操作。通道可分为三类：

- (1) 字节多路通道；
- (2) 选择通道；
- (3) 数组多路通道。

18. 为什么说有了通道技术和中断技术才真正做到了 CPU 与外设的并行操作？

答：通道是负责外围设备与主存之间进行数据交换，能单独完成输入输出操作的处理机。有了通道，主存和外围设备之间的数据交换就不要 CPU 干预了，CPU 可以做与输入输出无关的其他工作，从而使计算机系统获得了 CPU 与外围设备之间并行工作的能力。

I/O 中断是通道和 CPU 协调工作的一种手段。如果没有中断技术，CPU 就要不断去查询通道以及设备执行的情况，这样一来，CPU 还是把大量的时间花在了查询上，不能很好地为其他进程服务。使用中断技术，CPU 可以完全不管通道和设备的执行情况，因为无论操作正常结束或操作异常结束，通道都会发出中断，通知 CPU 来处理。

综上所述，通道技术和中断技术的出现，使得主存可以直接与外设交换数据，而 CPU 得以与外设并行地工作，大大提高了 CPU 的使用效率。

19. 分析磁盘数据块的一次读/写操作所需要花费的时间。

答：磁盘数据块的一次读/写操作所需要花费的时间分成三个部分：寻找时间、延迟时间和传送时间。寻找时间是指磁头在移动臂带动下移动到指定柱面所需的时间。延迟时间为指定扇区旋转到磁头位置所需的时间。传送时间是由磁头把扇区中信息读到主存或把主存中信息写到扇区中所需的时间。

寻找时间和延迟时间与信息在磁盘上的位置有关，传送时间是硬件设计就固定的。

三. 应用题

1. 若干个等待访问磁盘者依次要访问的柱面为 20，44，40，4，80，12，76，假设每移动一个柱面需要 3 毫秒时间，移动臂当前位于 40 号柱面，请按下列算法分别计算为完成上述各次访问总共花费的寻找时间。(1) 先来先服务算法；(2) 最短寻找时间优先算法。

解：(1) $3 \text{ 毫秒} \times 292 = 876 \text{ 毫秒}$

(2) $3 \text{ 毫秒} \times 120 = 360 \text{ 毫秒}$

(注：各算法使移动臂的移动次序和移动的柱面数如下：

(1) $40 \rightarrow 20 \rightarrow 44 \rightarrow 40 \rightarrow 4 \rightarrow 80 \rightarrow 12 \rightarrow 76$

(20) (24) (4) (36) (76) (68) (64)

共移动 292 柱面

(2) $40 \rightarrow 44 \rightarrow 20 \rightarrow 12 \rightarrow 4 \rightarrow 76 \rightarrow 80$

(4) (24) (8) (8) (72) (4)

共移动 120 柱面

2. 有如下请求磁盘服务的队列，要访问的磁道分别是 98、183、37、122、14、124、65、67。现在磁头在 53 道上，若按最短寻道时间优先法，磁头的移动道数是多少？

解：最短寻道时间优先法总是让查找时间最短的那个请求先执行，而不考虑请求访问者到来的先后时间。即靠近当前移动臂位置的请求访问者将优先执行。当前磁头在 53 道上

则总的移动道为： $12+2+30+23+84+24+2+59=236$

3. 若磁头的当前位置为 100 磁道，磁头正向磁道号增加方向移动。现有一磁盘读写请求队列：23，376，205，132，19，61，190，398，29，4，18，40。若采用先来先服务、最短寻道时间优先和扫描算法，试计算出平均寻道长度各为多少？

解：(1) 采用先来先服务磁盘调度算法，进行调度的情况为：（从 100 磁道开始）

下一磁道	移动磁道数
23	77
376	353
205	171
132	73
19	113
61	42
190	129
398	208
29	369
4	25
18	14
40	22

移动磁道数总数为 1596，平均寻道长度为 133。

(2) 采用最短寻道时间优先磁盘调度算法，进行调度的情况为（从 100 磁道开始）

下一磁道	移动磁道数
32	32
190	58
205	15
61	144
40	21
29	11
23	6
19	4
18	1
4	14
376	372
398	22

移动磁道总数为 700，平均寻道长度为 58.3。

(3) 采用扫描算法，进行调度的情况为：（从 100 磁道开始，磁头向磁道号增加方向移动）

下一磁道	移动磁道数
132	32
190	58
205	15
376	171
398	22
61	337
40	21
29	11
23	6
19	4
18	1
4	14

移动磁道总数为 692，平均寻道长度为 57.7。

4. 假定在某磁盘上移动臂刚从 58 号柱上完成任务，目前正在 96 号柱面上读信息，并有下列请求序列等待访问磁盘：175，52，157，36，159，106，108，72

(1) 请分别使用先来先服务调度算法、最短寻找时间优先算法、电梯调度算法，排出实际处理上述请求的次序。

(2) 计算(1)中三种算法下移动臂需要移动的距离。

解：(1) 使用先来先服务调度算法，按照访问者的先后次序进行访问，处理次序依次为：

96，175，52，157，36，159，106，108，72。

使用最短寻找时间优先调度算法，总是让查找时间最短的那个请求先执行，处理次序依

次为：96，106，108，72，52，36，157，159，175。

使用电梯调度算法，处理次序依次为：96，106，108，157，159，175，72，52，36。

(2) 先来先服务调度算法，移动臂需要移动的距离为：

$(175-96)+(175-52)+(157-52)+(157-36)+(159-36)+(159-106)+(108-106)+(108-72)=642$ 。最短寻找时间优先调度算法，移动臂需要移动的距离为：

$(106-96)+(108-106)+(108-72)+(72-52)+(52-36)+(157-36)+(159-157)+(175-159)=223$ 。

电梯调度算法，移动臂需要移动的距离为：

$(106-96)+(108-106)+(157-108)+(159-157)+(175-159)+(175-72)+(72-52)+(52-36)=128$ 。

6. 若现在磁盘的移动臂处于第 15 号柱面，下列六个请求者等待访问磁盘，如何响应这些访问才最节省时间？

序号	1	2	3	4	5	6
柱面号	12	5	16	6	16	12
磁头号	2	3	8	4	7	5
扇区号	6	2	7	1	3	6

解：分析各种移臂调度算法，发现以最短寻找时间优先调度算法使移臂距离最短，最节省时间，这时访问柱面的次序为：16，12，6，5。

考虑到旋转调度时无法确定同一柱面上的扇区先响应哪个扇区更快，所以最省时间的调度次序有以下四种：

(1) 3，5，1，6，4，2

(2) 3，5，6，1，4，2

(3) 5，3，1，6，4，2

(4) 5，3，6，1，4，2

第六章 文件管理

一. 选择题

1. 操作系统中对数据进行管理的部分叫做_____。

- A. 数据库系统 B. 文件系统
- C. 检索系统 D. 数据存储系统

答: B

2. 文件系统是指_____。

- A. 文件的集合 B. 文件的目录
- C. 实现文件管理的一组软件 D. 文件、管理文件的软件及数据结构的总体

答: D

3. 从用户角度看, 引入文件系统的主要目的是_____。

- A. 实现虚拟存储 B. 保存系统文档
- C. 保存用户和系统文档 D. 实现对文件的按名存取

答: D

4. 文件的逻辑组织将文件分为记录式文件和_____文件。

- A. 索引文件 B. 流式文件
- C. 字符文件 D. 读写文件

答: B

5. 文件系统中用_____管理文件。

- A. 作业控制块 B. 外页表
- C. 目录 D. 软硬件结合的方法

答: C

6. 为了对文件系统中的文件进行安全管理, 任何一个用户在进入系统时都必须进行注册, 这一级安全管理是_____安全管理。

- A. 系统级 B. 目录级
- C. 用户级 D. 文件级

答: A

7. 为了解决不同用户文件的“命名冲突”问题, 通常在文件系统中采用_____。

- A. 约定的方法 B. 多级目录
- C. 路径 D. 索引

答: B

8. 一个文件的绝对路径名是从_____开始, 逐步沿着每一级子目录向下追溯, 到指定文件的整个通路上所有子目录名组成的一个字符串。

- A. 当前目录 B. 根目录
- C. 多织目录 D. 二级目录

答：B

9. 对一个文件的访问，常由_____共同限制。

- A. 用户访问权限和文件属性
- B. 用户访问权限和用户优先级
- C. 优先级和文件属性
- D. 文件属性和口令

答：A

10. 磁盘上的文件以_____为单位读写。

- A. 块
- B. 记录
- C. 柱面
- D. 磁道

答：A

11. 磁带上的文件一般只能_____。

- A. 顺序存取
- B. 随机存取
- C. 以字节为单位存取
- D. 直接存取

答：A

12. 一般来说，文件名及属性可以收纳在_____中以便查找。

- A. 目录
- B. 索引
- C. 字典
- D. 作业控制块

答：A

13. 最常用的流式文件是字符流文件，它可看成是_____的集合。

- A. 字符序列
- B. 数据
- C. 记录
- D. 页面

答：A

14. 在文件系统中，文件的不同物理结构有不同的优缺点。在下列文件的物理结构中，_____不具有直接读写文件任意一个记录的能力。

- A. 顺序结构
- B. 链接结构
- C. 索引结构
- D. Hash 结构

答：B

15. 在下列文件的物理结构中，_____不利于文件长度动态增长。

- A. 顺序结构
- B. 链接结构
- C. 索引结构
- D. Hash 结构

答：A

16. 如果文件采用直接存取方式且文件大小不固定，则宜选择_____文件结构。

- A. 直接
- B. 顺序
- C. 随机
- D. 索引

答：D

17. 文件系统采用二级目录结构，这样可以_____。

- A. 缩短访问文件存储器时间
- B. 实现文件共享

- C. 节省主存空间 D. 解决不同用户之间的文件名冲突问题

答: D

18. 下列叙述中正确的五项是_____。

- A. 在磁带上的顺序文件中插入新的记录时，必须复制整个文件。
- B. 由于磁带的价格比磁盘便宜，用磁带实现索引文件更经济。
- C. 在索引顺序文件的最后添加新的记录时，必须复制整个文件。
- D. 在磁带上的顺序文件的最后添加新的记录时，不必须复制整个文件。
- E. 顺序文件是利用磁带的特有性质实现的，因此顺序文件只有存放在磁带
- F. 索引顺序文件既能顺序访问，又能随机访问。
- G.. 直接访问文件也能顺序访问，但一般效率较差。
- H. 变更磁盘上的顺序文件的记录内容时，不一定要复制整个文件。
- L. 在磁盘上的顺序文件中插入新的记录时，必须复制整个文件。
- J. 索引顺序文件是一种特殊的顺序文件，因此通常存放在磁带上。

答: ADFGH

19. 以下叙述中正确的是_____。

- A. 文件系统要负责文件存储空间的管理，但不能完成文件名到物理地址的转换。
- B. 多级目录结构中，对文件的访问是通过路径名和用户目录名来进行的。
- C. 文件被划分成大小相等的若干个物理块，一般物理块的大小是不固定的。
- D. 逻辑记录是对文件进行存取操作的基本单位。

答: D

20.常用的文件存取方法有两种：顺序存取和_____存取。

- A. 流式 B. 串联
C. 顺序 D. 随机

答: D

21.文件管理实际上是管理_____。

- A. 主存空间 B. 辅助存储空间
C. 逻辑地址空间 D. 物理地址空间

答: B

22.操作系统是通过_____来对文件进行编排、增删、维护和检索。

- A. 按名存取 B. 数据逻辑地址
C. 数据物理地址 D. 文件属性

答: A

23. 操作系统实现按名存取进行检索的关键在于解决_____。

- A. 文件逻辑地址到文件具体的物理地址的转换
- B. 文件名称与文件具体的物理地址的转换
- C. 文件逻辑地址到文件名称的转换

D. 文件名称到文件逻辑地址的转换

答：B

24. 下列描述不是文件系统功能的是_____。

- A. 建立文件目录
- B. 提供一组文件操作
- C. 实现对磁盘的驱动调度
- D. 实现从逻辑文件到物理文件间的转换

答：C

25. 有关操作系统和其他系统程序组成的文件是_____。

- A. 系统文件
- B. 档案文件
- C. 用户文件
- D. 顺序文件

答：A

26. 逻辑文件是_____的文件组织形式。

- A. 在外部设备上
- B. 从用户观点看
- C. 虚拟存储
- D. 目录中

答：B

27. 文件的逻辑结构是由_____决定的。

- A. 用户
- B. 操作系统
- C. 编译程序
- D. 装入程序

答：A

28. 与文件物理组织形式有关的是_____。

- A. 文件长度
- B. 记录的个数
- C. 文件目录结构
- D. 用户对文件的存取方式

答：D

29. 数据库文件的逻辑结构形式是_____。

- A. 字符流式文件
- B. 档案文件
- C. 记录式文件
- D. 只读文件

答：C

30. 记录式文件内可以独立存取的最小单位是由_____组成的。

- A. 字
- B. 字节
- C. 数据项
- D. 物理块

答：C

31. 根据辅存设备不同，文件被划分为若干个大小相等的物理块,它是_____的基本单位。

- A. 存放文件信息或分配存储空间
- B. 组织和使用信息
- C. 表示单位信息
- D. 记录式文件

答：A

32. 文件系统为每个文件各建立一张指示逻辑记录和物理记录之间的对应关系表，由此表和文件本身构成的文件是_____。

- A. 顺序文件 B. 链接文件
- C. 索引文件 D. 逻辑文件

答：C

33. 在以下的文件物理存储组织形式中，常用于存放大型系统文件的是_____。

- A. 顺序文件 B. 链接文件
- C. 索引文件 D. 多重索引文件

答：A

34. 在文件系统中，索引文件结构中的索引表是用来_____。

- A. 指示逻辑记录逻辑地址的 B. 存放部分数据信息的
- C. 存放关键字内容的 D. 指示逻辑记录和物理块之间对应关系的

答：D

35. 若用户总是要求用随机存取方式查找文件记录，则采用索引结构比采用链接结构_____。

- A. 麻烦 B. 方便
- C. 一样 D. 有时方便有时麻烦

答：B

36. 记录的成组和分解提高了存储介质的利用率和减少启动设备的次数，但代价是_____。

- A. 要设立主存缓冲区 B. 操作系统增加成组和分解功能
- C. A 和 B D. 没有额外系统开销

答：C

37. _____的有序集合称为文件目录。

- A. 文件符号名 B. 文件控制块
- C. 文件内部名 D. 文件占用的存储块

答：B

38. 文件名与_____的转化是通过文件目录来实现的。

- A. 逻辑地址 B. 物理地址
- C. 文件内部名 D. 文件记录

答：B

39. 多级目录结构形式为_____。

- A. 线性结构 B. 散列结构
- C. 网状结构 D. 树形结构

答：D

40. 文件系统采用二级文件目录可以_____。

- A. 缩短访问存储器的时间 B. 解决同一用户间的文件命名冲突
- C. 节省内存空间 D. 解决不同用户间的文件命名冲突

答：D

41. 在操作系统中，将文件名转换为文件存储地址，对文件实施控制管理都是通过_____来实现的。

- A. 文件目录 B. PCB 表
- C. 路径名 D. 文件名

答：A

42. 允许多个用户同时使用同一个共享文件时，下列做法不正确的是_____。

- A. 允许多个用户同时打开共享文件执行读操作
- B. 允许读者和写者同时使用共享文件
- C. 不允许读者和写者同时使用共享文件
- D. 不允许多个写者同时对共享文件执行写操作

答：B

43. 采取哪种文件存取方式，主要取决于_____。

- A. 用户的使用要求 B. 存储介质的特性
- C. A 和 B D. 文件的逻辑结构

答：C

44. 下列物理文件中，不要求物理块连续存放，但却不方便随机存取的是_____。

- A. 顺序文件 B. 链接文件
- C. 索引文件 D. 逻辑文件

答：B

二. 简答题

1. 什么是文件?它包含哪些内容及特点?

答：文件是信息的一种组织形式，是存储在外存上的具有标识名的一组相关信息集合。

文件包含的内容有：源程序、二进制代码、文本文档、数据、表格、声音和图像等。

文件的特点如下：

- (1)文件具有保存性，它被存储在某种存储介质上，长期保存和多次使用。
- (2)文件是按名存取的，每个文件具有惟一的标识名，通过标识名(文件名)来存取文件中的信息，而不需了解文件在存储介质上的具体物理位置。
- (3)文件的内容是一组信息的集合，信息可以是源程序、二进制代码、文本文档、数据、表格、声音和图像等。

2. 文件系统要解决哪些问题?

答：文件系统的主要目标是提高存储空间的利用率，它要解决的主要问题有：完成文件存储空间的管理，实现文件名到物理地址的转换，实现文件和目录的操作，提供文件共享能力和安全措施，提供友好的用户接口。文件系统向用户提供了有关文件和目录操作的各种功能接口和系统调用，如命令接口、程序接口和交互接口等。

3. 文件系统中常采用的物理结构有哪些?

答：文件的物理结构侧重于提高存储空间的利用率和减少存取时间，它对文件的存取方法有较大的影响。由于外存设备的不同，文件被划分为大小相等的物理块，它是存放文件信息或分配存储空间的基本单位，也是文件系统与主存之间传输和交换信息的基本单位。物理块大小一般是固定的，物理块与逻辑记录的关系可以是：一个物理块可以存放一个或

多个逻辑记录，或者多个物理块存放一个逻辑记录。 目前操作系统中常采用如下物理结构文件：

(1) 顺序文件：它是按照逻辑文件中的记录顺序，依次把逻辑记录存储到连续的物理块中而形成的文件。

(2) 链接文件：它的物理块不是连续的，也不必顺序排列，但每个物理块中设置一个指针，指向下一个物理块的地址，这样，所有的物理块被链接起来，形成一个物理文件，称为链接文件或串联文件。

(3) 索引文件：它是文件系统为每个文件另外建立一张指示逻辑记录和物理块之间的对应关系表，此表称为索引表，文件本身和索引表组成的文件称为索引文件。

4. 什么是文件的逻辑组织和物理组织？

答：文件的逻辑组织——用户对文件的观察和使用是从自身处理文件中数据时采用的组织方式来看待文件组织形式。这种从用户观点出发所见到的文件组织形式称为文件的逻辑组织。

文件的物理组织——文件在存储设备上的存储组织形式称为文件的物理组织。

5. 一个含五个逻辑记录的文件，系统把它以链接结构的形式组织在磁盘上，每个记录占用一个磁盘块，现要求在第一记录和第二记录之间插入一个新记录，简述它的操作过程。

答：从文件目录中找到该文件，按址读出第一个记录；取出第一个记录块中指针，存放到新记录的指针位置；把新记录占用的物理块号填入第一个记录的指针位置；启动磁盘把第一个记录和新记录写到指定的磁盘块上。

6. 为什么文件系统要具有为文件分配存储空间的能力？

答：要把文件保存到存储介质上时，必须要记住哪些存储空间已经被占用，哪些存储空间是空闲的，文件只能保存到空闲的空间去，否则会破坏已经保存的信息。

7. 当用户自愿撤离或终止一个作业时，某些系统自动删除所有相关的文件，除非用户明显地请求保留它们；而另一些系统则保留所有文件，除非用户明显的请求删除它们。试叙述这两种途径的特点。

答：对于前一种途径，由于不必保留不再需要的文件，可使每一用户所需的文件空间变成最小。后一种途径则对用户更为安全，因为用户不会由于疏忽没有保存文件而使文件丢失。

8. 设置文件目录的目的是什么？文件目录组织都有哪几种形式？

答：由于系统中文件的种类多、数量大，所以设置文件目录的目的是为了便于查找文件、操作文件、共享文件及保护文件。

文件目录组织有三种形式：(1)简单(一级)目录，(2)二级目录；(3)树型目录

9. 简述文件系统按名存取的含义。

答：文件系统为用户提供“按名存取”，即用户不必考虑文件存储在哪里，怎样组织输入输出等工作，只要使用文件名，操作系统通过查找目录，就能对存储介质上的信息进行相应的操作。

10. 什么是流式文件？什么是记录式文件？

答：逻辑文件有两种形式：流式文件和记录式文件。流式文件对文件内的信息不再划分单位，它是由一组相关信息组成的有序字符流。长度直接按字节计算。记录式文件允许用户把信息按逻辑上独立的涵义划分为信息单位(记录)，文件由若干个相关的记录组成，每个记录又由若干个数据项组成。

11. 怎样实现文件的保护和保密？

答：文件的保护是防止文件被破坏。为防止系统故障造成的破坏应采用建立副本和定时转储的方法保护文件；为防止用户共享文件时造成的破坏，应规定用户使用文件的权限。 文件的保密是指防止他人未经文件主授权而窃取文件。常见的文件保密措施有：设置口令、使用密码等。

12. UNIX 系统如何防止用户共享文件时造成的破坏?

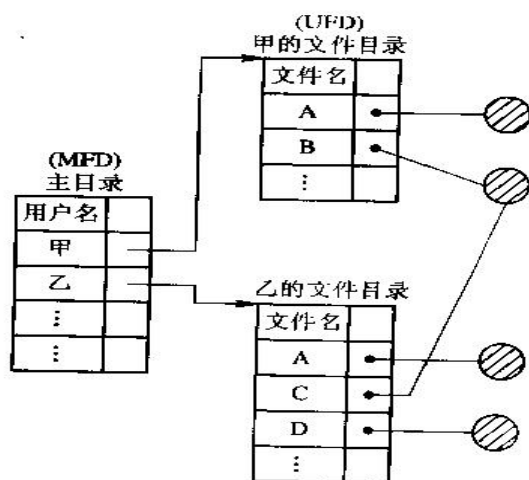
答: UNIX 系统将用户分成三类: 文件主、同组用户和其他用户。文件主根据情况规定其同组用户和其他用户对文件的使用权限(读、写、可执行), 用 3 位二进制数表示某类用户对某个文件的存取权限, 如 111 表示可读、可写、可执行, 而相应位为 0 则表示不允许该操作。

某文件的保护信息是 111 110 100, 就表示文件主可读、写、执行, 同组用户可读、写, 而其他用户仅能读。

三. 应用题

1. 若有甲、乙两个用户, 甲用户有文件 A、B, 乙用户有文件 A、C、D, 甲用户的文件 A 与乙用户的文件 A 不是同一个文件。甲用户的文件 B 与乙用户的 C 是同一个文件。请设计一个目录组织方案, 并画图说明。

解: 由于本问题是有两个用户, 并且存在文件重名和别名问题, 因此目录组织方案应采用二级目录结构, 如下图所示。



2. 有 3 个学生使用一个分时计算机系统, 程序和数据同时存放在同一磁盘上, 他们在各自的终端上如果都为自己的程序取了同一个名字 WJ1, 请问:

(1) 系统应采用何种的目录结构, 才能区别这些学生的程序?

(2) 简单阐述系统怎样为这 3 个学生索取他们各自的程序 WJ1。

解: (1) 考虑到不同文件的重名问题, 该系统应采取二级或多级文件目录。由于重名只发生在不同用户之间, 所以可以采用二级文件目录。在主目录下, 建立 3 个学生的学生目录, 然后将他们的程序 WJ1 放到他们各自的学生目录中。

(2) 任何一个学生要索取自己的文件时, 先给出程序名 WJ1, 系统检查主目录, 根据学生名找到该学生的学生目录, 然后找到程序名为 WJ1 的文件存放地址, 按地址启动磁盘就可以读出学生所需要的程序。

3. 一个文件的逻辑记录大小为 125 字节, 共有 20 个逻辑记录, 文件系统采用链接方式将这个文件存储到磁盘上, 磁盘分块大小为 512 字节, 请问:

(1) 采用什么方法可有效地利用磁盘空间?

(2) 若用户要读包含第 1285 字节的逻辑记录, 文件系统将如何工作?

答: (1) 可采用记录的成组方法, 因 $\lceil 512/125 \rceil = 4$, 所以成组的逻辑记录个数应为 4。20 个逻辑记录共需要 5 个存储块(0~4)。

(2) 首先, 由 $\lceil 1285/(125 \times 4) \rceil = 2$, 可知包含 1285 字节的逻辑记录在链接结构的第 2 块上。将该块读入主存缓冲区; 然

后由 $1285 \bmod (125 \times 4) = 285$ ，且 $\lceil 285/125 \rceil = 2$ ，可知文件系统从主存缓冲区中取出第 2 个记录传输给用户。
(注意：记录号和存储块号都是从 0 开始排序的。)

第七章 操作系统接口

一. 选择题

1. 只允许操作系统使用，不允许用户程序中直接使用的指令称为_____。

- A. 访管指令 B. 中断指令
- C. 特权指令 D. 调用指令

答：C

2. 当处理器处于管态时，处理器可以执行的指令应该是_____。

- A. 非特权指令 B. 仅限于特权指令
- C. 一切指令 D. 访管指令

答：C

3. 访管指令运行在_____。

- A. 管态 B. 目态
- C. 操作系统核心层 D. 硬件固化

答：B

4. 当操作系统退出执行，让用户执行时，系统会_____。

- A. 继续保持管态 B. 继续保持目态
- C. 从管态变为目态 D. 从目态变为管态

答：C

5. CPU 状态分为管态和目态，从目态转换到管态的惟一途径是_____。

- A. 运行进程修改程序状态字 B. 中断屏蔽
- C. 系统调用 D. 进程调度程序

答：C

6. 系统调用的目的是_____。

- A. 请求系统服务 B. 终止系统服务
- C. 申请系统资源 D. 释放系统资源

答：A

7. 用户要在程序一级获得系统帮助，必须通过_____。

- A. 进程调度 B. 作业调度
- C. 键盘命令 D. 系统调用

答：D

8. 设置时钟指令属于_____。

- A. 特权指令 B. 非特权指令
- C. 特权和非特权指令 D. 都不是

答：A

9. 系统调用工作于_____。

- A. 管态 B. 目态
- C. 用户态 D. 均可

答：A

10. _____程序可以执行特权指令。

- A. 一般用户 B. 同组用户
- C. 特权用户 D. 操作系统

答：D

二. 填空题

1. 操作系统为用户提供两个接口。一个是__①__，用户利用它来组织和控制作业的执行或管理计算机系统。另一个是__②__，编程人员使用它们来请求操作系统提供服务。

答：①命令接口 ②程序接口

2. 所谓系统调用，就是用户在程序中调用_____所提供的一些子功能。

答：操作系统

3. 按命令接口对作业控制方式的不同可将命令分为__①__和__②__。

答：①联机命令接口 ②脱机命令接口

4. 为了保证计算机系统能可靠地工作，把 CPU 执行的指令分成两类：_____和_____，其中前者是不允许用户程序中直接使用的指令。

答：特权指令、非特权指令

5. 为避免用户程序使用特权指令，CPU 的工作状态分为_____和_____。操作系统在_____下工作，用户程序在下工作。

答：管态、目态、管态、目态

6. 当中央处理器处于目态时不允许执行_____指令，当中央处理器处于管态时可以执行_____指令，以保证系统的安全可靠。

答：特权、一切

7. 操作系统编制了许多不同功能的子程序，供用户程序执行中调用，这些由操作系统提供的子程序简称为_____。

答：系统调用

8. 当处理器执行到用户的访管指令时就产生一个_____，实现用户程序与_____程序之间的转换。

答：中断事件、系统调用

三. 简答题

1. 在用户与操作系统之间存在哪几种类型的接口?它们的主要功能是什么?

答：用户与操作系统之间的接口主要有系统调用、键盘命令和作业控制命令。系统调用可直接写在用户程序中，它是为了方便的使用系统资源而提供的系统调用命令。键盘命令用于联机交互环境，用户从键盘上敲入命令后操作系统解释并执行这个命令。作业控制命令则用于脱机批处理环境下的作业说明书中，用户用这些命令编写作业说明

书，以告诉系统对作业的控制意图和处理方式。

2. 对于一个正确运转的计算机系统，保护 OS 是很重要的。但为了向用户提供更大的灵活性，应尽可能少地对用户加以限制。试问下面至少有哪几条指令需加以保护？

- (1) 改变成用户方式；
- (2) 改变成系统方式；
- (3) 从存放 OS 的存储区上读取数据
- (4) 将数据写到存放 OS 的存储区上
- (5) 从存放 OS 的存储区取指令
- (6) 打开计时器；
- (7) 关闭计时器。

答：至少第(2)、(4)操作应当加以保护。因为这两条操作是对操作系统本身的内容进行修改，如果交由用户任意操作，会对系统产生破坏，最终可能导致运行出错甚至系统瘫痪，所以应在任何时刻都加以保护。其它几种操作在一般情况下也应当保护，但至少不会像以上两种操作那样若不保护可能会对系统产生极大的破坏性，而本题问的是在最少保护的情况下，因此不做选择。

3. 什么是特权指令？

答：计算机系统中不允许用户程序直接使用的指令称为特权指令。例如，I/O 指令、设置时钟、设置控制寄存器等指令都是特权指令。

4. CPU 在目态和管态下工作有何不同？

答：当中央处理器(CPU)处于管态时，可以执行包括特权指令在内的一切机器指令；而中央处理器在目态下工作时，不允许执行特权指令。操作系统在管态下工作，用户程序在目态下工作，这样做保证了计算机系统的安全可靠。

5. 用户如何使用系统调用？

答：计算机系统一般都提供一条访管指令，它可以在目态下执行，用户要使用系统调用时必须使用访管指令；当 CPU 执行到访管指令时就产生一个中断事件，实现用户程序与系统调用程序之间的转换；由系统调用程序分析访管指令的参数，让相应的系统调用子程序为用户服务。

6. 系统调用与一般过程调用的主要区别是什么？

答：所谓系统调用，就是用户在程序中调用操作系统所提供的一些子功能。它是通过系统调用命令，中断现行程序而转去执行相应的子程序，以完成特定的系统功能。完成后，控制又返回到发出系统调用之后的一条指令，被中断的程序将继续执行下去。系统调用与一般过程调用的主要区别是：

(1) 运行的状态不同。一般的过程调用，其调用程序和被调用程序都运行在相同的状态—系统态或用户态；而系统调用的调用过程是用户程序，运行在用户态，而其被调用过程是系统过程，运行在系统态。

(2) 进入的方式不同。一般的过程调用可以直接由调用过程转向被调用过程。而执行系统调用时，由于调用过程和被调用过程处于不同的系统状态，因而不允许由调用过程直接转向被调用过程，通常是通过软中断（访管中断），先进入操作系统，经分析后，才能转向相应的系统调用处理子程序。

(3) 代码层次不同。一般过程调用中的被调用程序是用户级程序，而系统调用是操作系统中的代码程序，是系统级程序。

7. 系统调用是 OS 与用户程序的接口，库函数也是 OS 与用户程序的接口。这句话对吗?为什么?

答：不正确，系统中的各种资源都由操作系统统一掌管，故在用户程序中凡是与资源有关的操作都必须通过某种方式向 OS 提出服务请求，由 OS 代为完成。于是 OS 必须提供某种形式的接口，以便让用户程序通过接口使用系统提供的各种功能。这种接口称为系统调用。可见系统调用可以看成是用户在程序一级请求 OS 为之服务的一种手段。而库函数则是在程序设计语言中，将一些常用的功能模块编写成函数，放在函数库中供公共选用。函数库的使用与系统的资源分配并无关系，仍属用户程序而非 OS 程序，其功能的实现并不由 OS 完成，且运行时仍在用户状态而非系统状态。