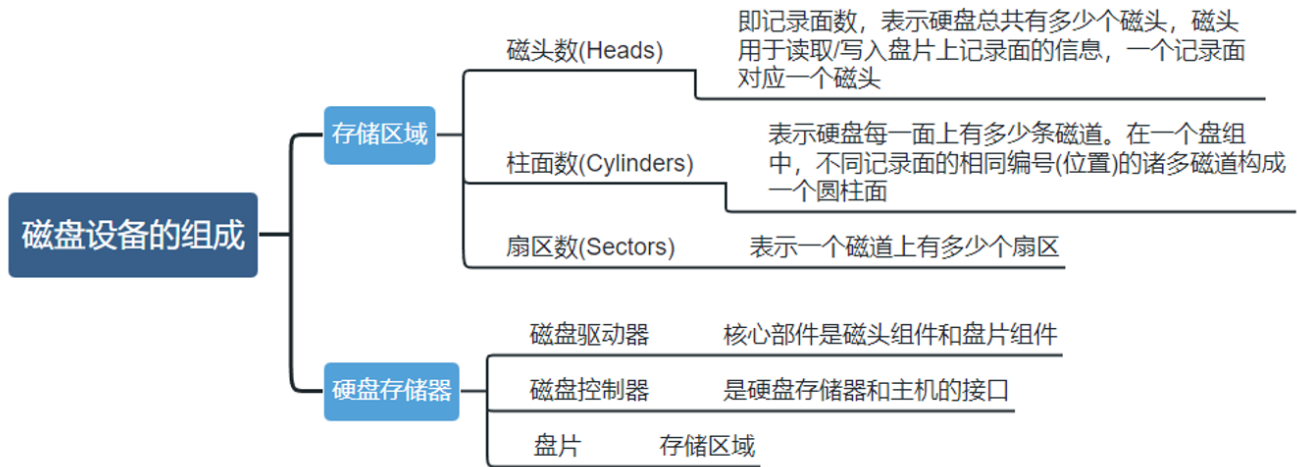


主题3

1、解析磁盘访问时间的组成，举例说明如何计算。

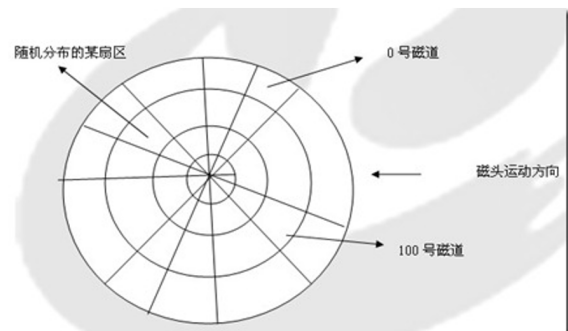


假设计算机系统采用CSCAN（循环扫描）磁盘调度策略

(1) 设某单面磁盘旋转速度为每分钟6000转，每个磁道有100个扇区，相邻磁道间的平均移动时间为1ms

若在某时刻，磁头位于100号磁道处，并沿着磁道号增大的方向移动（如下图所示），磁道号请求队列为50.90.30.120对请求队列中的每个磁道需读取1个随机分布的扇区，则读完这个扇区共需要多少时间？

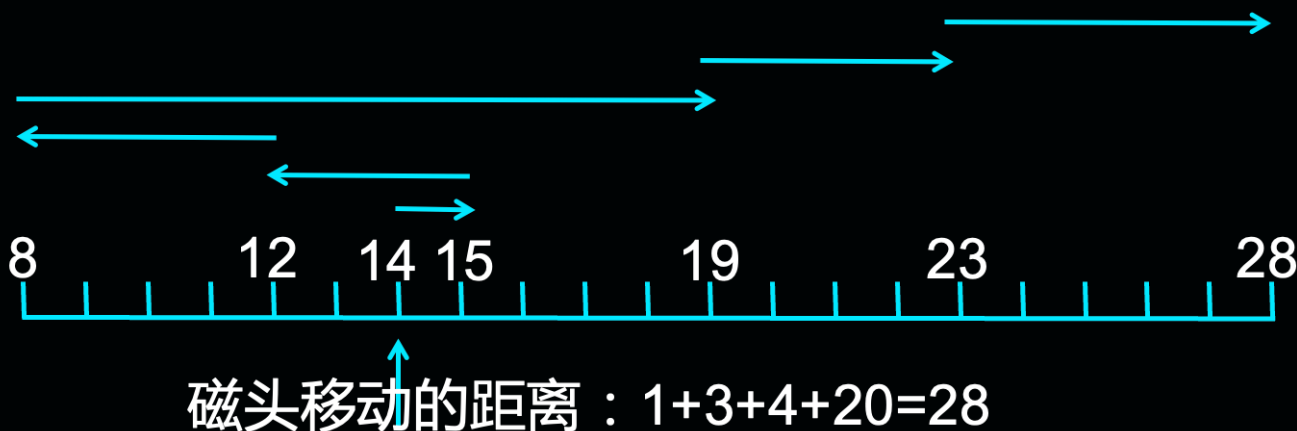
(1) 每分钟6000转，转一圈时间是0.01，同时因为是100个扇区，通过一个扇区的时间是0.0001s。
SCAN算法是单方向的，因限定了只能沿着磁道号增大的方向移动所以得出如下数序：100→120→30→50→90
寻道时间： $(120-100)+(120-30)+(50-30)+(90-50)) * 1ms$
 $=20+90+20+40=110+60=170ms$
旋转延迟时间： $0.01 * 0.5 * 4 = 20ms$
读取时间： $0.0001 * 4 = 0.4ms$
磁盘访问时间是 $170+20+0.4=192.4ms$



2、某时刻的磁盘访问请求序列：14、19、8、15、23、12、28，磁头当前在14磁道，向磁道号增加的方向移动。分别用SSTF和SCAN算法计算磁头移动的距离。通常情况下为何不用性能更好的SSTF算法？

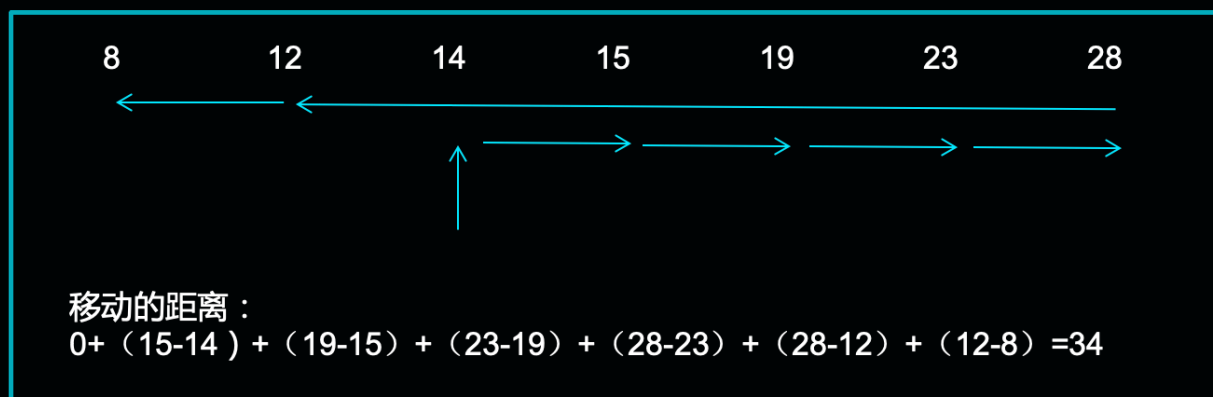
SSTF算法：

某时刻的磁盘访问请求序列：14、19、8、15、23、12、28，磁头当前在14磁道，向磁道号增加的方向移动。使用SSTF 算法计算磁头移动的距离。



SCAN算法：

某时刻的磁盘访问请求序列：14、19、8、15、23、12、28，磁头当前在14磁道，向磁道号增加的方向移动。使用SCAN 算法计算磁头移动的距离。



SSTF是基于贪心策略的，其优点为平均寻道长度大大减少，但是其缺点为可能出现“饥饿”现象

假设有一个硬盘，磁头数为16，每个磁道的扇区数为64，每个扇区存放512个字节，磁盘转速为每分钟7200转。操作系统限定每一个分区大小不超过65536个柱面。

3、题干如上描述请分析解决如下问题：

(1) 可设置分区的最大容量为多少 (KB) ？

存储容量 = 磁头数 × 磁道(柱面)数 × 每道扇区数 × 每扇区字节数

题目中：

- 磁盘磁头数为16
- 磁道数=柱面数，故磁道数最大为65536，即 2^{16} 个
- 每道扇区数为64
- 每扇区字节数为512B

故分区的最大容量：

$$16 * 2^{16} * 64 * 512B = 2^{35}B = 32GB$$

老式的硬盘每个磁道的扇区数一样，外圈的密度小，内圈的密度大，每圈可存储的数据量是一样的。新式的硬盘数据的密度都一致，这样磁道的周长越长，扇区就越多，存储的数据量就越大。

(2) 访问磁盘的平均旋转延迟时间是多少(ms)?

数据存储于盘片上，盘片旋转。读写头不能同时定位在所有数据上，因此盘片以较快的速度旋转，以便将数据放在读写头下。盘片在磁头下旋转数据所花费的时间就是旋转延迟。

旋转延迟：盘片旋转将请求数据所在扇区移至读写磁头下方所需要的时间，旋转延迟取决于磁盘转速，平均旋转延迟时间通常使用磁盘旋转一周所需时间的1/2表示。

$$1/2 * 1/(7200/60) = 4.167ms$$

(3) 访问磁盘的传输时间是多少? (ms)

磁盘访问时间=寻道时间+旋转时间+传输时间

传输时间

传输时间指的是从磁盘读出或将数据写入磁盘的时间。指把数据从磁盘读出或向磁盘写入数据所经历的时间若每次读/写的字节数为b，磁盘每秒钟的转速为r，一条磁道上的字节数为N

$$t=b/(rN)$$

$$\text{传输速度} = \frac{7200 \times 64 \times 512B}{60s} = 3840KB/s$$

$$\text{传输时间} = \frac{b}{3840KB/s}$$

(4) 若用位示图管理磁盘空闲分区，需要用多少个扇区存放位示图？

首先需要算出有多少个位示图，即先算出有多少个物理块需要管理

由于题目没有说明一个物理块含有多少个扇区，所以设一个物理块为 2^n 个扇区

总扇区数 = 磁头数 × 磁道（柱面）数 × 每道扇区数

由于一个扇区是512个字节，所以

所需扇区数 = 总扇区数 / 2^n / 8 / 512B

故所需扇区数：

$$16 * 2^{16} * 64 / 2^n / 8 / 512B = 2^{14-n}$$

位示图是利用二进制的一位来表示磁盘中一个盘块的使用情况。当其值为“0”时，表示对应的盘块空闲；为“1”时表示已分配。由所有盘块对应的位构成一个集合，称为位示图。

4、题干如上描述请分析解决如下问题：

(5) 若文件系统采用索引分配方式，则需要用几个字节表示索引项？

假设：磁盘块大小为 n B，且磁盘块指针全部用于标识物理块。

补充练习1

一个索引节点所表示的UNIX文件的组织：有12个直接块指针，在每个索引节点中有一个一级、二级和三级间接指针。此外，假设系统块大小和磁盘扇面大小都是8K。如果磁盘块指针是32位，其中8位用于标识物理磁盘，24位用于标识物理块，那么

磁盘块数 = 磁盘总容量 / 磁盘块大小

磁盘块数：33554432KB / n B = $2^{35} / n$ 块

索引项所需要的二进制位数： $\log_2 \frac{2^{35}}{n} = 35 - \log_2 n$ 位

因此，需要 $\frac{35}{8} \log_2 n$ 个字节表示索引项

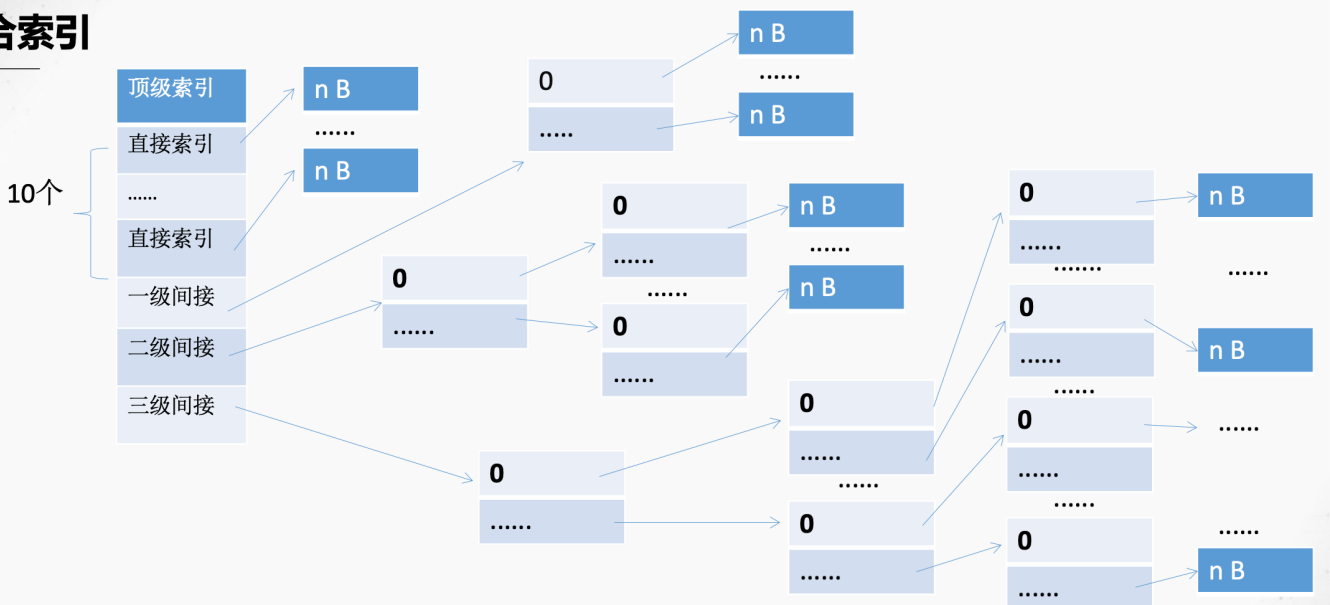
(6) 若采用UNIX混合索引方式，（索引表为13个单元，前10个为直接索引，这13个单元存放在磁盘inode中），则在这个系统中可以建立的最大文件的长度是多少（KB）？

索引块所存储的索引项的个数=磁盘块大小/索引项大小

因为磁盘块大小为 nB ，索引项大小为 $\frac{35-1}{8}\log_2 n$ 个字节

索引块所存储的索引项的个数： $n / (\frac{35-1}{8}\log_2 n) = \frac{8n}{35-\log_2 n}$ 个

混合索引



最大文件长度=索引项的总数×磁盘块大小

$$\text{最大文件长度: } [10 + \frac{8n}{35-\log_2 n} + (\frac{8n}{35-\log_2 n})^2 + (\frac{8n}{35-\log_2 n})^3] \times n = 10n + \frac{8n^2}{35-\log_2 n} + \frac{64n^3}{(35-\log_2 n)^2} + \frac{512n^4}{(35-\log_2 n)^3} B$$

$$\text{即 } \frac{5n}{512} + \frac{n^2}{128(35-\log_2 n)} + \frac{n^3}{16(35-\log_2 n)^2} + \frac{n^4}{2(35-\log_2 n)^3} \text{KB}$$

5、题干如上描述请分析解决如下问题：

(7) 若要存放一个文件长度为64208 KB的文件，此文件的文件体（文件本身的内容）需要多少个磁盘块？

| 磁盘块大小 | 所需磁盘块个数 |
|-------|---------|
| 512B | 128416 |
| 1KB | 64208 |
| 2KB | 32104 |
| 4KB | 16052 |
| 8KB | 8026 |
| 16KB | 4013 |
| 32KB | 2007 |
| 64KB | 1004 |
| 128KB | 502 |

(8) 设某个文件长度为64208 KB，需要多少个磁盘块存放这个文件的索引表？

| 磁盘块大小 | 索引项大小 | 存放文件磁盘块个数 | 存放索引表磁盘块个数 |
|-------|-------|-----------|--------------|
| 512B | 4B | 128416 | 1018(三级索引) |
| 1KB | 4B | 64208 | 252 (二级索引) |
| 2KB | 3B | 32104 | 48 (二级索引) |
| 4KB | 3B | 16052 | 22 (二级索引) |
| 8KB | 3B | 8026 | 3 (二级索引) |
| 16KB | 3B | 4013 | 1 (一级索引) |
| 32KB | 3B | 2007 | 1 (一级索引) |
| 64KB | 3B | 1004 | 1 (一级索引) |
| 128KB | 3B | 502 | 1 (一级索引) |

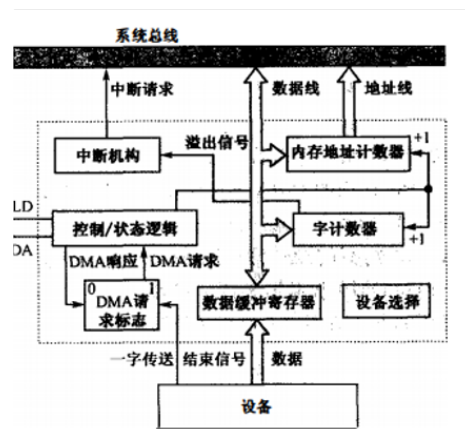
6、 比较书中讲述的I/O控制方式。PC中哪些设备使用了DMA方式？它们是如何工作的？

选择和衡量 I/O 控制方式有如下三条原则：

1. 数据传送速度足够快，能满足用户的需求但又不丢失数据；
2. 系统开销小，所需的处理控制程序少
3. 能充分发挥硬件资源的能力，使 I/O 设备尽可能忙，而 CPU 等待时间尽可能少。

使用DMA的设备：当下计算机中主要有磁盘控制器、显卡、声卡和网卡使用DMA技术。

声卡录制或播放数字音频时，将使用DMA通道，在其本身与RAM之间传送音频数据，而无需CPU干预，以提高数据传输率和CPU的利用率。16位声卡有两个DMA通道，一个用于8位音频数据传输，另一个则用于16位音频数据传输。



7、为什么要进行缓冲管理？讲述Unix字符设备和块设备的缓冲管理方法，讲述缓冲池的工作原理。

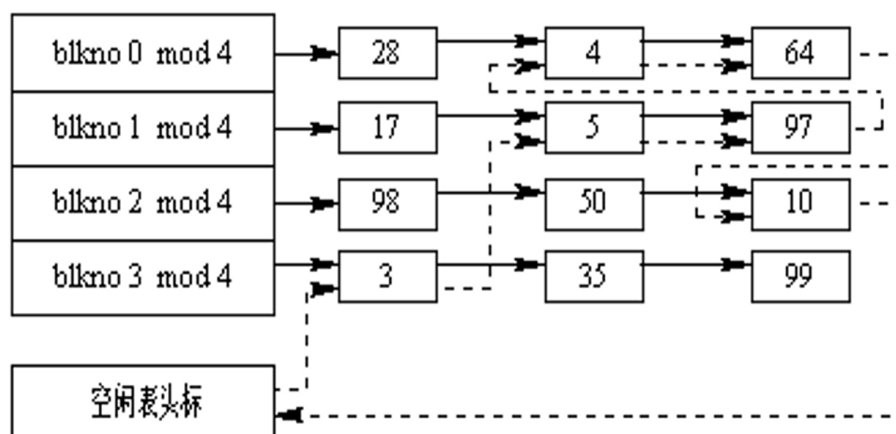
缓冲管理作用：

1. 为了缓和解决CPU与I/O设备传输速度不匹配的问题，使信息得以在系统中平滑传输
2. 提高了CPU、通道和设备之间的并行性，从而使系统的资源利用率及吞吐量增高
3. 减少中断次数和CPU的中断处理时间，放宽对中断响应时间的限制。没有缓冲，慢速I/O设备每传一个字节就要产生一个中断，CPU必须处理该中断。缓冲技术则让慢速I/O设备将缓冲区填满时，才向CPU发出中断，从而减少了中断次数和CPU的中断处理时间，降低了CPU的开销

设备分为字符设备和块设备两类：

1. 字符设备只能以字节为最小单位访问。常见的如键盘、控制台。
而块设备以块为单位访问，例如512字节，1024字节等。常见的如磁盘、光盘。
2. 块设备进行随机访问，而字符设备进行顺序访问。

| |
|-----------|
| 设备号 |
| 块号 |
| 状态 |
| 缓冲区指针 |
| 散列队列的前向指针 |
| 散列队列的后向指针 |
| 空闲表上的前向指针 |
| 空闲表上的后向指针 |



缓冲首部

空闲队列 (链) 及散列队列

8、 回顾脱机I/O技术，谈谈SPooling技术的实现过程。

脱机I/O技术：

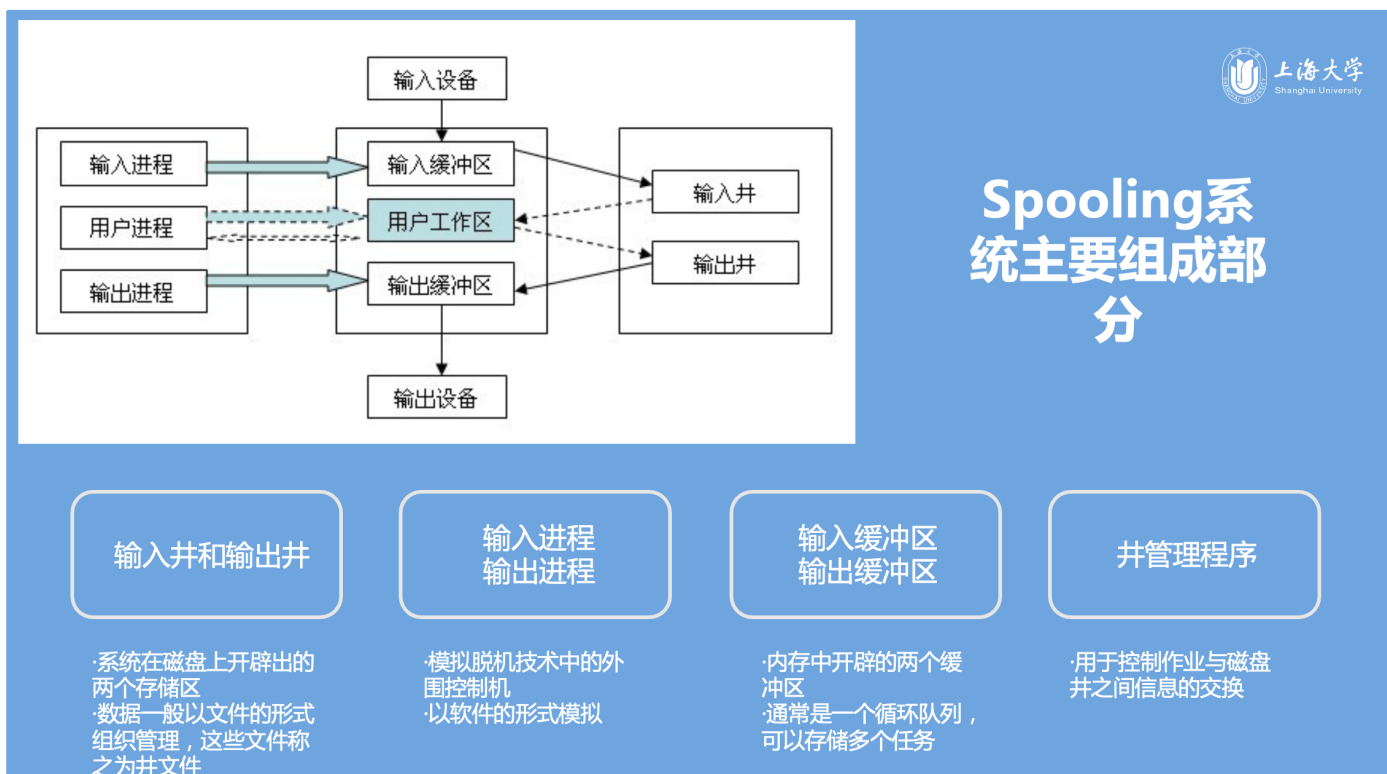
为了解决人机矛盾及CPU和I/O设备之间速度不匹配的矛盾，20世纪50年代末出现了脱机I/O技术。该技术是将事先装有用户程序和数据的纸袋装入纸带输入机，在一台外围机的控制下，把纸带（卡片）上的数据（程序）输入道磁带上。当CPU需要这些程序和数据时，再从磁带上高速地调入内存。

假脱机管理进程：

1. 在磁盘缓冲区中为要打印的数据申请空闲盘块，并将要打印的数据送入盘块中暂存。
2. 为用户进程申请一张空白的用户请求打印表，并将用户的打印要求填入其中，再将该表挂入假脱机文件队列上。

假脱机打印进程：负责真正的打印输出。

SPooling技术实现过程：

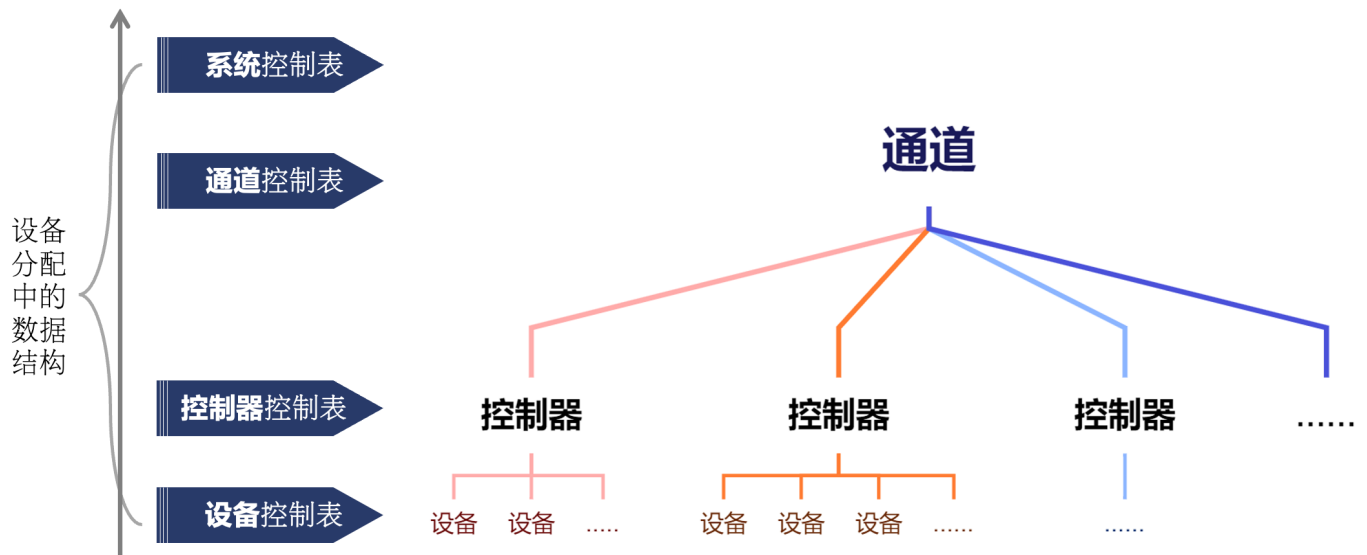


9、设备分配时的数据结构有哪些？分配时会出现死锁吗？Linux中有哪些就数据结构？它是如何分配的？

操作系统中，设备分配时的数据结构主要有以下几种：

- 系统控制表
- 通道控制表
- 控制器控制表
- 设备控制表

其关系如下图所示：



从进程运行的安全性上考虑，设备分配有以下两种方式：

1. 安全分配方式

- 每当进程发出 I/O 请求后，便进入阻塞状态，直到其 I/O 操作完成时才被唤醒。
- 在采用该策略时，一旦进程已经获得某种设备后便阻塞，不能再请求任何资源，而在它阻塞时又不保持任何资源。

2. 不安全分配方式

- 进程在发出 I/O 请求后仍继续运行，需要时又发出第二个 I/O 请求、第三个 I/O 请求等。
- 仅当进程所请求的设备已被另一进程占用时，才进入阻塞状态。