



# 操作系统内核的 IO子系统

# 内核I/O子系统

- ◆ 内核与I/O有关服务：**I/O调度、缓冲、高速缓冲、spooling、设备预订、错误处理。**
- ◆ 内核I/O子系统负责：
  - | 文件和设备命名空间的管理
  - | 文件和设备访问控制
  - | 操作控制（for example, a modern cannot seek()）
  - | 文件系统空间的分配
  - | 设备分配
  - | 缓冲、高速缓存、假脱机
  - | I/O调度
  - | 设备状态监控、错误处理、失败恢复
  - | 设备驱动程序的配置和初始化

# I/O调度

◆ **I/O调度**：调度一组I/O请求就是确定一个好的顺序来执行这些请求。

- | 某些I/O需要按设备队列的顺序——先来先服务
- | 某些操作系统尝试着公平——优先级高者优先
- | **磁盘I/O调度**

◆ **实现**

- | OS通过为每个设备维护一个请求队列来实现调度。
- | 可以试图公平，也可以根据不同的优先级进行I/O调度。
- | 其他方法：缓冲、高速缓冲、假脱机

# 缓冲buffer

- ◆ **缓冲 Buffering**—用来保存在两设备之间或在设备和应用程序之间所传输数据的内存区域。
- ◆ **缓冲区管理**：为了解决**CPU与I/O之间速度不匹配的矛盾**，在它们之间配置了缓冲区。这样设备管理程序又要负责管理缓冲区的建立、分配和释放。
- ◆ **缓冲作用**：
  - | 解决设备速度不匹配
  - | 解决设备传输块的大小不匹配
  - | 为了维持**拷贝语义 “copy semantics”** 要求
- ◆ **单缓冲、双缓冲、多缓冲、缓冲池**

# 高速缓存

- ◆ **高速缓存 Caching**—高速缓存 (cache)：是可以保留数据拷贝的高速内存。
- ◆ **缓冲与高速缓存的差别**是缓冲只是保留数据仅有的一个现存拷贝，而根据定义高速缓存只是提供了一个驻留在其他地方的数据的一个高速拷贝。
- ◆ 高速缓存和缓冲是两个不同的功能，但有时一块内存区域也可以同时用于两个目的。
  - | 当内核接收到I/O请求时，内核首先检查高速缓存以确定相应文件的内容是否在内存中。如果是，物理磁盘I/O就可以避免或延迟。

# 假脱机技术

- ◆ **SPooling** (Simultaneous Peripheral Operation On Line)，称为**假脱机**技术。：用来保存设备输出的缓冲，这些设备如打印机不能接收交叉的数据流。
  - 1 操作系统通过截取对打印机的输出来解决这一问题。应用程序的输出先是假脱机到一个独立的磁盘文件上。当应用程序完成打印时，假脱机系统将相应的待送打印机的假脱机文件进行排队
- ◆ **Printing**：打印机虽然是独享设备，通过SPooling技术，可以将它改造为一台可供多个用户共享的设备。

# 设备预订和错误处理

## ◆ 设备预订—提供对设备的独占访问

- | 分配和再分配的系统调用
- | 有可能产生死锁

## ◆ 错误处理 Error Handling

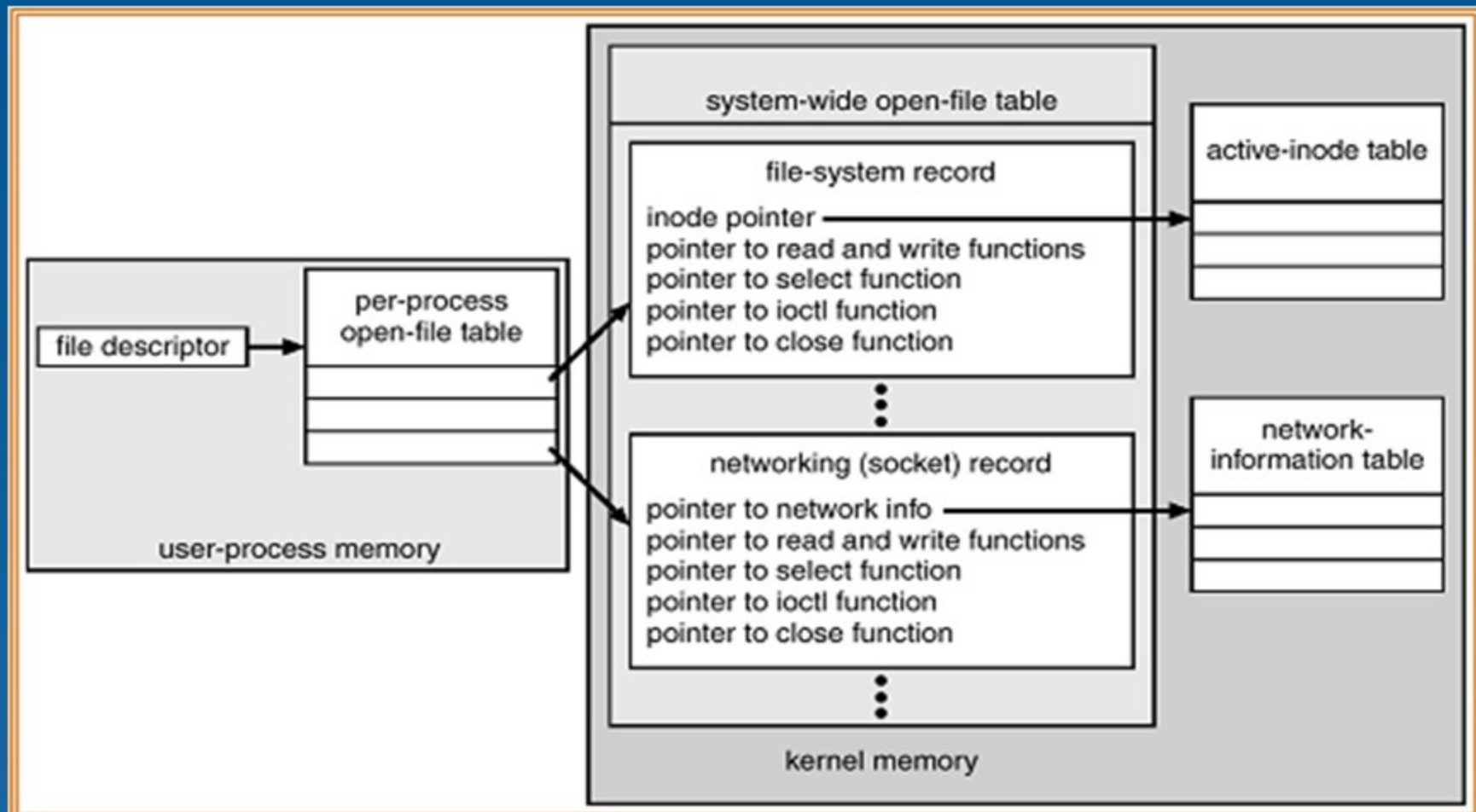
- | 操作系统可以恢复磁盘读，设备无效，暂时的失败
- | 当I/O失败时，大多数返回一个错误码
- | 系统日志记录了出错报告

# 内核数据结构

- ◆ 内核需要保存留I/O组件使用的状态信息，包括打开文件表，网络连接，字符设备状态等
- ◆ 许多复杂的数据结构用来跟踪缓冲，内存分配，及“脏”块
- ◆ 某些OS用面向对象的方法和消息传递的方法来实现I/O



# UNIX I/O内核结构



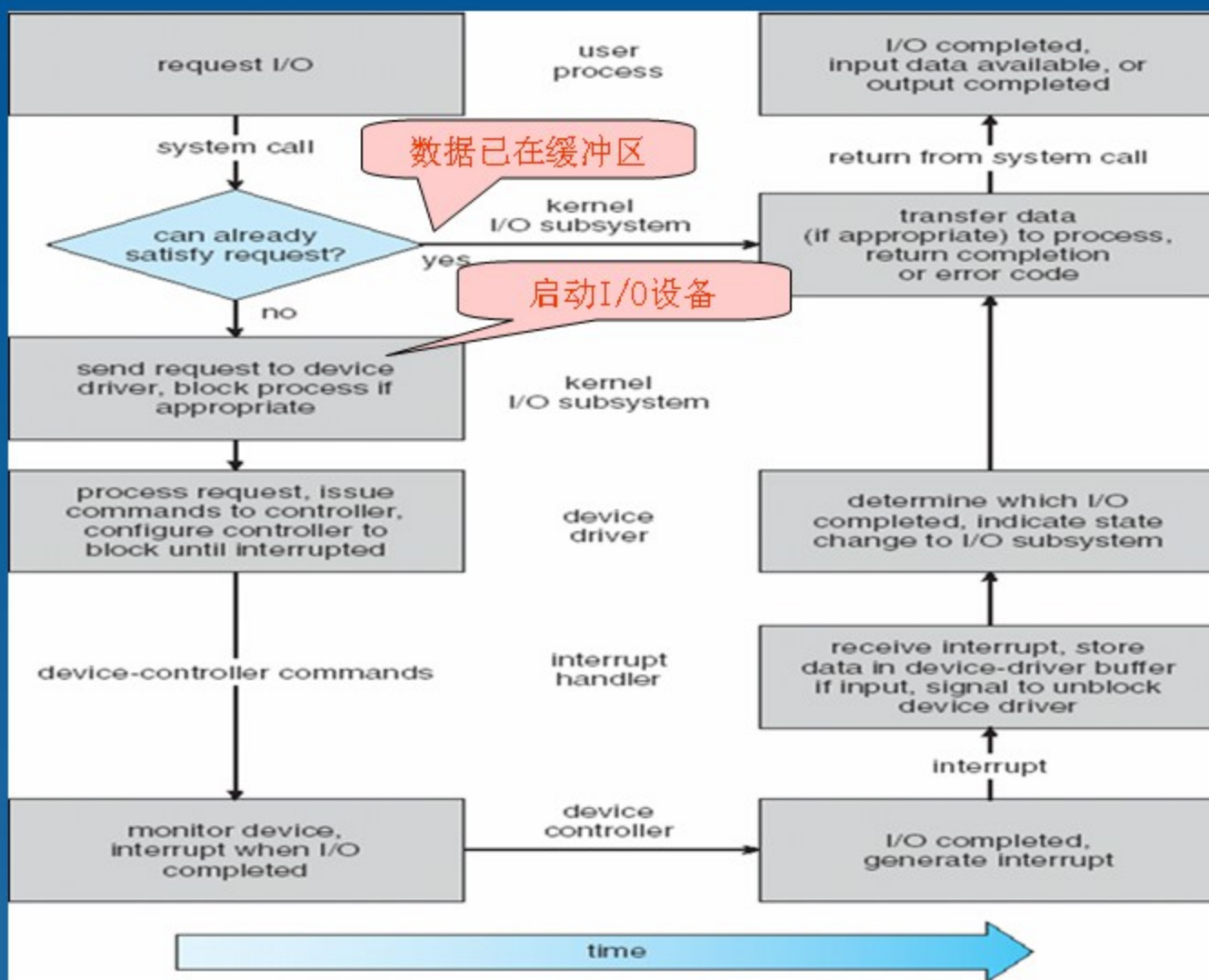
# 转换I/O请求为硬件操作

## I/O请求到硬件操作

### ◆ 考虑一个进程从磁盘中读取一个文件过程

- | 确定保存文件的设备
- | 转换名字到设备的表示法
- | 把数据从磁盘读到缓冲区中
- | 通知请求进程数据现在是有效的
- | 把控制权返回给进程

# I/O请求的周期



# 习题分析

1. 在操作系统中，用户在使用I/O设备时，通常采用\_\_\_\_\_。  
A. 设备的绝对号      B. 设备的相对号  
C. 虚拟设备号      D. 设备名
2. 在现代操作系统中采用缓冲技术的主要目的是\_\_\_\_\_。  
A. 改善用户编程环境      B. 提高CPU的处理速度  
C. 提高CPU和设备之间的并行程度      D. 实现与设备无关性
3. 设备的打开、关闭、读、写等操作是由\_\_\_\_\_完成的。  
A. 用户程序      B. 编译程序  
C. 设备分配程序      D. 设备驱动程序

# 习题分析

4. I/O系统有三种常用方式来与主机交换数据，它们是程序轮询方式、中断方式和DMA方式，其中DMA方式主要由硬件来实现，此时高速外设和内存之间进行数据交换\_\_\_\_\_。
- A. 不通过CPU的控制，不利用系统总线
  - B. 不通过CPU的控制，利用系统总线
  - C. 通过CPU的控制，不利用系统总线
  - D. 通过CPU的控制，利用系统总线



**End**