2021级期末考试复习版本

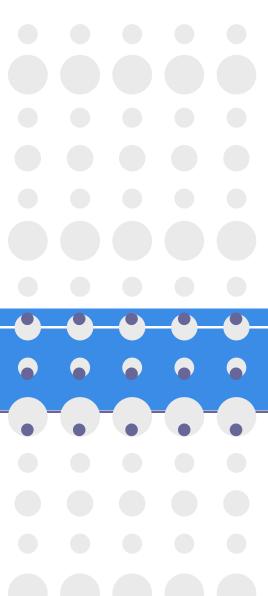
《操作系统原理》

第8章 设备管理(输入/输出管理)

教师: 邹德清,李珍,苏曙光 华中科技大学网安学院 2023年10月-2024年01月

第8章 设备管理(输入/输出管理)

- ●内容
 - ■设备管理概述
 - ■缓冲技术
 - ■设备分配
 - ■SPOOL技术
 - ■设备驱动程序
- ●重点
 - ■理解缓冲机制的作用
 - ■理解"设备是文件"的概念
 - ■掌握设备驱动程序开发



8.1 设备管理概念

设备类型和特征

- 按交互对象分类
 - ■人机交互:显示设备、键盘、鼠标、打印机
 - ■与CPU交互:磁盘、磁带、传感器、控制器、
 - 计算机间交互: 网卡、调制解调器
- 按交互方向分类
 - ■输入设备:键盘、扫描仪
 - ■输出设备:显示设备、打印机
 - ■双向设备:输入/输出:硬盘、软盘、网卡
- 按外设特性分类
 - ■使用特征:存储设备、输入设备、输出设备
 - ■数据传输率:低速(键盘)、中速(打印机)、高速(网卡、磁盘)

设备类型和特征

- 按信息组织特征分类
 - ■字符设备
 - ◆传输的基本单位是字符。例:键盘、串口
 - ■块设备
 - ◆传输的基本单位是块。例: 硬盘, 磁盘
 - ■网络设备
 - ◆采用socket套接字接口访问
 - ◆在全局空间有唯一名字,如eth0、eth1

设备管理功能

- 设备管理的目标
 - (1) 提高设备读写效率
 - ◆设备缓冲机制
 - (2) 提高设备的利用率
 - ◆设备分配(设备调度)
 - (3) 为用户提供统一接口
 - ◆实现设备对用户透明

设备管理功能

- 设备管理的功能
 - ■1) 状态跟踪
 - ■2)设备分配
 - ■3)设备映射
 - ■4)设备控制/设备驱动
 - ■5)缓冲区管理

设备管理功能>状态跟踪

● 状态跟踪

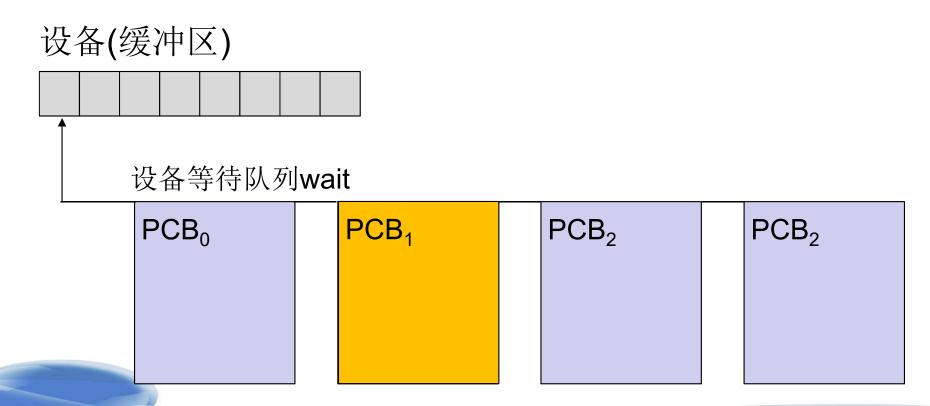
- ■记录设备的基本属性、状态、操作接口及进程访问 信息。
 - ◆设备控制块(Device Control Block, DCB)

设备名 设备属性 命令转换表 在I/0总线上的设备地址 设备状态 当前用户进程指针 I/0请求队列指针

- 设备名
 - 设备的物理名
- 设备属性
 - 设备当前状态(一组属性)
- 命令转换表
 - 设备操作接口
 - ◆设备I/O例程(可为NULL)

设备管理功能>设备分配

- ●功能
 - ■按一定策略安全地分配和管理各种设备。
 - ◆按相应算法把设备分配给请求该设备的进程,并 把未分到设备的进程放入设备等待队列。



- ●设备逻辑名/友好名(Friendly Name)
 - ■用户编程时使用的名字(文件名/设备文件名)
 - ■例: Linux: /dev/test

```
int testdev = open("/dev/test",O RDWR);
if ( testdev == -1 )
    printf("Cann't open file ");
    exit(0);
Read(testdev, lpBuffer, .....);
Write(testdev, lpBuffer, .....);
```

- ●设备逻辑名/友好名(Friendly Name)
 - ■用户编程时使用的名字(文件名/设备文件名)
 - ■例: Windows: \\.\MyDevice

```
hDevice = CreateFile("\\\.\\MyDevice",
                  GENERIC WRITE|GENERIC_READ,
                  FILE SHARE WRITE | FILE_SHARE_READ,
                  NULL,
                  OPEN EXISTING,
                  0.
                  NULL);
ReadFile(hDevice, lpBuffer, .....);
WriteFile(hDevice, lpBuffer, .....);
```

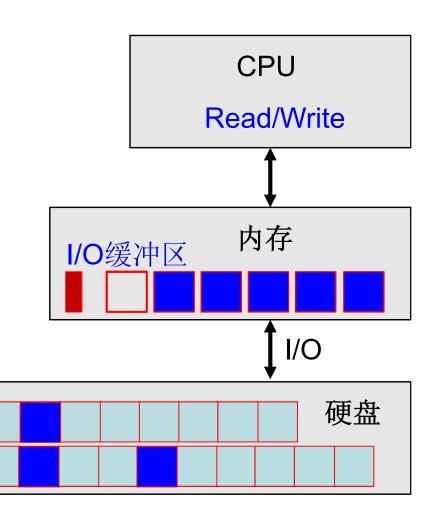
- 设备物理名
 - ■I/0系统中实际安装的设备
 - ■物理名:设备端口、中断号或主/次设备号等
- ●设备映射
 - ■逻辑设备到物理设备的转换
 - ◆逻辑名到物理名的转换

- 设备独立性/设备无关性
 - ■用户程序中使用统一接口访问逻辑设备,而不用考虑对应物理设备的特殊结构和操作方式。

```
#include "stdio.h"
3
     Evoid main (void)
5
            write (0, "Hello World!/n")
6
            printf("Hello World!/n");
```

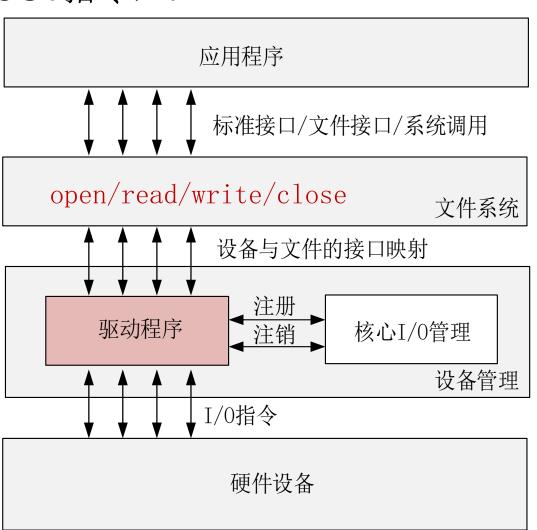
设备管理功能>I/O缓冲区管理

- I/O缓冲区管理
 - ■开辟和管理I/O缓冲区
 - ■提高读写效率



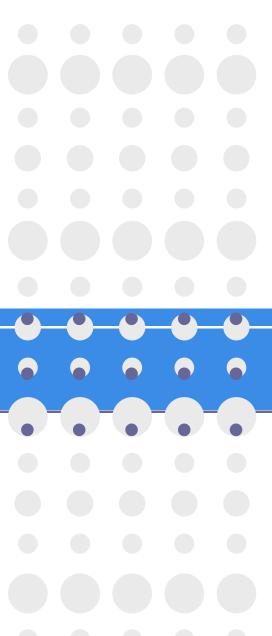
设备管理功能>设备驱动

- 设备驱动
 - ■对物理设备进行I/O操作(IN/OUT指令)。
 - ■把应用对设备的读/写请求 转换为对设备I/O操作。
 - ■应用读写请求采用文件接口
 - ◆open/read/write/close
 - ◆设备是文件。



设备管理功能>设备驱动

- 设备驱动程序的特点
 - ■设备驱动程序与硬件密切相关。
 - ■设备必须要配置驱动程序
 - ■驱动程序一般由设备厂商根据操作系统要求编写。

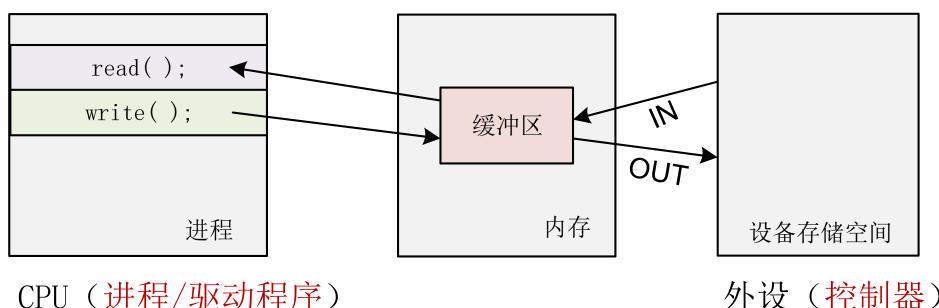


8.2 缓冲技术

缓冲技术

- 缓冲作用
 - ■1)连接不同数据传输速度的设备
 - ■2) 协调数据记录大小的不一致
 - ■3) 正确执行应用程序的语义拷贝

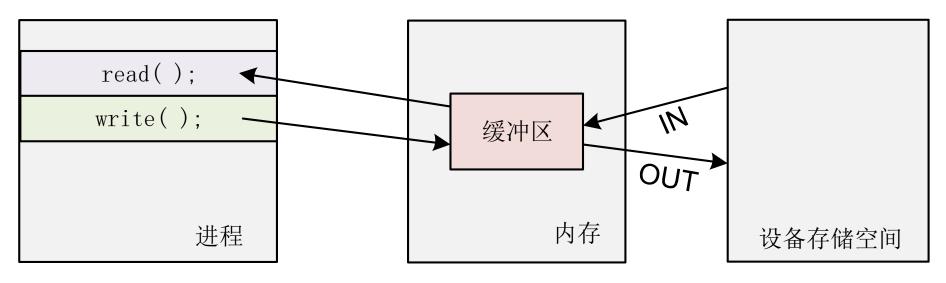
- 1)连接不同数据传输速度的设备
 - ■例子: CPU(设备驱动)与设备(控制器)之间传输数据
 - ■改进:内存中增加缓冲区



CPU(进程/驱动程序)

外设(控制器)

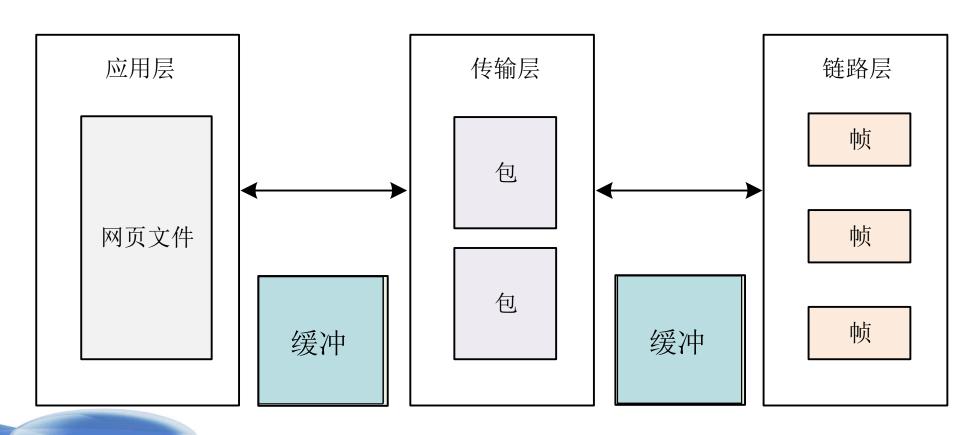
- 2) 协调数据记录大小的不一致
 - ■进程之间或CPU与设备之间的数据记录大小不一致
 - ■例: 进程(结构): 设备(字节)



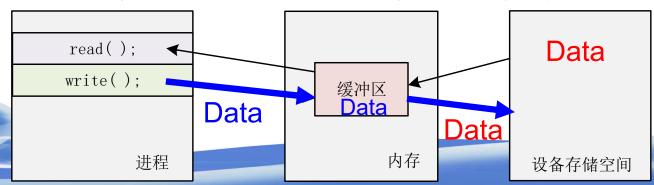
CPU(进程/驱动程序)

外设(控制器)

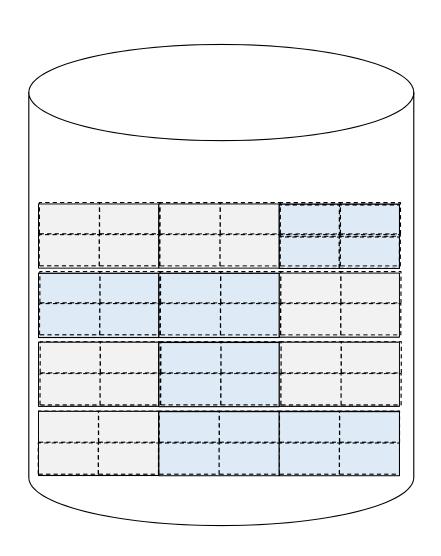
- 2)协调数据记录大小的不一致
 - ■进程之间或CPU与设备之间的数据记录大小不一致
 - ■例:不同网络层之间的数据记录:帧、包、文件



- 3)正确执行应用程序的语义拷贝
 - ■例子: 利用write(Data, Len)向磁盘写入数据Data
 - ◆确保写入磁盘的Data是调用时刻的Data
 - ■方法:
 - ◆方法1
 - □应用等待内核写完磁盘再返回。(实时性差)
 - ◆方法2
 - □应用仅等内核写完内存即返回
 - ▲事后由内核把<mark>缓冲区</mark>写到磁盘。(实时性好)
 - □语义拷贝:确保事后拷贝的数据是正确版本

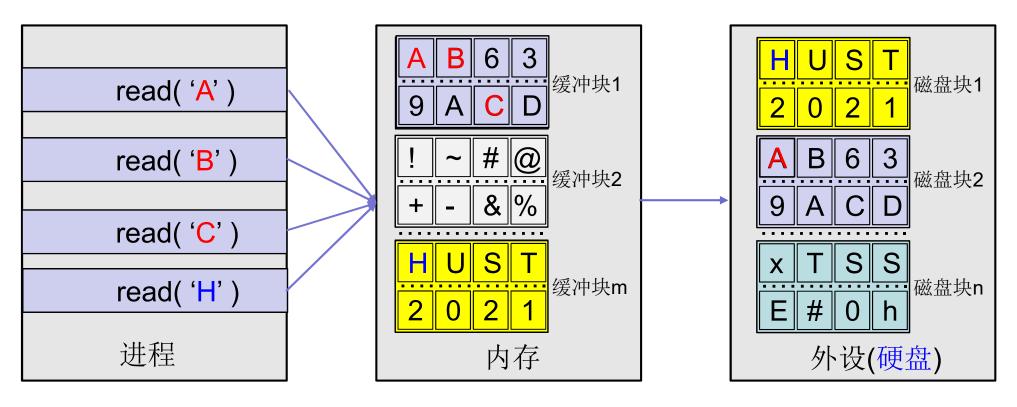


- ●典型的块设备
 - ■硬盘、软盘、RAM DISK等
 - ■块(block)和扇区
 - ◆硬盘读/写/寻址:扇区
 - ◆文件读/写/寻址:块
 - □块 = 2ⁿ ×扇区
 - □Linux块 = 1KB (n=1)
- Linux缓冲机制
 - ■内存开辟高速缓冲区
 - ■提前读/延后写



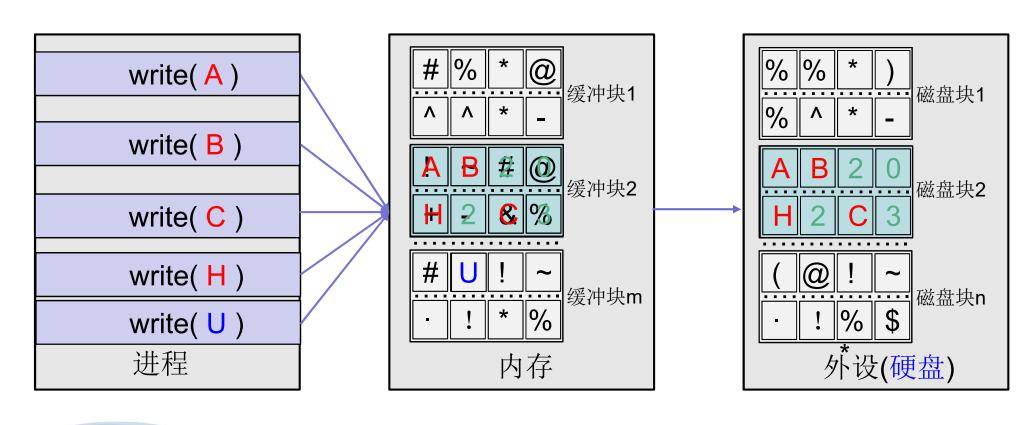
● 提前读

■进程读时,其所需数据已被<mark>提前读</mark>到了缓冲区中,不需要启动外设去执行读操作。

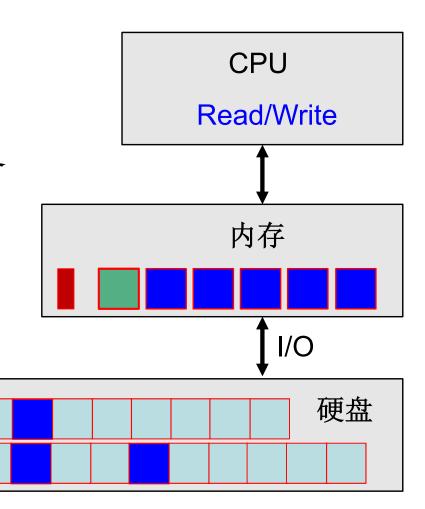


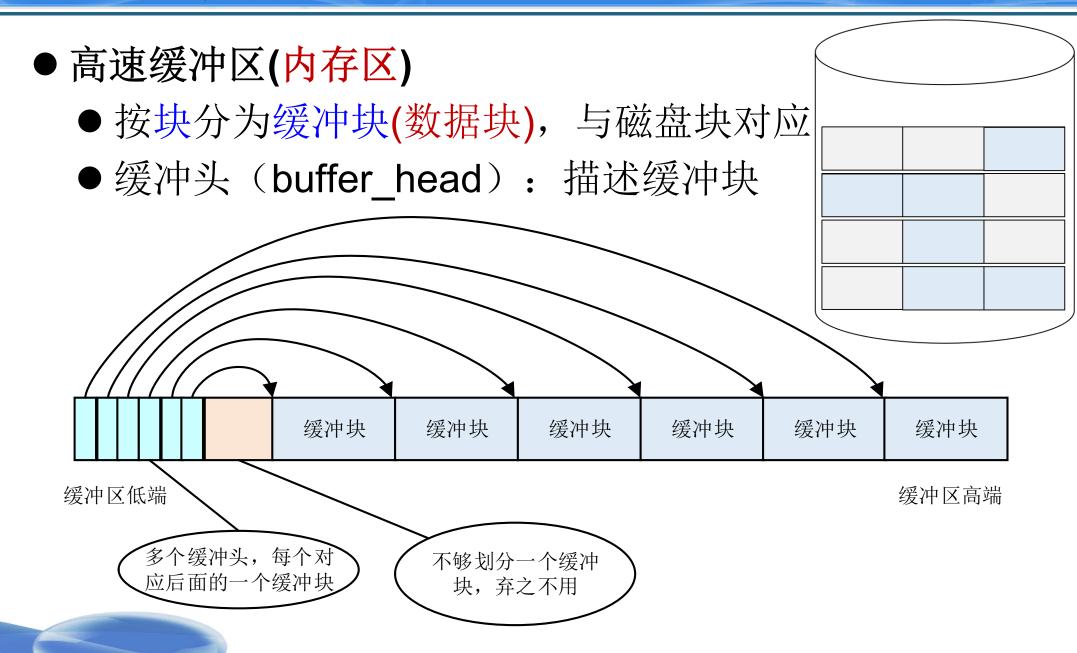
m << n

- 延后写
 - ■进程写时,数据先存在缓冲区,等到特定事件发生 或足够时间后(已延迟),再启动外设完成写入。



- 提前读与延后写
 - ■磁盘类的块设备
 - ■提高进程与外设数据传输效。
 - ■减少访问设备次数,提高设备 访问的效率。
 - ■内存开辟高速缓冲区



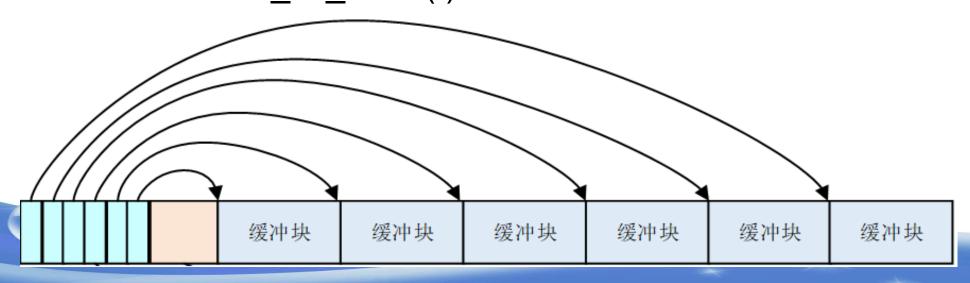


- 高速缓冲区(内存区)
 - 按块分为缓冲块(数据块),与磁盘块对应
 - 缓冲头(buffer_head):描述缓冲块

```
□struct buffer head {
        char * b data;
                      /* pointer to data bloc
        unsigned long b blocknr; /* block number */
        unsigned short b dev; /* device (0 = free) */
        unsigned char b uptodate;
        unsigned char b dirt; /* 0-clean,1-dirty */
        unsigned char b count; /* users using this block */
        unsigned char b lock; /* 0 - ok, 1 -locked */
        struct task struct * b wait;
        struct buffer head * b prev;
10
        struct buffer head * b next;
11
        struct buffer head * b prev free;
12
13
        struct buffer head * b next free;
14 };
```

- 高速缓冲区(内存区)
 - ■缓冲头buffer_head
 - ◆b_data: 指向缓冲块对应的数据区
 - ◆b_blocknr: 设备中的块号
 - ◆b_dev: 设备号
 - ◆b_lock:表示该缓冲块是否已被锁定
 - ◆b count:缓冲块被多少个进程引用
 - ◆b_dirt: 延迟写字段,即脏位字段
 - ◆b_uptodate: 数据有效位字段
 - ◆b wait: 指向访问缓冲块的等待队列

- 进程读写设备数据
 - 进程read/write →文件访问请求 → 块读取bread()
- 块读取函数bread(设备号,块号)
 - ■以(设备号, 块号)为索引搜索高速缓冲区, 查找对应的缓冲块
 - ◆若找到,直接读回
 - ◆若没有找到
 - □分配一个新缓冲块
 - □调用II rw block()读相应磁盘块到新分配的缓冲块



缓冲的组成

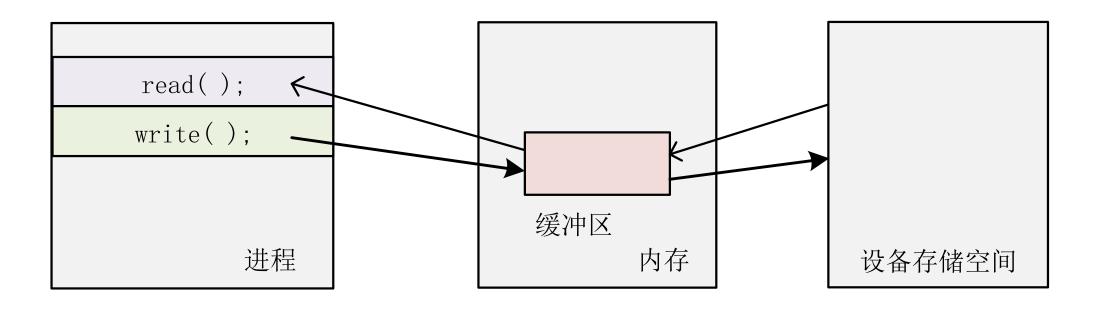
- 缓冲的组成形式
 - Cache
 - ◆高速缓冲寄存器【CPU ↔ 内存】
 - ■设备内部缓冲区
 - ◆外设或I/O接口的内部缓冲区【端口】
 - ■内存缓冲区
 - ◆应用广泛,使用灵活【CPU ↔ 接口/外设】
 - ◆应用开辟 | 内核开辟
 - ■辅存缓冲区
 - ◆开辟在辅存上【 暂存内存数据,SWAP】

缓冲的实现

- ●単缓冲
- ●双缓冲
- ●环形缓冲
- 缓冲池

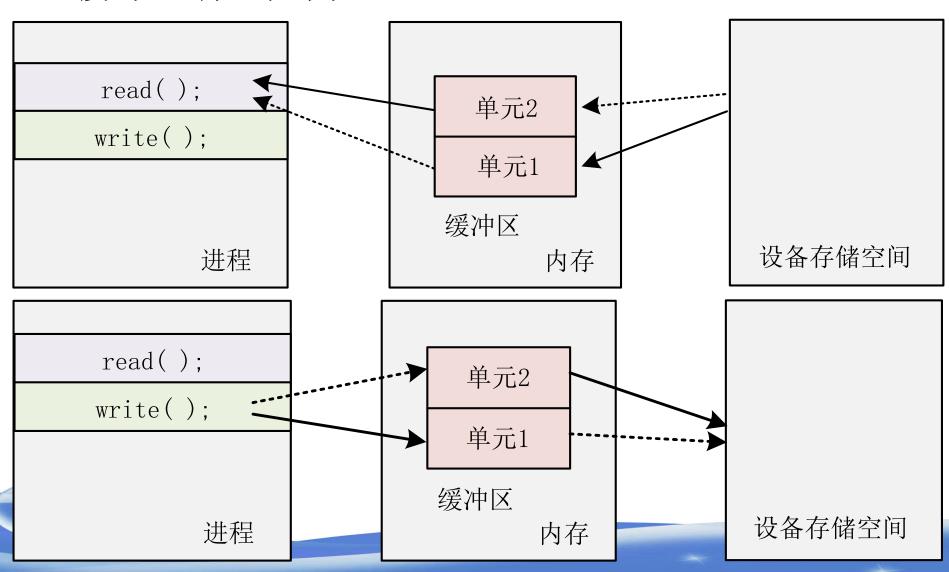
缓冲的实现>单缓冲

- 单缓冲
 - ■缓冲区仅有1个单元



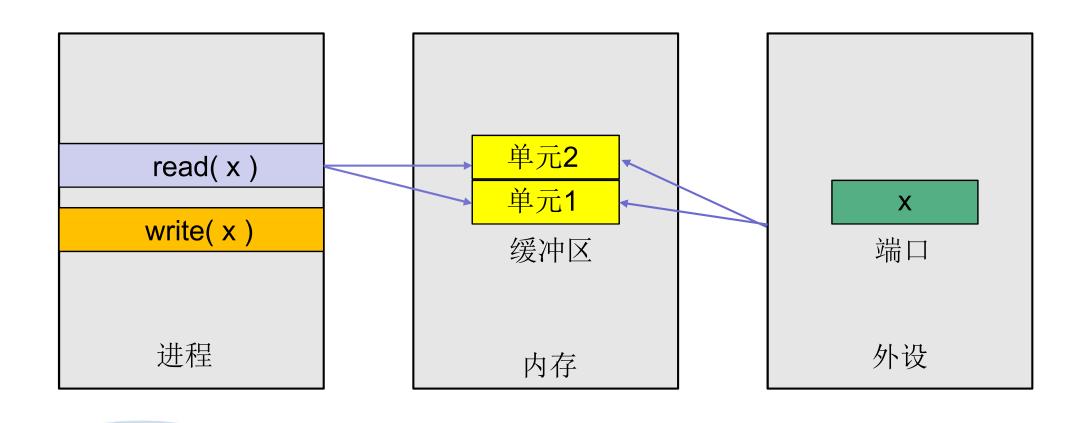
缓冲的实现>双缓冲

- ●双缓冲
 - ■缓冲区有2个单元



缓冲的实现>双缓冲

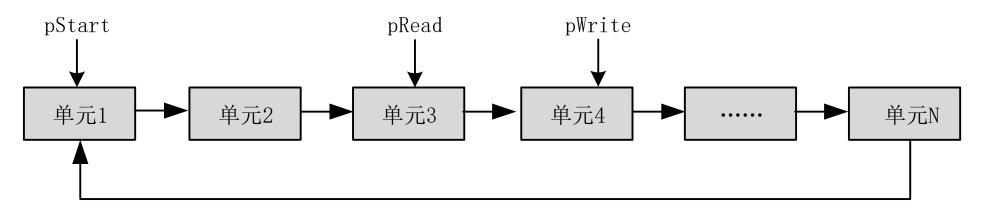
- ●双缓冲
 - ■缓冲区有2个单元



缓冲的实现>环形缓冲

●环形缓冲

■在双缓冲的基础上增加了更多的单元,并让首尾两个单元在逻辑上相连。

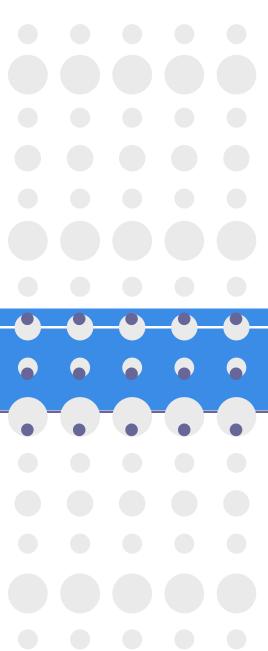


- ■起始指针pStart
- ■输入指针pWrite
- ■输出指针pRead

缓冲的实现>缓冲池

- 缓冲池
 - ■多个缓冲区
 - ■可供若干个进程共享
 - ■可以支持输入,也可以支持输出
 - ■提高缓冲区利用率,减少内存浪费





8.3 设备驱动程序

● 例:应用程序(读/写设备的状态)

```
main()
   ₽ {
        int fd, nDevState;
        //打开设备,设备文件名是: /dev/RWDevState
 4
 5
        fd = open("/dev/RWDevState", O RDWR);
 6789
        //读设备的状态DevState
        read(fd, &nDevState, sizeof(int));
        printf ("The Steate of Device is %d\n", nDevState);
        //更新设备的状态DevState
        nDevState = 100;
        write (fd, &nDevState, sizeof(int));
11
        //读设备的状态DevState
12
        read(fd, &nDevState, sizeof(int));
13
14
        printf ("The Steate of Device is %d\n", nDevState);
        //关闭设备
15
16
        close (fd);
```

● 例: 驱动程序(读/写设备的状态)

```
static int chr open (struct indoe * pinode, struct file *pfile)
   □{ // chr_open(): 打开设备 //打开文件 fd = open("/dev/RWDevState")
15
       MOD INC USE COUNT;
16
       return 0;
17
18
    static int chr read(struct file *pfile, char *buf, int len, int *off)
   □{ // chr_read(): 读设备的状态 //读文件 read(fd)
19
       //将 nDevState 从内核空间复制到用户空间
20
21
       copy to user (buf, &nDevState, sizeof(int));
22
       return 0;
23
    static int chr write(struct file *pfile, const char *buf, int len, int *off)
24
   ■{ // chr_write(): 写设备的状态 //写文件 write(fd)
       //将 nDevState 从用户空间复制到内核空间
26
       copy from user(&nDevState, buf, sizeof(int));
27
28
       return 0;
29
30
    static int chr release(struct indoe * pinode, struct file *pfile)
   □{ // chr_release(): 关闭设备 //关闭文件 close(fd)
31
32
       MOD DEC USE COUNT;
33
       return 0;
34
```

- ●例:驱动程序(读/写设备的状态)
 - ■定义设备操作接口与文件操作接口之间的映射

```
static const struct file operations
                                         MyFops
  ₽ {
3
        .read
                    = chr read,
       .write = chr write,
5
       .release = chr release,
       .open = chr open,
        .unlocked ioctl = chr ioctl,
   };
  pstruct file operations { // 文件操作结构体 (POSIX)
2
       struct module *owner;
       int (*llseek) (struct file *, int, int);
       int (*read) (struct file *, char user *, int, loff_t *);
4
       int (*write) (struct file *, char user *, int, loff t *);
5
       int (*poll) (struct file *, struct poll table struct *);
6
       int (*ioctl) (struct inode *, struct file *, int, long);
       int (*mmap) (struct file *, struct vm area struct *);
8
9
       int (*open) (struct inode *, struct file *);
       int (*flush) (struct file *, fl owner t id);
10
       int (*release) (struct inode *, struct file *);
13
```

- 例: 驱动程序(读/写设备的状态)
 - ■实现设备的注册函数和注销函数

```
//定义注册函数
35
   static int DevInit (void)
36
37
   ₽ {
38
       int ret;
39
       ret = register chrdev (MAJOR NUM, "RWDevState", &MyFops);
40
       printk ("RWDevState register success\n");
41
       return ret:
42
    //定义注销函数
43
44
    static void DevExit (void)
45
   □ {
46
       int ret;
47
       ret = unregister chrdev (MAJOR NUM, "RWDevState");
48
       printk ("RWDevState unregister success\n");
49
50
    module init (DevInit);
51
    module exit (DevExit);
```

- 例: 驱动程序(读/写设备的状态)
 - ■设备注册
 - ◆将用户定义的设备加入到系统的设备数组

```
//定义注册函数
35
36
    static int DevInit (void)
37
   ₽ {
38
       int ret;
39
        ret = register chrdev (MAJOR NUM, "RWDevState", &MyFops);
        printk("RWDevState register success\n");
40
41
        return ret;
42
      int register chrdev(
         int major,
         char *name,
         struct file operations *fops);
```

- 例: 驱动程序(读/写设备的状态)
 - ■设备注销
 - ◆释放设备,将设备从系统的设备数组删除

```
//定义注销函数
static void DevExit(void)

{
    int ret;
    ret = unregister_chrdev(MAJOR_NUM, "RWDevState");
    printk("RWDevState unregister success\n");
}

□unregister_chrdev()
□unregister_blkdev()
```

- 例: 驱动程序(读/写设备的状态)
 - ■实现设备的注册函数和注销函数

```
//定义注册函数
35
    static int DevInit (void)
36
37
   ₽{
38
       int ret;
39
       ret = register chrdev (MAJOR NUM, "RWDevState", &MyFops);
40
       printk ("RWDevState register success\n");
41
       return ret:
42
    //定义注销函数
43
44
    static void DevExit (void)
45
   □ {
46
       int ret;
47
       ret = unregister chrdev (MAJOR NUM, "RWDevState");
48
       printk("RWDevState unregister success\n");
49
    - }
50
    module init (DevInit);
51
    module exit (DevExit);
```

● 编译驱动程序

```
//Makefile
bj-m += RWDevState.ko
all:
make -C /lib/modules/$(shell uname -r)/build M=$(PWD) modules
clean:
make -C /lib/modules/$(shell uname -r)/build M=$(PWD) clean
```

● 安装/删除驱动程序

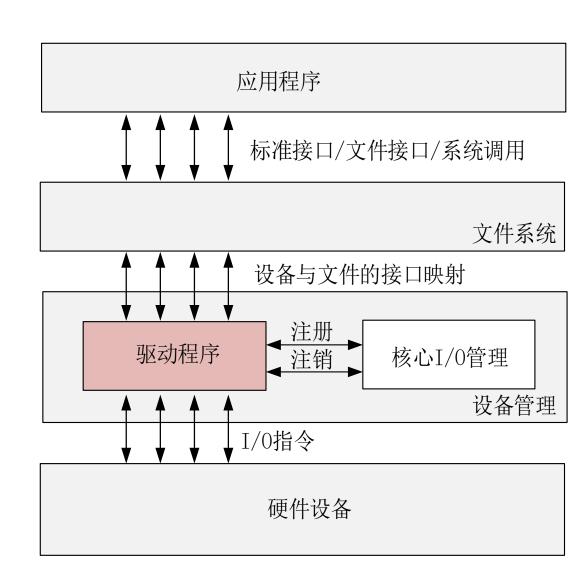
```
# insmod RWDevState.ko # rmmod RWDevState
```

● 创建设备文件

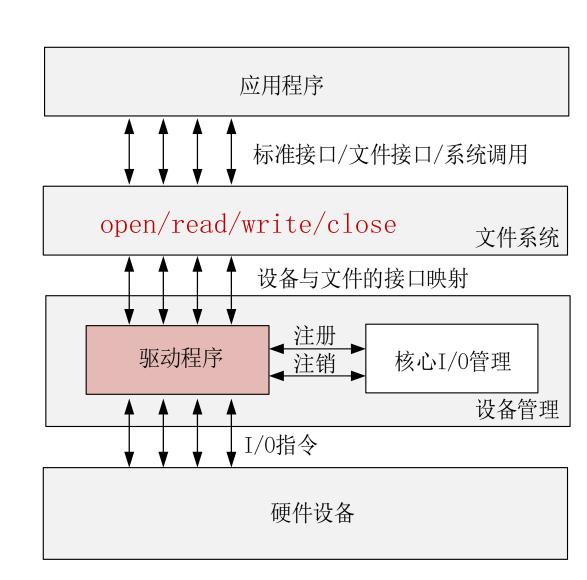
```
mknod /dev/RWDevState c 252 0
```

```
//打开设备,设备文件名是: /dev/RWDevState fd = open("/dev/RWDevState", O RDWR);
```

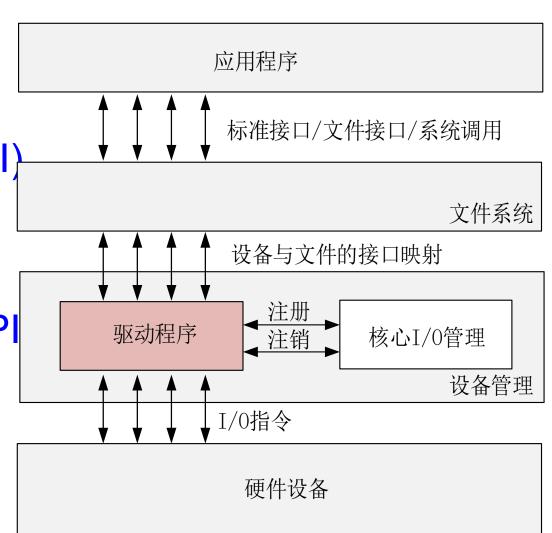
- 驱动程序的基本接口
 - ■面向用户程序的接口
 - ■面向I/O管理器的接口
 - ■面向设备的接口



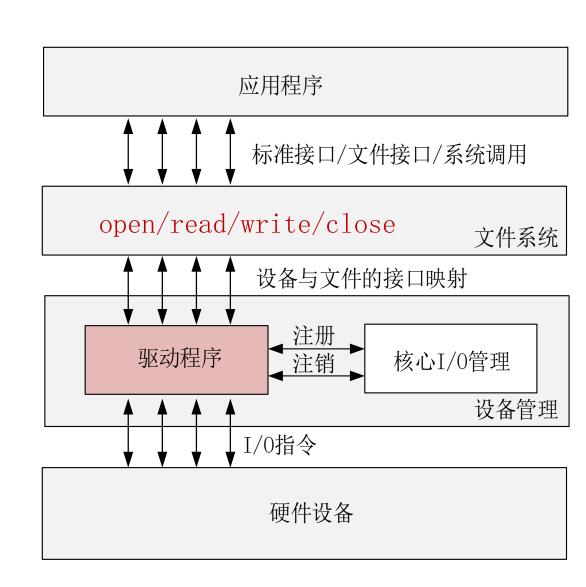
- ●面向用户程序的接口
 - ■设备的打开与释放
 - ■设备的读写操作
 - ■设备的控制操作
 - ■设备的中断处理
 - ■设备的轮询处理



- 面向I/O管理器的接口
 - ■注册函数
 - ◆insmod(命令)
 - □module_init() (API
 - ■注销函数
 - ◆rmmod (命令)
 - ■module_exit() (API
 - ■必需的数据结构



- ●面向设备的接口
 - ■实现设备的端口操作
 - ◆无条件传送
 - ◆查询传送
 - ◆中断传送
 - ◆DMA传送
 - 例:
 - ◆打开设备 chr_open()
 - ◆读设备 chr_read()
 - ◆写设备 chr_write()
 - ◆关闭设备chr_release()



Linux设备的分类

- 字符设备
 - ■传输的基本单位是字符。例:键盘、串口。
- 块设备
 - ■传输的基本单位是块。例: 硬盘, 磁盘。
- 网络设备
 - ■采用socket套接字接口访问
 - ■在全局空间有唯一名字,如eth0、eth1。

驱动程序工作在核态

- ●用户态与内核态
 - ■驱动程序工作在内核态
 - ■应用程序和驱动程序之间传送数据
 - ◆get_user()
 - put_user()
 - copy_from_user()
 - copy_to_user()

设备文件

- 设备文件
 - ■硬件设备作为文件看待
 - ■使用文件操作接口来完成设备的打开、关闭、读写和I/O控制等操作。
 - ■仅字符设备和块设备通过设备文件访问
 - ◆创建设备文件: mknod

设备文件(Is -I /dev)

```
susq : bash
                 Bookmarks
      Edit
          View
                           Settings
                                   Help
[susg@localhost ~]$ ls -l /dev/
total 0
                                          1 17:06 autofs
crw-----. 1 root root
                            10, 235 5月
drwxr-xr-x. 2 root root
                                280 5月
                                          1 17:06 block
                                 80 5月
drwxr-xr-x. 2 root root
                                             2017 bsg
crw-----. 1 root root
                           10, 234 5月
                                         1 17:06 ptrf5-qentrol
drwxr-xr-x 63 root root
                                 60 5月 2
                                            26 17 Jus
1rwxrwxrwx. 1 root root
                                            17:06 cdrom -> sr0
                               3400 5月
                                          1 17:06 char
drwxr-xr-x. 2 root root
                                  1.5月
                                            17:06 console
crw-----. 1 root root
                                             2017 core - proc/kcore
                                 11 5月
lrwxrwxrwx. 1 root root
drwxr-xr-x, 6 root root
                                140 5月
                                         1 17:06 cpu_dma_latency
            1 routing
                            10, 62 5月
                            10, 203 5月
crw----- 1 root root
                                          1 17:06 cuse
drwxr-xr-x. 4 root root
                                 80 5月
                                             2017 disk
                                  0 5月
brw-rw----. 1 root disk
                                          1 17:06 dm-0
                           253,
                                 1 5月
brw-rw----. 1 root disk
                           253,
                                          1 17:06 dm-1
brw-rw----. 1 root disk
                           253,
                                  2 5月
                                          1 17:06 dm-2
                           253.
                                 3 5月
                                          1 17:06 dm-3
brw-rw----. 1 root disk
                                100 5月
                                          2 2017 dri
drwxr-xr-x, 2 root root
                            29, 05月
                                         1 17:06 fb0
crw-rw----. 1 root video
                                          2 2017 fd -> /proc/self/fd
lrwxrwxrwx, 1 root root
                                 13 5月
                   susg: bash
```

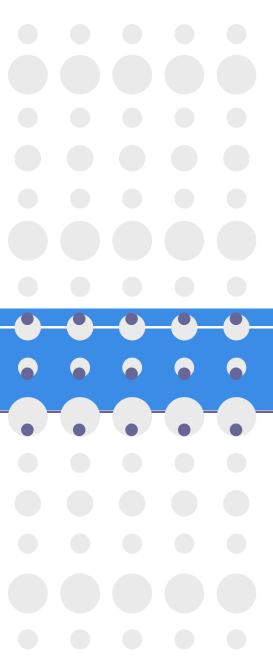
设备文件

- 主设备号和次设备号
 - ■主设备号
 - ◆标识该设备种类,标识驱动程序
 - ◆主设备号的范围: 1-255
 - ◆Linux内核支持动态分配主设备号
 - ■次设备号
 - ◆标识同一设备驱动程序的不同硬件设备

Linux 2.6之后的内核

● 驱动注册过程发生变化

```
//V2.4 字符设备注册
Led_Major = register_chrdev(0, DEVICE_NAME, &my_fops);
//V2.6字符设备注册
cdev_add(&cdev, MKDEV(major_No,0), DEV_NR);
```



8.5 设备分配

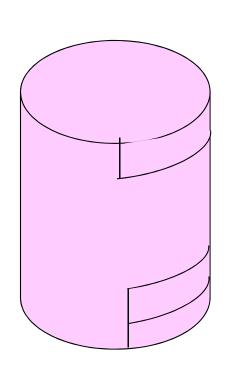
设备分类

- 独占设备
 - ■不可抢占设备(普通外设或资源)
 - ◆使用时独占,释放后才能被其它进程申请到。
 - ◆先申请,后使用(主动)
- 共享设备
 - ■可抢占设备(CPU,内存,硬盘)
 - ◆允许多个作业或进程同时使用。
 - ◆不申请,直接用(被动+主动)
- ●虚拟设备
 - ■借助虚拟技术,在共享设备上模拟独占设备。

- 独享分配
- ●共享分配
- ●虚拟分配

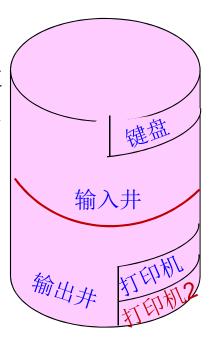
- 独享分配
 - ■针对独占设备
 - ■流程:申请→占用→释放
 - ◆指进程使用设备之前先申请,申请成功开始使用,直到使用完再释放。
 - ■若设备已经被占用,则进程会被阻塞,被挂入设备对应的等待队列等待设备可用之时被唤醒。

- 共享分配
 - ■针对共享设备
 - ◆典型共享设备: 硬盘
 - ■当进程申请使用共享设备时,操作系统能 立即为其分配共享设备的一块空间(空分方式),不让进程产生阻塞。
 - ■共享分配随时申请,随时可得。



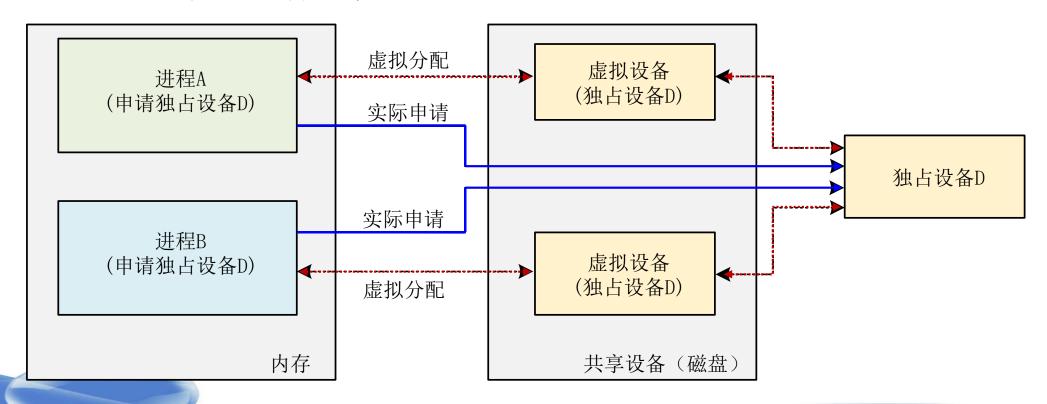
●虚拟分配

- ■虚拟技术
 - ◆在一类物理设备上模拟另一类物理设备的技术
 - ◆通常借助辅存部分区域模拟独占设备,将独占设备转化为共享设备。
- ■虚拟设备
 - ◆用来模拟独占设备的**辅存区域**称为虚拟设备 □具有独占设备的逻辑特点
 - ◆输入井: 模拟输入设备的辅存区域
 - ◆输出井: 模拟输出设备的辅存区域

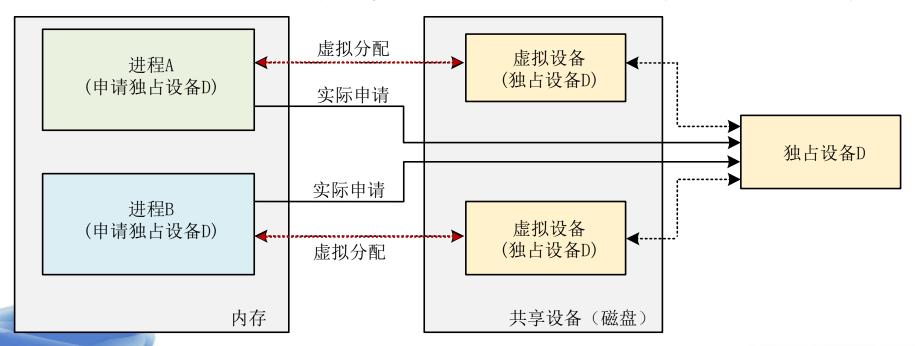


●虚拟分配

- ■当进程申请独占设备时将对应虚拟设备分配给它。
 - ◆首先,采用共享分配为进程分配虚拟设备;
 - ◆其次,将虚拟设备与对应的独占设备关联。



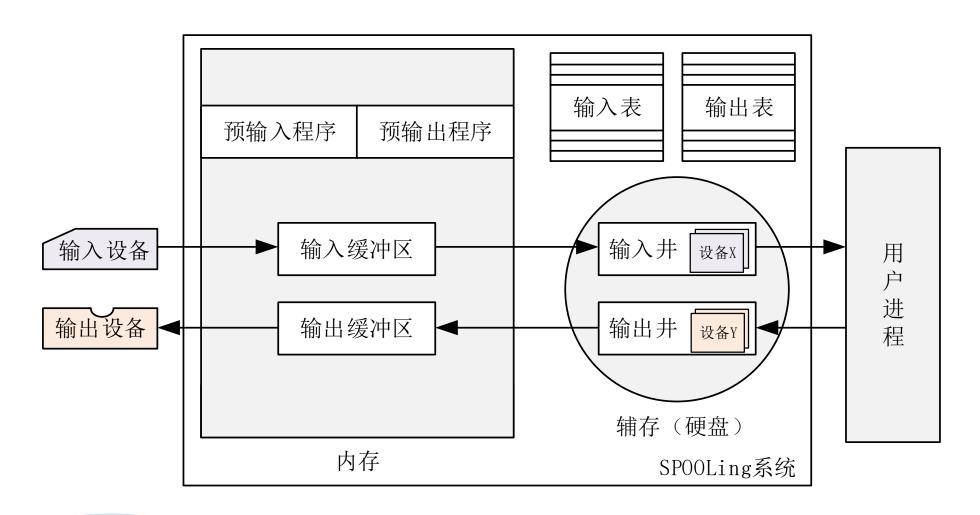
- ●虚拟分配
 - ■当进程申请独占设备时将对应虚拟设备分配给它。
 - ◆首先,采用共享分配为进程分配虚拟设备;
 - ◆其次,将虚拟设备与对应的独占设备关联。
 - ■进程运行中仅与虚拟设备交互,提高了运行效率



- ●虚拟分配
 - ■当进程申请独占设备时将对应虚拟设备分配给它。
 - ■例: SPOOLing系统
 - ◆Simultaneaus Periphernal Operations OnLine
 - ◆SPOOLing是虚拟技术和虚拟分配的实现
 - ◆外部设备同时联机操作 | 假脱机输入/输出

虚拟分配

● SPOOLing系统的结构



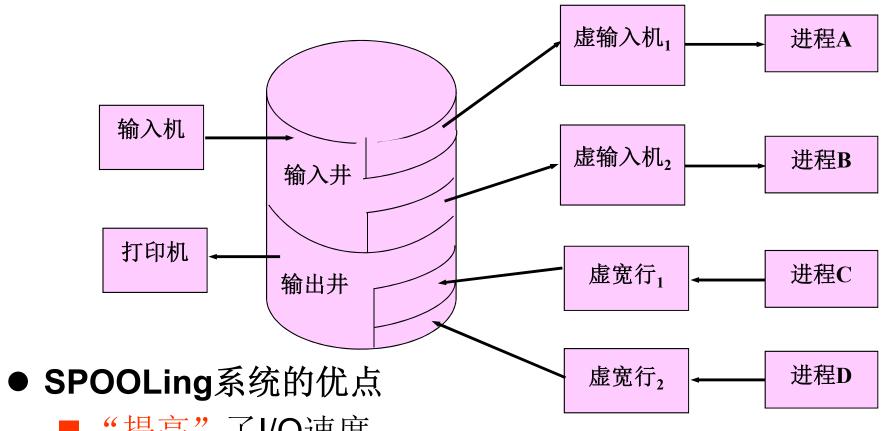
SPOOLing系统的结构 (硬件)

- 输入井和输出井
 - ■磁盘上开辟的两个存储区域
 - ◆输入井模拟脱机输入时的磁盘
 - ◆输出井模拟脱机输出时的磁盘
- 输入缓冲区和输出缓冲区
 - ■内存中开辟的存储区域
 - ◆输入缓冲区: 暂存输入数据,以后再传送到输入井。
 - ◆输出缓冲区: 暂存输出数据,以后再传送到输出设备。

SPOOLing系统的结构(软件)

- 预输入程序
 - 控制信息从独占设备输入到辅存,模拟脱机输入的卫星机;
- 输入表
 - 独占设备↔虚拟设备
- 缓输出程序
 - 控制信息从辅存输出到独占设备,模拟脱机输出的卫星机;
- 输出表
 - ■独占设备↔虚拟设备
- 井管理程序
 - 控制用户程序和辅存之间的信息交换

SPOOLing的例子



- "提高"了I/O速度
- ■将独占设备改造为"共享"设备
 - ◆实现了虚拟设备功能