

华中科技大学

## 操作系统原理

20级考试复习版本，以本课件为准

## 第 09 章 文件系统

教师：苏曙光

华中科技大学软件学院

2021年10月-2021年12月

---

## ● 内容

- ◆ 文件和文件系统的概念
- ◆ 文件的逻辑结构
- ◆ 文件的物理结构
- ◆ 文件存储空间管理
- ◆ 文件目录
- ◆ 文件和目录的操作

## ● 重点

- ◆ 文件逻辑结构
- ◆ 文件物理结构
- ◆ 文件分配表

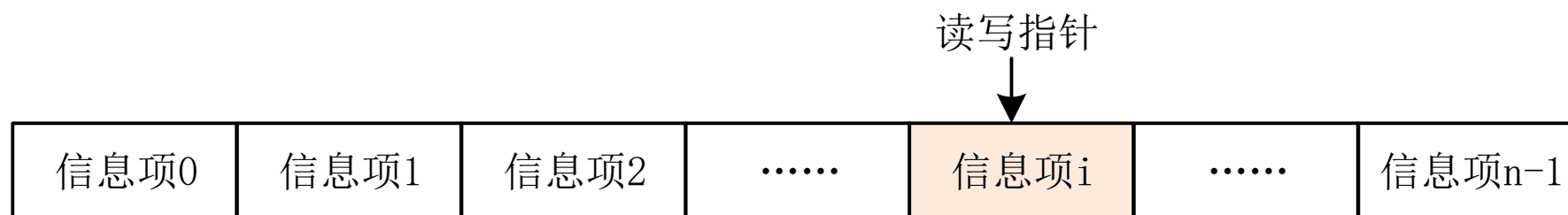
## 9.1 文件和文件系统概念

---

- 定义
- 分类
- 属性
- 文件系统

# 文件的定义

- 文件是系统中信息存放的一种组织形式
  - ◆ 文件是若干信息项的构成。
    - 信息项可以是字节，可以是结构化数据。



- ◆ 用户通过读写指针来存取文件的信息项。
- ◆ 文件具有文件名。用户通过文件名存取文件。

# 文件分类

---

- 分类标准1：文件的用途

- ◆ 系统文件

- ◆ 包括操作系统的可执行程序和数据文件。这种文件不对用户开放，仅供系统使用。

- ◆ 库文件

- ◆ 系统为用户提供的各种标准函数库和实用程序等。用户只能使用这些文件，而无权对其进行修改。

- ◆ 用户文件

- ◆ 用户创建的文件，如用户可执行程序，源程序，数据文件等。这种文件的使用和修改权均属于用户。

# 文件分类

---

- 分类标准2：文件的操作权限

- ◆ 只读文件

- 只允许进行读操作。

- ◆ 读写文件

- 允许进行读写操作。

- ◆ 不保护文件

- 不作任何操作限制。

# 文件分类

---

## ● 分类标准3：文件的性质

### ◆ 普通文件

- 指一般的用户文件和/或系统文件。

### ◆ 目录文件

- 由目录项组成的文件。
- 目录项: 文件名, 文件属性, 文件存放地址, ...

### ◆ 设备文件

- 把设备作为文件管理和使用

# 文件分类

---

- 按文件的存储时间
  - ◆ 永久文件
  - ◆ 临时文件



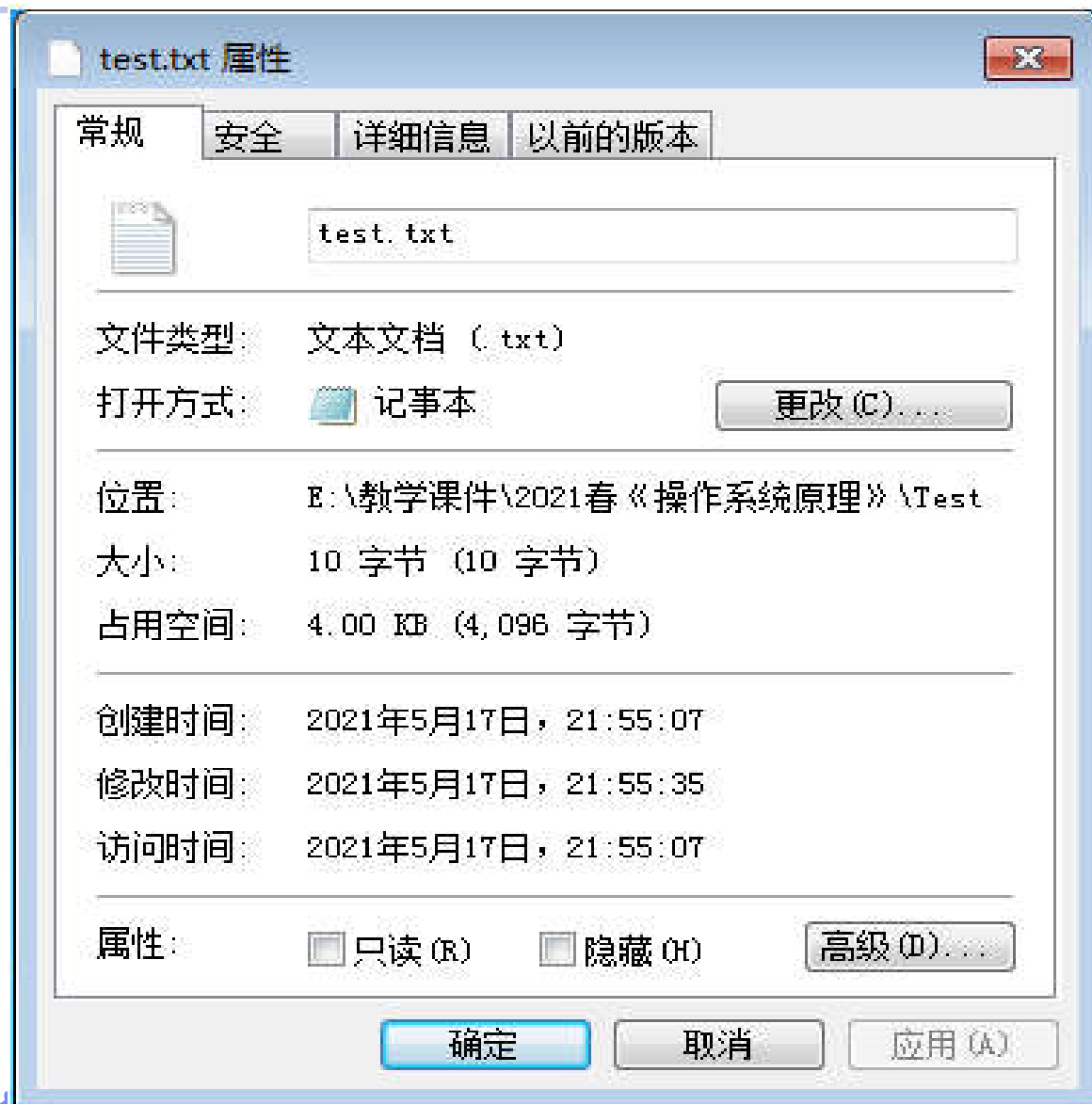
# 文件属性

## ● 文件属性

- ◆ 指定文件的类型、操作特性和存取保护等一组信息。
- ◆ 文件的属性一般存放在文件的（目录/文件）中。
- MS-DOS系统中，文件属性占目录项的一个字节。
  - 00000001表示文件仅读，00000010表示隐含文件等。



# 文件属性



# 文件属性

---

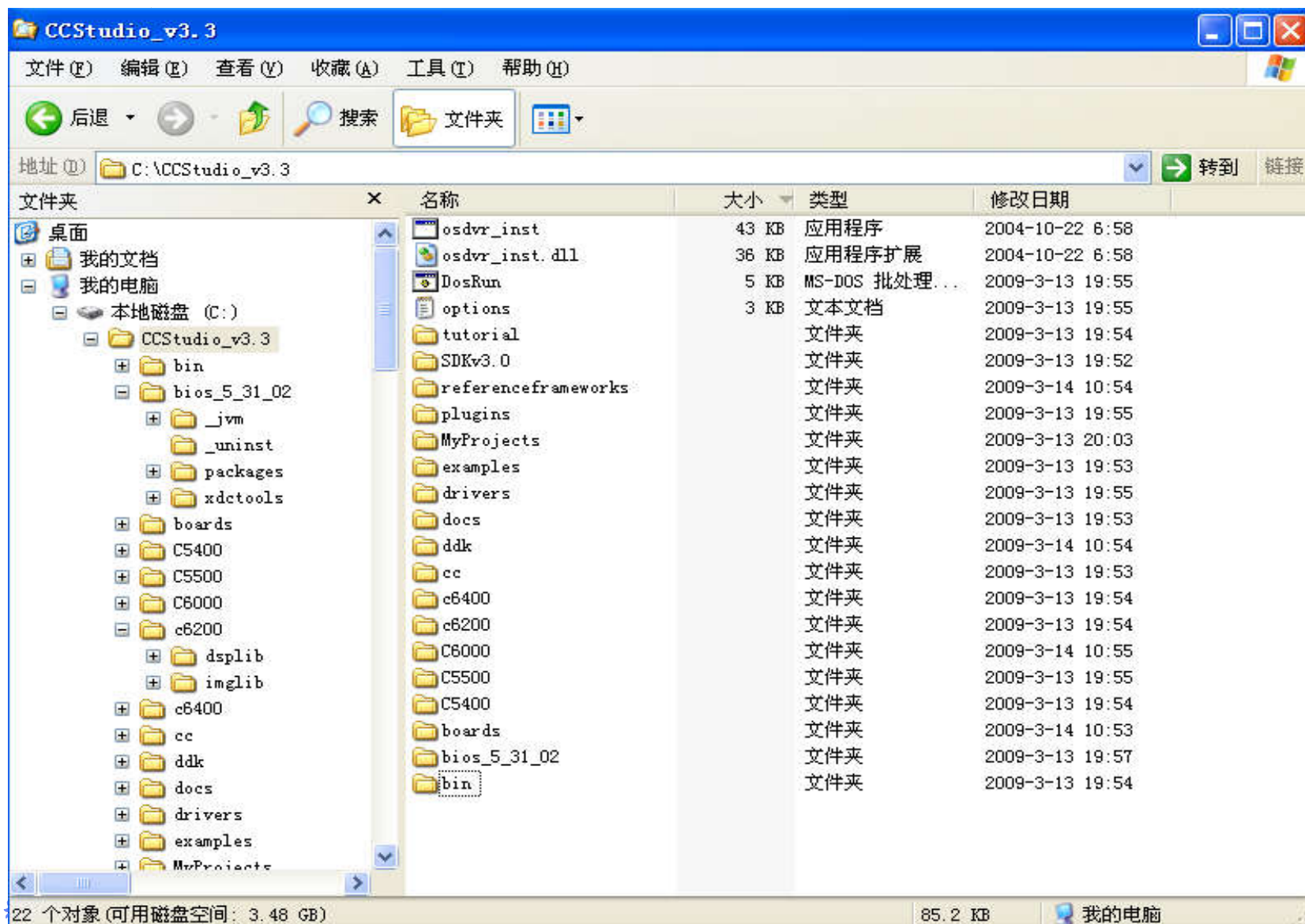
- **Linux**文件属性

- **ls -l**命令

- ◆ 类型和权限属性

- 文件类型：第0位
- 属主权限：第1-3位
- 属组权限：第4-6位
- 其他用户权限：第7-9位

# 文件系统



## 9.2 文件的逻辑结构和存取方法

- 文件的结构

- ◆ 逻辑结构

- (用户的观点)

- 为用户提供逻辑结构清晰、使用方便的文件。

- 强调文件信息项的构成方式和用户的存取方式。

- ◆ 物理结构

- (系统的观点)

- 文件在存储设备（例：硬盘）上的存储结构

- 强调合理利用储存空间，缩短I/O存取时间。

# 文件的逻辑结构

- 流式文件

- ◆ 信息项是字节，文件长度就是字节的数量。

- ◆ 优点

- 文件无需额外的说明信息或控制信息
    - 节省存储空间

- 记录式文件

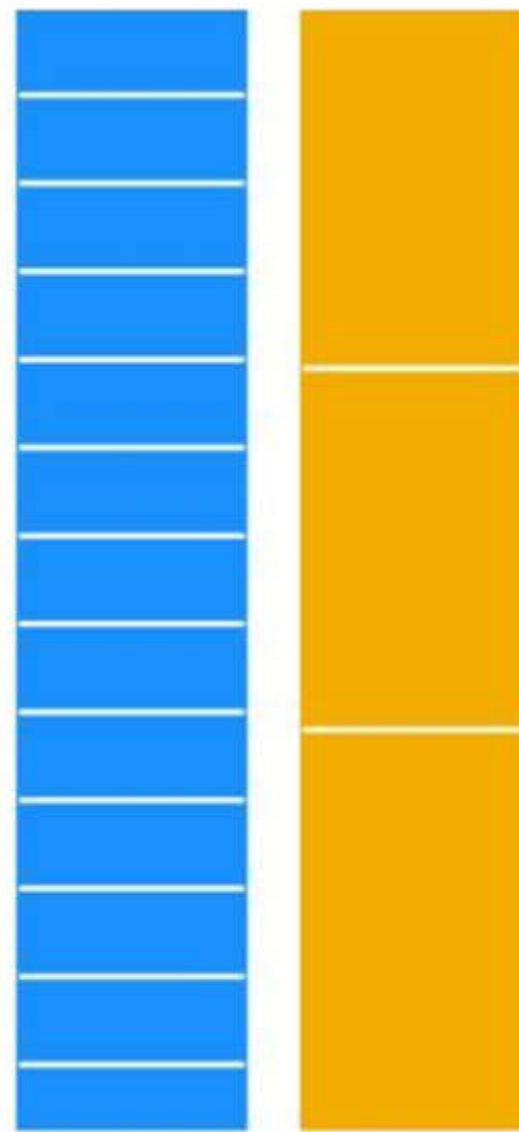
- ◆ 信息项由记录组成，一个记录包含若干成员。

- 学生记录：姓名，学号，性别，成绩
    - 学生花名册文件：包含若干个学生记录

- ◆ 特点

- 文件中需保存记录长度和数量等说明信息
    - 浪费存储空间

- 现代OS中文件是流式文件，由应用解释文件。



流式文件    记录文件



---

- 文件的存取方法

- ◆ 顺序存取

- ◆ 随机存取

---

## ● 顺序存取

◆ 按文件信息单位排列的顺序依次存取。

### □ 读写指针

□ 当打开文件时，文件的读写指针指向第1个信息单位(字节或记录)，每存取1个信息单位后读写指针自动加1而指向下一个信息单位。



## ● 随机存取

### ◆ 概念

- 直接存取

- 每次存取操作时先确定存取位置。

### ◆ 特点

- 对流式或定长记录文件比较容易确定存取位置。

- 对不定长的记录式文件比较麻烦

  - 从首记录顺序查询，直到找到目的记录为止。

  - 建立索引

    - 索引可作为文件一部分也可单独建索引文件。

## 文件读写示例(利用读写指针)

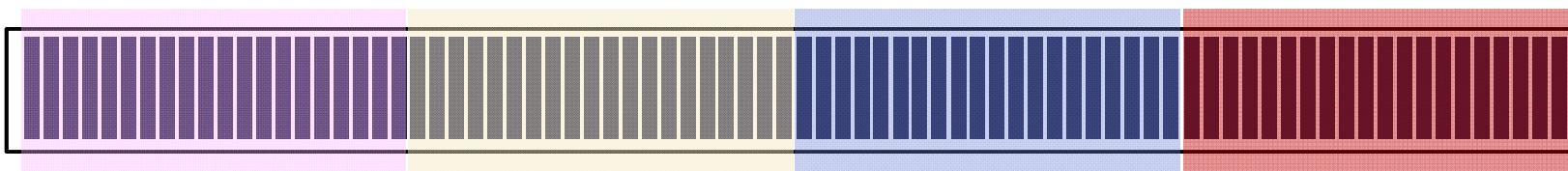
```
1 //打开文件
2 FILE *pFile=fopen("MyTestFile.txt","rb");
3 char *pBuf;
4 //移动文件指针到文件末尾
5 fseek(pFile,0,SEEK_END);
6 //获取文件指针的偏移量
7 int len=ftell(pFile);
8 pBuf=new char[len];
9 //将指针移动到文件头
10 rewind(pFile);// = fseek(pFile,0,SEEK_SET);
11 //读文件的内容
12 fread(pBuf,1,len,pFile);
13 fclose(pFile);
```



### 3 文件的物理结构

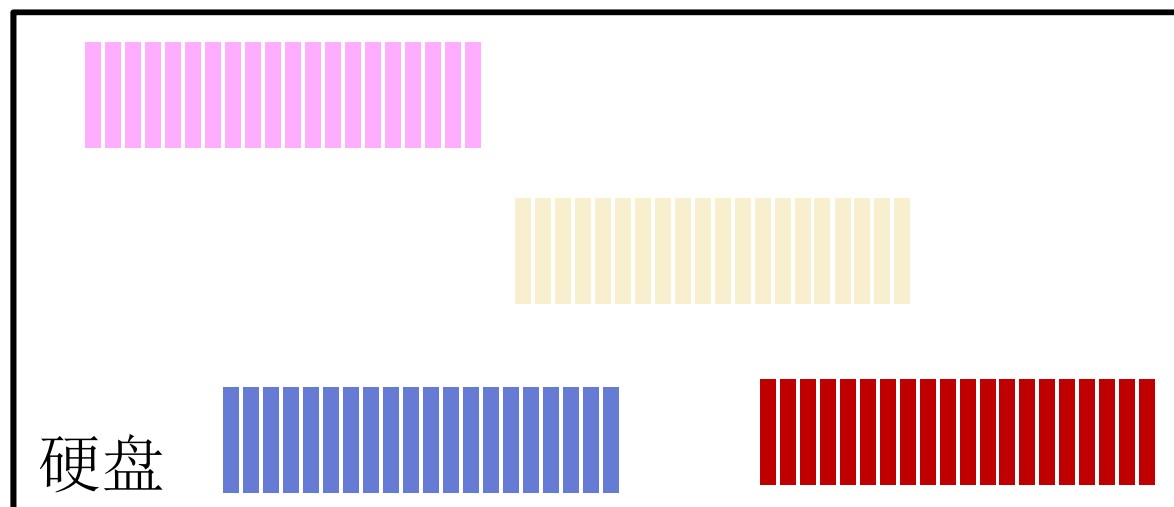
#### ● 概念

- ◆ 文件的物理结构是指文件在存储设备上（例如硬盘）的存储方式。
- ◆ 强调合理利用储存空间，缩短I/O存取时间。



#### ● 类型

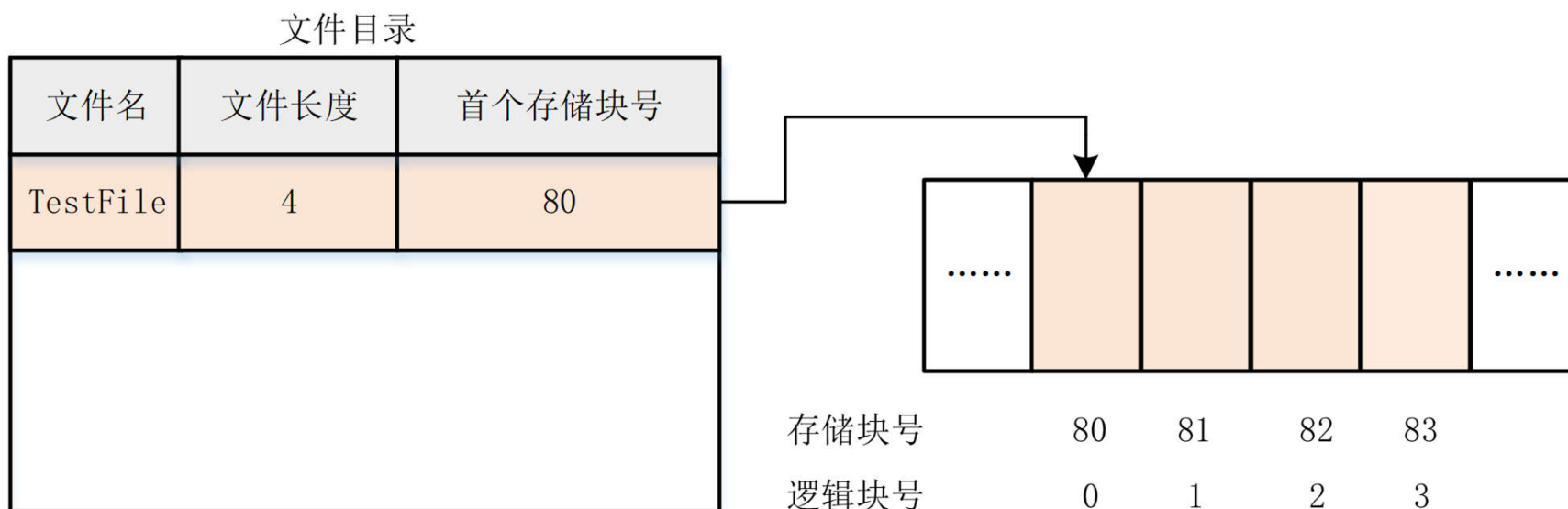
- ◆ 连续文件
- ◆ 串联文件
- ◆ 索引结构



# 连续文件

## ● 概念

- ◆ 文件按逻辑顺序存放在存储设备的连续物理块中。
- ◆ 文件目录
  - 记录文件长度（块数）和首个物理块号



# 连续文件

- 特点

- ◆ 文件建立时给出文件最大长度并登记文件起始位置。

- 优点

- ◆ 简单
- ◆ 支持顺序存取和随机存取
- ◆ 顺序存取速度快：所需磁盘寻道次数和寻道时间最少

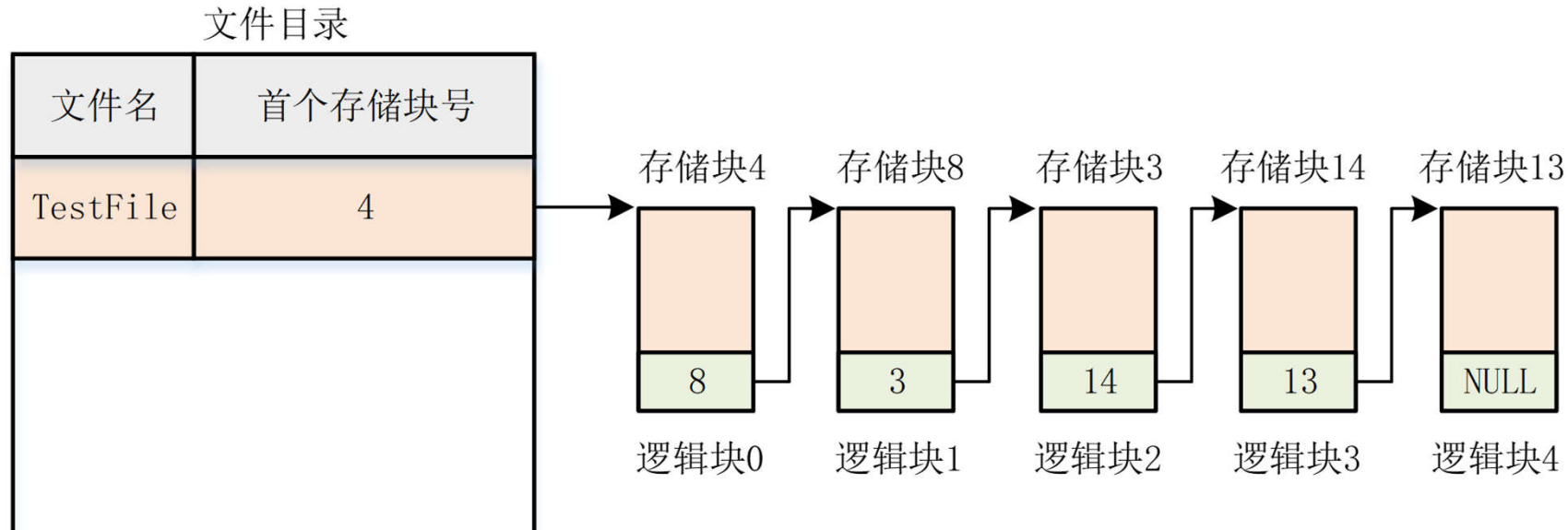
- 缺点

- ◆ 文件不易动态增长
  - 预留空间：浪费
  - 重新分配和移动
- ◆ 不利于文件插入和删除
- ◆ 外部碎片问题

# 串联文件

- 文件信息存放在不连续的存储块中
  - ◆ 每个存储块有一个指针(**next**)，指向文件下一个逻辑块所在的存储块。
  - ◆ 文件目录
    - 文件名+存储指针（指向第一个存储块）

文件目录	
FileA	0002
FileB	0005
TestFile	0004



## ● 特点

- ◆ 串联文件适用于顺序存取。随机存取较为困难。

## ● 优点

- ◆ 提高了磁盘空间利用率，不存在外部碎片问题
- ◆ 有利于文件插入和删除
- ◆ 有利于文件动态扩充

## ● 缺点

- ◆ 存取速度慢，不适于随机存取
- ◆ 可靠性问题：例如某个next链接指针出错
- ◆ next链接指针占用一定的空间

## ● 串联文件的应用

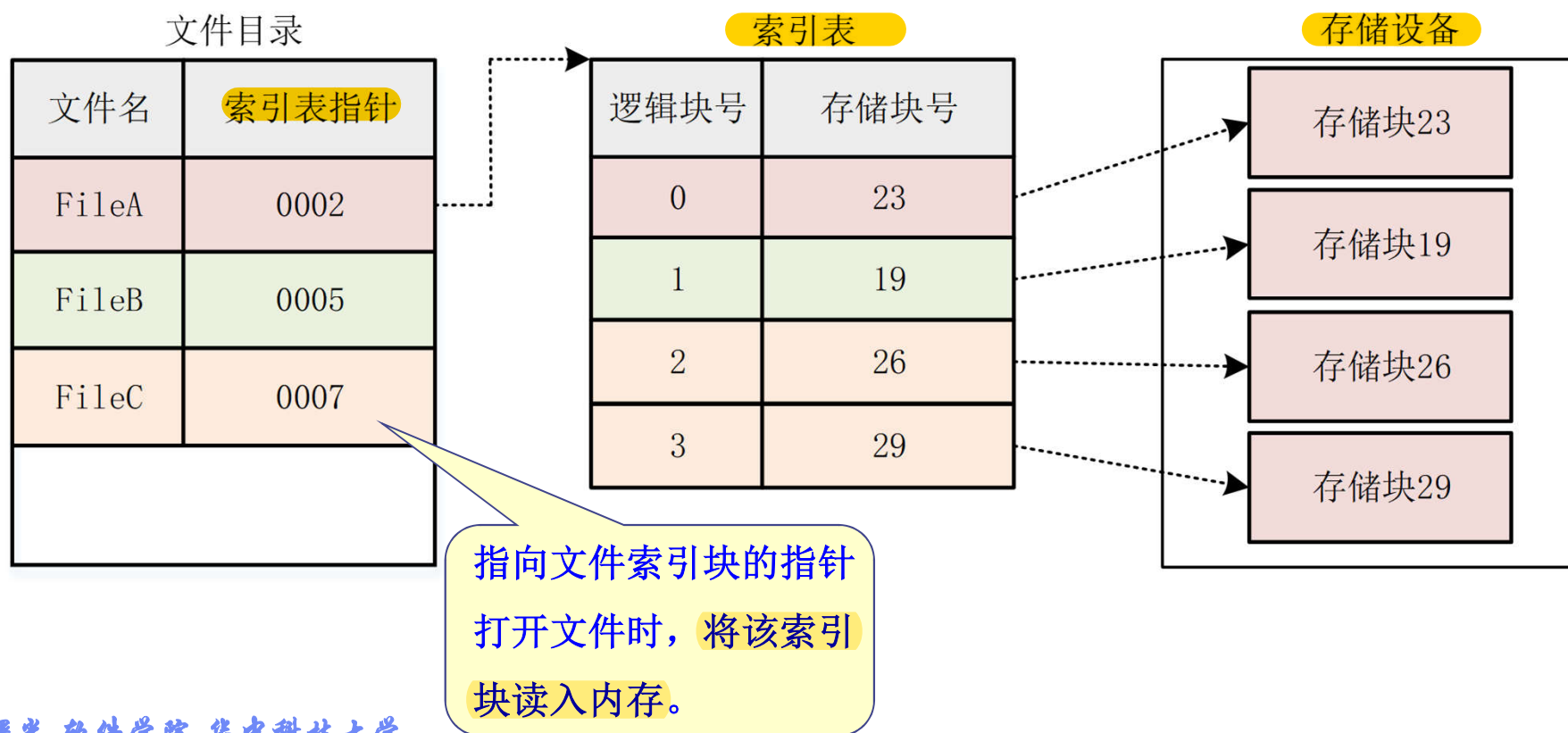
- ◆ FAT文件系统
  - File Allocation Table



# 索引文件

## ● 概念

- ◆ 文件存放在不连续的物理块中，系统建立索引表记录文件的逻辑块和存储块的对应关系。





# 索引文件

## ● 概念

- ◆ 文件存放在不连续的物理块中，系统建立索引表记录文件的逻辑块和存储块的对应关系。
- ◆ 索引类型的文件 = 索引表 + 数据区
- ◆ 索引表类似页表
- ◆ 索引表本身要占据额外的存储区域

---

## ● 索引文件的操作

- ◆ 查文件索引，由逻辑块号查得物理块号
- ◆ 由此磁盘物理块号而获得所要求的信息

---

## ● 优点

- ◆ 保持了链接结构的优点，又解决了其缺点：
- ◆ 即能顺序存取，又能随机存取
- ◆ 满足了文件动态增长、插入删除的要求
- ◆ 能充分利用外存空间

## ● 缺点

- ◆ 索引表本身带来了系统开销
  - 如：内外存空间，存取时间

## 5 文件存储空间管理

---

### ● 概念

- ◆ 记录当前磁盘的使用情况，创建文件时分配存储空间，删除文件时收回存储空间。

### ● 记录磁盘空闲块的方法

- ◆ 空闲文件目录
- ◆ 空闲块链
- ◆ 位示图

## ● 空闲文件目录

- ◆ 一片连续空闲区当作一个特殊文件：**空闲文件**。该文件由多个连续的**空闲**存储块组成。
- ◆ 所有的空闲文件代表存储设备的空闲空间。
- ◆ **空闲文件目录**
  - 记录所有空闲文件目录，每个表项对应一个空闲文件
  - 表项：文件号、第一个空闲块号、空闲块个数

序号	第1个空闲块号	空闲块个数	物理块号
1	2	4	2,3,4,5
2	18	9	18,19,20,21,22,23,24,25,26
3	59	5	59,60,61,62,63
⋮	⋮	⋮	⋮

---

## ● 空闲块链

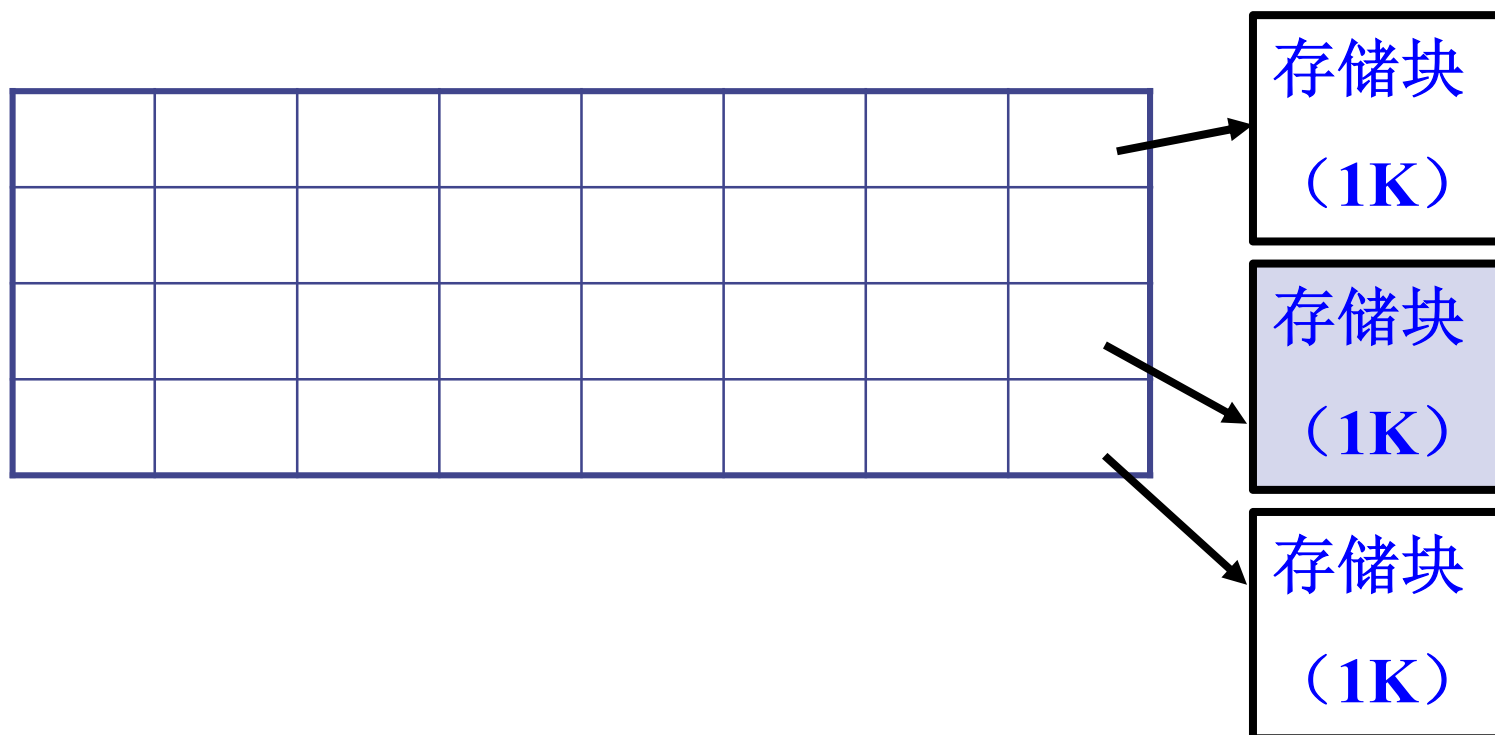
- ◆ 把存储设备上的所有空闲块链接在一起，当申请者需要空闲块时，分配程序从链头开始摘取所需要的空闲块，然后调整链首指针。反之，当回收空闲块时，把释放的空闲块逐个加在链尾上。

## ● 位示图

◆ 从内存中划出若干个字节，每位对应1个存储块。

■ 该位为1：对应存储块空闲

■ 该位为0：对应存储块已分配

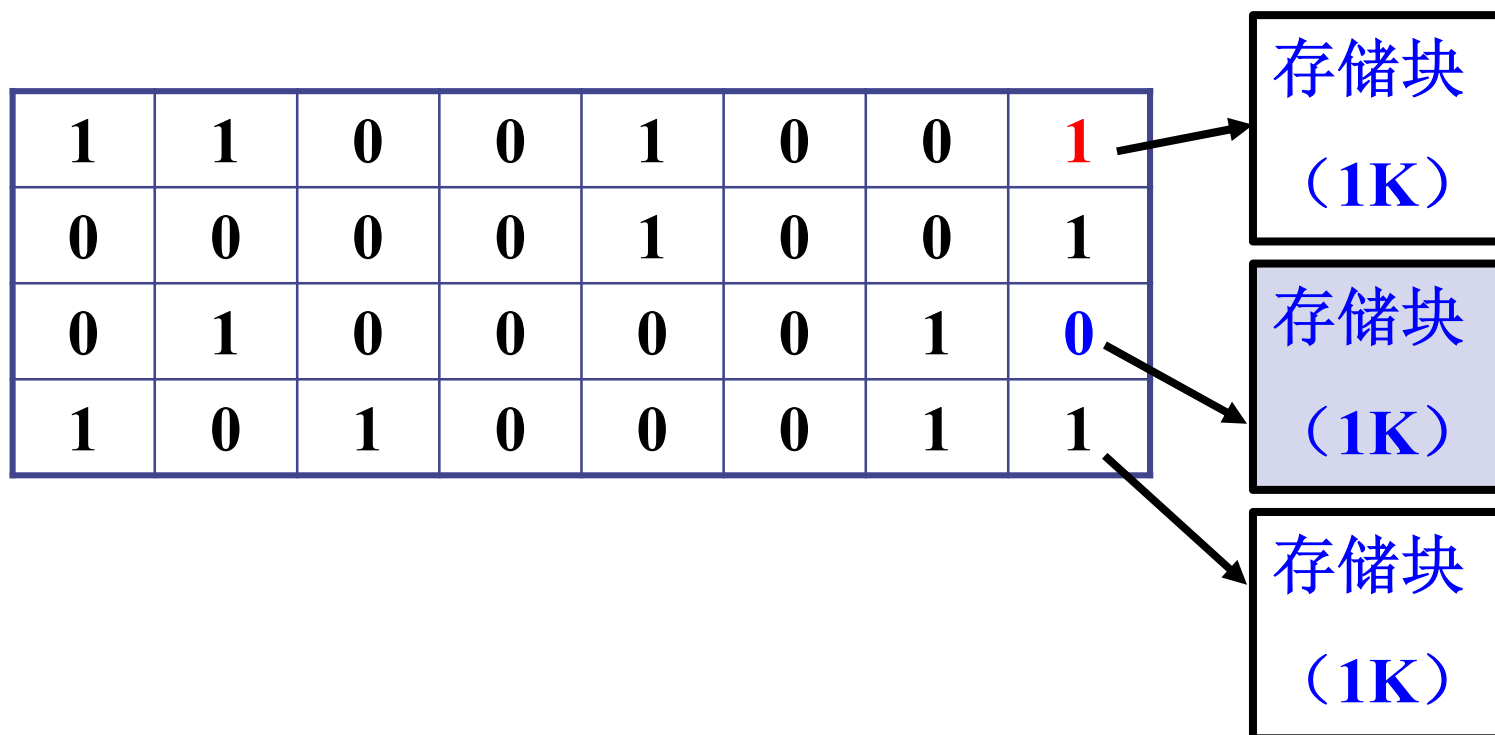


## ● 位示图

◆ 从内存中划出若干个字节，每位对应1个存储块。

■ 该位为1：对应存储块空闲

■ 该位为0：对应存储块已分配





## 6 文件目录管理

### ● 文件目录

- ◆ 文件名址录，记录文件名和存放地址的目录表
- ◆ 为了对大量文件进行分门别类的管理，提高文件检索的效率，现代操作系统往往将文件的一些属性也记录在目录中。

### ● 目录文件

- ◆ 文件目录以文件形式存于外存，这个文件叫目录文件。

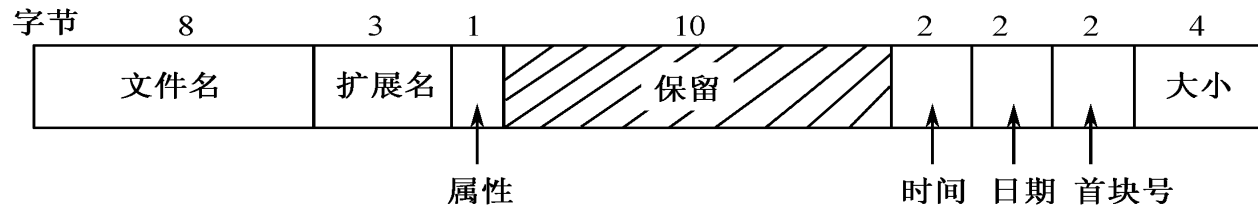
### ● 文件目录的功能

- ◆ 将文件名转换为外存物理位置的功能

## ● 目录结构 不做考察要求

◆ 不同的系统，文件目录的组织也不完全相同。

### ■ DOS系统（32字节）



### ■ UNIX/Linux系统

#### ■ 索引节点

- 文件目录项中的文件名和其他信息分开。后者单独组成一个定长数据结构：索引节点 **i\_node**。

## ● 索引结点 不做考核要求

### ◆ 磁盘索引结点

- 存放在磁盘上的索引结点。每个文件有惟一的一个磁盘索引结点。

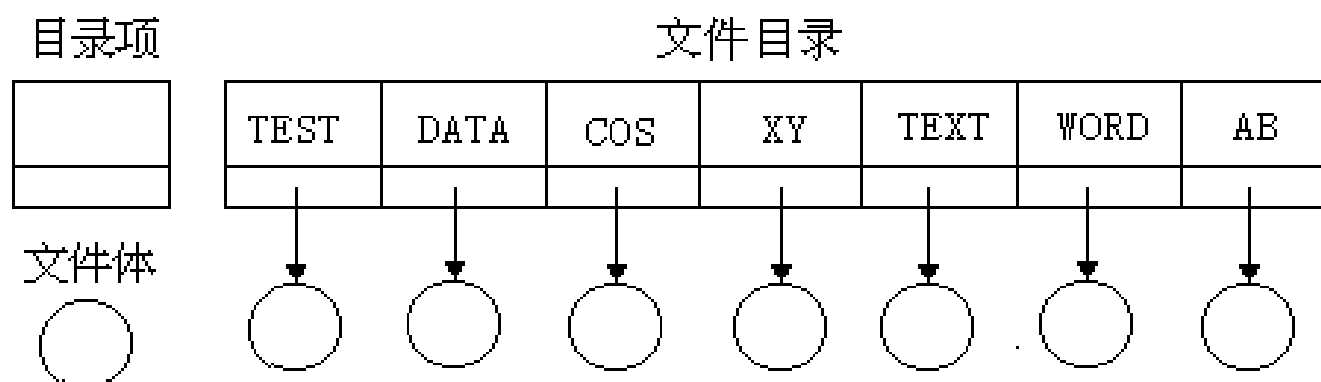
### ◆ 内存索引结点

- 存放在内存的索引结点。当文件打开时，要将磁盘索引结点拷贝到内存索引结点中。

- 
- 目录结构
    - ◆ 一级目录
    - ◆ 二级目录
    - ◆ 多级文件目录（树型目录）
  - 文件全名

## ● 单级目录

- ◆ 单级目录是最简单的目录结构。在这种组织方式下，全部文件都登记在同一目录中。

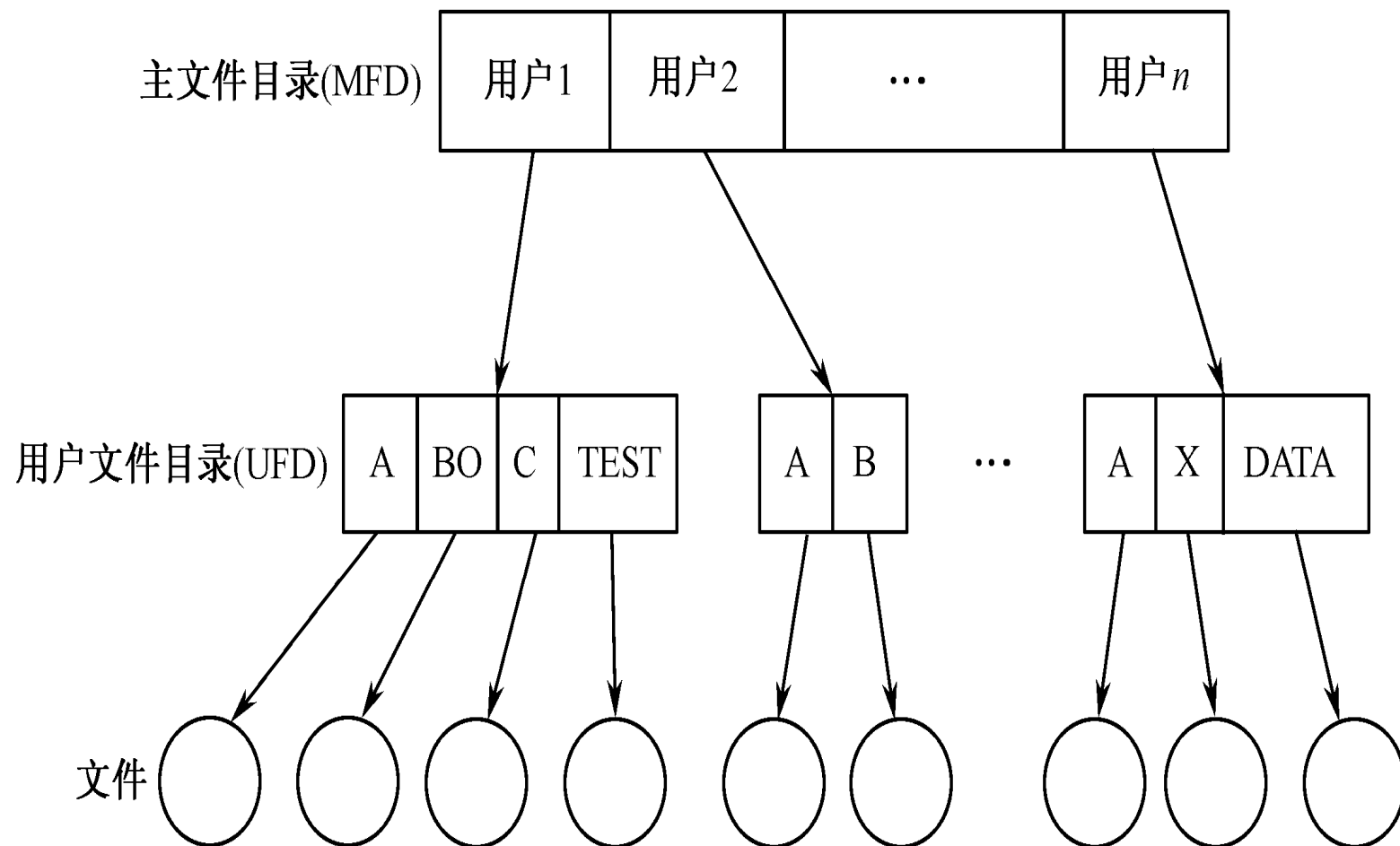


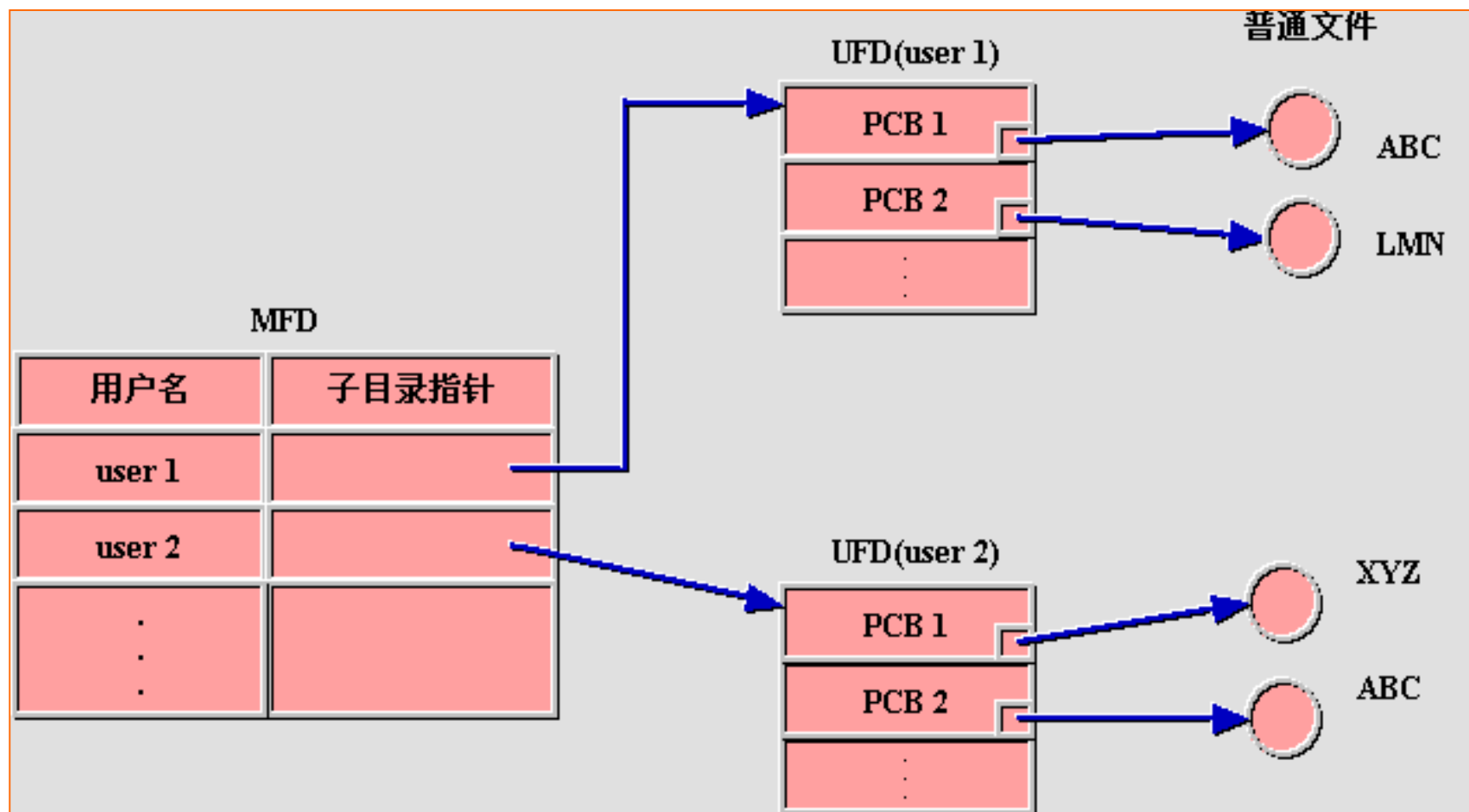
- ◆ 特点：简单、易于理解和实现
- ◆ 缺点：查找速度慢、不允许重名、不便于文件共享

## ● 两级目录

- ◆ 每个用户使用一个相对独立的目录，在所有用户的目录上层再建一层目录来管理各个用户目录。
- ◆ 二级目录结构把文件目录分成二级，第一级称为主目录，第二级称为子目录或次目录。
- ◆ 系统允许每个用户有一个子目录。也称为用户目录。
- ◆ 二级目录结构有效地解决文件重名的问题，不同用户的文件，使用相同名字也不会导致混乱。

## ● 两级目录





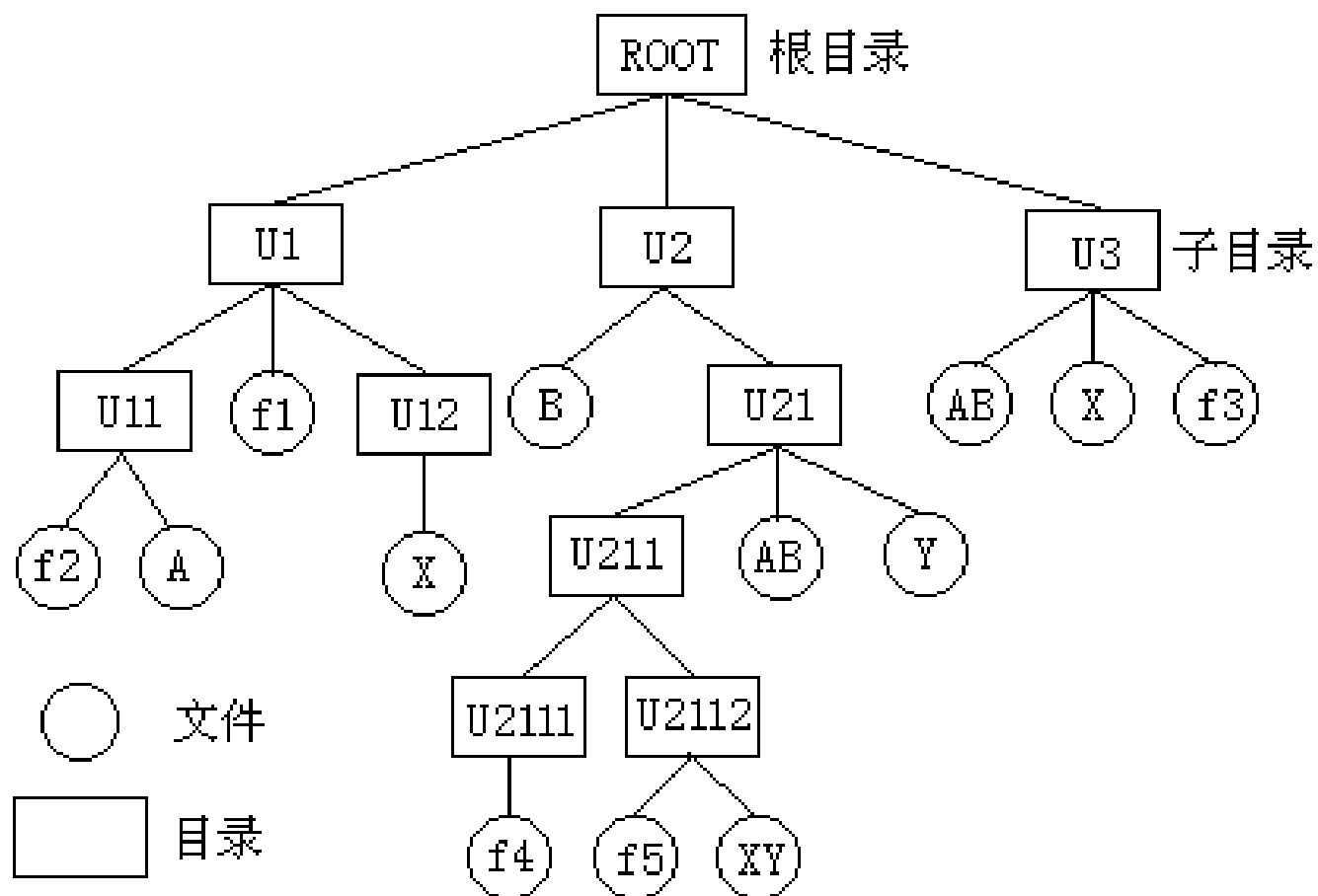


---

## ● 树型目录

- ◆ 多级目录结构
- ◆ 二级目录结构的扩充。
- ◆ 目录结构如同倒置的树，主目录是树根，称根目录
- ◆ 枝结点是子目录，树叶描述一个文件。

## ● 树型目录



# 文件全名和路径

- 文件的全名

- ◆ 包括从根目录开始到文件为止的通路上所有子目录路径。

- 子目录名之间用正斜线 “/”或反斜线 “\”隔开

- 子目录名组成的部分又称为路径名。

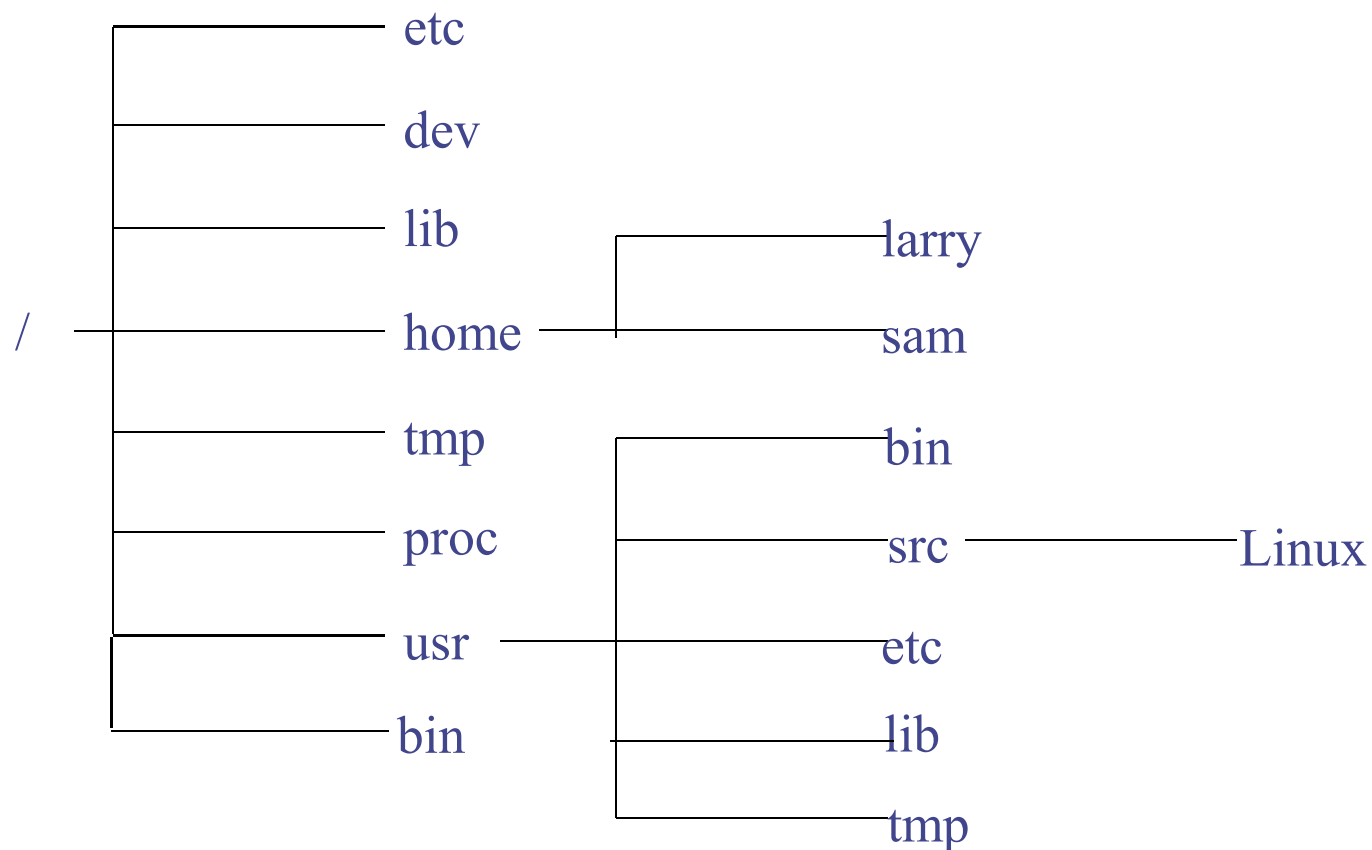
- 每个文件都有惟一的路径名。

- 两种路径名形式

- ◆ 绝对路径名：从根目录直到指定的文件

- ◆ 相对路径名：从当前目录直到指定的文件

## ● LINUX文件目录



Linux系统的树形文件目录结构

## 7 文件的保护

- 对文件的访问系统首先要检查访问权限

- ◆ 仅允许执行 (E)。
- ◆ 仅允许读 (R)。
- ◆ 仅允许写 (W)
- ◆ 仅允许在文件尾写 (A)
- ◆ 仅允许对文件进行修改 (U)
- ◆ 允许改变文件的存取权限 (C)
- ◆ 允许取消文件 (D)
- 权限可进行适当的组合。