

# 第六章 文件管理

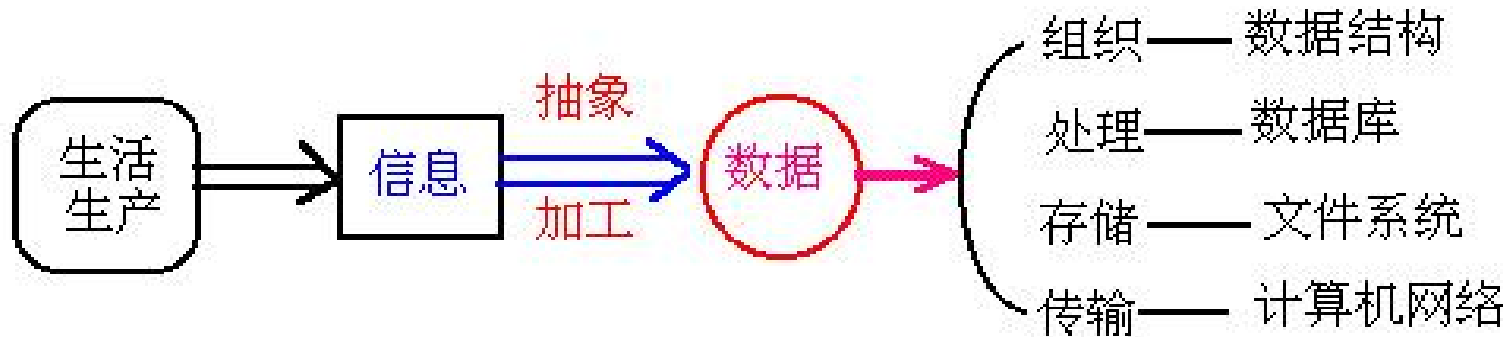
6.1 文件

6.2 文件目录

6.3 文件组织与数据存储

6.4 文件共享

# 文件系统概念



- 文件系统: 操作系统中负责存取和管理信息的模块, 它用**统一的方式**管理用户和系统信息的**存储、检索、更新、共享和保护**, 并为用户提供一整套方便有效的文件使用和操作方法。
- 文件的逻辑结构和存储结构

# 文件系统面向用户的功能

- 文件的**按名存取**
- **文件目录**建立和维护
- 实现**逻辑文件到物理文件**的转换
- 文件**存储空间**的分配和管理
- 提供合适的**文件存取**方法
- 实现文件的**共享、保护和保密**
- 提供一组可供用户使用的**文件操作**

## 6.1 文件

6.1.1 文件概念

6.1.2 文件命名

6.1.3 文件类型

6.1.4 文件属性

6.1.5 文件存取方法

## 6.1.1 文件的概念

文件：逻辑上具有完整意义的信息集合。它有一个**名字**以供标识。文件名是以字母开头的**字母数字串**

文件的基本单位：**数据项、记录**

### ◆ **数据项**

- 1. 基本数据项：用于描述一个对象的某种属性的字符集，是数据组织中可以命名的最小逻辑数据单位，即原子数据，又称为**数据元素或字段**
- 2. 组合数据项：由若干个基本数据项组成的，简称**组项**。

- ◆ **记录**：一组相关数据项的集合，用于描述一个对象在某方面的属性。一个记录应包含哪些数据项，取决于需要描述对象的哪个方面。而一个对象，由于他所处的环境不同可把他作为不同的对象。

## 6.1.2 文件的命名

**文件**是一个抽象机制，提供了把文件保存在磁盘上，用户不必了解信息存储细节且便于读取的方法，z  
重点：文件命名

DOS、Windows系统中文件名：

**文件名. [文件扩展名]**

文件名：不大于8个字符的字符串，但WINDOWS系统中允许超长文件名，最多可有256个字符，不区分英文字母的大小写。

- 不同的操作系统对文件名的字符长度限制可能不同。
- 不同的操作系统对文件名的英文字母的大小写处理可能不同。

## 6.1.3 文件的类型

UNIX系统中文件属性：

普通文件：

— 普通文件

目录文件：

d 目录文件

特别文件：

p 管道文件

c 字符型设备文件

b 块设备文件

块设备文件、字符设备文件、管道文件

用 `ls -l` 命令列文件清单，首项特征为：

“c”，即 **c**rw-rw-rw- （character, 字符型）

“b”，即 **b**rw-rw-rw- （block, 块型）

## 6.1.4文件的属性

- 文件基本属性
- 文件类型属性
- 文件保护属性
- 文件管理属性
- 文件控制属性

- UNIX/Linux把用户分为文件主、同组用户、其他用户三类，
- 定义存取权限可读r、可写w、可执行x，文件属性共有10位：

-**rwX**rwXrwx

文件中的数据形式：源文件、目标文件、可执行代码文件  
存取属性：只执行文件、只读文件、读写文件

文件逻辑结构：

有结构文件（记录、数据项）、无结构文件（流式文件）

文件结构类型：顺序文件、链接文件、索引文件



## 6.2 文件目录

6.2.1 文件控制块、文件目录与目录文件

6.2.2 层次目录结构

6.2.3 目录的检索

# 文件系统的分层结构

- **文件管理**--实现文件的逻辑结构，为用户提供各种文件系统调用，及文件访问权限的设置等工作；
- **目录管理**--负责查找**文件描述符**，进而找到需要访问的文件，及进行访问权限检查等工作；
- **磁盘主存映射管理**--将文件的逻辑地址转换成磁盘的物理地址，即由逻辑块号找到柱面号、磁道号和扇区号，具体的数据传输操作由设备管理实现。

## 6.2.1 文件目录与文件目录项

- 文件目录：支持“按名存取”
- 文件系统基本功能：文件目录的建立、维护和检索
- 文件目录项又称文件控制块FCB (File Control Block)

- 有关文件存取控制的信息
- 有关文件结构的信息
- 有关文件使用的信息
- 有关文件管理的信息

- 文件目录项
- 文件目录
- 目录文件

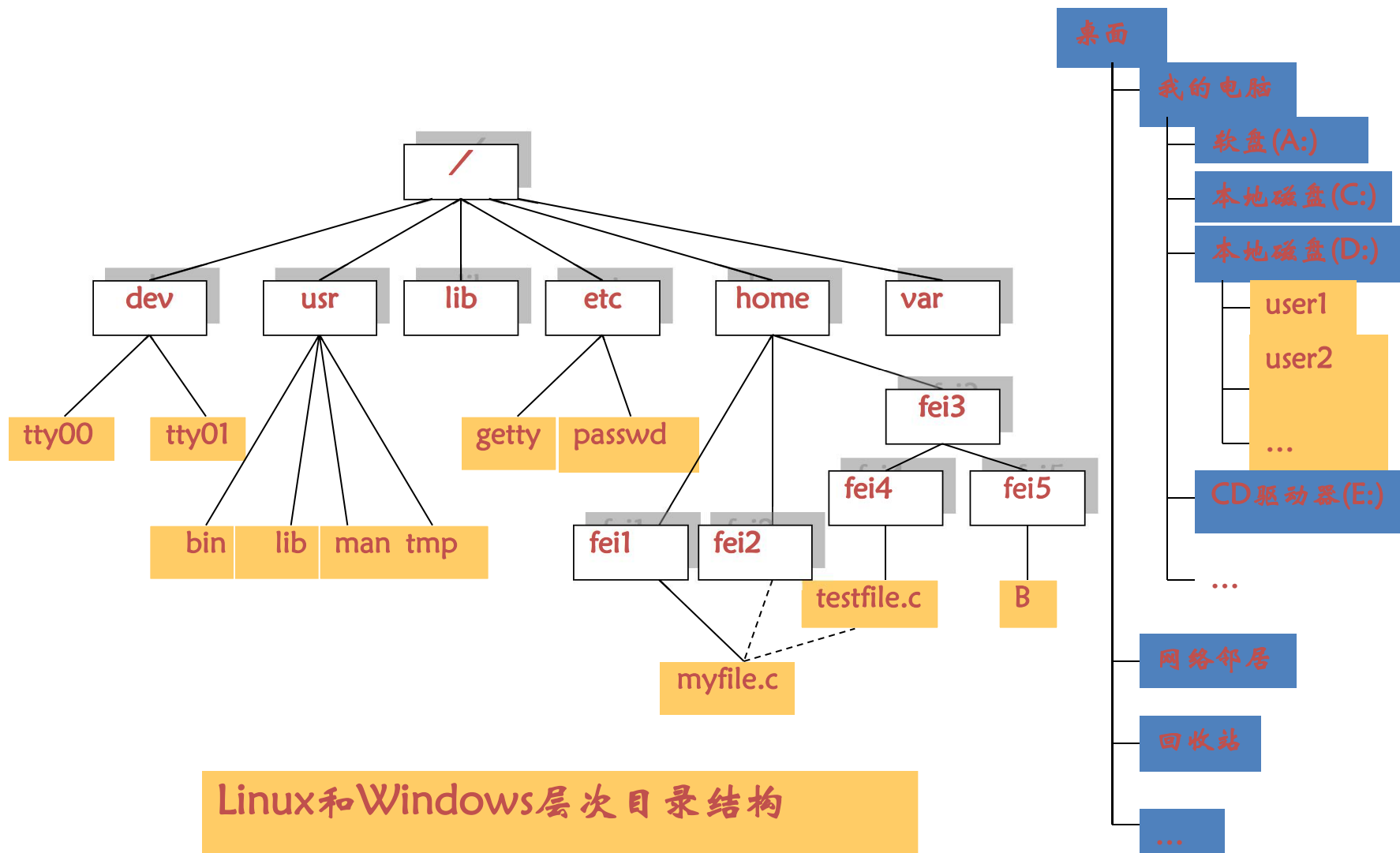
# UNIX/Linux中的基本目录项

文件名	inode号
-----	--------

14个字节

2个字节

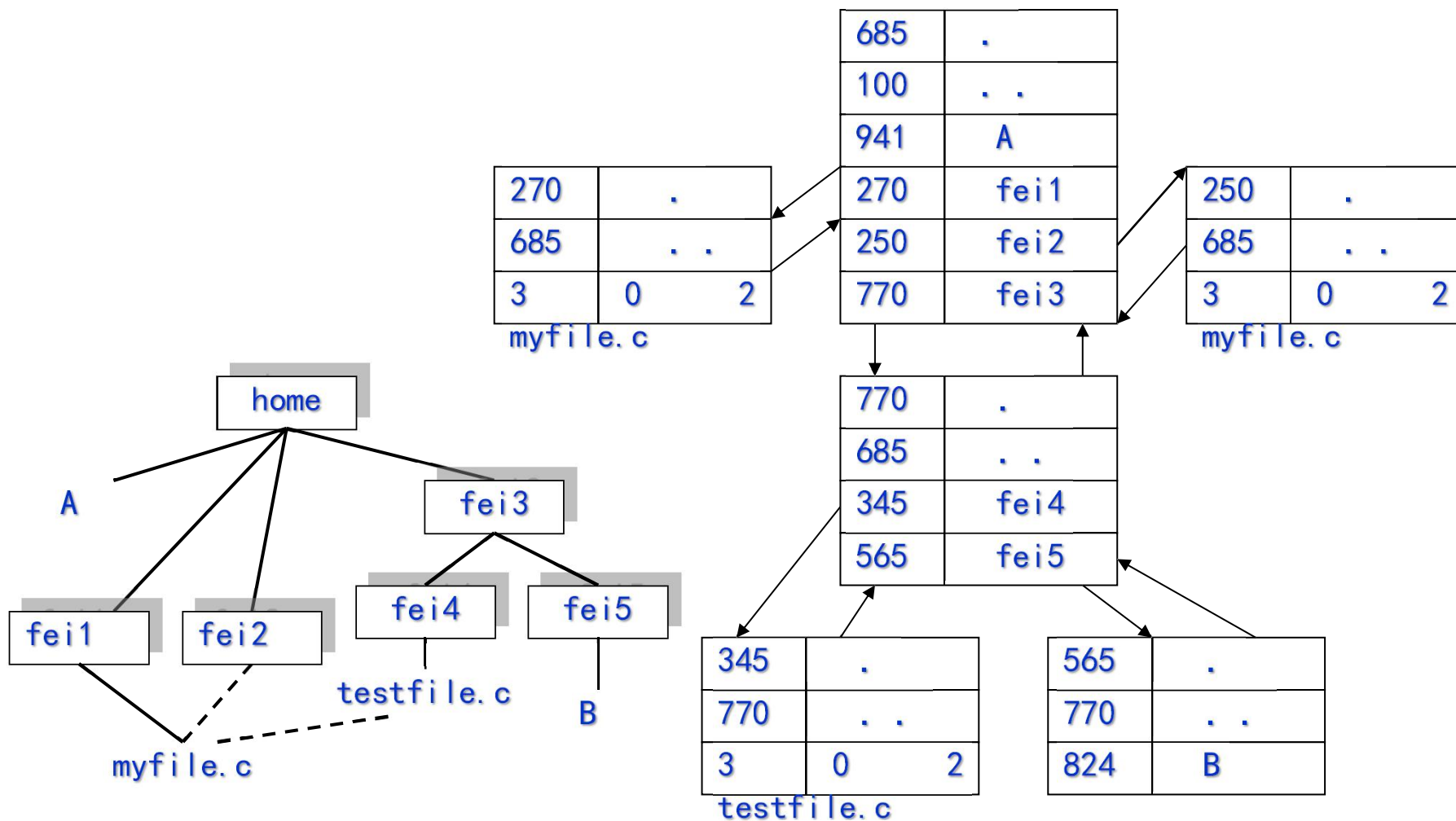
## 6. 2. 2层次目录结构



# 层次目录结构

- 路径名
- 目录路径名
- 文件路径名
- 绝对路径名
- 相对路径名

## 6.2.3 文件目录的检索



(a) 用户角度目录结构

(b) 系统角度目录链接

不同角度的目录结构

## **6.3 文件组织与数据存储**

### **6.3.1 文件的逻辑结构**

### **6.3.2 顺序文件**

### **6.3.3 文件的物理结构**



## 6.3.1 文件的逻辑结构类型

### 1. 有结构文件

– 分类：定长记录、变长记录

– 记录的组织方式

- (1) **顺序文件**：由一系列记录按某种顺序排列形成的文件。记录通常是定长记录，查找迅速。
- (2) **索引文件**：当记录为**可变长度**时，应建立索引表，每个记录对应一个表项，以加快记录检索的速度。
- (3) **索引顺序文件**：上述两种文件构成方式的结合。它为文件建立一张索引表，为每一组记录中的第一个记录设置一个表项。

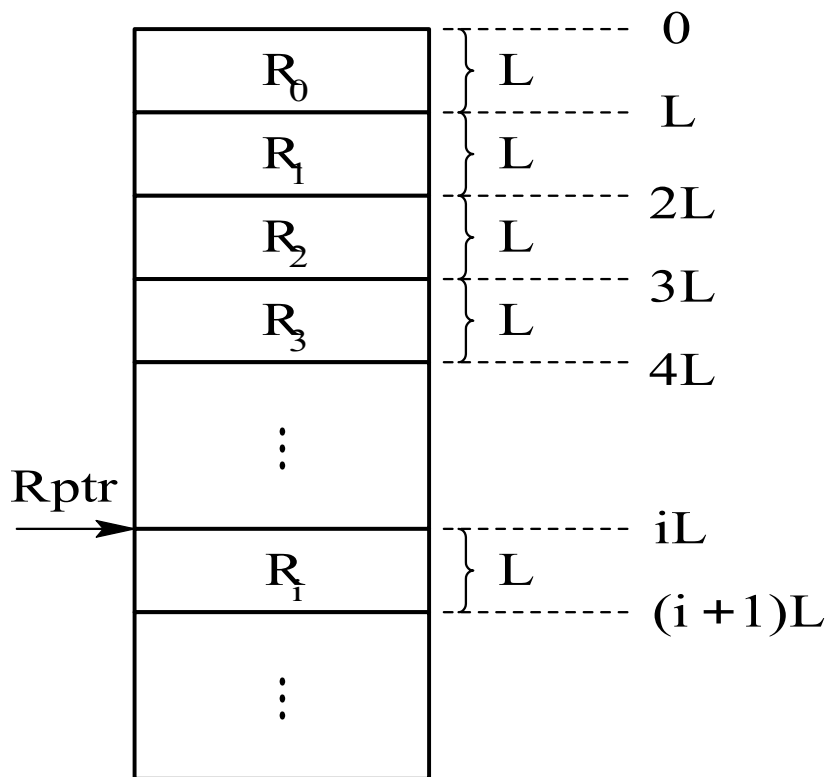
- 2. 无结构文件

- 源程序、可执行文件、库函数，即流式文件
- 长度以字节为单位
- 访问：采用读/写指针来指出下一个要访问的字符。可以把流式文件看做是记录式文件的一个特例。

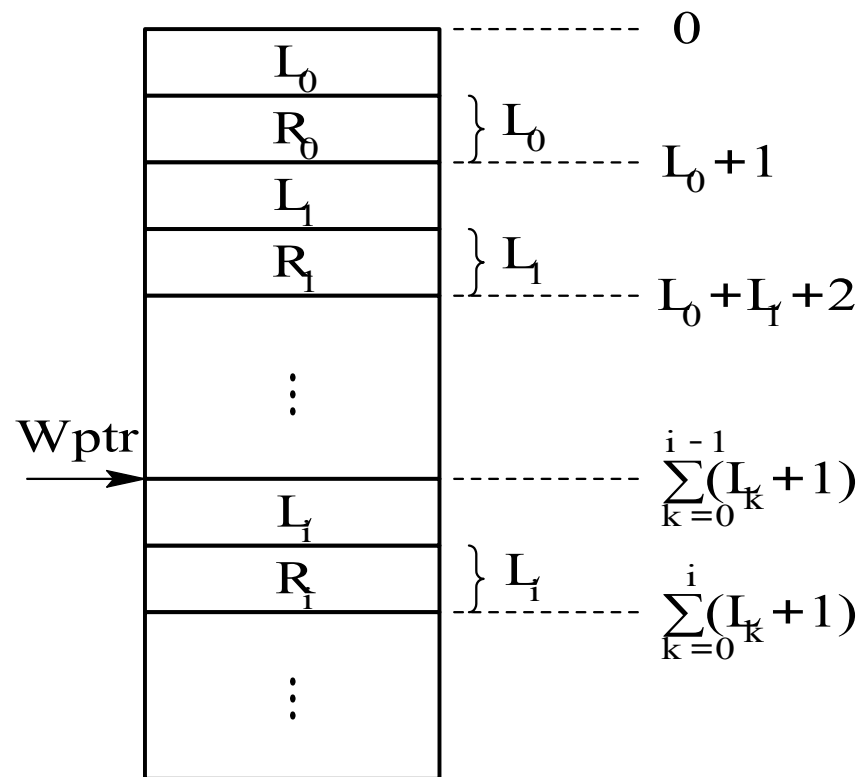
## 6.3.2 顺序文件

### 1. 逻辑记录的排序

- **串结构**：各记录之间的顺序与关键字无关。典型是按记录建立时间从早到晚来排列。检索效率较低。
- **顺序结构**：指文件中的所有记录按关键字(词)排列(升序、降序)。检索效率较高，可采用折半查找、插值查找、跳步查找等算法。
- **对顺序文件(Sequential File)的读/写操作**
- **索引文件**
- **索引顺序文件**



(a) 定长记录文件



(b) 变长记录文件

便于随机访问

定长和变长记录文件

不便于随机访问

索引号	长度 $m$	指针 $ptr$
0	$m_0$	●
1	$m_1$	●
$\vdots$		
$i$	$m_i$	●
$\vdots$		

索引表

$R_0$
$R_1$
$\vdots$
$R_i$
$\vdots$

逻辑文件

索引表已排序，  
便于随机访问

索引文件的组织

键	逻辑地址
An Qi	●
Bao Rong	●
Chen Lin	

姓 名	其它属性
An Qi	
An Kang	
Bao Rong	
⋮	

逻辑文件

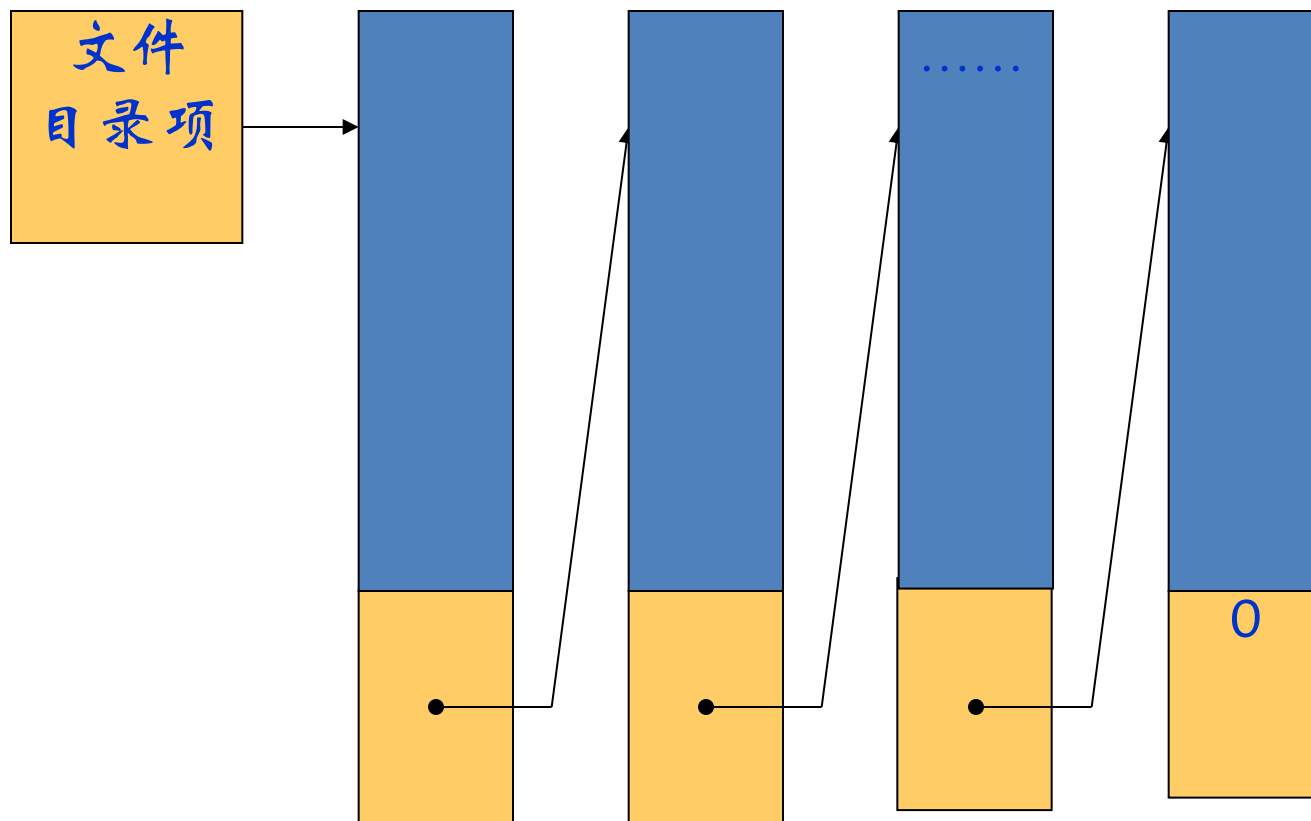
索引顺序文件

## 6.3.3 文件的物理结构

### 1、顺序文件(连续文件)

- 文件中逻辑上连续的信息存放到存储介质的依次相邻的块上便形成顺序文件
- 逻辑记录顺序和物理记录顺序完全一致的文件，通常，记录按出现的次序被读出或修改。
- 顺序文件的优点/缺点：
- 顺序文件变种：
  - 紧凑顺序文件、扩展顺序文件
  - 连接顺序文件、划分顺序文件

## 2、连接文件(串联文件)





### 3、直接文件(哈希文件)

- 记录的**关键字**与其**地址**间可通过某种方式建立对应关系，利用这种关系实现存取的文件叫直接文件。
- hash技术：建立hash表，hash表是一个指针数组，数组通过索引访问，找到的指针便指向数据记录。索引是与数据记录有关的关键字或其变换，
- 描述一座城市人口的hash文件举例。

## 直接文件(哈希文件)

### 步1 构造转换(hash)函数

- 设文件名为8个ASCII字符。构造的hash函数为模2加“ $\oplus$ ”，求已知文件名的ASCII字符值的模2加值作为该文件的FCB所在物理块在目录文件中的索引A，那么，

$$A = (a_1 \oplus a_2 \oplus \dots \oplus a_8)$$

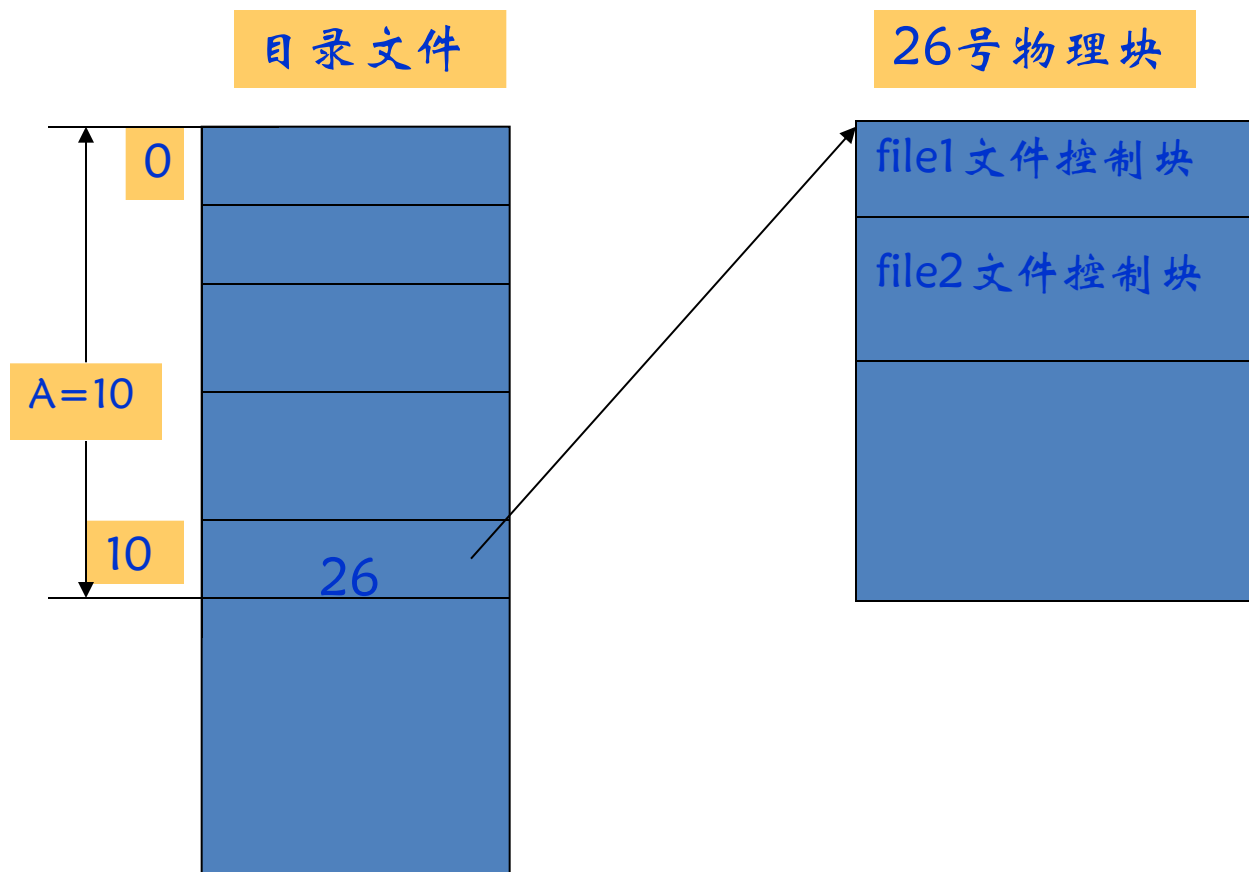
# 直接文件(哈希文件)

## 步2 建立目录文件

- 目录文件采用索引结构，建立文件时由步1求出文件名的hash值A，
- 凡A值相同的文件的FCB都存放在同一个物理块。磁盘的物理块号存放在索引表中的相对位置应等于A值。

# 直接文件(哈希文件)

## 步2 建立目录文件



# 直接文件(哈希文件)

## 步3 查找文件

- 根据给定文件名，由步1算出该文件的FCB所在物理块号在索引表中的相对位置A。根据A就可找到该FCB所在物理块号，
- 把这个物理块读入主存缓冲区，用文件名逐个比较，找出要求的FCB。

# 直接文件(哈希文件)

## 步4 溢出处理

- 物理块中存放的FCB是有限的，建立目录文件时，如果A值相同的文件数目超过物理块能容纳数时，产生溢出。
- 溢出时，系统再申请一个盘区，该区物理块号放在 $A+k$ 的索引表目中， $k$ 是质数作为位移常数。
- 第二块盘区也溢出，则申请第三块，块号放在 $A+2\times k$ 表目中，依此类推。
- 查找目录时，如第一块找不到可找 $A+k$ 表目中的物理块号，读出后继续比较，依次类推。

## 4、索引文件(1)

- 索引结构是实现非连续存储的另一种方法，适用于数据记录保存在随机存取存储设备上的文件。
- 使用索引表，**每个表目包含一个记录的键及其记录数据的存储地址**，这类文件称索引文件。

# 索引文件(2)

