



张涛

# Review

存储管理的基本概念

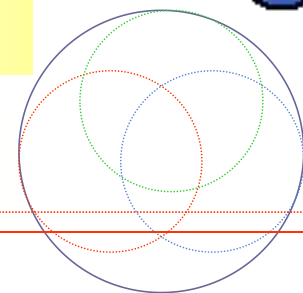
分区存储管理

分页存储管理

虚拟存储

请求分页存储管理

分段/段页式存储管理



# 本章内容

文件系统概述

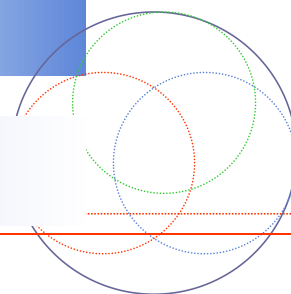
文件的结构和存取法

文件目录

文件系统的使用

文件存储空间的管理

文件的共享与保护



# Today we focus on...

---

**文件系统概述**

**文件和文件系统**

**文件类型**

**文件系统的基本功能**

**文件的结构和存取法**

# 5.1 文件系统概述

- 所有的计算机应用程序都要：
  - 存储信息，检索信息
- 三个基本要求：
  - 能够存储大量的信息
  - 长期保存信息
  - 可以共享信息
- 解决方法：
  - 把信息以一种单元——即文件的形式存储在磁盘或其他介质上
- 文件是通过操作系统来管理的
- 文件的结构，命名，存取，使用，保护和实现方法

## ■ 两种观点：

### ■ 用户观点：

- 文件系统如何呈现在其面前：一个文件由什么组成，如何命名，如何保护文件，可以进行何种操作等等

### ■ 操作系统观点：

- 文件目录怎样实现，怎样管理存储空间，文件存储位置，磁盘实际运作方式(与设备管理的接口)，文件系统性能等等

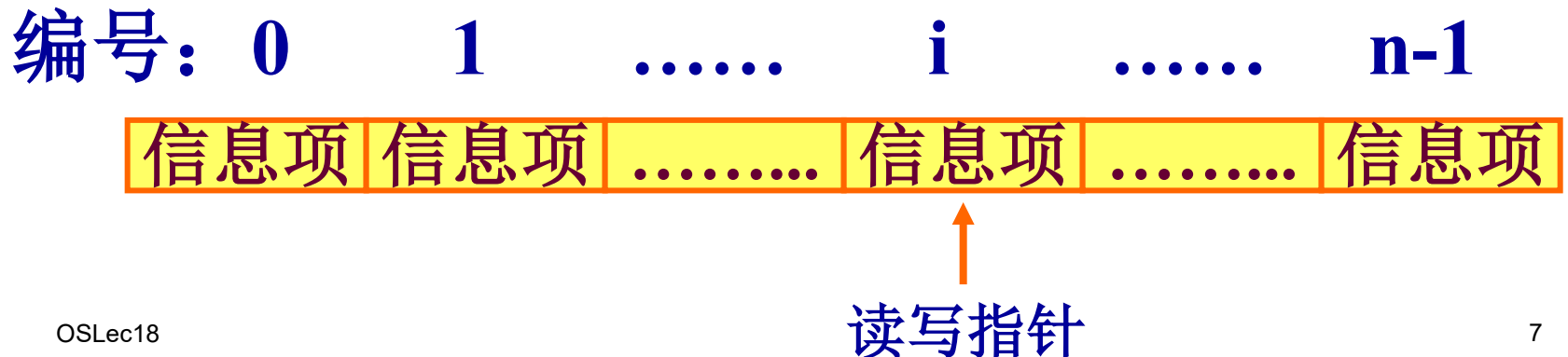
# 5.1.1 文件与文件系统

## ■ 什么是文件？

- 文件是赋名的信息 (数据)项的集合。
- 文件是赋名有关联的信息单位 (记录)的集合。

## ■ 什么是文件名？

- 文件的标识符号，一个用来标识文件的有限长度的字符串。



# 文件

- 一组带标识的在逻辑上有完整意义的信息项的序列，这个标识为文件名
  - 信息项：构成文件内容的基本单位
  - 长度：单个字节，或多个字节
  - 文件内容的意义：由文件的建立者和使用者解释
- 文件的组成：
  - 文件体：文件本身的信息；
  - 文件说明：文件存储和管理信息；如：文件名、文件内部标识、文件存储地址、访问权限、访问时间等；
- 文件属性：
  - 用一组信息指定文件的类型、操作特性和存取保护等，把这组信息称为文件的属性。
  - 文件的属性一般存放在文件的目录项中。



# 文件系统

## ■ 什么是文件系统？

- 操作系统中负责管理相关文件信息的软件机构。

## ■ 文件系统管理的对象：

- (1) 文件
- (2) 目录
- (3) 磁盘空间

## ■ 文件系统的组成：

- 被管理的文件
- 与文件管理相关的软件
- 实施文件管理所需的数据结构

# 5.1.2 文件的类型

## ■ 按照性质和用途分类

- 系统文件：只能通过系统调用为用户服务；
- 库文件：允许用户调用但不允许用户修改；
- 用户文件：用户委托操作系统保存的文件，又分为：
  - 1) 临时文件；2) 永久文件；3) 档案文件。

## ■ 按文件的保护方式分类

- 只读文件、读写文件、不保护文件

## ■ 按文件信息的流向

- 输入文件、输出文件、输入输出文件

## ■ 按文件的逻辑结构分类

- 流式文件；记录式文件

## ■ UNIX系统将文件分为三类：

### ■ 普通文件：

- 包含的是用户的信息，一般为ASCII或二进制文件

### ■ 目录文件：

- 管理文件系统的系统文件

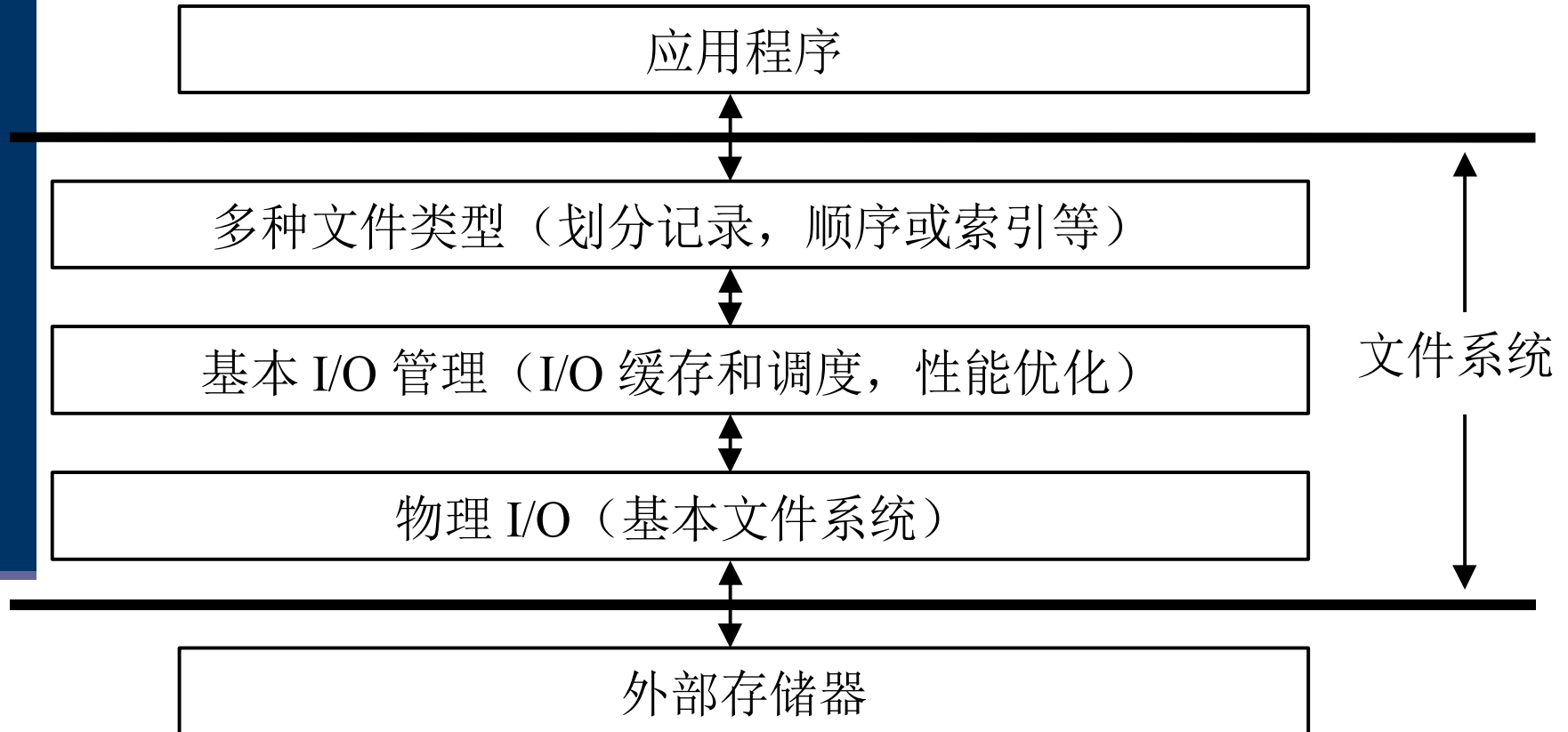
### ■ 特殊文件（设备文件，外部设备也看作文件）

- 字符设备文件：和输入输出有关，用于模仿串行I/O设备，例如终端，打印机，网络等
- 块设备文件：模仿磁盘

## 5.1.3 文件系统的基本功能

- 五个基本功能：
  - 文件的结构及有关存取方法；
  - 文件的目录结构和有关处理
  - 文件存储空间的管理
  - 文件的共享和存取控制
  - 文件操作和使用

# 文件系统的结构和功能元素



# 文件系统要实现的功能模块

- 文件的分块存储：与外存的存储块相配合；
- I/O缓冲和调度：性能优化；
- 文件定位：在外存上查找文件的各个存储块；
- 外存存储空间管理：如分配和释放。主要针对可改写的外存如磁盘；
- 外存设备访问和控制：包括由设备驱动程序支持的各种基本文件系统如硬盘，软盘，CD ROM等。

# 文件系统的接口

- 为方便用户使用文件系统，文件系统通常向用户提供两种类型的接口：
- (1) 命令接口。这是指作为用户与文件系统交互的接口。用户可通过键盘终端键入命令，取得文件系统的服务。
- (2) 程序接口。这是指作为用户程序与文件系统的接口。用户程序可通过系统调用来取得文件系统的服务。

## 5.2 文件的结构和存取法

- 文件结构分别从以下观点研究和设计文件的结构和存取方法：

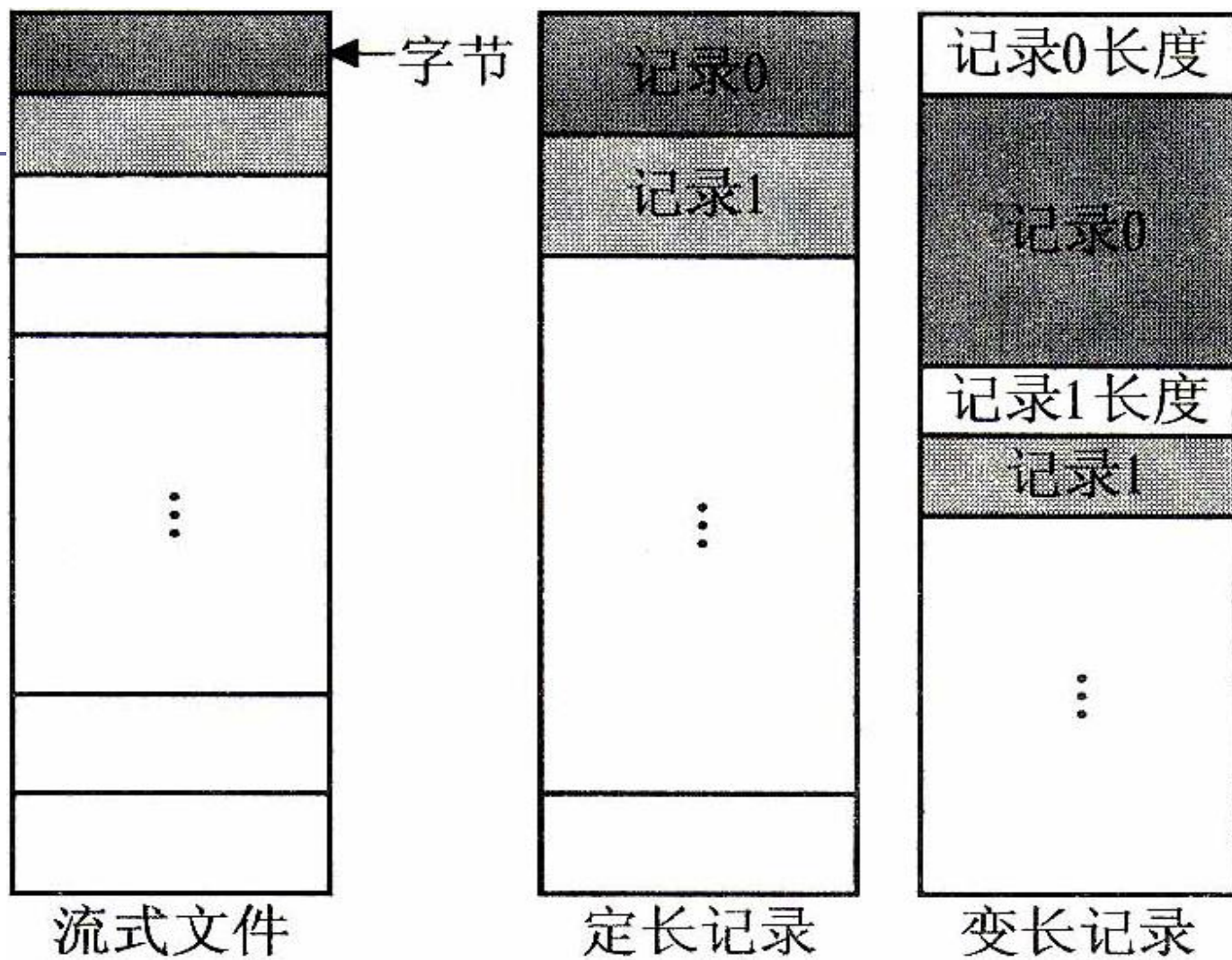
- 1) **用户观点**      研究用户“思维”中的抽象文件即逻辑文件。
- 2) **系统观点**      研究驻留在设备介质的实际文件即物理文件。

- 文件系统的重要作用之一：用户逻辑文件和相应设备上的物理文件之间建立映射，实现二者之间的转换。
- 文件的存取方法是由文件的性质和用户使用文件的要求决定的。



## 5.2.1 文件的逻辑结构

- **逻辑结构**：从用户角度看文件，研究文件的组织形式
- **无结构的流式文件**
  - 基本信息单位是字节或字，其长度是所含字节的数量。
  - 优点：节省存储空间，无需额外的说明和控制信息。
- **有结构的记录式文件**：由若干个记录组成，文件中的记录可按顺序编号为记录1，记录2，……，记录n。
  - **定长记录文件**：文件中所有记录的长度相等。文件的长度为记录个数与记录长度的积。
  - **变长记录文件**：文件中的记录长度不相等。文件长度为所有记录长度之和。

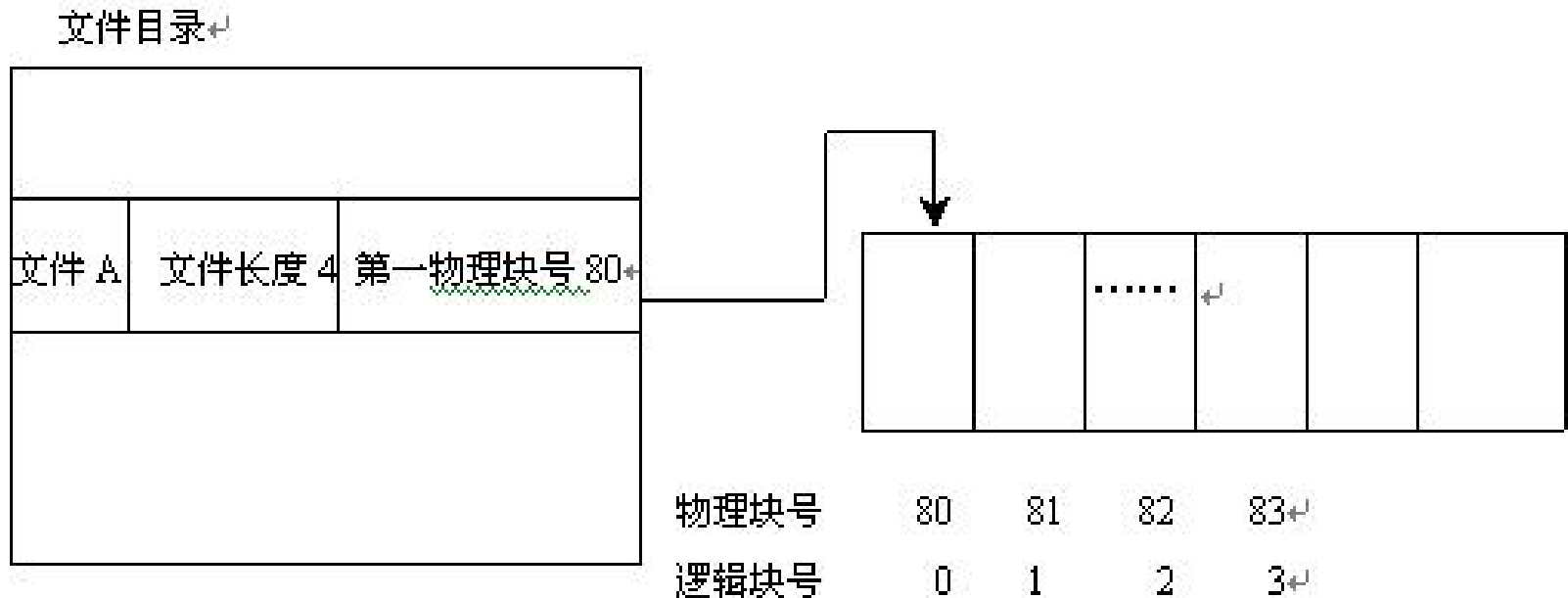


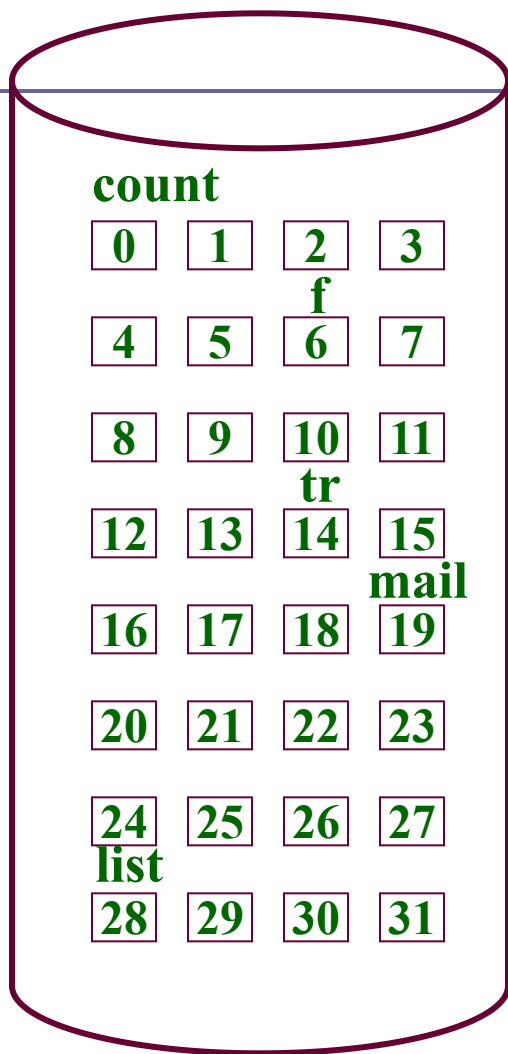
## 5.2.2 文件的物理结构

- 文件的物理结构是指文件在物理存储介质上的存储结构。
- 为了有效的分配文件存储器的空间，通常把它们分成若干块，并以块为单位进行分配和传送，称为物理块。而块中的信息称为物理记录。
- 基本的文件物理结构
  - 连续结构
  - 链接结构
  - 索引文件

# 连续结构

- 一个逻辑文件的信息存放在存储器上的相邻物理块中，该文件为连续文件，这样结构称为连续结构。





## 文件目录

文件名	始址	块数
count	0	2
tr	14	3
mail	19	6
list	28	4
f	6	2

# 连续结构的优缺点

## ■ 优点：

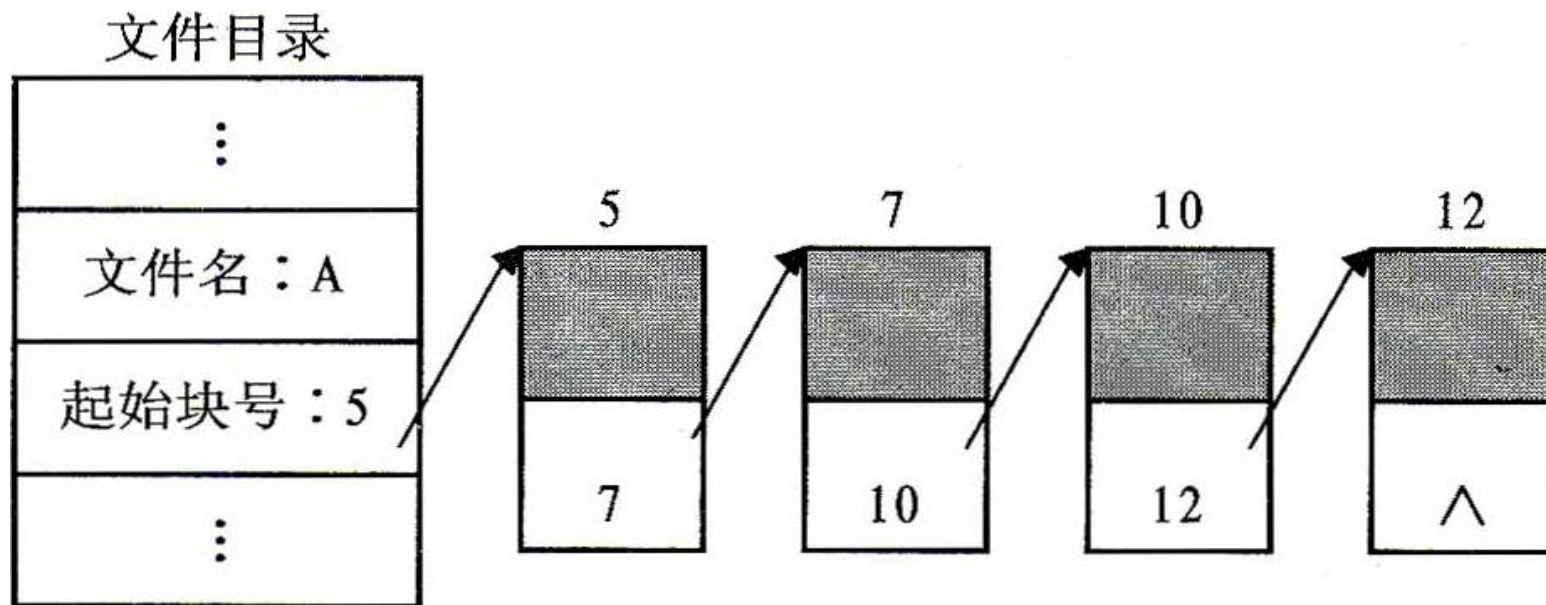
- 顺序存取速度快，所需的磁盘寻道次数和寻道时间最少。知道文件存储的起始块号和文件块数，就可以立即找到所需要的信息。
- 简单，支持顺序存取和随机存取。

## ■ 缺点：

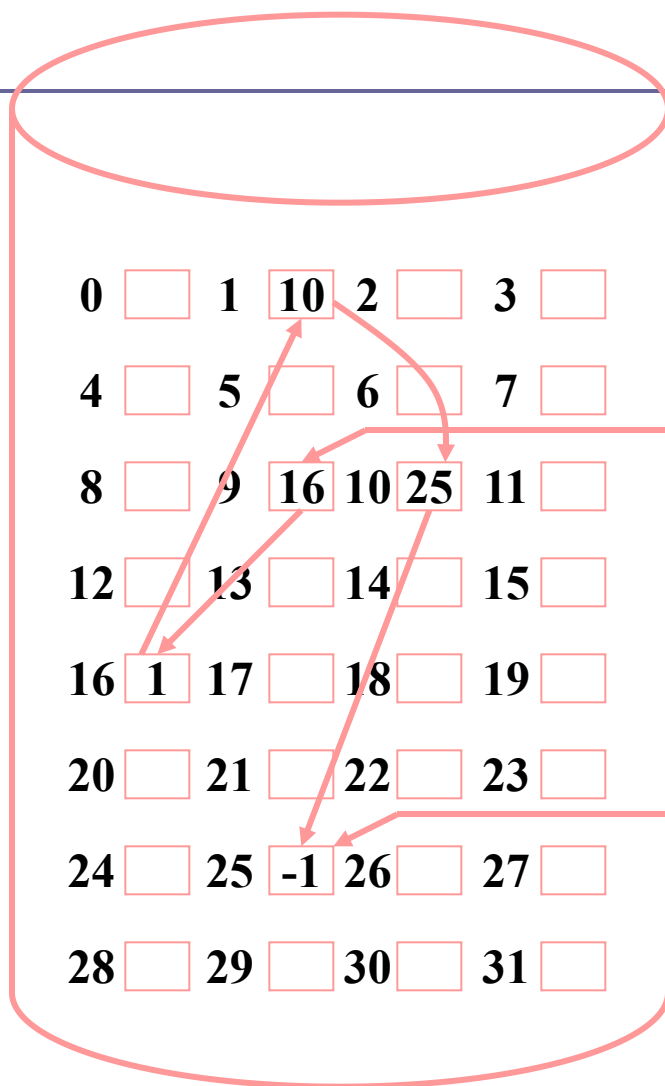
- 在建立连续结构文件时，要求用户给出文件的最大长度，以便系统分配足够的存储空间，但这个有时候难以办到；
- 不便记录的增删操作，一般只能在末端进行。

## 串连结构 〈链接结构〉

- 在每个物理块中设置一指针，指向该文件的下一个物理块号，文件的末尾块存放结束标记“NULL”。



## 文件目录



文件名	始址	末址
jeep	9	25



# 链接结构的优缺点

## ■ 优点：

- 文件可以动态扩充，也不必事先提出文件的最大长度。
- 由于不连续分配，不存在外部碎片问题，所以不会造成几块连续区域的浪费。
- 有利于文件插入和删除

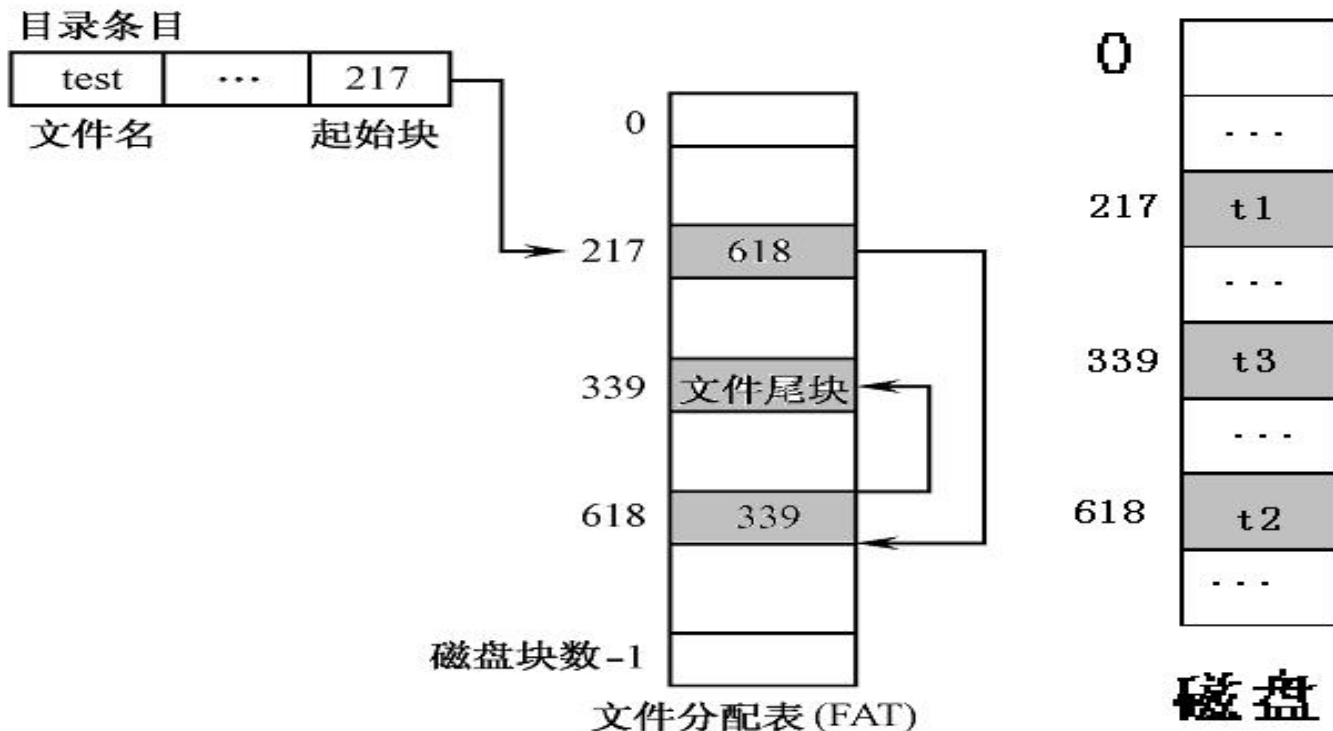
## ■ 缺点：

- 存取速度慢，不适于随机存取，只适合顺序存取，
- 每块设置链接字破坏物理信息的完整性
- 链接指针占用一定的空间

# 链接结构的变形

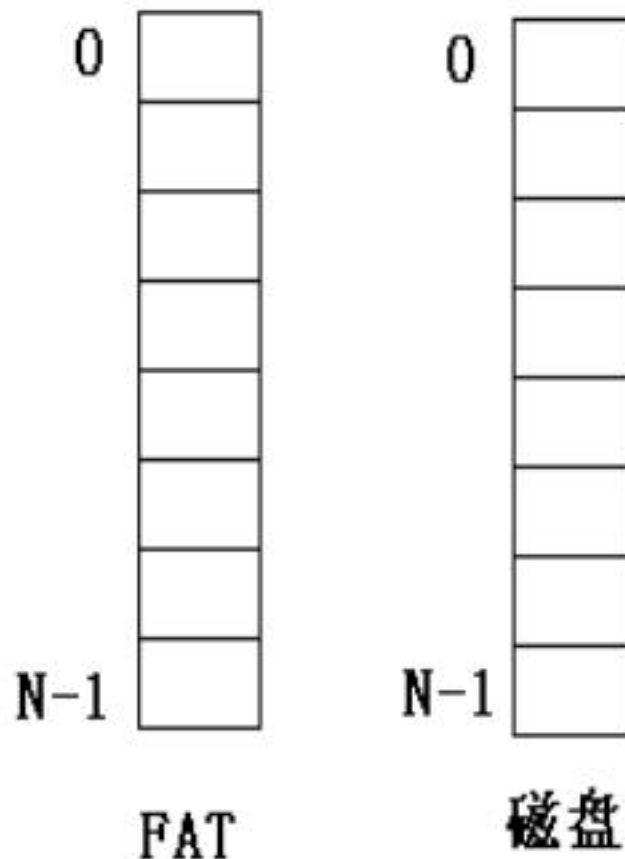
## ■ 文件分配表 (FAT)

- 将盘块中的链接字按盘块号的顺序集中起来，构成盘文件映射表/文件分配表FAT。



# FAT的实例

- 在MS-DOS和Windows系统中，文件的物理结构使用的是FAT（File Allocation Table）结构。
- 将磁盘空间划分为块，每块大小为扇区的整数倍。在FAT文件系统中块称为簇
- 一个磁盘分区能分为多少簇则FAT就有多少表项



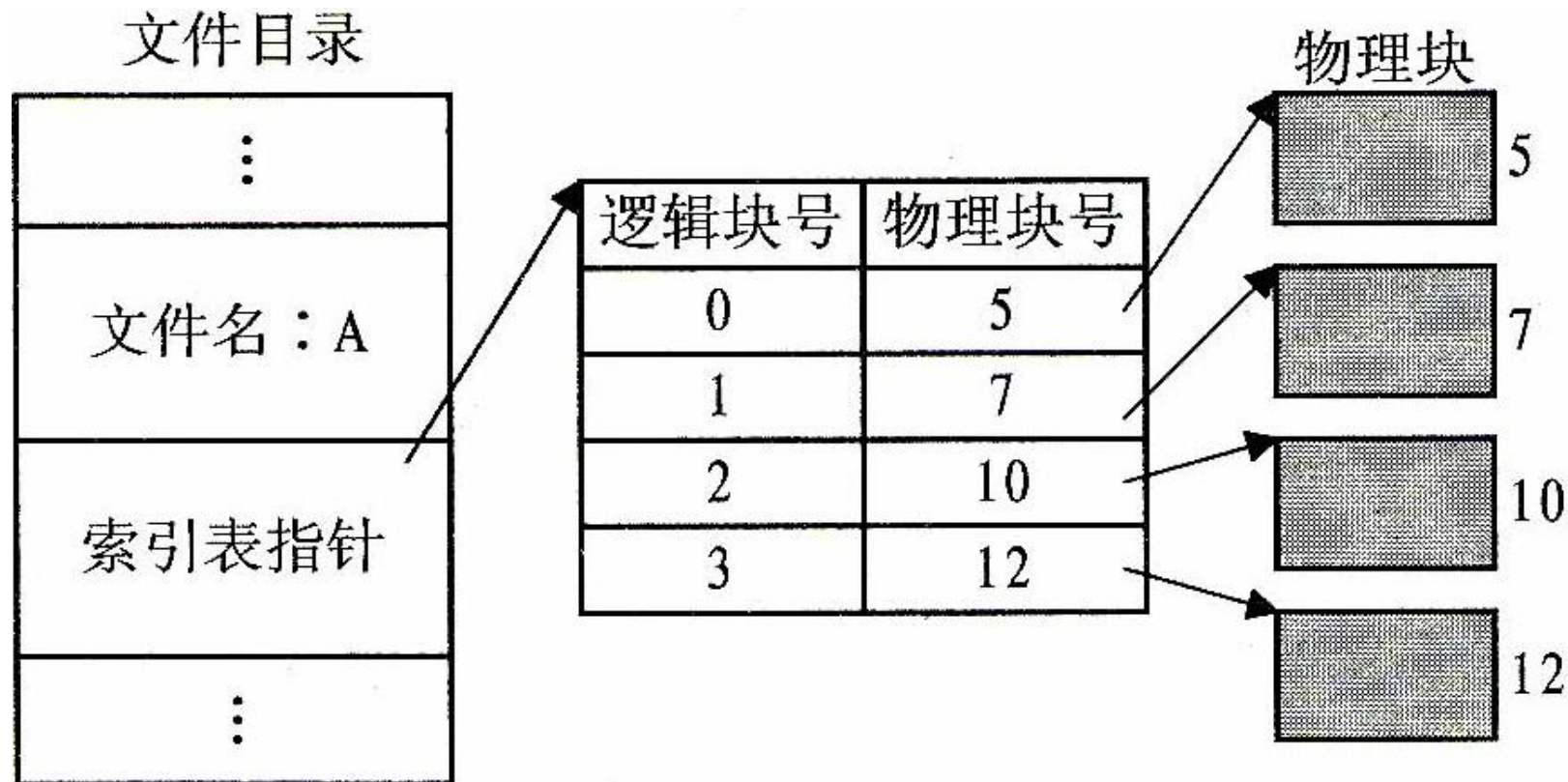
# 思考

- 什么叫FAT16、FAT32？
- 在FAT16中一簇最大64个扇区，为什么FAT16能管理的磁盘分区为2G？
- FAT32同FAT16相比有什么优点？
- 对于FAT16文件系统，若一个磁盘分区的大小为512M，问一个簇最少要为多少个扇区？
- 簇是大点好，还是小点好？

# 索引文件

- 为文件建立一张索引表，每个记录设置一个表项。索引表按记录关键字排序，本身是顺序文件。在对索引文件进行检索的时候，首先按照顺序文件检索方法查找索引表，从中找到相关表项，然后直接访问该记录。
- 当文件较大的时候，索引表也会较大。如果索引表的大小超过一个物理块，索引表的存取就成为新问题。一种较好的办法是采取间接索引，甚至多重索引。
- 索引文件既可以满足文件动态增长的要求，也可以较方便和迅速的实现随机存取。既适用顺序存取也适合随机存取。缺点是索引文件本身增加了存储开销，而且多重索引的访问时间开销也较大。

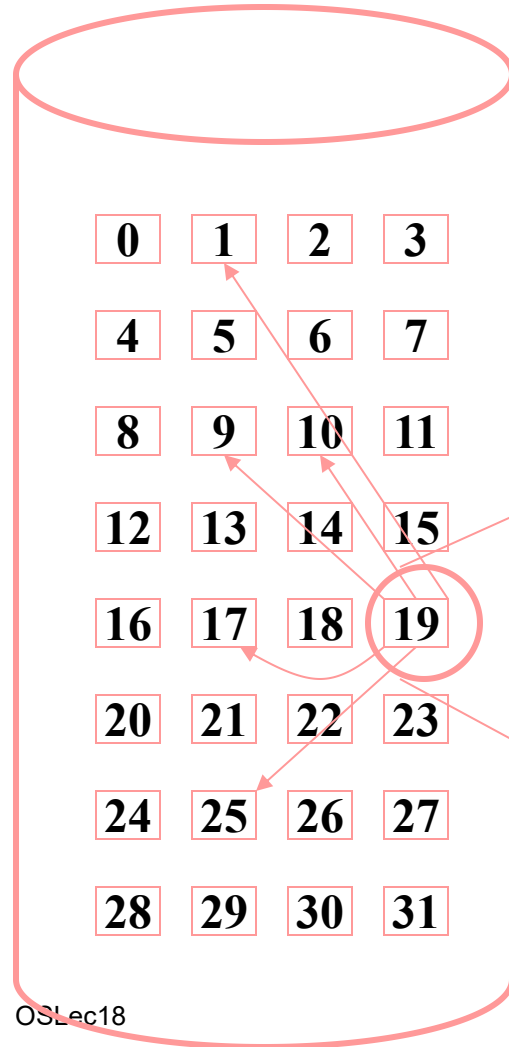
## 索引文件：



## 文件目录

文件名	索引表地址
-----	-------

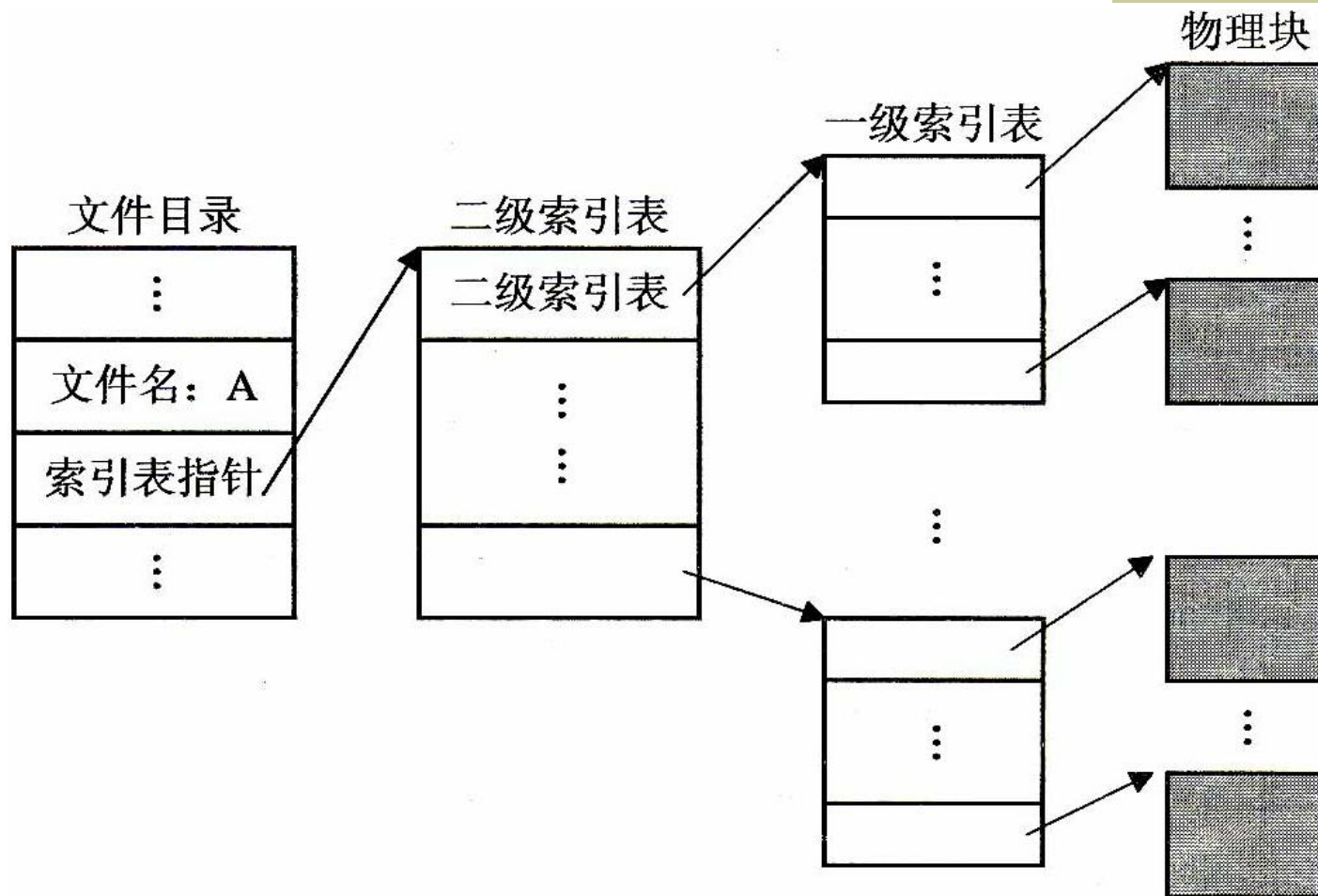
Jeep	19
------	----



19

9  
16  
1  
10  
25  
-1  
-1  
-1

# 二级索引文件





# 索引结构的特点

## ■ 优点：

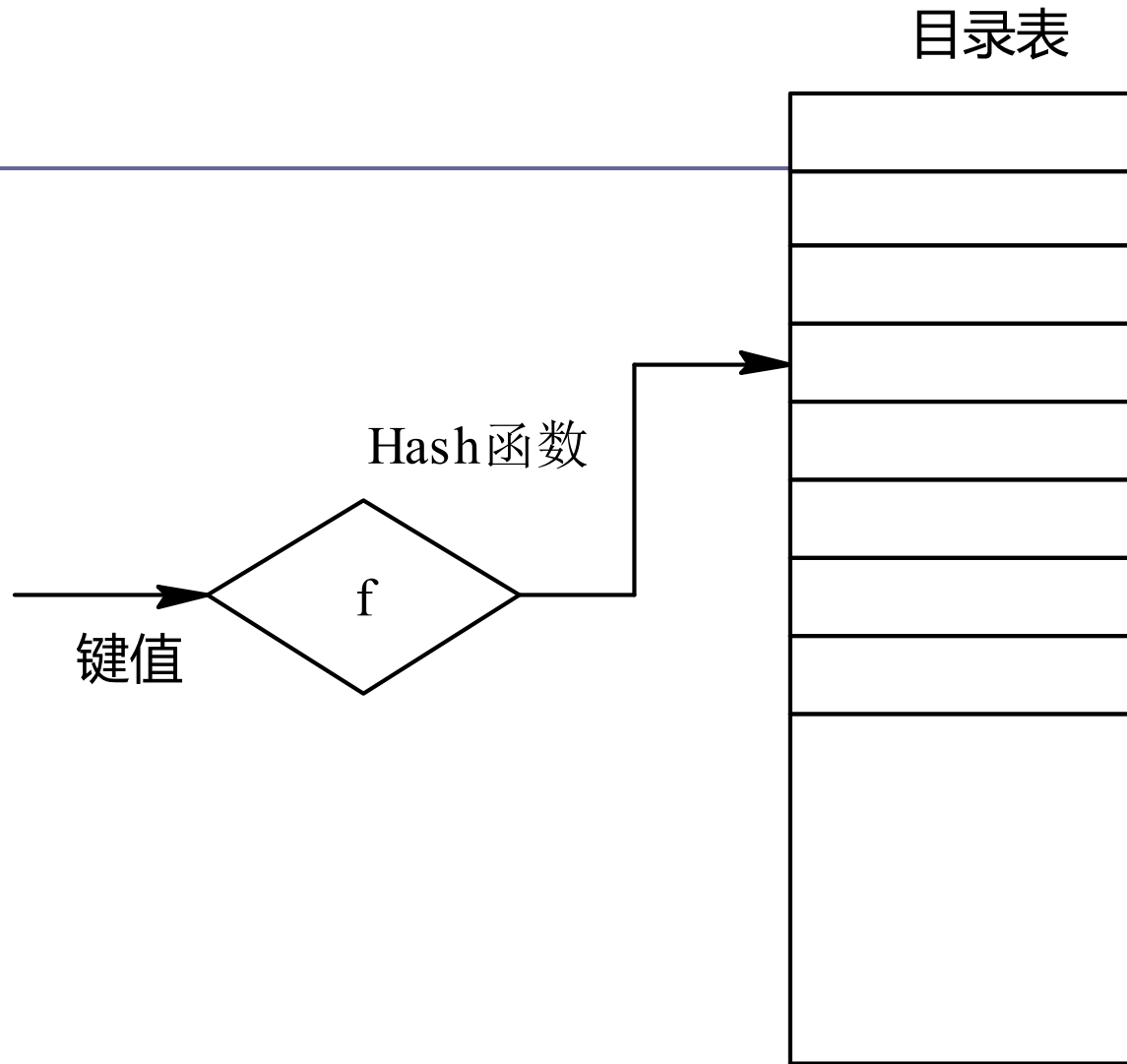
- 保持了链接结构的优点，又解决了其缺点：
- 即能顺序存取，又能随机存取
- 满足了文件动态增长、插入删除的要求
- 能充分利用外存空间

## ■ 缺点：

- 索引表本身带来了系统开销，如：内外存空间，存取时间

# 直接文件和哈希文件

- 对于直接文件，则可根据给定的记录键值，直接获得指定记录的物理地址。换言之，记录键值本身就决定了记录的物理地址。这种由记录键值到记录物理地址的转换被称为键值转换(Key to address transformation)。组织直接文件的关键，在于用什么方法进行从记录值到物理地址的转换。
- 记录大小相同。由主文件和溢出文件组成。记录位置由哈希函数确定。检索时给出记录编号，通过哈希函数计算出该记录在文件中的相对位置。访问速度快，但在主文件中有空闲空间。



## Hash文件的逻辑结构

# 物理结构比较

## ■ 从查寻时间来看

- 连续文件最快, 索引文件和文件映照次之, 串联文件最慢

## ■ 从空间开销来看

- 连续文件不需要额外的空间开销; 串联文件的每个物理块中需要存放链接字; 文件映照需存放文件映照表; 索引文件为每个文件建立一张索引表。

## ■ 从适宜设备和存取方法来看

- 连续文件可用于磁带和磁盘; 串联文件、索引文件只适用于磁盘; 串联文件只适合顺序存取; 而索引文件和磁盘上的连续文件, 除了能进行顺序存取外, 也能实现随机存取。

## ■ 从文件增删来看

- 连续文件不能动态增长, 其他三种都可较容易实现文件的动态改变。

## 5.2.3 文件的存取方法

- **存取方法**：读写文件存储器上的一个物理块的方法，通常有3种方法。
  - **顺序存取法**：严格按文件信息单位排列的顺序依次存取。
  - **直接存取法**：也称随即存取法，每次存取操作时必须先确定存取的位置。对流式文件或定长记录的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。
  - **按键存取法**：文件的组织按照逻辑记录中的某个数据项的内容来存放，根据记录内容进行存取。

■ 存储设备、文件物理结构和存取方法的关系：

存储设备	磁带	磁盘		
文件结构	连续	连续	串联	索引
存取方法	顺序	顺序、随机	顺序	顺序、随机