



## 第8章 文件系统

---



# 文件系统

---

- 介质需求
  - 内存：通常太小，且为易失存储，无法永久存储数据和程序
  - 非易失存储(Non-volatile)：磁盘、固态硬盘、磁带等
- 非易失存储的管理
  - 文件：由创建者定义的相关信息集合
  - OS要提供映射机制：文件→物理设备
- 文件系统
  - 是OS对于用户最明显的部分
  - 方便对OS与用户的数据与程序进行在线存储和访问
  - 是OS提供的关于信息存储的统一逻辑视图
  - 是对OS中各类大容量非易失存储资源的抽象
  - 是操作系统中与管理文件有关的软件和数据集合



# 文件系统

---

- 文件系统的组成
  - 管理文件所需的数据结构：
    - 如目录：用于组织系统内所有文件并提供文件信息
  - 被管理的文件集合
    - 每个文件存储相关数据
  - 相应的管理软件
    - 存储设备空间的分配与回收
    - 目录与文件管理
    - 用户接口
    - 共享与保护



---

## 第8章 文件系统

# 8.1 文件的概念



# 1.文件的定义

---

## ■ 文件

- 是记录在外部存储上，具有名字的，一组相关记录的集合，文件由若干记录组成。
- 记录是一些相关数据项的集合。
- 数据项是数据组织中可以命名的最小逻辑单位。
- 从用户角度看，文件是逻辑外存的最小分配单元。
- 是文件系统的基本数据单位，即数据保存在文件中



## 2. 文件分类

---

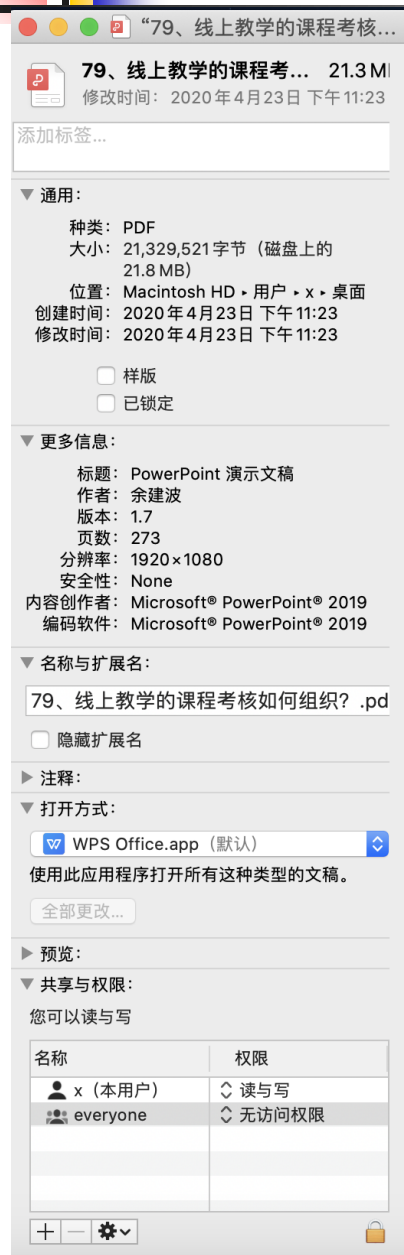
- 文件的分类
  - 按用途
    - 系统文件、库文件、用户文件
  - 按保护级别
    - 只读文件、读写文件、执行文件、不受保护文件
  - 按信息流向
    - 输入文件、输出文件、输入输出文件
  - 按数据形式
    - 源文件、目标文件、可执行文件



## 3. 文件属性

- 文件的属性，因OS而异
  - 名称：符号文件名，唯一的，且按照方便人类读取的形式保存
  - 标识符：文件系统中对应文件的唯一标签，数字形式为主，不易于人类读取
  - 类型：支持不同类型文件的系统所需要
  - 位置：在设备上的文件位置指针
  - 大小：当前文件的大小，以及可能允许的最大尺寸
  - 保护：访问控制信息确定谁能读、写和执行
  - 时间、日期、用户标识：文件创建、最后修改、最后使用等信息可被保存，用于保护、安全和使用监控
- 所有关于文件的信息都保存于目录结构中，并  
在外存上长期存储

# Mac OS X中的文件信息



## ■ 文件的安全属性

- 用于防止文件被破坏，称为文件保护。包括：
  - 问题1：防止系统崩溃所造成的文件破坏；
  - 问题2：防止文件所有者和其他用户有意或无意的非法操作所造成的文件不安全性。
- 防止问题一：
  - 定时转储
  - 多副本
- 防止问题二
  - 访问控制——基本思想是建立三元组：（用户、对象、存取权限）





## 4. 文件操作

- 文件是一个抽象数据类型，定义必须的操作
  - Create：
    - 在文件系统中找到空间，在目录中创建条目
  - Write：
    - 维护写位置指针，并通过系统调用向相应位置写入内容
  - Read：
    - 维护读位置指针，并通过系统调用读取相应内容
  - Seek：搜索目录条目，并设置文件指针为给定值
  - Delete：给定文件，删除文件，释放外存空间，删除目录条目
  - Truncate：删除文件内容，但保留文件属性
- 为避免频繁搜索文件，系统提供统一的系统调用
  - Open()：在磁盘上搜索目录结构，并将所搜索到的文件目录内容移入内存
  - Close ()：将对应文件的目录内容移出内存，回到磁盘目录结构中



# 打开文件Open()

- 在打开文件过程中需要维护许多的数据
  - 打开文件表(Open-file Table):
    - 跟踪打开的文件
    - 当涉及多进程访问环境，系统会维护进程表（局部）+系统表（全局）
  - 文件指针：
    - 指向最近的读写位置，每个进程都维护自己的文件指针
  - 打开文件计数：
    - 文件被打开次数的计数，允许当最后一个进程关闭文件时从文件打开表中移除相关数据
  - 文件的磁盘位置：
    - 便于数据访问，避免每次修改操作时从磁盘中获取相关信息
  - 访问权限：
    - 为每个进程提供访问模式



# 打开文件锁

---

- 由系统提供
  - 类似读者-写者模型
  - 共享锁—类似读锁：提供多进程并发读
  - 独占锁—类似写锁
- 强制机制or建议机制
  - 强制机制Mandatory：当锁定后拒绝其他所有请求锁的操作
  - 建议机制Advisory：进程能够发现锁状态并决定后续操作

## 5. 文件类型

- OS需要识别和支持文件类型，方便以合理方式操作文件。

file type	usual extension	function
executable	exe, com, bin or none	ready-to-run machine-language program
object	obj, o	compiled, machine language, not linked
source code	c, cc, java, pas, asm, a	source code in various languages
batch	bat, sh	commands to the command interpreter
text	txt, doc	textual data, documents
word processor	wp, tex, rtf, doc	various word-processor formats
library	lib, a, so, dll	libraries of routines for programmers
print or view	ps, pdf, jpg	ASCII or binary file in a format for printing or viewing
archive	arc, zip, tar	related files grouped into one file, sometimes compressed, for archiving or storage
multimedia	mpeg, mov, rm, mp3, avi	binary file containing audio or A/V information



## 6. 文件结构

---

- 文件结构指文件的组织形式，文件有两种形式的结构：
  - 逻辑结构：
    - 又称文件组织，是**从用户观点出发所看到的文件组织形式**。
  - 物理结构：
    - 又称文件的存储结构，是**文件在外存上的存储组织形式**。它与存储设备特性、外存分配方式有关。
    - 数据块是逻辑存储单元，扇区是物理存储单元。



# (1) 文件的逻辑结构

---

- 文件的逻辑结构可分为两类：
  - 记录式文件：是一种有结构文件，由一组相关记录组成。又分为：
    - 等长记录文件：又称定长记录文件，是指文件中所有记录的长度相等。
    - 变长记录文件：是指文件中各记录长度不相等。
  - 流式文件：是一种无结构文件，由字符序列构成。



# 记录式文件的组织方式

- 根据用户和系统管理的需要，可以采用多种方式来组织记录式文件：
  - ① 顺序文件：一组记录按关键字的大小顺序排列所形成的文件。其中的记录通常是定长的。
  - ② 索引文件：为文件设置一个索引表，文件中的每个记录在索引表中有一个表项，用于存放记录的存放地址及长度。
  - ③ 索引顺序文件：是前两者的结合。它将**顺序文件**中的所有记录**分成若干组**，为顺序文件建立一张**索引表**，为**每组中的第一个记录建立一个索引项**，其中含有该记录的键值和指向该记录的指针。



## (2) 文件的物理结构

- 文件的物理结构与存储介质的存储特性及外存分配方式有关。
- 文件存储设备通常划分为大小相等的物理块，物理块是分配及传输信息的基本单位。**物理块的大小与设备有关**，但**与逻辑记录的大小无关**，因此一个物理块中可以存放若干个逻辑记录，一个逻辑记录也可以存放在若干个物理块中。
- 为有效地利用外存设备和便于系统管理，**一般**也把文件信息划分为与物理存储块**大小相等**的逻辑块。





# 物理结构类型

---

- 常见的文件物理结构有以下几种形式：
  - 顺序结构
  - 链接结构
  - 索引结构



# 顺序结构

---

- 顺序结构又称连续结构，它将一个在逻辑上连续的信息存放在外存连续的物理块中。
- 以顺序结构存放的文件称为顺序文件或连续文件。
- 特点：顺序存取速度较快；对等长记录文件**支持随机访问**。但因要求连续存放，**会产生碎片**，同时也不利于文件的动态扩充。



# 链接结构

---

- 链接结构又称串联结构，它将一个逻辑文件的信息存放在外存**不连续物理块**中，且在每个物理块中设置一个指向下一个物理块的指针。
- 采用链接结构存放的文件称为链接文件或串联文件。
- 特点：可**解决碎片问题**，便于文件动态增长。但只能**顺序访问**，因而查找效率较低，指针占用存储空间。



# 索引结构

---

- 索引结构：将一个逻辑文件的信息存放于外存的若干个物理块中，并为每个文件**建立一个索引表**，其中的每个表目存放文件信息所在的逻辑块号和与之对应的物理块号。
- 采用索引结构存放的文件称为**索引文件**。
- 特点：既可以顺序访问也可以随机访问，但增加了存储空间开销，且要**两次访问外存**。



---

第8章 文件系统

## 8.2 文件访问方法



# 文件访问方法

---

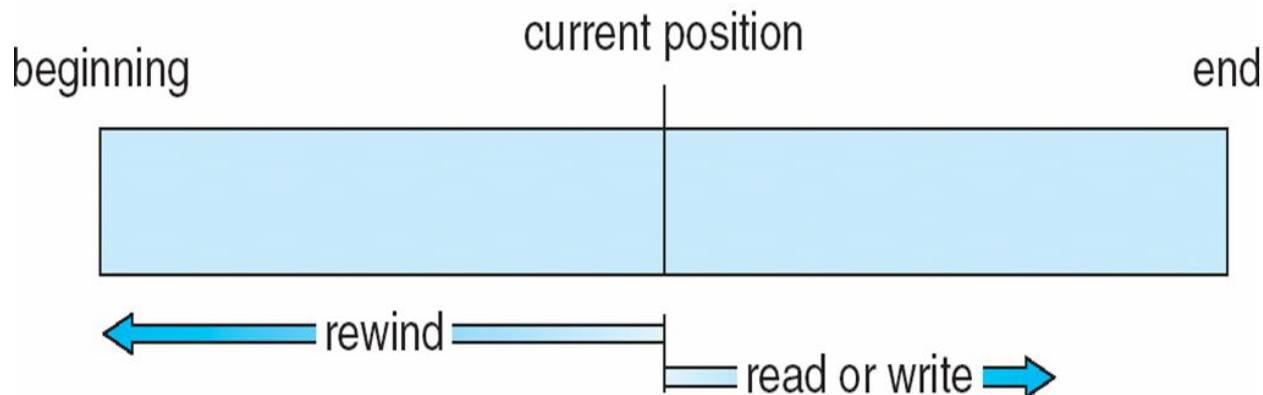
- 常用的文件访问方法有：
  - 顺序访问法
  - 直接访问法
  - 索引访问法

# (1) 顺序访问法

- 也称顺序存取法，是按照文件信息的逻辑顺序依次存取。
- 在记录式文件中，顺序存取反映为按记录的排列顺序来存取；在流式文件中，顺序存取反映为当前读写指针的变化。
- 对定长记录的顺序文件，若知道当前记录地址，则很易确定下一个记录地址。

$$rptr = rptr + L$$

- 其中L为文件记录的长度，rptr为读写指针。



## (2) 直接访问法

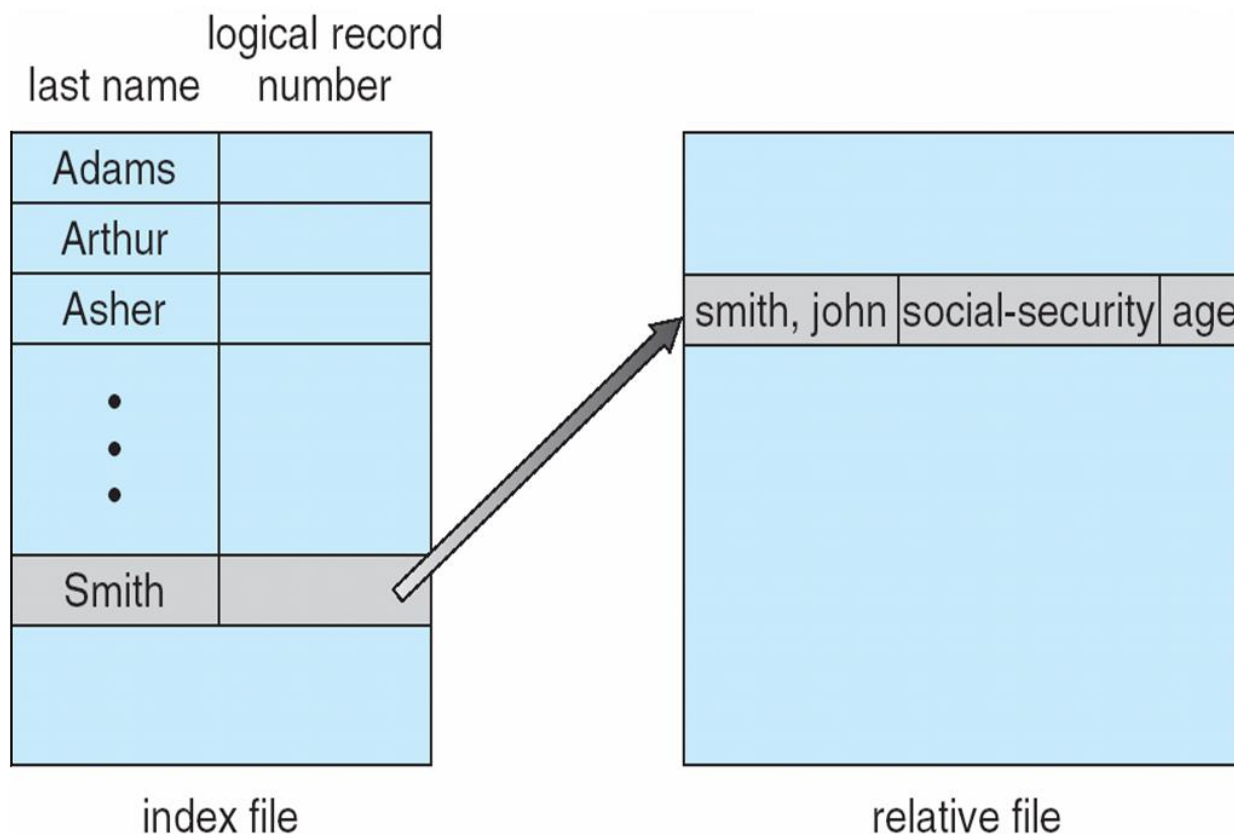
- 又称随机存取法，允许按任意顺序存取文件中的任何一个物理记录。
- 对于定长记录的顺序文件，若知道文件的起始地址和记录长度，则第*i*个记录 ( $i = 0, 1, 2, \dots$ ) 的首地址为
  - $$\text{rptr} = \text{addr} + i \times L$$
  - 其中addr是该文件的首地址，L为记录长度。
- 直接访问法可模拟顺序访问法

sequential access	implementation for direct access
<i>reset</i>	$cp = 0;$
<i>read next</i>	$\text{read } cp;$ $cp = cp + 1;$
<i>write next</i>	$\text{write } cp;$ $cp = cp + 1;$



### (3) 索引访问法

- 索引访问法实质上也是直接存取法，它根据文件记录中数据项（通常称为关键字），经过某种方法计算处理，转换成相应的物理地址后进行存取。
- 也可通过关键字建立索引，按照索引来进行访问。





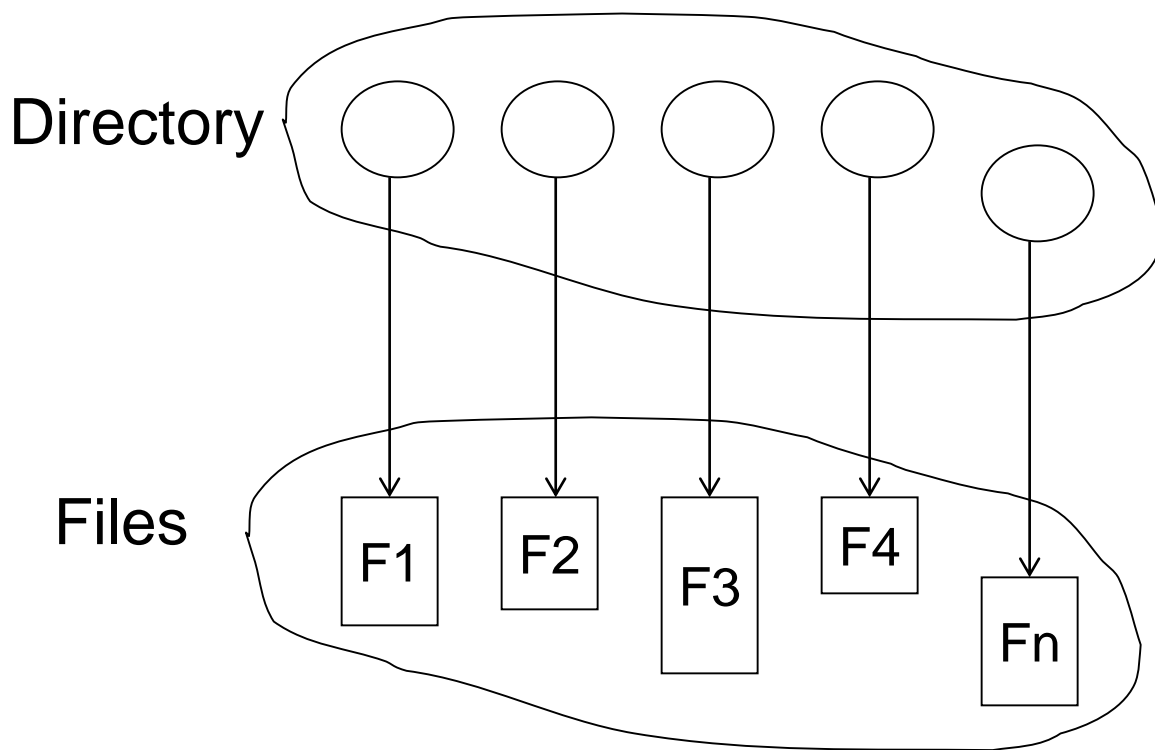
---

第8章 文件系统

## 8.3 磁盘和文件目录结构

# 磁盘和文件目录结构

- 计算机系统中的文件种类繁多，数量庞大，为了有效地管理这些文件，方便用户查找所需的文件，应对它们加以适当的组织。文件的组织可以通过目录来实现。



目录结构和文件都驻留在磁盘上



# 1. 磁盘结构的几个常识

## ■ 分区 partitions

- 磁盘能被分割成分区

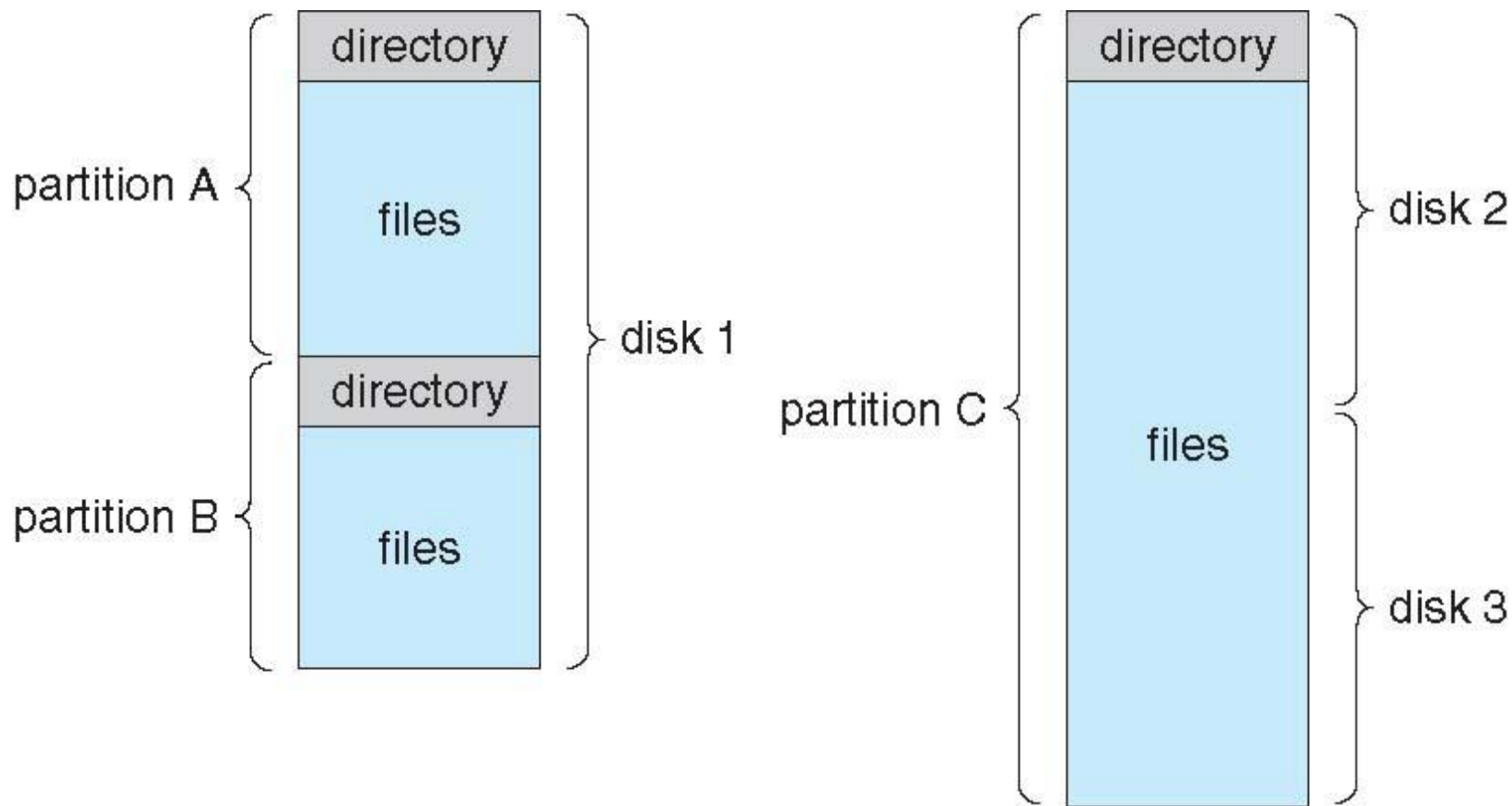
## ■ 磁盘或分区

- 能通过RAID技术实现容错
- 在没有文件系统时，被作为raw形态使用
- 当格式化formatted后，磁盘/分区则具有一个文件系统

## ■ 卷， volume： 包含文件系统的分区实体

- 卷可以是设备一部分、整个设备，或多个设备组成的RAID集合。每个卷可作为虚拟磁盘，卷可以存储多个OS
- 每个卷都包含系统中的文件信息，信息保存在设备目录(Device Directory)、卷目录表(Volume Table of Content)

# 典型的文件系统结构





# 文件系统类型

- 本课程仅考虑通用文件系统
- 但，还有一些专用文件系统，如Solaris中：
  - tmpfs – 临时文件系统，建立于易失内存中，用于快速I/O，重启或系统崩溃后即消失
  - objfs – 虚拟文件系统，本质上是一个内核接口用于调试器访问内核符号
  - ctfs – 合约文件系统，用于管理后台进程的启动与运行
  - lofs – 环回文件系统，允许一个文件系统代替另一个被访问
  - procfs – 用于将进程信息作为文件信息呈现
  - ufs, zfs – 通用文件系统



## 2. 目录上的基本操作

---

- 目录可视为符号表，可将文件名称转成目录条目
- 目录上的基本操作
  - 搜索文件
  - 创建文件
  - 删除文件
  - 查看目录
  - 重命名文件
  - 遍历文件系统



# 目录组织的目标

---

- 高效性 —— 快速定位一个文件
- 命名 —— 方便用户使用
  - 两个用户对于不同文件可以起相同的名字
  - 相同的文件也可以有多个不同的名字
- 分组 —— 文件可根据不同属性进行逻辑分组





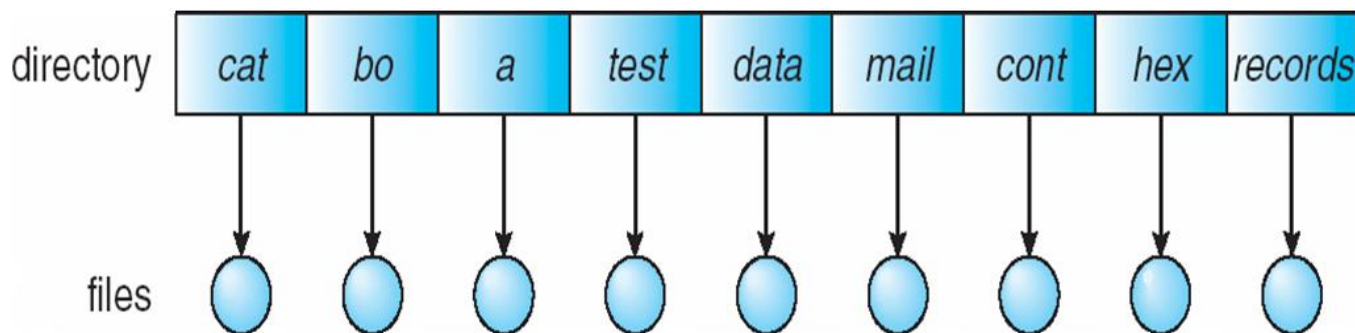
## 3. 目录结构

---

- 常用的文件目录结构有：
  - 单级目录结构
  - 二级目录结构
  - 多级目录 结构

# (1) 单级目录结构

- 单级目录结构又称一级目录结构。在这种结构中，整个文件系统只建立一个目录，从实现上，每个文件占据一个目录的项。





# 单级目录结构的操作

## ■ 建立

- 先应确定该文件名在目录中是否惟一，若惟一则找出一个空表目，将新文件的相关信息填入其中。

## ■ 删除

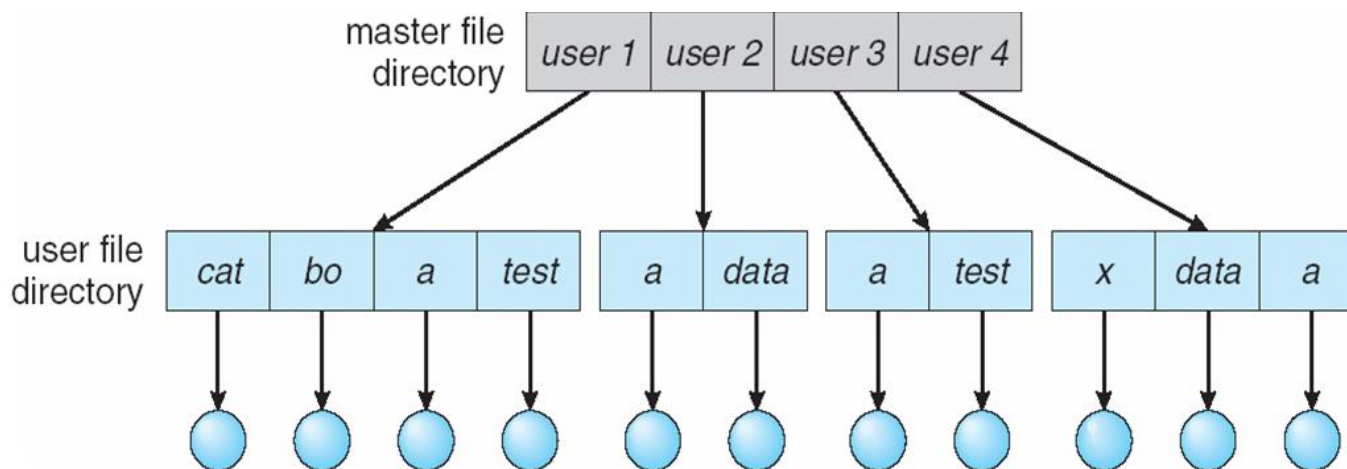
- 先从目录表中找到文件的目录项，从中找到该文件的物理地址，对其占用空间进行回收，然后再清除其所占用的目录项。

## ■ 特点

- 易于实现，管理简单；
- 当文件数增多时，查找时间较长
- 不支持重名问题

## (2) 二级目录结构

- 二级目录结构将文件目录分成：
  - 主文件目录MFD：记录用户名及相应用户文件目录所在的存储位置。
  - 用户文件目录UFD：记录该用户文件的有关信息。
  - 查找路径：只面向用户自己的UFD搜索



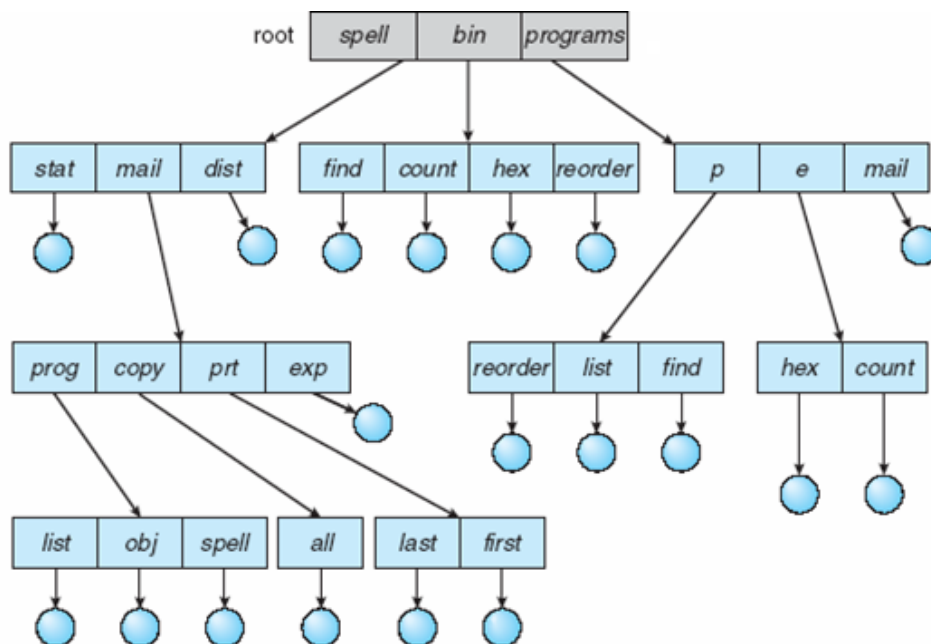


# 二级目录结构的特点

- 优点：有效实现了用户之间的隔离
  - 重命名问题
    - 不同MFD下，文件名可重复
    - 相同UFD下，文件名不可重复
- 缺点：隔离不适合合作共享
  - 用户文件的共享：需了解完整路径，支持额外语法
    - /MFD/UFD/Filename,
    - [partition name]:[directory name]:[Filename]
  - 系统文件的共享：加载器、编译器、库函数、工具程序等
    - 访问方式：给定一条命令，即文件名称，而非整条路径
    - 方法1：将所有系统文件拷贝到每个UFD
    - 方法2：建立特殊目录，建立搜索路径

### (3)多级目录结构

- 多级目录结构是二级目录结构层次关系的推广，也称为**树形目录结构**。
- 在多级目录结构中，第一级目录称为根目录（树根），目录树中的非叶节点均为目录文件（又称子目录），叶结点为文件。





# 多级目录结构 (Cont.)

---

- 路径名：
  - 是一个字符串
  - 从根目录出发到所找文件的通路上所有各级子目录名和该文件名用分隔符连接起来构成
- 绝对路径
  - 从根目录出发的路径, /spell/mail/prog/filename
- 当前目录
  - 由用户在一定时间内指定某个目录为当前目录, 或称**工作目录**。
  - `cd /spell/mail/`



# 多级目录结构 (Cont.)

## ■ 相对路径

- 进程对各文件的访问相对于当前目录进行 ./prog/filename
- 从当前目录出发到所找文件的通路上的所有目录名与数据文件名用分隔符连接起来而形成

## ■ 有两个特殊目录：

- “..”：表示给定目录的父目录
- “.”：表示当前目录

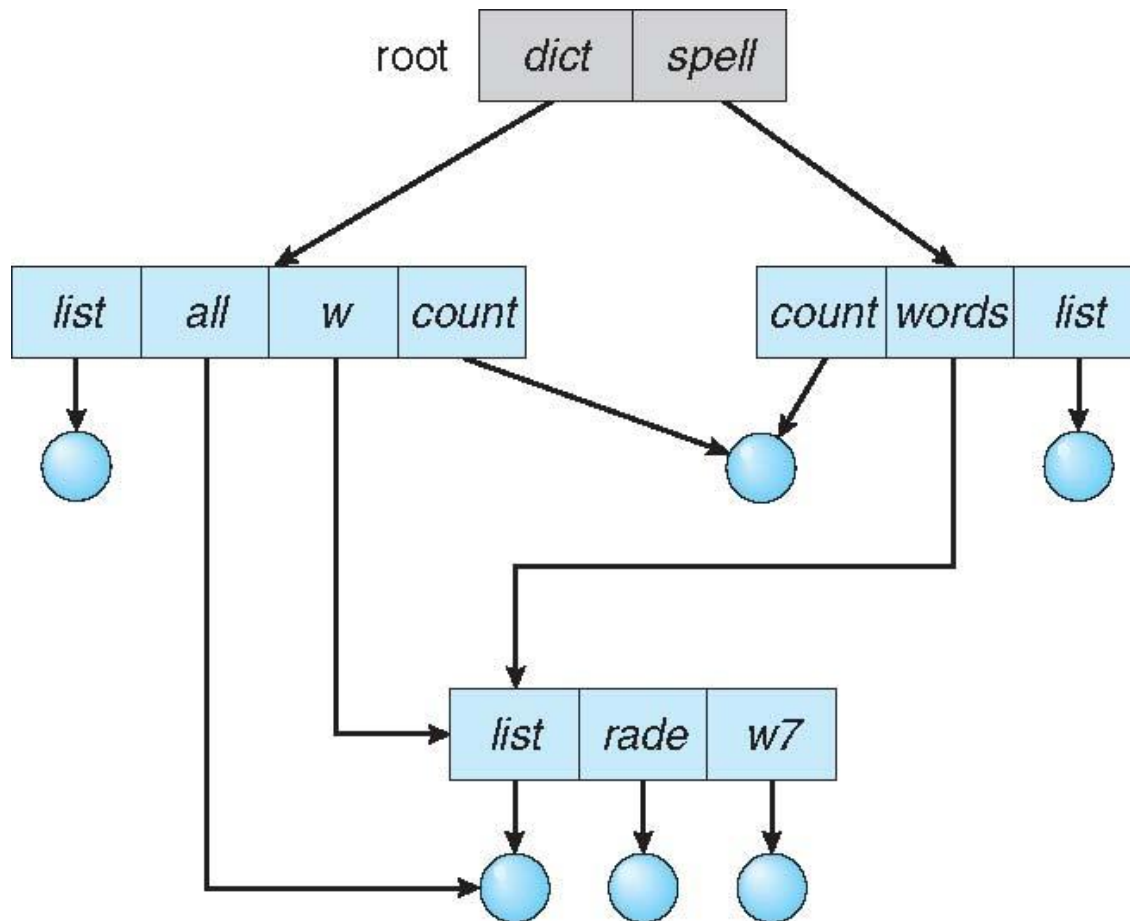
## ■ 删除策略：如何删除目录

- 不为空不能删除，用户必须先删除其中内容：MS-DOS
- 提供选择，可删除子目录与目录下的文件：UNIX rm



## (4) 无环图目录

- 树型不允许共享文件和目录，无环图允许共享子目录和文件



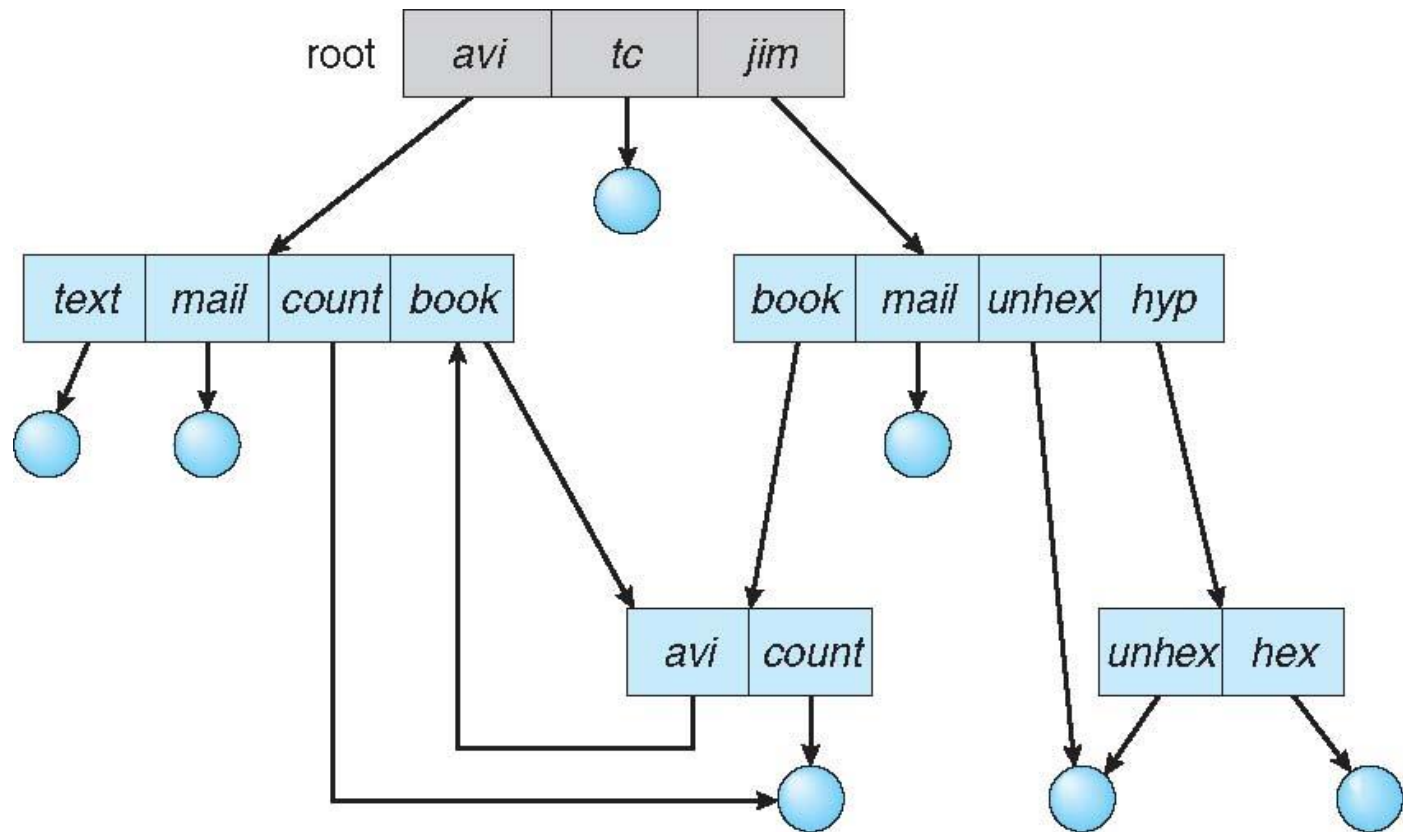


# 无环图目录 (Cont.)

- 支持别名
- 共享方法：Link,
  - 链接：创建一个指向已存在文件的指针
  - 解析：沿指针定位文件
- 共享文件的删除
  - 问题：
    - 容易产生指针悬挂：从/dict/conunt下删除文件
  - 方法一：对于符号链接实现的共享
    - 只删除链接指针，不删除原始文件
    - 若原文件被删除了，指针依然悬挂，则需要检查机制
  - 方法二：保留文件，直到删除所有引用
    - 需引入引用列表，或引用计数

## (5) 通用图目录

- 当对已存在的树状结构不断增加链接，就形成了图结构



- 如何避免循环搜索，限制搜索深度
- 如何删除：建立垃圾收集机制 or 遍历目录时规避链接



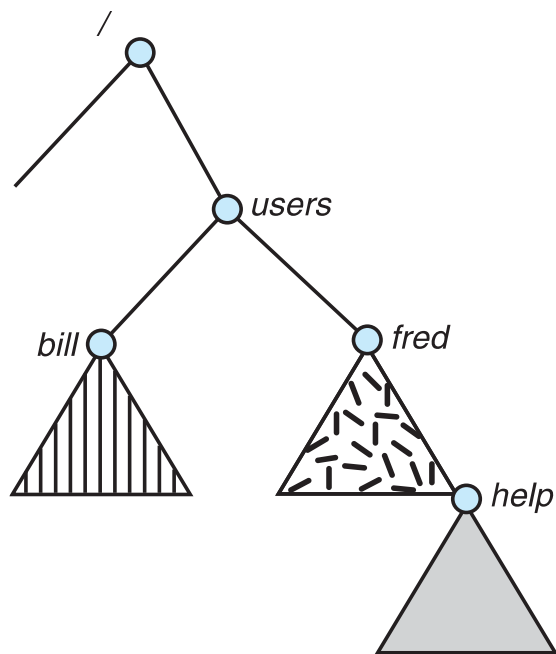
---

第8章 文件系统

## 8.4 文件系统的安装

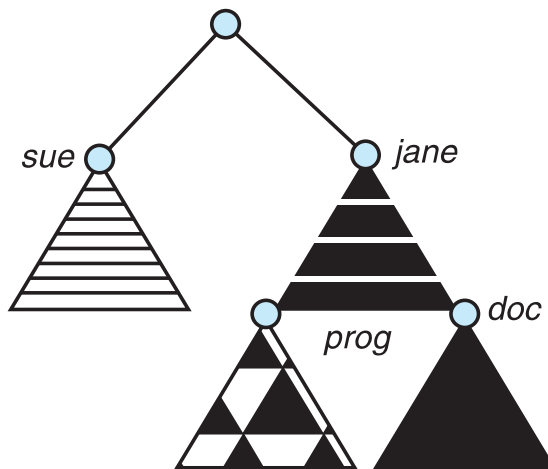
# 文件系统的安装

- 一个文件系统在被访问前需要安装挂载, mount



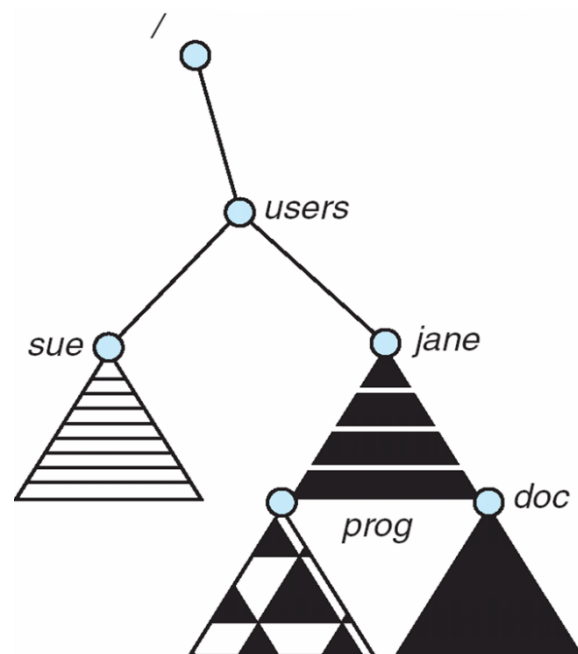
(a)

已有文件系统



(b)

未安装的卷



安装点



---

第8章 文件系统

## 10.5 文件系统的共享



# (1) 多用户

---

- 目标：
  - 多用户系统共享文件
  - 文件共享时需要考虑保护机制
  - 对于分布式系统，文件可以通过网络共享，如NFS
- 对于一个多用户系统，具有：
  - UserID
  - GroupID
  - 文件Owner
  - 文件Group



## (2) 远程文件系统

- 使用网络实现系统间文件访问的方法
  - 手工程序实现，如FTP
  - 自动化的，无缝实现：分布式文件系统DFS
  - 半自动化的通过万维网实现
- Client-Server 模型允许客户端挂载远程文件系统
  - Server能够服务多个客户端
  - 客户端的身份识别是不安全或者复杂的
  - NFS：标准的 UNIX client-server 文件共享协议
  - CIFS：标准的 Windows 协议
  - 可实现将标准操作系统文件调用翻译为远程调用
- 分布式信息系统（分布式命名服务）
  - LDAP、DNS、NIS、Active Directory





## (2) 远程文件系统

- 所有的文件系统都包含故障模式Failure Mode
  - 本地文件系统故障原因：磁盘、目录结构、磁盘管理信息、磁盘控制器、线缆、适配器故障、目录卷丢失等。
  - 远程文件系统引入了新故障模式
    - 网络故障、服务器故障
- 故障恢复
  - 涉及每个远程请求状态的状态信息
  - 无状态DFS
    - 自身并不记忆历史文件打开情况，每次都进行一次完整的信息，这样当服务器崩溃时候，就很简单，不需要复杂的故障恢复，直接重新做就行了。
    - 除非文件系统已经远程安装，且已经打开文件，否则所有请求都不会发生，即执行请求中需要携带必要的文件信息。In NFS v3



## (3) 共享中的一致性语义

- 描述多个用户如何同时访问共享文件
  - 规定了一个用户数据被修改，何时另一个用户可见？
  - 进程间同步算法不适合解决此问题
    - 磁盘I/O和网络延迟所导致，复杂算法无法有效工作
  - Andrew File System (AFS)会话语义
    - 写操作，对于打开同一文件的其他用户不可见
    - 一旦文件关闭，写操作仅对后来打开的会话可见，已打开的文件实例不反映变化
  - Unix file system (UFS)
    - 写操作对其他打开同一文件的用户是立即可见的
    - 允许多用户在读写时共享文件指针
  - 不可变共享文件语义
    - 一旦一个文件由创建者声明为共享，他就不能被修改。
    - 不可变：名称不可重用、内容不可改变，即共享是只读的



---

第8章 文件系统

## 8.6 文件保护



# 文件保护

---

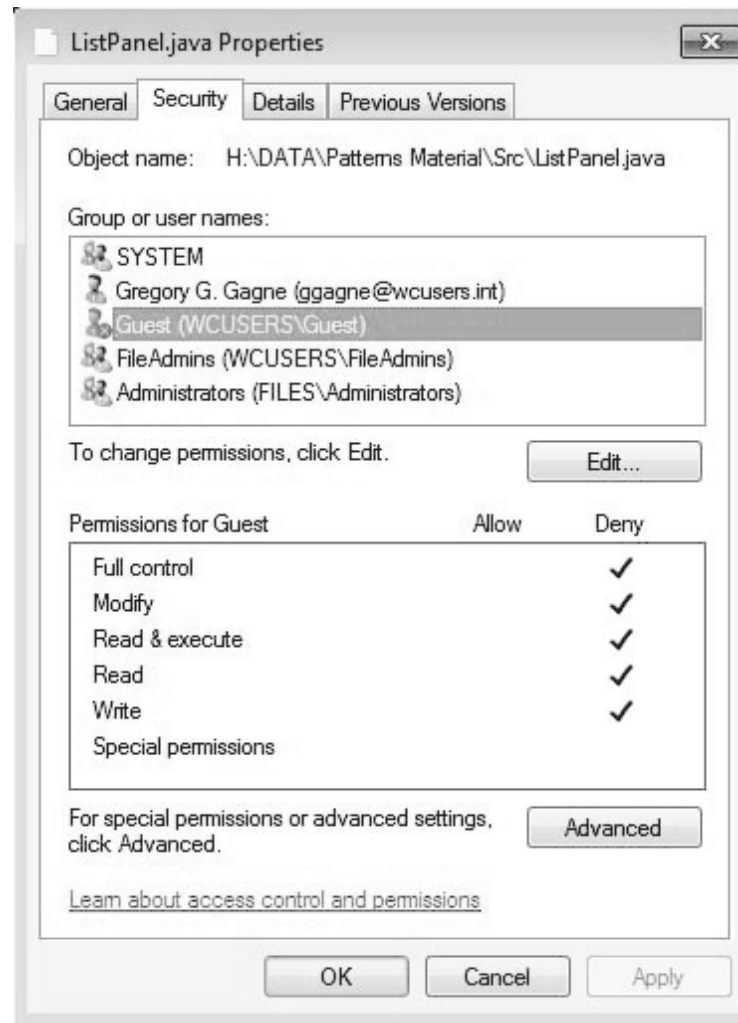
- 系统中的文件既存在保护问题，又存在访问控制问题。
- 文件保护是指
  - 避免文件拥有者或其他用户因有意或无意的错误操作使文件受到破坏。
- 文件访问控制是指
  - 文件本身不得被未授权的用户访问。
- 文件的保护与访问控制涉及：
  - 存取访问控制
  - 认证与授权机制
  - 密码机制

- 
- 访问模式: read, write, execute
  - Unix / Linux三类用户

a) <b>owner access</b>	7	⇒	RWX 1 1 1 RWX
b) <b>group access</b>	6	⇒	1 1 0 RWX
c) <b>public access</b>	1	⇒	0 0 1

- 修改权限: chmod
- 修改组关系: chgrp

# Windows 7 Access-Control List Management





# Unix 目录列表

---

-rw-rw-r--	1	pbg	staff	31200	Sep 3 08:30	intro.ps
drwx-----	5	pbg	staff	512	Jul 8 09:33	private/
drwxrwxr-x	2	pbg	staff	512	Jul 8 09:35	doc/
drwxrwx---	2	pbg	student	512	Aug 3 14:13	student-proj/
-rw-r--r--	1	pbg	staff	9423	Feb 24 2003	program.c
-rwxr-xr-x	1	pbg	staff	20471	Feb 24 2003	program
drwx--x--x	4	pbg	faculty	512	Jul 31 10:31	lib/
drwx-----	3	pbg	staff	1024	Aug 29 06:52	mail/
drwxrwxrwx	3	pbg	staff	512	Jul 8 09:35	test/



本章结束!

---