

# 第十章 文件系统接口

---

1. 解释文件系统功能
2. 描述文件系统接口

1. 文件概念
2. 访问方法
3. 目录结构
4. 文件系统安装
5. 文件共享
6. 文件一致性语义与保护

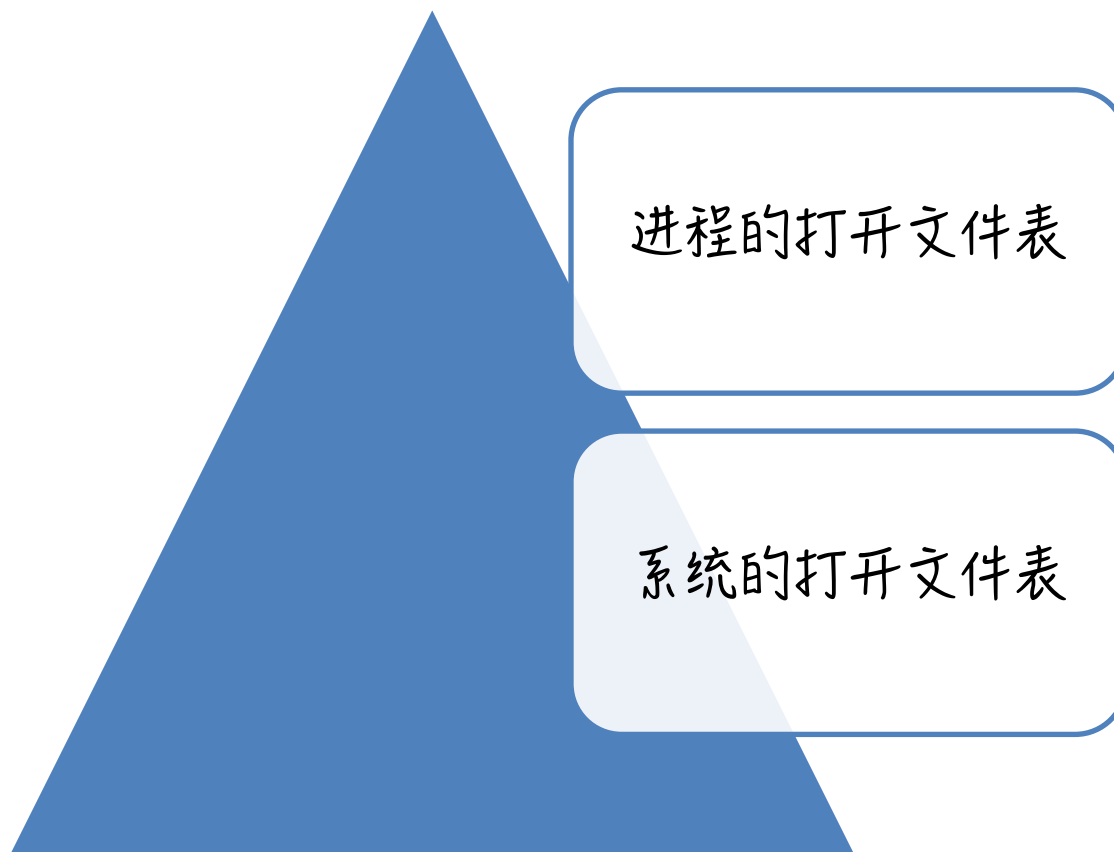
1. 操作系统提供信息存储的、统一的逻辑接口。
2. 文件是逻辑外存的最小分配单元，它可以大分为数据和程序。
3. 根据文件类型分为不同文件，如文本文件，可执行文件，图像文件，声音文件等。
4. 文件根据其类型具有不同的结构（格式）

1. 名称
2. 标识符
3. 类型
4. 位置 :设备上文件位置的指针
5. 大小
6. 保护:访问控制信息
7. 时间, 日期和用户标识

```
hbpark@hbpark-VirtualBox: ~  
-rw-r--r-- 1 hbpark hbpark 220 3月 20 2015 .bash_logout  
-rw-r--r-- 1 hbpark hbpark 3637 3月 20 2015 .bashrc  
drwx----- 22 hbpark hbpark 4096 3月 8 12:24 .cache  
drwx----- 3 hbpark hbpark 4096 3月 20 2015 .compiz  
drwx----- 24 hbpark hbpark 4096 3月 14 09:01 .config  
drwx----- 3 hbpark hbpark 4096 3月 27 2015 .dbus  
drwxr-xr-x 3 hbpark hbpark 4096 3月 13 22:33 Desktop  
-rw-r--r-- 1 hbpark hbpark 25 2月 23 14:57 .dmrc  
drwxr-xr-x 2 hbpark hbpark 4096 3月 20 2015 Documents  
drwxr-xr-x 3 hbpark hbpark 4096 4月 21 11:12 Downloads  
-rw-r--r-- 1 hbpark hbpark 8980 3月 20 2015 examples.desktop  
drwx----- 3 hbpark hbpark 4096 5月 5 09:30 .gconf  
drwxrwxr-x 4 hbpark hbpark 4096 3月 13 13:29 GNUstep  
-rw----- 1 hbpark hbpark 17740 5月 5 09:30 .ICEauthority  
-rw----- 1 hbpark hbpark 163 3月 13 23:22 .lessht  
drwxr-xr-x 3 hbpark hbpark 4096 3月 20 2015 .local  
drwx----- 4 hbpark hbpark 4096 3月 20 2015 .mozilla  
drwxr-xr-x 2 hbpark hbpark 4096 3月 20 2015 Music  
drwxrwxr-x 3 hbpark hbpark 4096 2月 23 17:31 .openoffice  
drwxr-xr-x 2 hbpark hbpark 4096 3月 20 2015 Pictures  
-rw-r--r-- 1 hbpark hbpark 675 3月 20 2015 .profile  
drwxr-xr-x 2 hbpark hbpark 4096 3月 20 2015 Public  
drwxrwxr-x 13 hbpark hbpark 4096 4月 17 20:24 src  
drwx----- 2 hbpark hbpark 4096 3月 20 2015 .ssh  
drwxr-xr-x 3 hbpark hbpark 4096 3月 13 22:33 Templates  
drwxrwxr-x 2 hbpark hbpark 4096 3月 2 22:27 test  
-rw-r----- 1 hbpark hbpark 5 5月 5 09:30 .vboxclient-clipboard.pid  
-rw-r----- 1 hbpark hbpark 5 5月 5 09:30 .vboxclient-display.pid  
-rw-r----- 1 hbpark hbpark 5 5月 5 09:30 .vboxclient-draganddrop.pid  
-rw-r----- 1 hbpark hbpark 5 5月 5 09:30 .vboxclient-seamless.pid  
drwxr-xr-x 2 hbpark hbpark 4096 3月 20 2015 Videos  
drwxr-xr-x 2 hbpark hbpark 4096 3月 30 22:05 .vim  
-rw----- 1 hbpark hbpark 14997 4月 21 10:15 .viminfo
```

1. 创建
2. 写文件
3. 读文件
4. 文件内重定位 (`seek`)
5. 删除文件
6. 截短文件 (`Truncate`) , 不改变文件属性,
  - I. If a FILE is larger than the specified size, the extra data is lost.
  - II. If a FILE is shorter, it is extended and the extended part (hole) reads as zero bytes.
  - III. `$truncate --help`

系统调用`open()`返回值是文件标识符（指针）



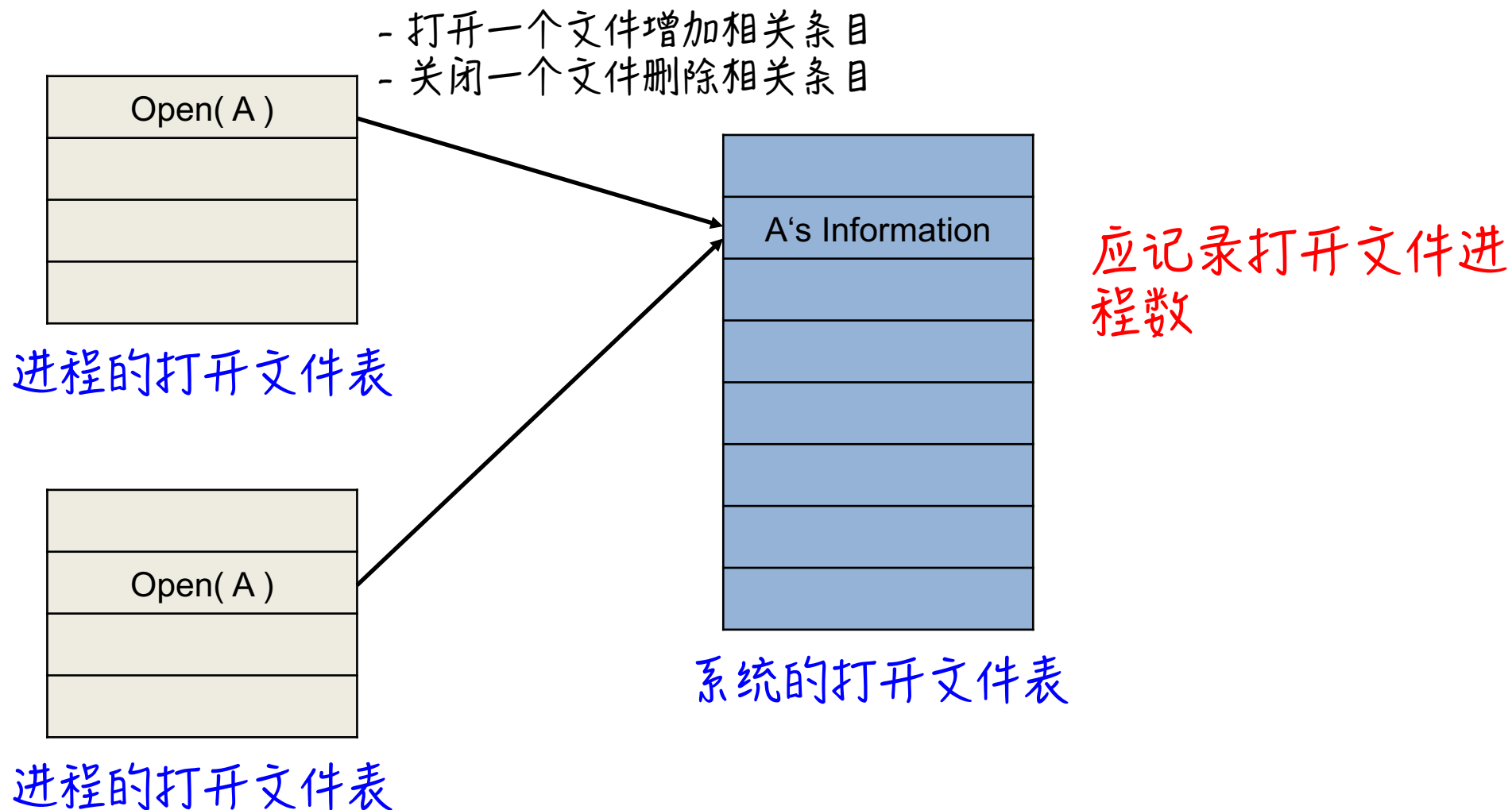
### 1. 进程的打开文件表

单个进程打开文件的信息表，每个条目内容包括文件指针，文件访问权限，记账信息等

### 2. 系统的打开文件表

整个系统内的打开文件信息表，每个条目内容包括文件在磁盘上的位置、访问日期、文件大小、文件打开次数等





### 1. 文件指针

- 文件最后读/写位置指针
- 每个进程打开的文件个数、指针

### 2. 文件打开计数器

- 一个文件被打开的次数，当文件关闭时，操作系统必须重设打开文件表的条目

### 3. 文件磁盘位置

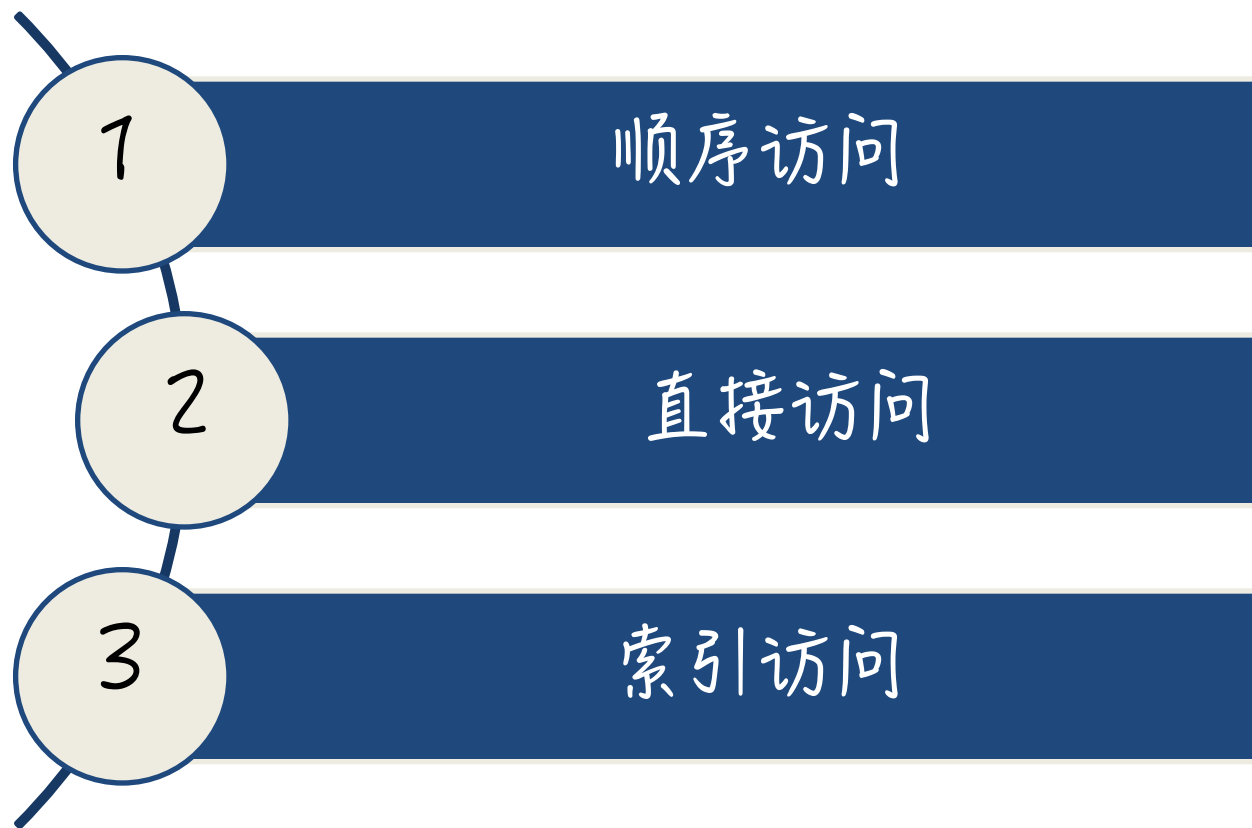
- 用于定位磁盘上文件位置信息保存在内存中

### 4. 访问权限

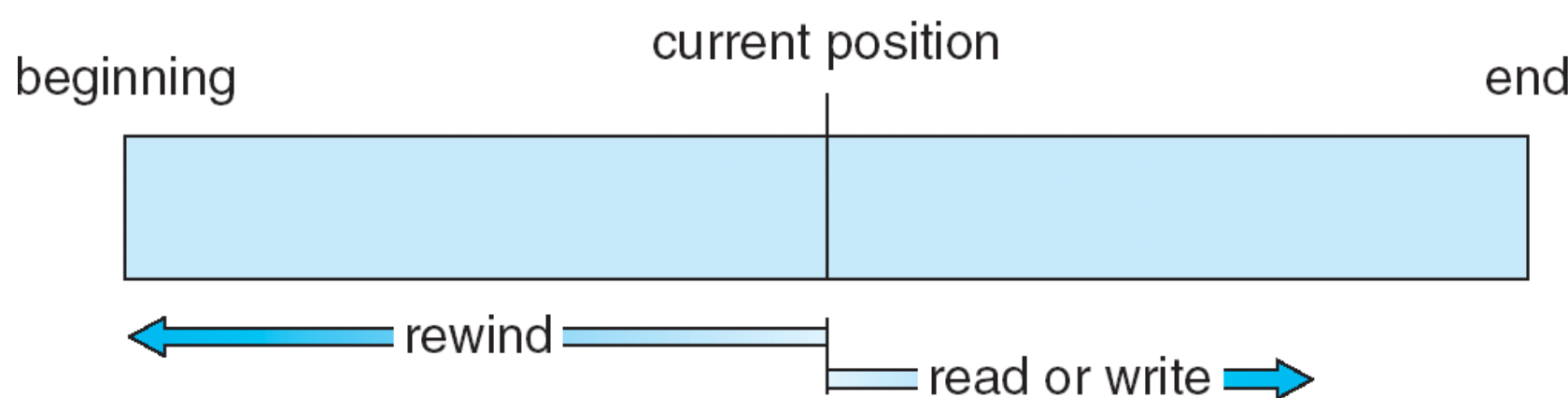
- 确定打开文件模式，以便操作系统允许或拒绝以后的I/O请求

1. 操作系统可提供**共享锁或专用锁**文件加锁机制
  - 共享锁 :读
  - 专用锁 :写
2. 操作系统可提供**强制或建议**文件加锁机制
  - 强制 :一旦被加锁，其他进程就不能访问
  - 建议 :根据程序员的设计需要确保适当的获取与释放锁

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



文件信息按顺序，一个记录接着一个记录地加以处理



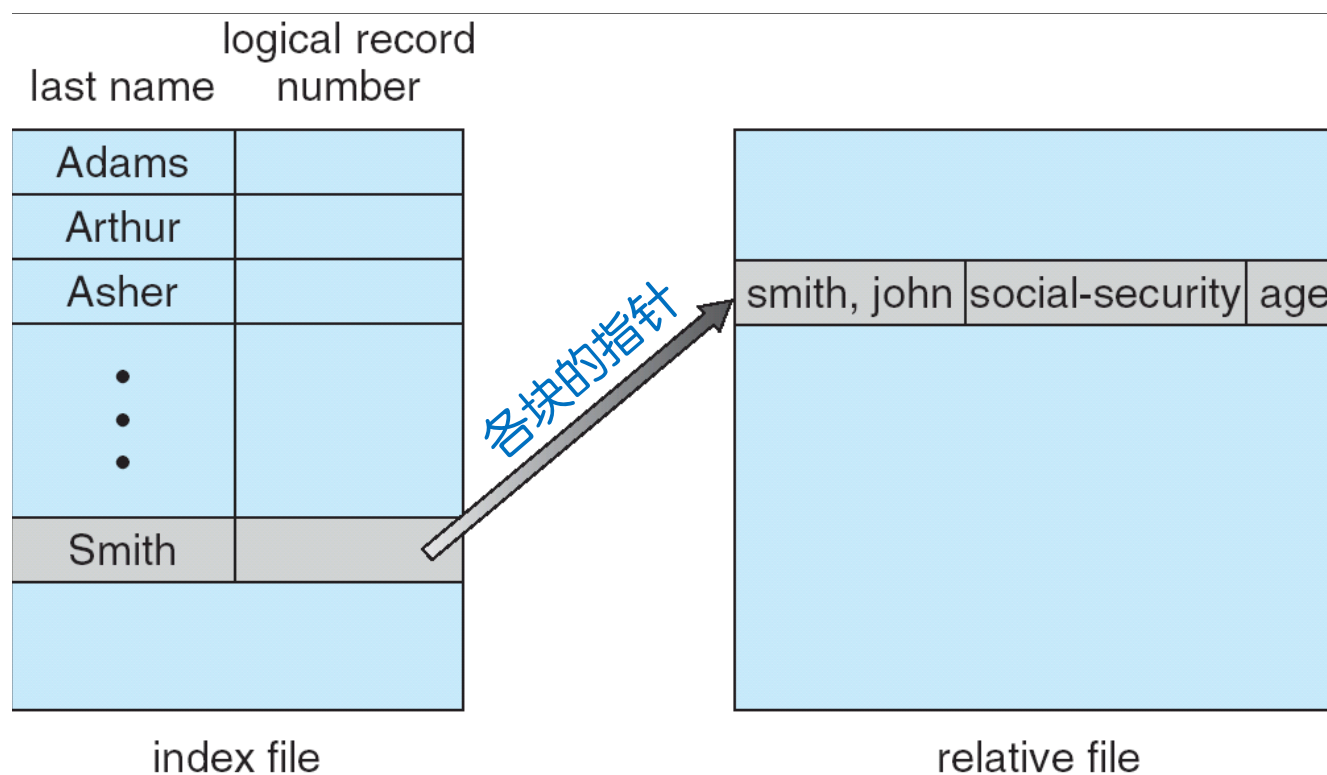
1. 文件由固定长度的逻辑记录组成
2. 操作系统所提供的块儿号通常为**相对块号**，它是相对于文件开始的索引

相对块号	物理块号
0	14703
1	2342
2	3192
...	...

sequential access	implementation for direct access
<i>reset</i>	<i>cp = 0;</i>
<i>read next</i>	<i>read cp;</i> <i>cp = cp + 1;</i>
<i>write next</i>	<i>write cp;</i> <i>cp = cp + 1;</i>

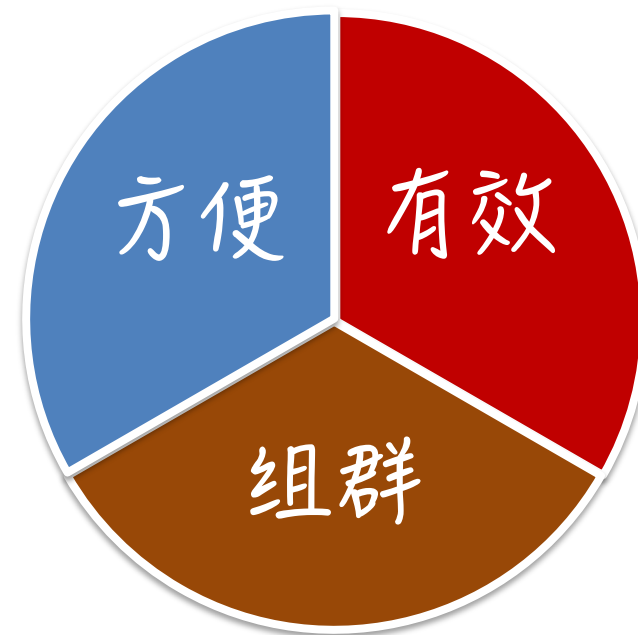


- 对文件创建索引文件，索引包括各块的指针
- 为查找文件中的记录，首先搜索索引，再根据指针直接访问文件



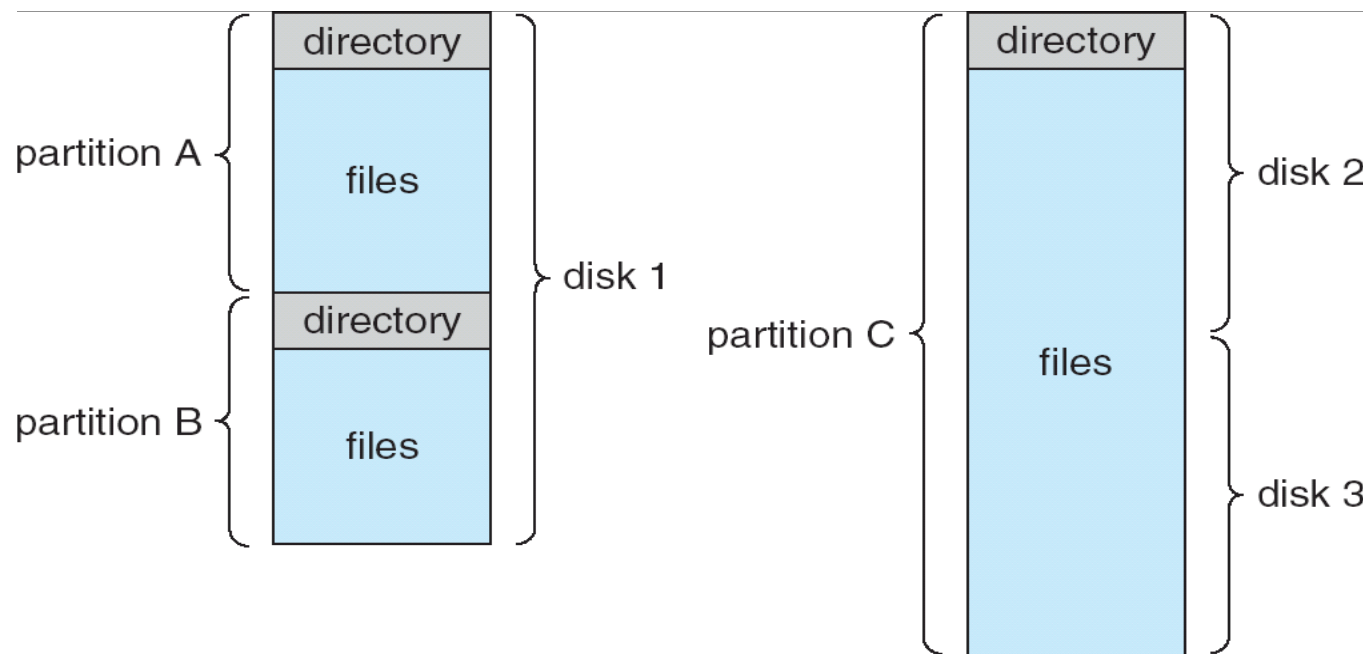
### 管理、组织文件的需要

1. 单层次目录结构
2. 双层次目录结构
3. 树状目录结构
4. 无环图目录结构
5. 通用图目录结构



用目录组织文件的目的

- 每个磁盘分区可以创建一个文件系统
- 可以将存储文件系统的一大块存储空间称为卷
- 每个卷必须包含其文件系统上文件的信息，这些信息保存在设备目录或卷表中



### \$df -Th 查看分区文件系统

```
root@hbpark-OptiPlex-390:~/Downloads# df -Th
Filesystem      Type      Size  Used Avail Use% Mounted on
/dev/sda1       ext4      222G   17G   194G   8% /
udev            devtmpfs  3.9G   4.0K   3.9G   1% /dev
tmpfs           tmpfs     1.6G   808K   1.6G   1% /run
none            tmpfs     5.0M    0    5.0M   0% /run/lock
none            tmpfs     4.0G   152K   4.0G   1% /run/shm
none            tmpfs     100M    40K   100M   1% /run/user
root@hbpark-OptiPlex-390:~/Downloads#
```

### \$Downloads 目录下文件

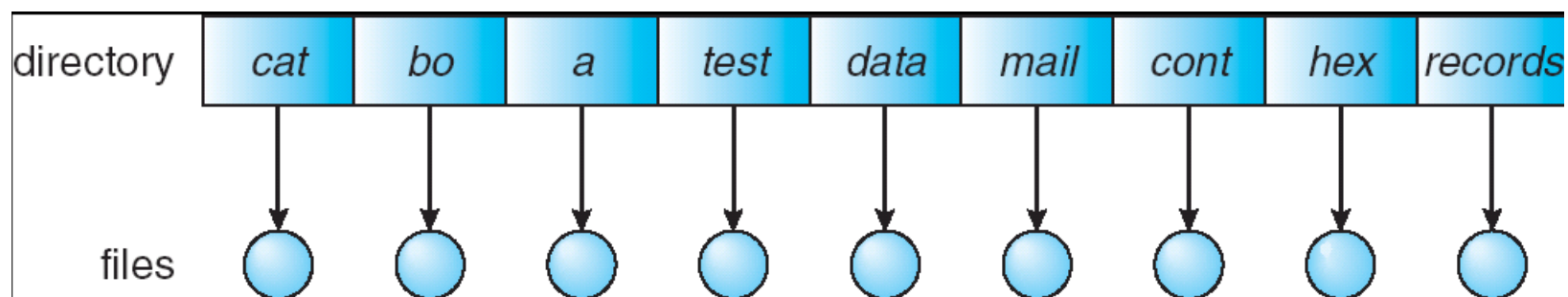
```
root@hbpark-OptiPlex-390:~/Downloads# ls
subversion-1.6.1  subversion-1.6.1.tar.gz  subversion-deps-1.6.1.tar.gz
root@hbpark-OptiPlex-390:~/Downloads#
root@hbpark-OptiPlex-390:~/Downloads#
root@hbpark-OptiPlex-390:~/Downloads#
root@hbpark-OptiPlex-390:~/Downloads# ls -al
total 11956
drwxr-xr-x  3 root root   4096 May  5 10:06 .
drwx----- 27 root root   4096 May  5 10:06 ..
drwxr-xr-x 16 1016 1016   4096 Aug 12  2014 subversion-1.6.1
-rw-r--r--  1 root root 7501026 Aug 12  2014 subversion-1.6.1.tar.gz
-rw-r--r--  1 root root 4726179 Aug 12  2014 subversion-deps-1.6.1.tar.gz
root@hbpark-OptiPlex-390:~/Downloads#
```

\$vi Downloads

```
"  Sort sequence: [\\/]$,\\<core\\%(\\.\\.d\\+\\)\\=\\>,\\.h$,\\.c$,\\.cpp$,\\
"  Quick Help: <F1>:help  _:go up dir  D:delete  R:rename  S:sort
"  =====
./
subversion-1.6.1/
subversion-1.6.1.tar.gz
subversion-deps-1.6.1.tar.gz
.swp
~
```

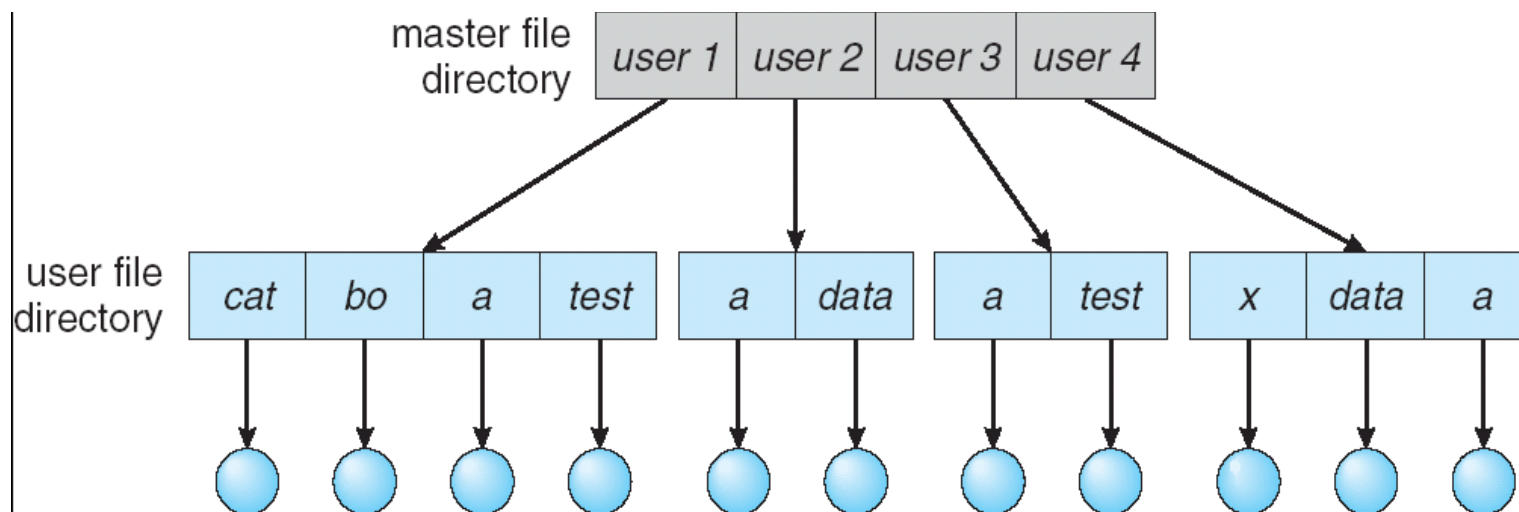
1. 搜索文件
2. 创建文件
3. 删除文件
4. 遍历目录 (List a directory)
5. 重命名文件
6. 跟踪文件系统 (备份)

所有文件都包含在同一个目录中



**问题:** 文件命名, 文件名称必须是唯一的多用户之间的文件名  
称也不能冲突

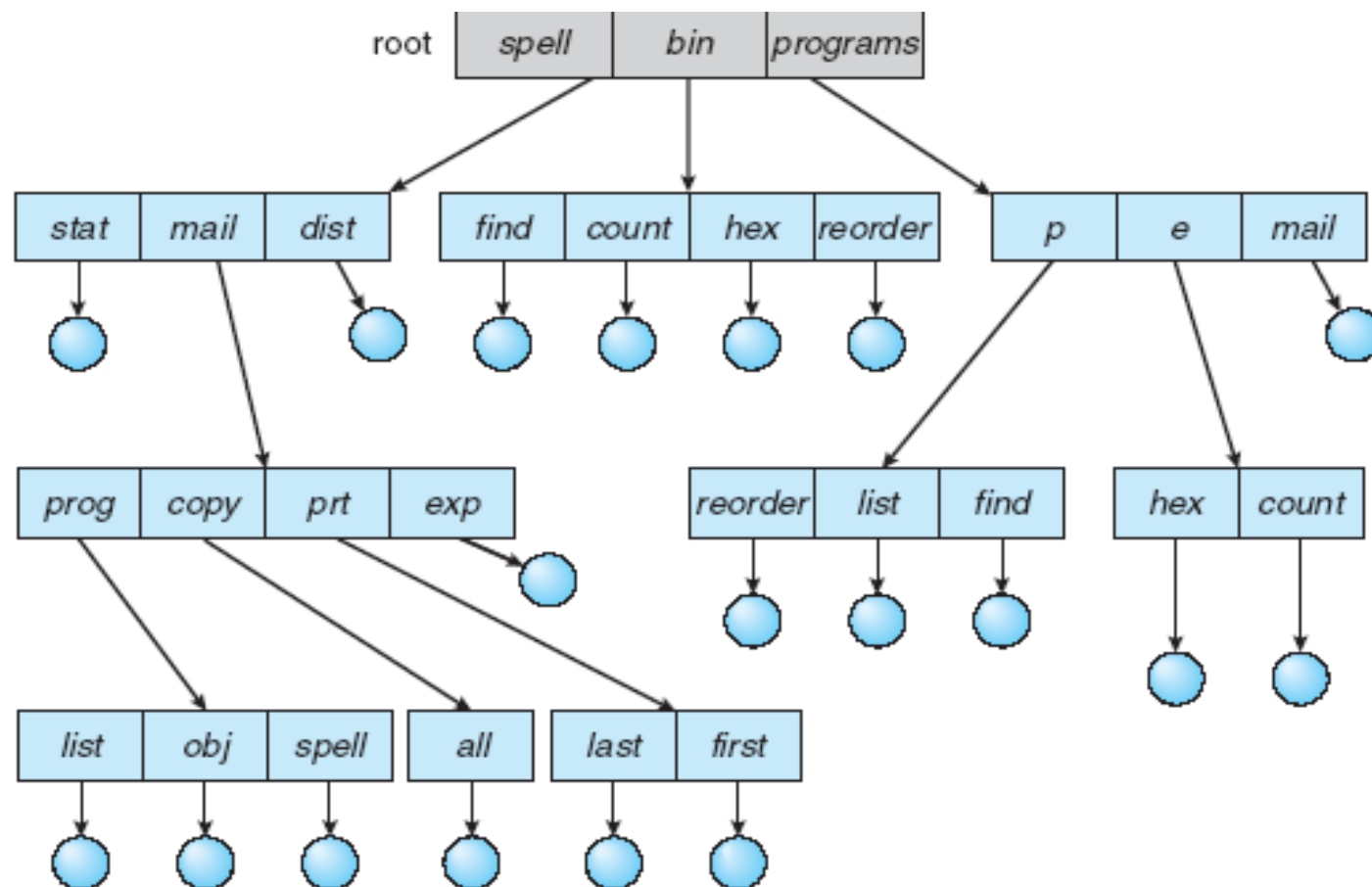
系统有主文件目录，每个目录是用户文件目录



- 路径是由用户名和文件名定义
- 不同的用户可以有相同的文件名



目录可以包含文件和子目录，当需要访问文件时，就从搜索当前目录



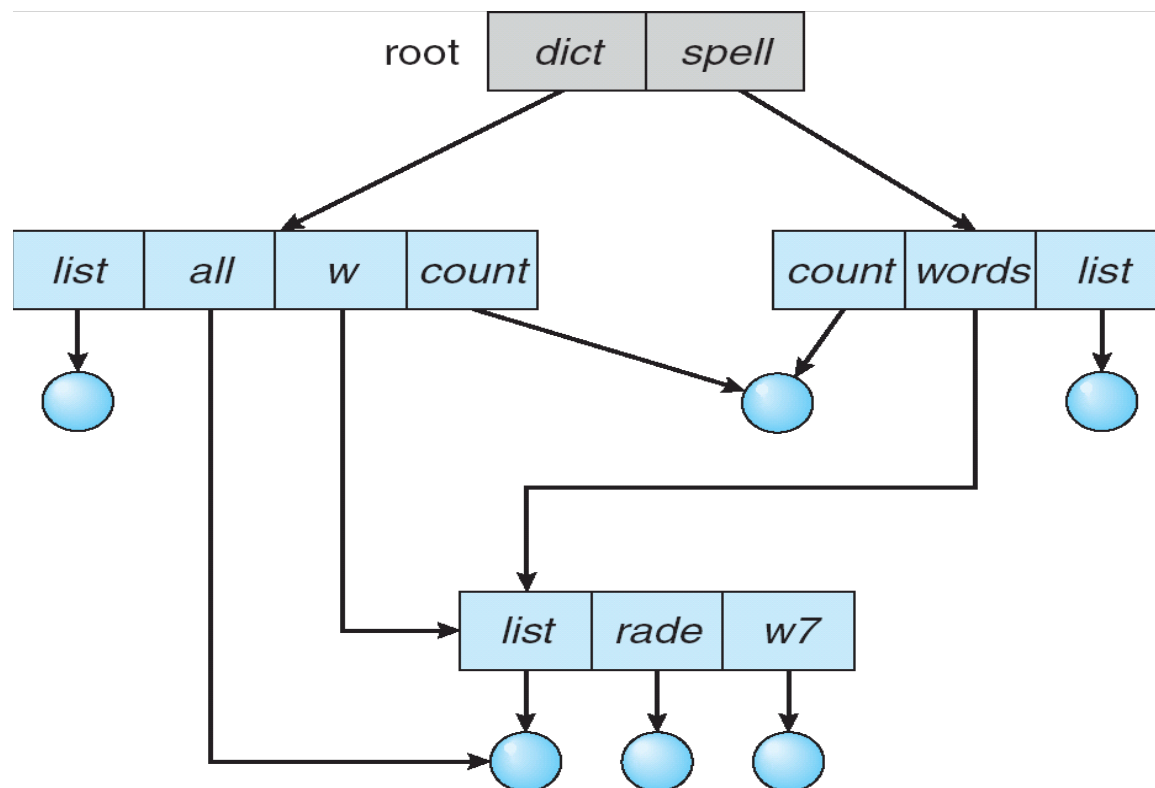
路径名有以下两种形式:

1. 相对路径:从当前目录开始
2. 绝对路径:从根目录开始

### 树状结构禁止共享文件和目录

- 用户的初始目录, 用户登录时搜索账户文件以得到该用户的初始目录
- 删除目录需要目录为空, 目录不为空不能删除

- 可共享文件和目录，允许目录含有共享子目录和文件
- 同一文件或子目录可出现在两个不同目录中



## 共享文件和目录的实现方法

1. 创建一个称为链接的新目录条目。链接实际上是另一文件或目录的指针  
→ 符号链接(symbolic link)
2. 共享目录中重复所有被共享文件的信息，但必须提供一致性（同步机制）  
→ 非符号链接(hard link)

`$ln -s source destination`

`$ln source destination`

## 问题

- 文件查找（重复查找）
- 删除（删除链接、删除原文件）

\$ln -s subversion-1.6.1.tar.gz symbolic

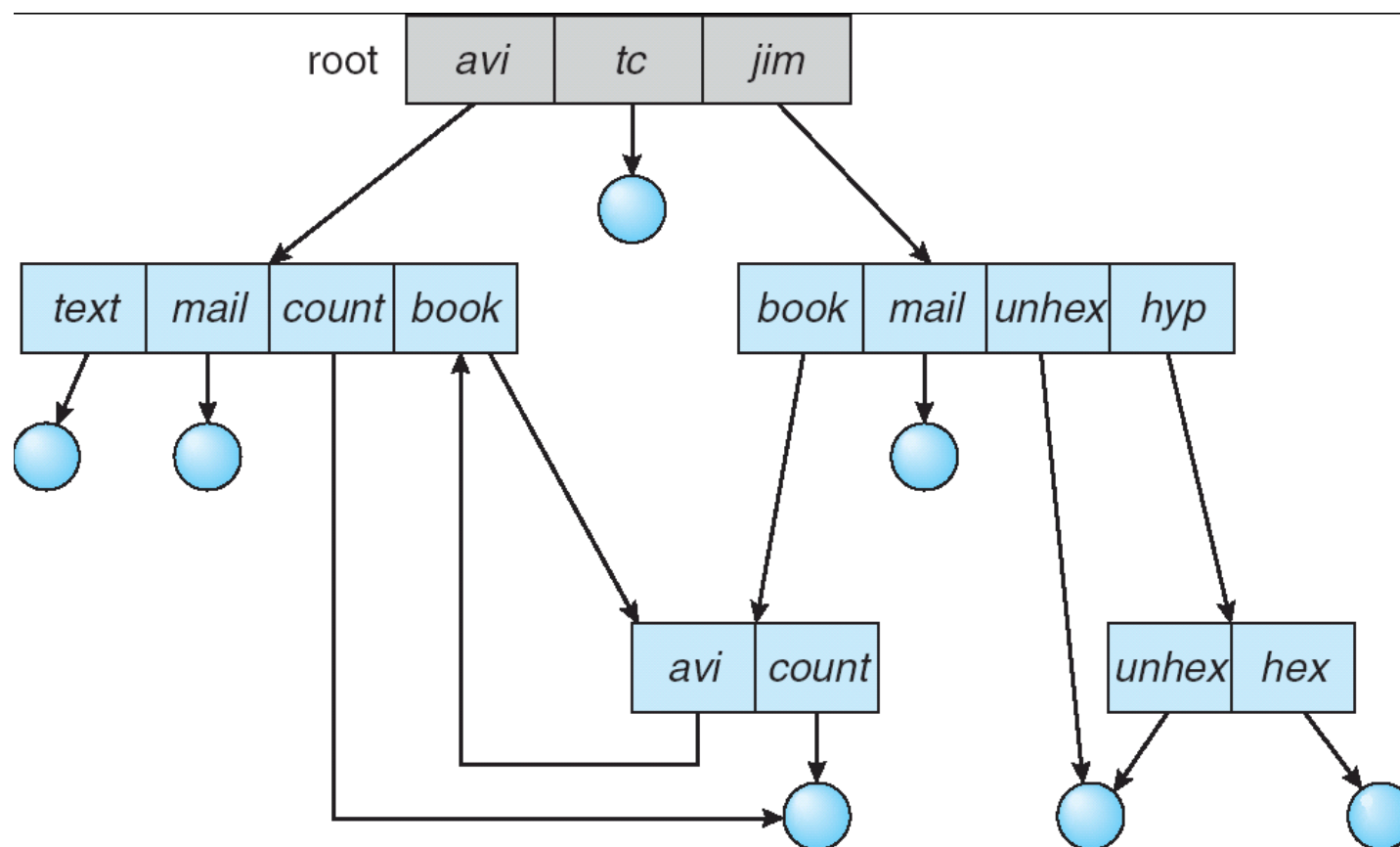
```
root@hbpark-OptiPlex-390:~/Downloads# ln -s subversion-1.6.1.tar.gz symbolic
root@hbpark-OptiPlex-390:~/Downloads# ls
subversion-1.6.1  subversion-1.6.1.tar.gz  subversion-deps-1.6.1.tar.gz  symbolic
root@hbpark-OptiPlex-390:~/Downloads# ls -al
total 11956
drwxr-xr-x  3 root root    4096 May  5 10:51 .
drwx----- 27 root root    4096 May  5 10:47 ..
drwxr-xr-x 16 1016 1016    4096 Aug 12 2014 subversion-1.6.1
-rw-r--r--  1 root root 7501026 Aug 12 2014 subversion-1.6.1.tar.gz
-rw-r--r--  1 root root 4726179 Aug 12 2014 subversion-deps-1.6.1.tar.gz
lrwxrwxrwx  1 root root      23 May  5 10:51 symbolic -> subversion-1.6.1.tar.gz
root@hbpark-OptiPlex-390:~/Downloads#
```

\$ln subversion-1.6.1.tar.gz hard

```
root@hbpark-OptiPlex-390:~/Downloads# ls -al
total 19284
drwxr-xr-x  3 root root    4096 May  5 10:54 .
drwx----- 27 root root    4096 May  5 10:54 ..
-rw-r--r--  2 root root 7501026 Aug 12 2014 hard
drwxr-xr-x 16 1016 1016    4096 Aug 12 2014 subversion-1.6.1
-rw-r--r--  2 root root 7501026 Aug 12 2014 subversion-1.6.1.tar.gz
-rw-r--r--  1 root root 4726179 Aug 12 2014 subversion-deps-1.6.1.tar.gz
lrwxrwxrwx  1 root root      23 May  5 10:51 symbolic -> subversion-1.6.1.tar.gz
root@hbpark-OptiPlex-390:~/Downloads#
```

允许有环的目录结构，需要避免重复搜索

删除文件需要确认引用计数，引用计数为0才能删除



如何确保不能形成环

1. 只允许链接文件，不允许链接子目录
2. 垃圾收集: Garbage collection (when link itself)
3. 每当建立新的链接时，利用某种检测算法检测是否形成环，如形成环就不允许链接

太费时，不常用

1. 文件系统被系统上的进程使用之前必须安装
2. 操作系统需要知道
  - + 设备名称和
  - + 文件系统的安装位置(mount point)



- /mnt/Lectures 目录下

```
hbpark@hbpark-VirtualBox: /mnt/Lectures
hbpark@hbpark-VirtualBox: /mnt/Lectures$ ls
hbpark@hbpark-VirtualBox: /mnt/Lectures$
```

- mount -t vboxsf 01.Lectures /mnt/Lectures

```
hbpark@hbpark-VirtualBox: /mnt/Lectures
hbpark@hbpark-VirtualBox: /mnt/Lectures$ ls
2012年课程      2014年课程      sources.list      课程资料
2013年课程      2015年春季学期课程  Webpages
2014年春季学期课程  2016年春季学期课程  培养方案
2014年秋季学期课程  sche.c           教学有关材料
hbpark@hbpark-VirtualBox: /mnt/Lectures$
```

```

MOUNT(8)                                System Administration                                MOUNT(8)

NAME
    mount - mount a filesystem

SYNOPSIS
    mount [-lhV]

    mount -a [-fFnrsvw] [-t vfstype] [-O optlist]

    mount [-fnrsvw] [-o option[,option]...] device|dir

    mount [-fnrsvw] [-t vfstype] [-o options] device dir

DESCRIPTION
    All files accessible in a Unix system are arranged in one big tree, the
    file hierarchy, rooted at /. These files can be spread out over several
    devices. The mount command serves to attach the filesystem found on some
    device to the big file tree. Conversely, the umount(8) command will detach
    it again.

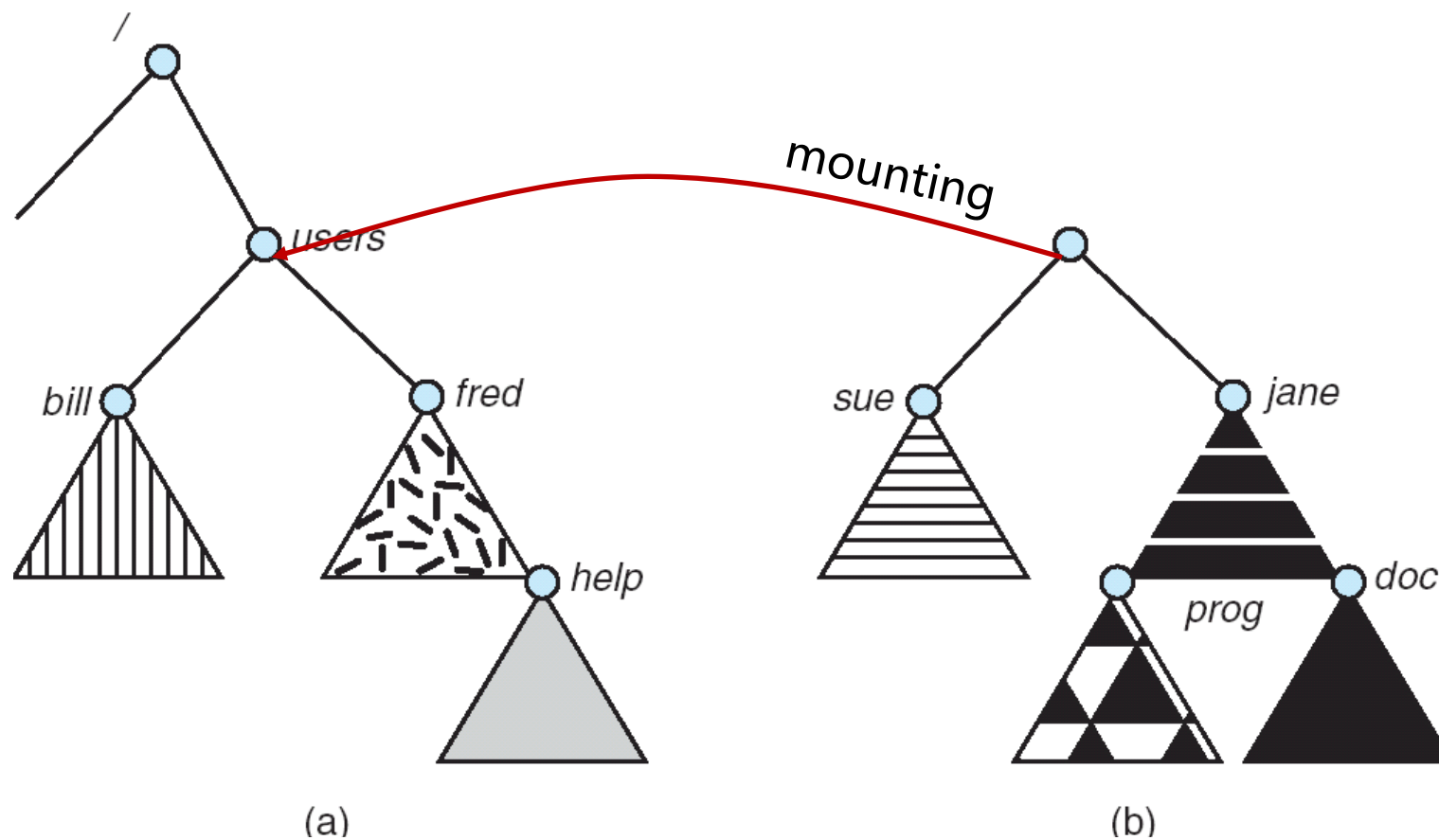
    The standard form of the mount command, is

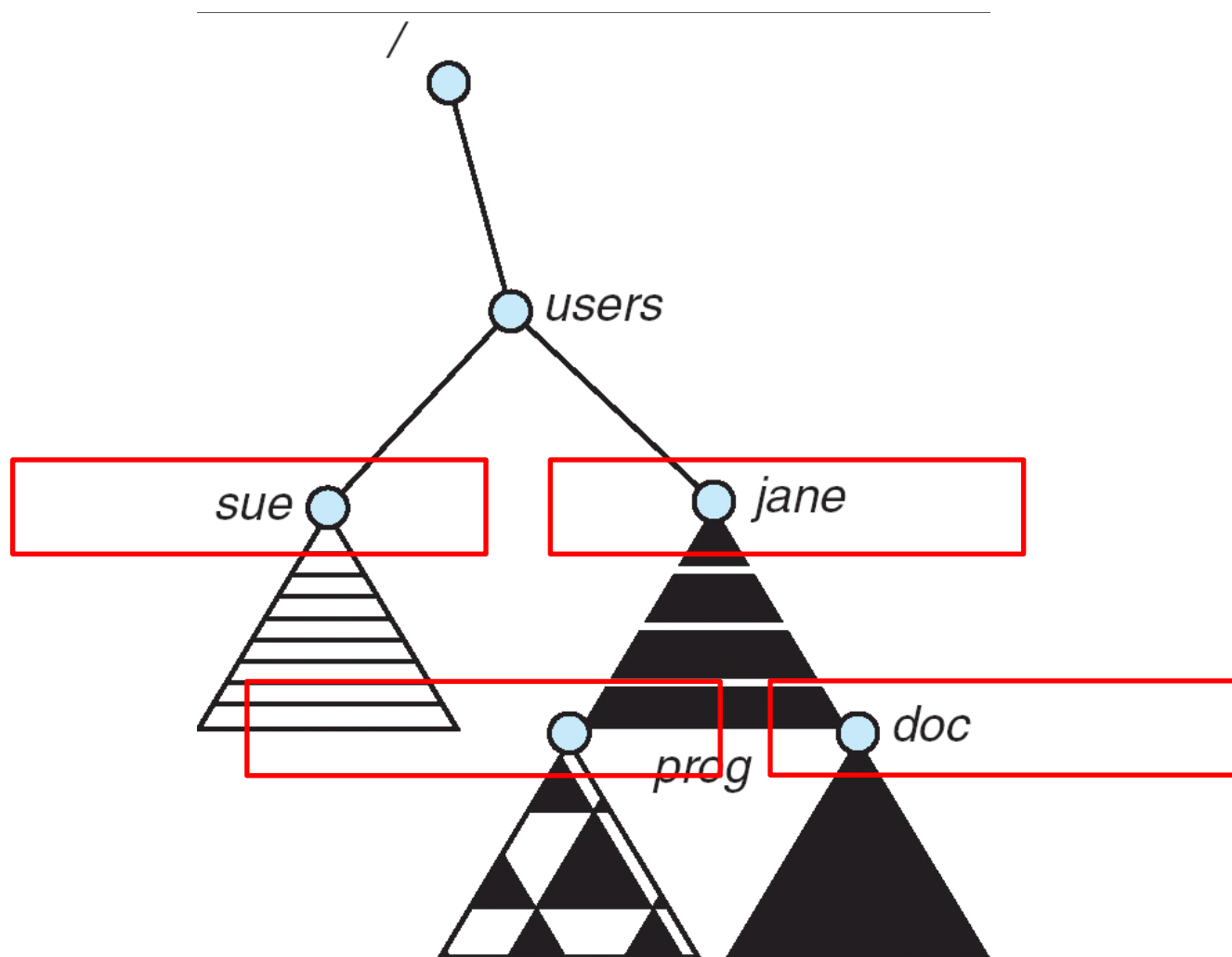
        mount -t type device dir

    This tells the kernel to attach the filesystem found on device (which is of
    type type) at the directory dir. The previous contents (if any) and owner
    and mode of dir become invisible, and as long as this filesystem remains
    mounted, the pathname dir refers to the root of the filesystem on device.

    If only directory or device is given, for example:

        mount /dir
    
```





多用户系统比单用户系统需要更多的文件和目录属性，即拥有者ID，组ID

1. 拥有者是目录最高控制权的用户，可以改变属性和授权访问
2. 组属性定义对文件拥有相同权限的用户子集

通过网络允许系统之间的文件系统之间的相互访问， 实现方法

1. 通过 FTP
2. 通过分布式文件系统
3. 通过浏览器（基本上FTP的包装）

1. 客户机-服务器模型，服务器可以为多用户服务
2. 一台机器上可以安装一台或多台远程机器上的一个或多个文件系统

举例:

- I. Unix 的客户机-服务器文件共享协议NFS(Network File System)网络文件系统
- II. Windows 的 CIFS(Common Interface File System) 公共网络文件系统

在远程文件系统环境下，对文件的标准调用会自动转变为远程调用

**故障模式**，远程文件系统故障与本地文件系统故障不同

1. 增加了**新的故障模式**
  2. 增加了每个远程请求的**状态信息**，通过调用状态信息恢复故障
- 如NFS文件系统对每个远程请求都会记录状态信息，但缺点是安全性能差



描述多用户同时访问共享文件时的语义，即规定了一个用户所修改的共享数据时，对另一个用户（共享数据的用户）是否可见的问题

1. 会话语义:Andrew File System(AFS)，一个用户的文件修改不能实时被其他用户所见。
2. Unix语义:Unix File System (UFS)，一个用户的文件修改可以实时被其他用户所见
3. 不可修改共享文件语义: 一旦共享，就不能改

- 可靠性（不受物理损坏）
  - 通过备份
- 保护（非法访问）
  1. 访问类型
    - 读，写，执行，添加，删除，列表清单
  2. 访问控制
    - 拥有者，组，其他

# Q&A