

存储区以及在一个逻辑磁盘块中有多少个扇区等细节并不是用户所关心的，当然对文件系统的设计者来说这些内容是相当重要的。正因为如此，本章将分为几节讲述，前两节分别叙述在用户层面的关注内容——文件和目录，随后是有关文件系统实现的详细讨论，最后是文件系统的一些实例。

4.1 文件

在本节中，我们从用户角度来考察文件，也就是说，用户如何使用文件，文件具有哪些特性。

4.1.1 文件命名

文件是一种抽象机制，它提供了一种在磁盘上保留信息而且方便以后读取的方法。这种方法可以使用户不用了解存储信息的方法、位置 and 实际磁盘工作方式等有关细节。

也许任何一种抽象机制的最重要的特性就是对管理对象的命名方式，所以，我们将从对文件的命名开始考察文件系统。在进程创建文件时，它给文件命名。在进程终止时，该文件仍旧存在，并且其他进程可以通过这个文件名对它进行访问。

文件的具体命名规则在各个系统中是不同的，不过所有的现代操作系统都允许用1至8个字母组成的字符串作为合法的文件名。因此，andrea、bruce和cathy都是合法文件名。通常，文件名中也允许有数字和一些特殊字符，所以像2、urgent!和Fig.2-14也是合法的。许多文件系统支持长达255个字符的文件名。

有的文件系统区分大小写字母，有的则不区分。UNIX是前一类，MS-DOS是后一类。所以在UNIX系统中maria、Maria和MARIA是三个不同的文件，而在MS-DOS中，它们是同一个文件。

关于文件系统在这里需要插一句，Windows 95与Windows 98用的都是MS-DOS的文件系统，即FAT-16，因此继承了其很多性质，例如有关文件名的构造方法。Windows 98对FAT-16引入了一些扩展，从而成为FAT-32，但这两者是很相似的。并且，Windows NT、Windows 2000、Windows XP和Windows Vista支持这两种已经过时的FAT文件系统。这4个基于NT的操作系统有着一个自带文件系统（NTFS），它具有很多不同的性质（例如基于Unicode的文件名）。在本章中，当提到MS-DOS或FAT文件系统的时候，我们指的是用在Windows上的FAT-16和FAT-32，除非特别指明。我们将晚一些在这章讨论FAT文件系统，并在第11章讨论NTFS，并细致地分析了Windows Vista。

许多操作系统支持文件名用圆点隔开分为两部分，如文件名prog.c。圆点后面的部分称为文件扩展名（file extension），文件扩展名通常表示文件的一些信息，如MS-DOS中，文件名由1至8个字符以及1至3个字符的可选扩展名组成。在UNIX里，如果有扩展名，则扩展名长度完全由用户决定，一个文件甚至可以包含两个或更多的扩展名。如homepage.html.zip，这里.html表明HTML格式的一个Web页面，.zip表示该文件（homepage.html）已经采用zip程序压缩过。一些常用文件扩展名及其含义如图4-1所示。

扩展名	含 义
file.bak	备份文件
file.c	C源程序文件
file.gif	符合图形交换格式的图像文件
file.hlp	帮助文件
file.html	WWW超文本标记语言文档
file.jpg	符合JPEG编码标准的静态图片
file.mp3	符合MP3音频编码格式的音乐文件
file.mpg	符合MPEG编码标准的电影
file.o	目标文件（编译器输出格式，尚未连接）
file.pdf	pdf格式的文件
file.ps	PostScript文件
file.tex	为TEX格式化程序准备的输入文件
file.txt	一般正文文件
file.zip	压缩文件

图4-1 一些典型的文件扩展名

在某些系统中（如UNIX），文件扩展名只是一种约定，操作系统并不强迫采用它。名为file.txt的文件也许是文本文件，这个文件名在于提醒所有者，而不是表示传送什么信息给计算机。但是另一方面，C编译器可能要求它编译的文件以.c结尾，否则它会拒绝编译。

对于可以处理多种类型文件的某个程序，这类约定是特别有用的。例如，C编译器可以编译、连接多种文件，包括C文件和汇编语言文件。这时扩展名就很必要，编译器利用它区分哪些是C文件，哪些是汇编文件，哪些是其他文件。

相反，Windows对扩展名赋予含义。用户（或进程）可以在操作系统中注册扩展名，并且规定哪个