

文章编号:1674-2869(2011)04-0094-03

linux 操作系统中 EXT2 文件的组成

朱 颂^{1,2}

(1. 武汉大学计算机学院,湖北 武汉 430072;2. 孝感学院计算机与信息科学学院,湖北 孝感 432000)

摘 要:EXT2 文件系统是 linux 操作系统的重要组成部分,其核心由超级块、组块描述表、索引节点表组成. 通过源代码,剖析了这几个关键部分的数据结构和作用. 介绍了 linux 操作系统寻找文件的过程.

关键词:超级块;块组描述符;索引节点;EXT2 文件系统

中图分类号:TP316 **文献标识码:**A **doi:**10.3969/j.issn.1674-2869.2011.04.024

0 引 言

随着 Linux 操作系统在服务器领域和嵌入式领域中的广泛应用,ext2 作为 linux 操作系统上的标准文件系统也得到了广泛的支持和考验. 其精巧、灵活的设计,在存取中小文件时的性能非常好,极大地满足了人们对数据存储的要求. 同时,ext2 文件系统也是其后继 ext3 和 ext4 日志文件系统的基础. 剖析 ext2 文件系统,对深入理解 linux 操作系统的文件存储方式很有帮助.

1 ext2 文件系统整体结构

硬盘首先分区,然后格式化,才能使用. 格式化的过程会在硬盘上建立很多块,为了管理这些块,文件系统将其分组,每个组称为块组(block group). 每个块组又由六部分组成(见图 1):超级块(super block)、块组描述表(group descriptor table)、块位图(block bitmap)、索引位图(inode bitmap)、索引表(inode table)和数据块(data block)^[1].

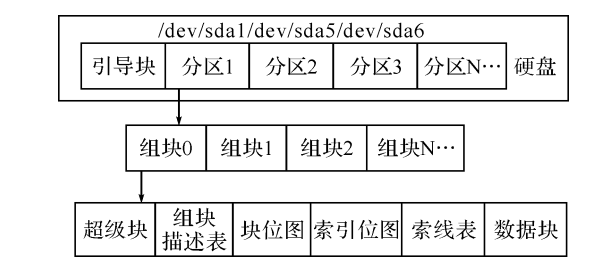


图 1 硬盘、分区、块组的组成

Fig. 1 Composition of hard drive, Zoning and block group

1.1 超级块

超级块是硬盘分区中最重要的结构,是一个分区的全局数据信息,描述文件系统的目录和文件的静态

分布情况以及各种组成结构的尺寸、数量. 一旦超级块损坏,操作系统将无法读取该分区的信息,整块分区的数据就会丢失. 当有程序对 ext2 文件系统执行一致性检查时,就会将块组 0 中的超级块拷贝到其他块组中,这些多的拷贝信息可以帮助系统恢复分区的信息,从而减少损失. 正常情况下,linux 系统只会读取块组 0 中包含的超级块.

在 linux 内核源代码中超级块的数据结构定义是 ext2_super_block,超级块的结构见图 2,其结构体较长,具体字段不罗列出来.

超级用户保留块(s_r_blocks_count)字段是为超级用户保留的总块数. 当硬盘分区被其他用户有意或无意的填满后,将无法登陆. 超级用户则可以利用这些保留的空间登陆,编辑配置文件,解决相应问题.

第一个数据块块号(s_first_data_block)的值在块大于 1 kB 的文件系统里总是 0,当块大小是 1 kB 时,这个数值为 1,因为超级块大小正好是 1 kB.

块大小(s_log_block_size)字段的值显示了块的大小,计算方法是 $1024 \leq s_log_block_size$,如果值是 0,块的大小就是 1024 bytes,这就是最小块. 通常这个字段的值是 2,块大小为 4096 bytes,正好是 4 K.

块总数(s_blocks_count)和每个块组中的块数(s_blocks_per_group)两字段可以计算出块组总数 $ceil((s_blocks_count - s_first_data_block - 1) / s_blocks_per_group)$ ^[2]. s_blocks_count 字段里面包含了整个分区所有块,包括描述信息所占用的块,例如:超级块、块组描述符、位图块都会在分区上占用一定数量的块,而这些块不在块组管理范围内,所以要减去 s_first_data_block. 分区划分块组时最后可能多出一些块,系

系统将最后多余的块划分到一个块组中,所以最后计算结果向上取整.



图 2 超级块的结构

Fig. 2 Structure of super block

块组中索引节点数(s_inodes_per_group)也很重要. 在寻找索引节点的过程中,通过此字段,可以计算出该索引节点所在的块组号: $\text{floor}((\text{ino}-1)/\text{s_inodes_per_group})$ 和在该块组中的偏移量: $(\text{ino}-1)\%\text{s_inodes_per_group}^{[2]}$,这样就可以迅速定位索引节点.

挂载计数(s_mnt_count)表示文件系统从最近一次完整校验后被安装的次数,最大挂载计数(s_max_mnt_count)记录了能够挂载最大次数,一旦达到最大次数,将进行文件系统检测.

1.2 块组描述符表

每一个块组都有一个块组描述符对应,若干个块组描述符集中在一起按顺序存放,就是块组描述符表. 在分区中的地位仅次于超级块,也是分区的全局数据信息,作用是描述整个分区中块组、索引节点和数据块的使用情况,在每个块组中也都有一份拷贝. bg_block_bitmap 和 bg_inode_bitmap 字段分别指出了块位图(block bitmap)和索引节点位图(inode bitmap)第一个块的位置.

硬盘格式化之后,硬盘上的数据块和索引节点的数量就确定下来,是一个固定值,只要存放文件,就会占用数据块和索引节点. 块位图和索引节点位图的作用就是表示数据块和索引节点使用的情况. 位图是位的序列,系统用 0 表示相应的块或者索引节点为空闲,1 表示占用^[3](见图 3). 每个位图数据必须存放在一个单独的块中,如果块大小设置为 4K,那么一个单独的块位图可以描述 $4 \times 1024 \times 8 = 32768$ 个块的状态.

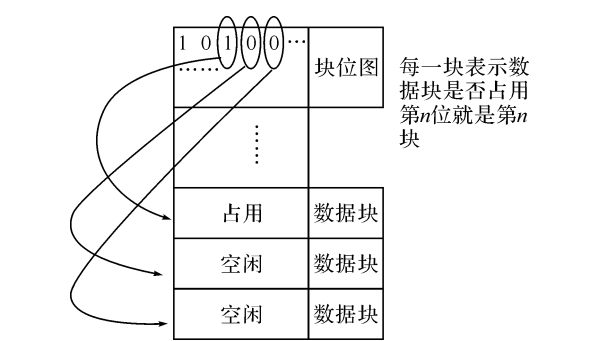


图 3 块位图描述

Fig. 3 Description of block bitmap

1.3 索引节点表

索引节点也称为 i 节点. 当建立和保存文件时,文件名和文件的内容会保存在数据块中,文件的属性以及数据块的地址保存在索引节点内. 从 ext2_inode 数据结构中,可以看到文件的属性包括:访问权限、文件所有者、文件修改时间和文件大小等项目. 在 linux 下建立一个一般文件时,系统会分配至少一个索引节点和相对于该文件大小数据块给文件. 索引节点的数量是一个固定值,其多少由格式化时的参数决定,默认两个块分配一个索引节点. 如果索引节点用完即便该分区还有空间,也不能保存文件.

重点要提出的是 ext2_inode 结构中 i_block [Ext2_N_BLOCKS]^[4] 字段,该字段保存了文件在数据块中的具体位置. 通常 Ext2_N_BLOCKS 的值是 15^[5],此字段的 i_block[0]~ i_block[11]是直接保存数据块的地址,i_block[12]、i_block[13]和 i_block[14]分别是一级、二级和三级指针,见图 4. 这样一个索引节点可以指向文件的大小是 $12 \times \text{block} + (\text{block}/4) \times \text{block} + (\text{block}/4)^2 \times \text{block} + (\text{block}/4)^3 \times \text{block}$,如果 block 大小为 4K,那么指向的这个文件相当大了,但实际上该结构中 i_size 字段限制了索引节点指向文件的大小,该字段是一个无符号的 32 位整数,所以最大文件只能是 $2^{32} = 4\text{G}$.

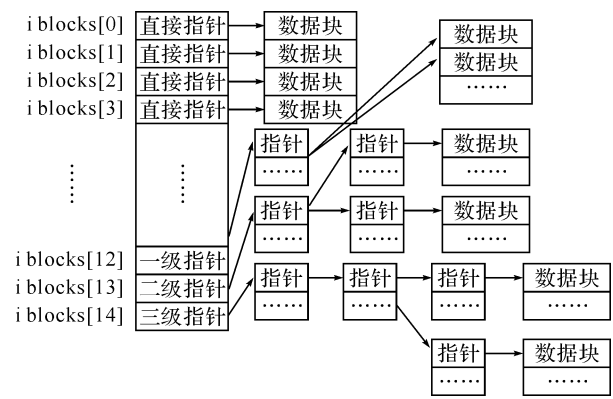


图 4 i_block 字段的作用
Fig. 4 Function of i_block

1.4 目录结构

在 linux 系统上,目录也是一种文件.目录内保存着其他文件或目录的名称和索引节点号.当建立一个目录时,系统会分配一个索引节点与至少一个数据块给该目录,见图 5.

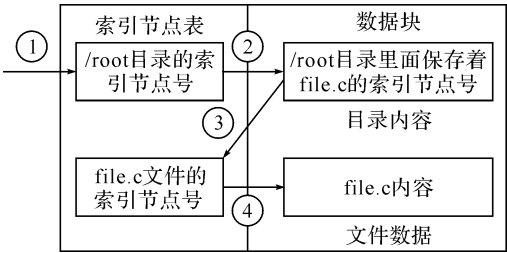


图 5 系统读取文件的过程
Fig. 5 Process of read file

系统读取文件/root/file.c 时的过程:

a. 系统从根目录信息可以得到/root 目录的

索引节点号,并且读取索引节点号的内容,得到/root 目录的属性和数据块中的具体位置.

b. 读取/root 目录数据块内容,得到 file.c 文件的索引节点号.

c. 读取 file.c 文件的索引节点号,得到 file.c 文件的属性和数据块具体位置.

d. 读取 file.c 文件的内容.

2 结 语

尽管 EXT2 文件系统设计目标明确,层次分明,但是仍然有不少缺陷,例如对大文件的读写性能,对文件的灾难恢复等,这些都不在本篇探讨范围内.在用户要求不高,处理数据量不大的情况下,EXT2 文件系统依然是最好的选择.

参考文献:

[1] Daniel P B,Marco C. 深入理解 Linux 内核[M]. 陈莉君,冯锐,牛欣源,译. 北京:中国电力出版社,2001.

[2] 包怀忠. ext2 文件系统分析[J]. 计算机工程与设计, 2005,26(4):1022 - 1024.

[3] 徐虹,张钦. EXT2 文件系统的分析与研究[J]. 成都信息工程学院报,2007,22(3):330 - 333.

[4] 唐靖飏,周良源. UNIX 平台下 C 语言高级编程指南 [M]. 北京:宇航出版社,2000.

[5] 郭学理,韦智,潘松. Linux 的 EXT2 文件系统[J]. 计算机应用研究,2001,18(5):128 - 130.

Construction of EXT2 File System in Linux operation system

ZHU Song^{1,2}

(1. Computer School, Wuhan University, Wuhan 430072, China;

2. Department of Computer and Information Science, Xiaogan University,Xiaogan 432000, China)

Abstract: The EXT2 file system is an important part of the Linux operating system, the key part consists of the super block, the group descriptor table and the inode table. The data structure and function on these important parts by source code are analyzed. A process about Linux search a file is shown.

Key words: super block; group descriptor; inode; EXT2 file system

本文编辑:陈锦华