FAT12文件系统之引导扇区结构

FAT12文件系统将按照下表所示的方式划分全部的容量，即文件系统数据结构：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| FAT12 2880扇区 (1474560B) | 逻辑扇区 | 占用扇区 | 内容 | 磁盘CHS参数(磁头/柱面/扇区) |
| 0 | 1(512B) | 引导程序 | 起：0/0/1 |
| 1 | 9(4608B) | FAT文件分配表1 | 起：0/0/2 止：0/0/10 |
| 10 | 9(4608B) | FAT文件分配表2 | 起：0/0/11 止：1/0/1 |
| 19 | 14(9728B) | 根目录 |  |
| 33 | 14(9728B) | 文件数据区 |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 标识 | 偏移量 | 类型 | 大小 | 说明 | 默认值 |
|  | 0 | db | 3 | 跳转指令 | Jmp short LABEL  nop |
| BS\_OEMName | 3 | db | 8 | OEM字符串，必须为8个字符，不足以空格填空 | MSWIN4.1 |
| BPB\_BytsPerSec | 11 | dw | 2 | 每扇区字节数 | 200h |
| BPB\_SecPerClus | 13 | db | 1 | 每簇占用的扇区数 | 1 |
| BPB\_RsvdSecCnt | 14 | dw | 2 | 保留扇区数 | 1 |
| BPB\_NumFATs | 16 | db | 1 | FAT表的记录数 | 2 |
| BPB\_RootEntCnt | 17 | dw | 2 | 最大根目录文件数 | 0e0h |
| BPB\_TotSec16 | 19 | dw | 2 | 逻辑扇区总数 | 0b40h |
| BPB\_Media | 21 | db | 1 | 媒体描述符 | 0f0h |
| BPB\_FATSz16 | 22 | dw | 2 | 每个FAT占用扇区数 | 9 |
| BPB\_SecPerTrk | 24 | dw | 2 | 每个磁道扇区数 | 12h |
| BPB\_NumHeads | 26 | dw | 2 | 磁头数 | 2 |
| BPB\_HiddSec | 28 | dd | 4 | 隐藏扇区数 | 0 |
| BPB\_TotSec32 | 32 | dd | 4 | 如果BPB\_TotSec16是0，则在这里记录扇区总数 | 0 |
| BS\_DrvNum | 36 | db | 1 | 中断13的驱动器号 | 0 |
| BS\_Reserved1 | 37 | db | 1 | 未使用 | 0 |
| BS\_BootSig | 38 | db | 1 | 扩展引导标志 | 29h |
| BS\_VolID | 39 | dd | 4 | 卷序列号 | 0 |
| BS\_VolLab | 43 | db | 11 | 卷标，必须是11个字符，不足以空格填充 |  |
| BS\_FileSysType | 54 | db | 8 | 文件系统类型，必须是8个字符，不足填充空格 | FAT12 |
|  | 62 |  |  | 引导代码，由偏移0字节处的短跳转而来 |  |
|  | 510 | dw | 2 | 系统引导标识 | 0aa55h |

    首先是跳转指令，偏移0处的跳转指令必须是合法的可执行的基于x86的CPU指令，如：jmp start，这样可以生成3字节长的指令，（加关键字short的短跳转指令的长度是2字节），指向操作系统引导代码部分。Windows和MS-DOS生成的FAT12启动扇区中的跳转指令是短跳转，如：jmp short start，然后加一个nop的空指令来保持3字节的长度。

    接着是位于偏移3处的OEM字符串，它必须是一个8字节长的字符串，标识了格式化此磁盘的操作系统的名称和版本号，为了保留与MS-DOS的兼容性，通常Windows 2000系统格式化的磁盘上在此记录中的字符串是“MSDOS5.0”，在Windows 95系统格式化的磁盘上在此记录中的字符串是“MSWIN4.0”，在Windows 95 OSR2和Windows 98系统上格式化的磁盘上在此记录中的字符串是“MSWIN4.1”。

    接下来是每扇区的字节数，类型是双字节长，标准分区上的每扇区字节数一般是512B，但也可以是其它的数字，如1024,2048和4096，FAT12的格式下设置为512(200h)。

    偏移13处的是每簇所占用的扇区，类型是字节，簇是数据存储的最小单位，此字段的值取决于分区的大小，在FAT12格式下一般为1，即每簇只有1个扇区(512字节)，簇越大，那么分区的容量也就越大，通过增加簇的扇区数，可以支持更大的磁盘分区，标准的簇大小为1、2、4、8、16、32、64和128，FAT12格式下只能管理2^12个簇(4096)，所以在FAT12格式下能管理和分配的最大空间为：4096\*1\*512=2097152B=2M，所以FAT12一般只适合3.5寸高密度软盘(1.44M)。

    保留扇区指的是在第一个FAT文件分配表之前的引导扇区，一般情况下只保留1个扇区(512B)。

    接下来是类型为1字节长的FAT表的总数，默认情况下此字段的值为2，也就是有两个FAT表，FAT1和FAT2的内容相同，当FAT1表出错的时候可以使用FAT2来恢复文件分配表。

    位于偏移17处的字段是类型为双字节长的能够储存在根目录下的最大文件(包含子目录)数量，默认为224，每个目录或文件名占用32B的空间，因此根目录的大小为：224\*32=7168B=7KB，如果使用长文件名的话，根目录文件数还可能无法达到224的数量。

    接下来是位于偏移19处的逻辑扇区总数，类型是双字节，如果此磁盘的逻辑扇区总数大于2^16位(65536)的话，就设置此字段为0，然后使用偏移32处的双字来表示逻辑总扇区数。

    位于偏移21处的是单字节长的磁盘类型标识符，使用0f0h表示3.5寸高密码软盘，用0f8h来表示硬盘。此字段主要用于FAT12或FAT16的分区格式中，在Windows 2000中未使用。

    偏移22处双字节长的是每个FAT文件分配表所占用的扇区数，操作系统用这个字段和FAT表数量以及隐藏扇区数量来计算根目录所在的扇区。还可以根据最大根目录数来计算用户数据区从哪里开始。

    根目录扇区位置=FAT表数量\*FAT表所占用的扇区数量+隐藏扇区数量

    用户数据开始位置=根目录扇区位置+根目录所占用扇区(FAT12格式下为224\*32/512)

    此处所说的扇区指的是逻辑(线性)扇区，需要通过转换才能得到CHS磁盘参数，然后通过CHS参数来读写磁盘扇区。

    接下来是位于偏移24处的每磁道扇区总数，类型是双字节长，软盘的默认值为18，即每个磁道有18个扇区。

    然后是双字节长的磁头数，磁头数指的是磁盘面数，每面都有一个磁头，软盘都是2面的，所以在FAT12格式下此字段固定为2。

    接下来是的位于偏移28处类型为双字(4B)长的隐藏扇区数，指的在引导扇区之前的隐藏扇区，在FAT12格式上此字段默认为0，即不隐藏任何扇区，此字段参与计算根目录区和用户数据区位置。

    偏移32处的是类型为双字(4B)长的逻辑扇区总数，如果此分区或磁盘的逻辑扇区总数大于65536则用这个字段来表示逻辑扇区总数，否则设置此字段为0后用位于偏移19处的双字节字段来表示。

    偏移36处的是物理驱动器号，类型是字节长，它与BIOS物理驱动器相关，在磁盘中断Int13h相关的操作中使用，第一个软盘驱动器设置为0，第一个硬盘驱动器设置为80h，第二个硬盘驱动器设置为81h，以此类推。此字段的值可以在系统引导时用dl寄存器得到。

    位于偏移37处的字节没有使用，保留并设置为0。

    位于偏移38处的是扩展引导标识，类型是字节，操作系统用它来识别引导信息，值可以是28h或29h。

    接下来的是位于偏移39处的卷标号，类型是双字(4B)长，在格式化磁盘时所产生的一个随机序号，有助于区分磁盘，可以为0。

    然后是位于偏移43处的卷标，长度必须是11字节长(不足以空格20h填充)，此字段只能使用一次，用来保存磁盘卷的标识符，再次设置的时候被保存到根目录中作为一个特殊的文件来储存。

    最后是位于偏移54处的是长度为8字节的文件系统类型标识符，不足8字节则以空格20h来填充。FAT12格式下此字段为“FAT12   ”，相应的还有“FAT16   ”和“FAT32   ”。但要注意的是，操作系统并不使用这个字段来识别此磁盘所用的文件系统。

FAT12文件系统之数据存储方式详解

    下面我们来详细了解一下FAT文件分配表、根目录、用户数据的数据结构，只有通过详细分析这些数据结构，才能自由的存取FAT12格式的磁盘文件。

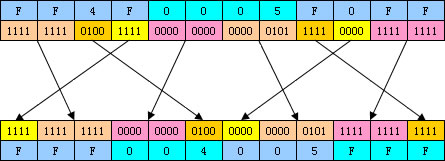
    我们先来看看文件分配表的数据格式，文件分配表所在的扇区应该是(隐藏扇区+保留扇区)=0+1=第1扇区处，从第1扇区开起到第9扇区结束，第一个文件分配表共占用9个扇区，第二个文件分配表从第10个扇区开始到第18扇区结束，在引导扇区的数据结构中明明确的指出了这些位置。

    文件分配表数据结构如下图所示：

[AT12文件系统 <wbr>数据存储方式详解](http://photo.blog.sina.com.cn/showpic.html%23blogid=3edcf6b80100crz1&url=http://s3.sinaimg.cn/orignal/3edcf6b8g68485aa62252)

    在FAT表开始扇区的第1字节是存储介质，0f0h代表软盘，0f8代表硬盘；第2、3这两个字节都是0ffh，代表了FAT文件分配表标识符，从第四个字节开始与用户数据区所有的簇一一对应，应该注意的是，用户数据区的第一个簇的序号是002，而不是000，因为储存介质和标识符占用了这两个序号。

    在FAT12格式中用12比特位来代表一个簇的序号，我们知道，每个字节有8位比特，所以每个簇要占用1.5个字节，也就是说，占用了第1字节和第2字节的一半才能表示一个簇的序号，半字节的拆分办法按照下图的方式进行：

[](http://photo.blog.sina.com.cn/showpic.html%23blogid=3edcf6b80100crz1&url=http://s3.sinaimg.cn/orignal/3edcf6b8g68487f250632)

    例如：在FAT表中开始位置储存的字节内容依次是F0 FF FF FF 4F 00 05 F0 FF，前面三个字节是储存介质和标识符，我们不管它，前面三个字节占用了0和1这两个簇序号，那么就应该从2簇开始了。经过转换得到的簇序号是：0fffh 004h 005h 0fffh，簇号是12位比特，第4字节(1111 1111)的作为第2簇号的低8位(0-7)，第5字节(4F)的低4位(1111)作为第2簇号的高4位(8-11)，这样就得到了第2簇号的内容为0fffh；然后第5字节的高4位(0100)作为第3簇号的低4位(0-3)，第6字节(0000 0000)作为第3簇号的高8位(4-11)，这样便得到了第3簇号的内容为004h；第7字节(0000 0101)作为第4簇号的低8位(0-7)，第8字节的低4位(0000)作为第4簇号的高4位(8-11)，这样可以得到第4簇号的内容为005h；第8字节的高4位(1111)作为第5簇号的低4位(0-3)，第9字节(1111 1111)作为第5簇号的高8位(4-11)，这样得到第5簇号的内容为0fffh。

    这里叙述得非常的繁琐，实际情形是，用汇编来读取双字节(16位)，然后执行AND或右移操作就可以得到相应的簇号，非常的简便，后文会讲述到，从这里可以清楚，使用汇编来操作字节或位，非常的简单和功能强大。

    最终我们得到了这4个簇号的内容，内容是0ff0-0ff7h代表坏簇，磁道或柱面损坏不可使用，在格式式磁盘时由系统自动填充；内容是0ff8-0fffh代表文件内容结束，到此簇为止；其它的值代表着下一个簇号，接着我们分析一下刚才得到的那4个簇号的内容代表的意义，第2簇号的内容是0fffh，代表这个文件只占用了第2簇，文件的大小在512B(每簇1扇区512字节)之内。第3簇号的内容是004h，代表它的下一个簇是第4簇；第4簇号的内容是005h,代表它的下一个簇是第5簇，第5簇号的内容是0fffh，代表文件内容结束，所以这个文件应该是从第3簇开始到第5簇结束，占据了3簇的空间，文件大小在1536B(每簇1扇区512字节)之内。

    这种在簇号内容中储存下一个簇号的存储方式称为簇链，从文件目录项中查找到文件的首簇号，再在FAT文件分配表中查找对应的下一个簇号内容，直到它的内容是0fffh为止，就可以得到整个文件占用的所有簇空间。

    文件目录项最开始位于扇区19(隐藏隐藏+保留扇区+FAT文件分配表占用扇区=0+1+9+9)处，它一共占用了14个扇区的空间，在引导扇区的数据结构中偏移17处的字段BPB\_RootEntCnt指出了最大根目录数，在FAT12中的默认值是224，每个文件目录项占用了32字节的空间，那么共占用224\*32/512=14扇区。

    文件目录项的32字节数据结构如下图所示：

[](http://photo.blog.sina.com.cn/showpic.html%23blogid=3edcf6b80100crz1&url=http://s13.sinaimg.cn/orignal/3edcf6b8g68498badf03c)

    每个文件目录项的长度是32字节，偏移0处的8字节长的文件名，必须是大写的ASCII字符串，不足8字节就以空格(20h)来填充，接着是3字节长的文件后缀名，同样是大写的ASCII字符串，不足3字节以空格填充。如果文件名的第1字节是0E5h表示此文件已经删除，第1字节是0表示此目录项可用。

    位于偏移11处的是文件属性，使用比特位来设置值：

    00000000：普通文件，可随意读写

    00000001：只读文件，不可改写

    00000010：隐藏文件，浏览文件时隐藏列表

    00000100：系统文件，删除的时候会有提示

    00001000：卷标，作为磁盘的卷标识符

    00010000：目录文件，此文件是一个子目录，它的内容就是此目录下的所有文件目录项

    00100000：归档文件

    文件的属性可以叠加使用，可以具有多重属性，即设置为只读的时候也可以同时隐藏。

    偏移12处的共10字节长的内容没有使用，保留。

    偏移22处的是双字节长的文件最后修改时间，使用的方式是分位压缩存储方式，两个字节共16位从高到低分别存储时分秒的数值，其中时占用5位(11-15)，值从0-23，代表小时；分占用6位(5-10)，值从0-59,代表分钟；秒占用5位(0-4)，值为0-29，它的倍数(值\*2)就得到秒数。

    偏移24处的是双字节长的文件最后修改日期，使用的方式是分位压缩存储方式，两个字节共16位从高到低分别存储年月日的数值，其中年占用7位(9-15)，值从0-199，此值加上1980就可以得到年份，代表从1980到2099年；月占用4位(5-8)，值从1-12，代表月份；日占用5位(0-4)，值从1-31，代表当月天数。

    位于偏移26处的是双字节长的文件首簇号，利用此值就可以得到文件内容占用的第一个簇，然后在文件分配表中就可以得到所有的文件簇了。

    最后位于偏移28处的双字(4B)长的文件长度(字节)值。

    以上据说的扇区都是逻辑扇区，而通过BIOS读取的扇区都是使用CHS参数Int13h功能实现的，逻辑扇区转换成磁盘CHS参数的计算方式如下所示：

    根目录所占扇区数=最大根目录文件数\*32/每扇区字节数

    注意，在FAT12中：0e0h\*32/512=14

    用户数据区起始扇区=隐藏扇区+保留扇区+FAT表数\*FAT表所占扇区+根目录所占扇区

    注意，在FAT12中：0+1+2\*9+14=33

    簇起始线性扇区=用户数据区起始扇区+(簇号-2)\*每簇所占扇区-1

    例如：第2号簇的起始线性扇区=33+(2-2)\*1-1=32，而第6号簇的起始线性扇区=33+(6-2)\*1-1=36

    这样我们便可以得到具体簇号所在的起始线性扇区了，将线性扇区转换为磁盘CHS参数如下所示：

    扇区=线性扇区 MOD 每磁道扇区数+1

    注意：MOD为整除取模(余数)；如：32 MOD 18+1=14+1=15

    磁道=线性扇区/每磁道扇区数/磁头数

    注意：除法为取整，在FAT12默认情况中，每磁道扇区数为18，磁头数为2；如：32/18/2=1/2=0

    磁头=(线性扇区/每磁道扇区数) MOD 磁头数

    注意：除法为取整，MOD为整除取模；如：(32/18) MOD 2=1 MOD 2=1

    在文件根目录项和FAT文件分配表中得到具体的簇号后就可以根据上面的公式计算出磁盘CHS参数后，就可以通过BIOS的Int13h功能来进行读取扇区内容的操作了。在后文中，将编写一个能够读取FAT12格式下的一个子目录中的引导加载文件到内存中，以验证这篇文章的有效性，敬请浏览。

    附录(小知识)：系统如何检测文件系统是FAT12还是FAT16？或者FAT32？

    千万不要依靠引导扇区数据结构中的BS\_FileSysType字段(在FAT12中位于偏移54处)来比较这个字符串，很多时候这个字段很不准确，或许压根就没有(全是0或空格20h)，这个字段存在与否或内容是什么，与文件系统一点关系也没有，采取什么样的文件系统只有一个计算标准：此卷所有簇的数量。

    簇总数=(逻辑扇区总数-(隐藏扇区+保留扇区+FAT表数\*FAT表所占用扇区+根目录所占用扇区))/每簇扇区数

    注：逻辑扇区总数位于引导扇区数据结构偏移19处(字段BPB\_TotSec16，长度为2字节)或32处(字段BPB\_TotSec32，长度为4字节)，例如，在FAT12默认值中：(2880-(0+1+2\*9+14))/1=2847

    卷中簇的总数小于4085的为FAT12，总数大于或等于4085并且小于65525的为FAT16，总数大于或等于65525的为FAT32。这就是Microsoft操作系统认可的区分文件系统的标准，必须遵循此标准来操作磁盘卷，否则，Microsoft的操作系统将认为此卷损坏导致不可使用。