磁盘及文件系统

磁盘结构

1.主引导记录MBR（Main Boot Record）

位于**整个硬盘的第一个扇区**（偏移量为0），共占512个字节（即一个扇区），由三部分构成：

1. 第一部分是446字节的引导代码，也就是上面提到的MBR；
2. 第二部分是DPT（Disk Partition Table，硬盘分区表），包含4个表项，每个表项16字节，共占用64字节；
3. 第三部分是2个字节的结束标志，0x55AA。

图 2-1　主引导扇区的结构

分区表结构

分区表由四个分区项构成，每项16个字节，结构如书中P15页所示

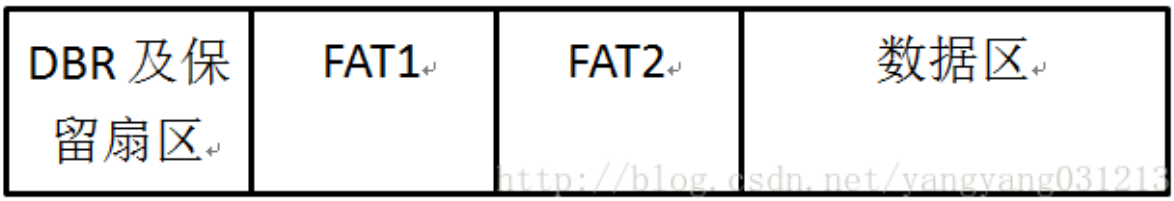
2.分区引导记录DBR（DOS Boot Record）

在对硬盘分区之后，每一个分区均有一个DBR与之对应。**DBR位于每个分区的第一个扇区，大小为512字节。**

其结构与文件系统有关。FAT32与NTFS的DBR格式不同。

FAT32文件系统-重点

结构：



1.引导扇区DBR即保留扇区

重点信息：

每扇区字节数、每簇扇区数

保留扇区数（第一个FAT表开始之前的扇区数，包括引导扇区），用于计算FAT表开始位置

根目录簇号（一般为2）

保留扇区

FAT32文件系统在DBR的保留扇区中安排了一个文件系统信息扇区，用以记录数据区中空闲簇的数量及下一个空闲簇的簇号，该扇区一般在分区的1号扇区，也就是紧跟着DBR后的一个扇区，

FAT32分区的BPB字段表			
字节位移	字段长度(字节)	图8对应取值	名称和定义
0x0B	2	0x0200	扇区字节数(Bytes Per Sector) 硬件扇区的大小。本字段合法的十进制值有512、1024、2048和4096。对大多数磁盘来说，本字段的值为512
0x0D	1	0x20	每簇扇区数(Sectors Per Cluster),一簇中的扇区数。FAT32文件系统只能跟踪有限个簇(最多为4 294 967 296个)，因此，通过增加每簇扇区数，可以使FAT32文件系统支持最大分区数。一个分区缺省的簇大小取决于该分区的大小。字段的合法十进制值有1、2、4、8、16、32、64和128。Windows 2000的FAT32实现只能创建最大为32GB的分区。但是，Windows 2000能够访问由其他操作系统(Windows 95、OSR2及其以后的版本)所创建的更大的分区
0x0e	2	0x0EA4	保留扇区数(Reserved Sector) 第一个FAT开始之前的扇区数，包括引导扇区。本字段的十进制值一般为32
0x10	1	0x02	FAT数(Number of FAT) 该分区上FAT的副本数。本字段的值一般为2
0x11	2	0x0000	根目录项数(Root Entries)只有FAT12/FAT16使用此字段。FAT32分区本字段必须设置为 0
0x13	2	0x0000	小扇区数(Small Sector)(只有FAT12/FAT16使用此字段)FAT32分区本字段必须设置为0
0x15	1	0xF8	媒体描述符(Media Descriptor)提供有关媒体被使用的信息。值0xF8表示硬盘，0xF0表示高密度的3.5寸软盘。媒体描述符要用于MS-DOS FAT16磁盘，在Windows 2000中未被使用
0x16	2	0x0000	每FAT扇区数(Sectors Per FAT)只被FAT12/FAT16所使用,FAT32分区本字段必须设置为0
0x18	2	0x003F	每道扇区数(Sectors Per Track) 包含使用INT13h的磁盘的“每道扇区数”几何结构值。该分区被多个磁头的柱面分成了多个磁道
0x1A	2	0x00FF	磁头数(Number of Head) 本字段包含使用INT 13h的磁盘的“磁头数”几何结构值。例如，在一张1.44MB 3.5英寸的软盘上，本字段的值为2

0x1C	4	0x04F4C000	隐藏扇区数(Hidden Sector) 该分区上引导扇区之前的扇区数。在引导序列计算到根目录的数据区的绝对位移的过程中使用了该值。本字段一般只对那些在中断13h上可见的媒体有意义。在没有分区的媒体上它必须总是为0
0x20	4	0x028B3000	总扇区数(Large Sector) 本字段包含FAT32分区中总的扇区数
0x24	4	0x000028AE	每FAT扇区数(Sectors Per FAT)(只被FAT32使用)该分区每个FAT所占的扇区数。计算机利用这个数和 FAT数以及隐藏扇区数(本表中所描述的)来决定根目录从哪里开始。该计算机还可以从目录中的项数决定该分区的用户数据区从哪里开始
0x28	2	0x0000	扩展标志(Extended Flag)(只被FAT32使用)该两个字节结构中各位的值为： 位0-3：活动 FAT数(从0开始计数，而不是1). 只有在不使用镜像时才有效 位4-6：保留 位7：0值意味着在运行时FAT被映射到所有的FAT 1值表示只有一个FAT是活动的 位8-15：保留
0x2A	2	0x0000	文件系统版本(File ystem Version)只供FAT32使用,高字节是主要的修订号，而低字节是次要的修订号。本字段支持将来对该FAT32媒体类型进行扩展。如果本字段非零，以前的Windows版本将不支持这样的分区
0x2C	4	0x00000002	根目录簇号(Root Cluster Number)(只供FAT32使用) 根目录第一簇的簇号。本字段的值一般为2，但不总是如此
0x30	2	0x0001	文件系统信息扇区号(File System Information SectorNumber)(只供FAT32使用) FAT32分区的保留区中的文件系统信息(File System Information, FSINFO)结构的扇区号。其值一般为1。在备份引导扇区(Backup Boot Sector)中保留了该FSINFO结构的一个副本，但是这个副本不保持更新
0x34	2	0x0006	备份引导扇区(只供FAT32使用) 为一个非零值，这个非零值表示该分区保存引导扇区的副本的保留区中的扇区号。本字段的值一般为6，建议不要使用其他值
0x36	12	12个字节均为0x00	保留(只供FAT32使用)供以后扩充使用的保留空间。本字段的值总为0

2.FAT表

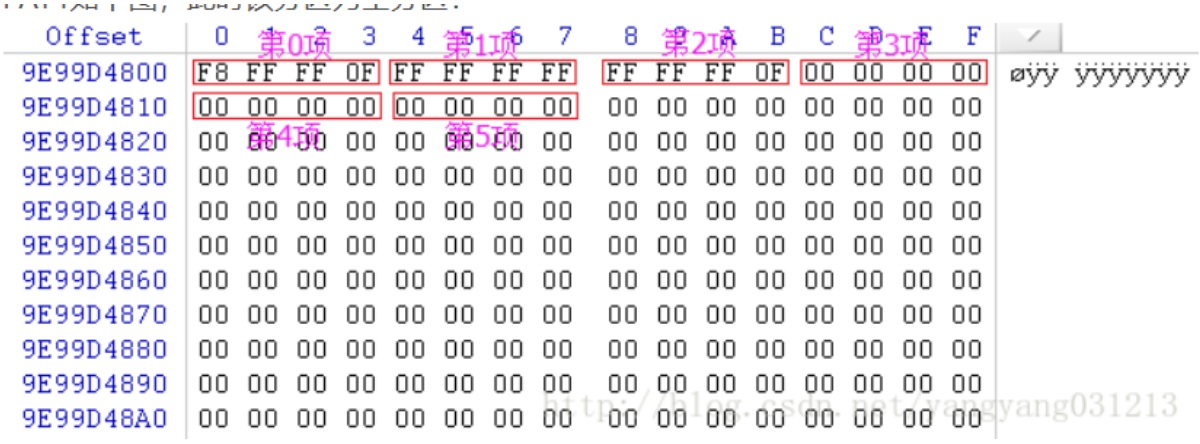
以簇链的形式记录每个文件所包含的簇号。FAT表共有两个，FAT2是FAT1的备份。

FAT1起始位置为：DBR所在位置+保留扇区数*扇区大小。

FAT表表项

FAT表项编号从0开始，但是编号0表示FAT介质类型，编号1表示FAT文件系统错误标志，所以实际存储从2号开始。大文件占用多个簇的话，则FAT项纪录下一个FAT项编号，依次类推直到最后以“0F FF FF FF“表示文件末尾。

FAT32中的32表示，FAT表项中的每一项占用32位。



3.数据区

起始位置：FAT2+FAT扇区数*扇区大小

数据区存储文件的目录项，目录项分为长目录项和短目录项

短目录项内容：

根据起始簇号的高低16位可以计算该文件的开始位置

表14　FAT32短文件目录项32个字节的表示定义			
字节偏移(16进制)	字节数	定义	
0x0~0x7	8	文件名	
0x8~0xA	3	扩展名	
0xB*	1	属性字节	00000000(读写)
			00000001(只读)
			00000010(隐藏)
			00000100(系统)
			00001000(卷标)
			00010000(子目录)
			00100000(归档)
0xC	1	系统保留	
0xD	1	创建时间的10毫秒位	
0xE~0xF	2	文件创建时间	
0x10~0x11	2	文件创建日期	
0x12~0x13	2	文件最后访问日期	
0x14~0x15	2	文件起始簇号的高16位	
0x16~0x17	2	文件的最近修改时间	
0x18~0x19	2	文件的最近修改日期	
0x1A~0x1B	2	文件起始簇号的低16位	
0x1C~0x1F	4	表示文件的长度	

长目录项内容

长文件名使用长目录项，采用Unicode存储，2字节为一个字，且为倒序存储文件名

表15　FAT32长文件目录项32个字节的表示定义

字节偏移	字节数	定义		
0x0	1	属性字节位意义	7	保留未用
			6	1表示长文件最后一个目录项
			5	保留未用
			4	顺序号数值
			3	
			2	
			1	
			0	
0x1~0xA	10	长文件名unicode码①		
0xB	1	长文件名目录项标志，取值0FH		
0xC	1	系统保留		
0xD	1	校验值(根据短文件名计算得出)		
0xE~0x19	12	长文件名unicode码②		
0x1A~0x1B	2	文件起始簇号(目前常置0)		
0x1C~0x1F	4	长文件名unicode码③		

FAT32中被删除文件的恢复

当文件被删除时

1. 目录项中文件名首字节被修改为E5，首簇高位被清零
2. FAT表簇链被清空
3. 文件内容无变化

恢复方法

1. 长文件名逆向定位完整文件名
2. 参考相邻目录项的首簇高位，还原首簇
3. 通过文件大小计算所占簇数，按照连续存储假设进行簇链修复，末簇为0FFFFFFF

NTFS文件系统-重点

NTFS文件系统详解_缘木之鱼的博客-CSDN博客_ntfs文件系统

NTFS （New Technology File System），是 Windows NT 环境的文件系统。新技术文件系统是Windows NT家族（如，Windows 2000、Windows XP、Windows Vista、Windows 7和 windows ...

https://blog.csdn.net/tianjin_ren/article/details/127241467

示意图如下：

1	2	MFT 分配的空间	文件存储区	3	文件存储区

注：

1. 1个引导扇区+15 个扇区的 NTLDR区域
2. MFT元数据文件
3. MFT前几个数据文件的备份

图 2-3 NTFS 文件系统示意图

1.引导扇区DBR

DBR包含3字节跳转指令，8字节的OEM代号，以及73字节的BPB（BIOS Paramter Block）信息

偏移（offset）	长度（字节）	含义
00-02H	3	跳转指令EB 52 90
03-0AH	8	文件系统的ASIIC码，表示形式NTFS
0B-0CH	2	每个扇区内的字节总数，一般为00 02H
0DH	1	簇大小
0E-0FH	2	保留扇区
10-12H	3	总为0
13H	1	不使用
14-15H	2	介质描述，硬盘为F8
16-17H	2	总为0
18-19H	2	每磁头扇区数
1A-1BH	2	每柱面磁头数
1C-1FH	4	隐含扇区（从MBR到DBR的扇区总数）
20-23H	4	不使用
24-27H	4	不使用，总为80 00 80 00
28-2FH	8	扇区总数，即分区大小
30-37H	8	$MFT的开始簇号
38-3FH	8	$MFTmirr的开始簇号
40-43H	4	每个MFT记录的簇数
44-47H	4	每索引的簇数
48-4FH	8	分区的逻辑序列号

2.主文档表MFT（Master File Table）

MFT由一个个文件记录组成，NTFS认为”磁盘上的一切皆文件“

MFT中前16个文件记录是元文件的记录，第17-23个记录是系统保留的记录，从第24个记录开始存放用户文件的记录。每个文件记录的大小为1kb(即偏移量为0x400)

定位MFT

MFT偏移量=MBR起始扇区*每扇区字节数+MFT起始簇号*每簇扇区数*每扇区字节数

MFT详解

文件记录由两部分构成，一部分是文件记录头，另一部分是属性列表，最后结尾是四个“FF”。在同一系统中，文件记录头的长度和具体偏移位置的数据含义是不变的，而属性列表是可变的，其不同的属性有着不同的含义。

偏移（offset）	长度（字节）	描述
0x0	4	固定值，一定是“FILE”
0x4	2	更新序列号的偏移
0x6	2	更新序列号与更新数组以字为单位大小（S）
0x8	8	日志文件序列号（每次记录被修改，都将导致该序列号加1）
0x10	2	序列号（记录本文件记录被重复使用的次数，每次文件删除时加1，跳过0值，如果为0，则保持为0）
0x12	2	硬连接数，只出现在基本文件记录中，目录所含项数要使用到它
0x14	2	第一个属性流的偏移地址
0x16	2	标志字节，1表示记录使用中，2表示记录为目录
0x18	4	文件记录实际大小（填充到8字节，即以8字节为边界）
0x1C	4	文件记录分配大小（填充到8字节，即以8字节为边界）
0x20	8	所对应的基本文件记录的文件参考号（扩展文件记录中使用，基本文件记录中为0，在基本文件记录的属性列表0x20属性存储中扩展文件记录的相关信息）
0x28	2	下一个自由ID号，当增加新的属性时，将该值分配给新属性，然后该值增加，如果MFT记录重新使用，则将它置0，第一个实例总是0。
0x2A	2	边界，windows xp中使用，也就是本记录使用的两个扇区的最后两个字节的值
0x2C	4	windows xp中使用，本MFT记录号
-	2	更新序号
-	2S-2	更新序列数组

在NTFS文件系统中所有与文件相关的数据结构均被认为是属性，包括文件的内容。文件记录是一个与文件相对应的文件属性数

属性有常驻与非常驻之分。

1. 当一个文件很小时，其所有属性体都可以存放在文件记录中，该属性就称为常驻属性。

2. 如果某个文件很大，1KB的文件记录无法记录所有属性时，则文件系统会在MFT元文件之外的区域（也称数据流）存放该文件的其他文件记录属性，这些存放在非MFT元文件之外的区域（也称数据流）存放该文件的其他文件记录属性，这些存放在非MFT元文件内的记录就称为非常驻属性。

常驻属性

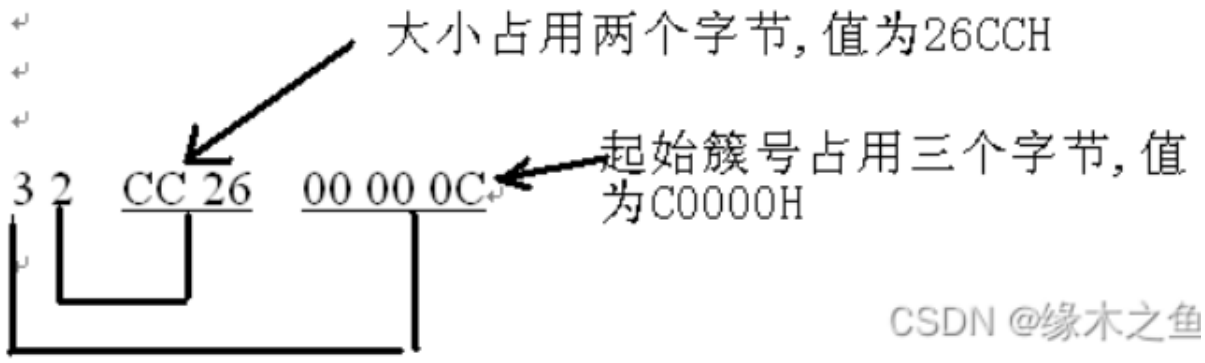
偏移（offset）	长度（字节）	常用值	含义
00-03	4	-	属性类型
04-07	4	-	属性的长度，8的整数倍 **(整个属性长度)**
08	1	00	是否为常驻属性，00表示为常驻属性
09	1	00	属性名的长度，00表示没有属性名
0A-0B	2	18 00	属性值的开始偏移
0C-0D	2	00	标志，如压缩、加密、稀疏等
0E-0F	2	00	标识
10-13	4	Length	属性长度
14-15	2	18	属性体开始位置
16	1	-	索引标志
17	1	-	填充
18	Length	-	属性体开始

非常驻属性

偏移（offset）	长度（字节）	常用值	含义
00-03	4	-	属性类型
04-07	4	-	属性的长度，8的整数倍（整个属性长度）
08	1	01	是否为常驻属性，01表示为非常驻属性
09	1	00	属性名的长度，00表示没有属性名
0A-0B	2	-	属性值的开始偏移
0C-0D	2	-	标志，如压缩、加密、稀疏等
0E-0F	2	-	属性ID
10-17	8	-	起始虚拟簇号VCN
18-1F	8	-	结束虚拟簇号VCN
20-21	2	40	Data Run的偏移地址
22-23	2	-	压缩单位大小，2的N次方
24-27	4	-	不使用
28-2F	8	-	属性分配大小
30-37	8	-	属性实际大小
38-3F	8	-	属性原始大小
40	-	-	Data Run信息

非常驻属性的位置计算方法，利用Data Run信息，也叫Run List

Data Run的第一个字节为压缩字节，高位表示该非常驻属性的起始簇号占用多少个字节，低位表示该属性的大小。第一个字节后的字节分别表示大小、起始簇号。



10H属性

10H属性包含文件的一些基本信息，如文件的传统属性，文件的创建时间和最后修改时间和日期，文件的硬链接数等等。

偏移(offset)	长度（字节）	操作系统	描述
-	-	-	标准属性头（已经分析过）
0x00	8	-	C TIME 文件创建时间
0x08	8	-	A TIME 文件修改时间
0x10	8	-	M TIME MFT变化时间
0x18	8	-	R TIME 文件访问时间
0x20	4	-	文件属性（按照DOS术语来称呼，都是文件属性）
0x24	4	-	文件所允许的最大版本号（0表示未使用）
0x28	4	-	文件的版本号（最在版本号为0，则他为0）
0x2C	4	-	类ID（一个双向的类索引）
0x30	4	2K	所有者ID（表示文件的所有者，是文件配额$QUOTA$中O和$Q索引的关键字，为0表示未使用磁盘配额）
0x34	4	2K	安全ID是文件$SECURE$中SII和$SDS数据流的关键字，注意不要与安全标识相混淆
0x38	8	2K	本文件所占用的字节数，它是文件所有流占用的总字节数，为0表示未使用磁盘配额
0x40	8	2K	更新系列号（USN），是到文件$USNJRNL的一个直接的索引，为0表示USN日志未使用

20H属性

20H类型属性既属性列表，当一个文件需要好几个文件记录时，才会用到20H属性。20H属性记录了一个文件的下一个文件记录的位置。

偏移(offset)	长度(字节)	描述
-	-	标准属性头（已经分析过）
0x00	4	类型
0x04	2	记录长度
0x06	1	属性名长度（N，为0表示没有属性名）
0x07	1	属性名偏移（如果没有属性名，则指向属性内容）
0x08	8	起始VCN（属性常驻时为0）
0x10	8	属性的基本文件记录中的文件参考号（所有MFT的文件都有一个文件索引号，引用到这个文件参考号，等价于引用这个文件记录，这个参考号在文件记录头中有定义）
0x18	2	属性ID（每个属性都有一个唯一的ID号）
0x1A	2N	Unicode属性名（如果有属性名）

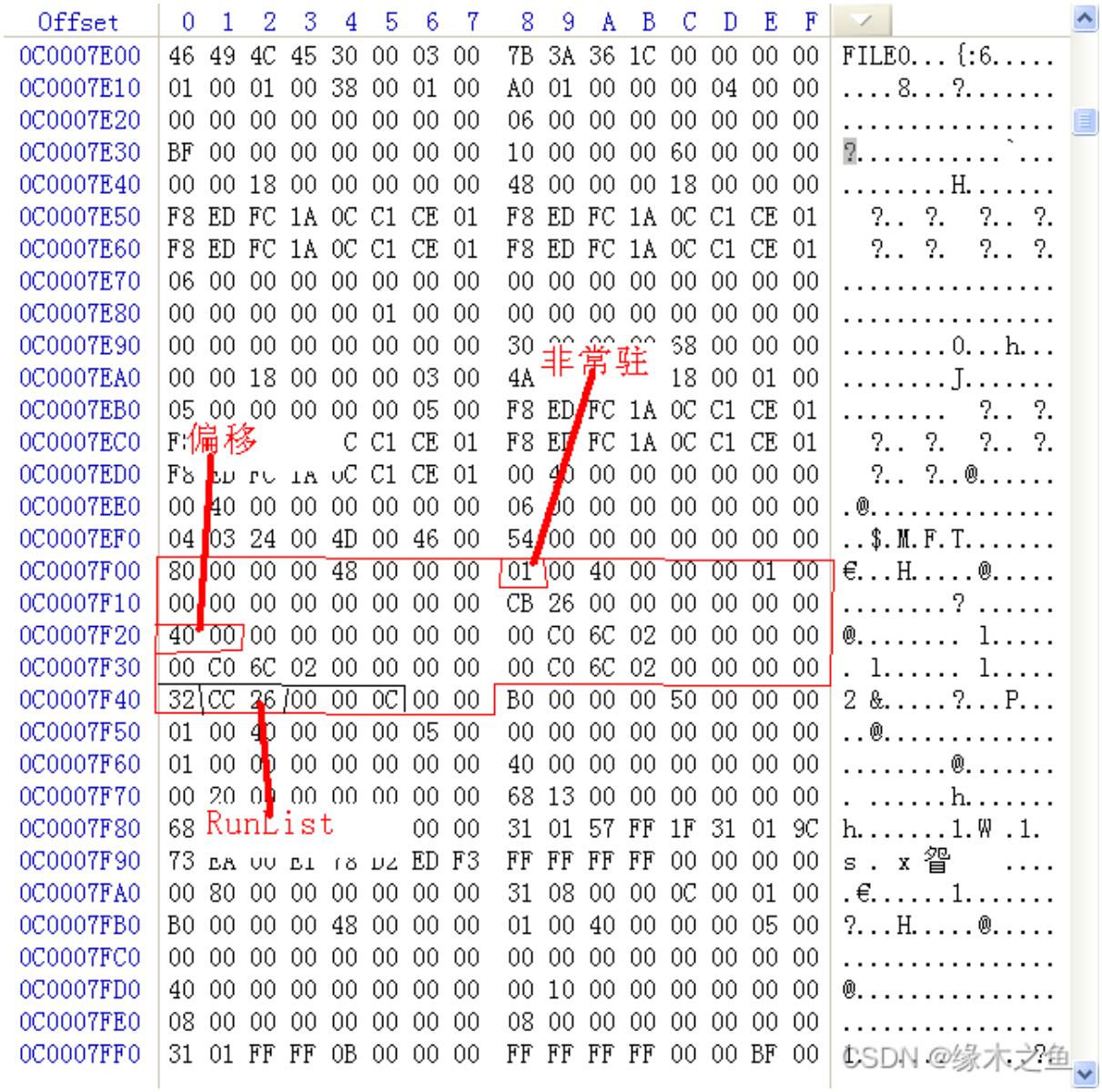
30H属性

该属性用于存储文件名 ，它总是常驻属性。最少68字节，最大578字节，可容纳最大Unicode字符的文件名长度。

偏移(offset)	长度(字节)	描述
-	-	标准属性头（已经分析过）
0x00	8	父目录的文件参考号（即父目录的基本文件记录号，分为两部分，前6个字节48位为父目录的文件记录号，此处为0x05，即根目录，所以$MFT的父目录为根目录，后2个字节为序列号）
0x08	8	文件创建时间
0x10	8	文件修改时间
0x18	8	最后一次MFT更新时间
0x20	8	最后一次访问时间
0x28	8	文件分配大小
0x30	8	文件实际大小
0x38	4	标志，如目录、压缩、隐藏等
0x3C	4	用于EAS和重解析点
0x40	1	以字符计的文件名长度，每字符占用字节数由下一字节命名空间确定，一个字节长度，所以文件名最长255字节。
0x41	1	文件名命名空间
0x42	2L	以Unicode方式表示的文件名

80H属性

80H属性是文件数据属性，该属性容纳着文件的内容，文件的大小一般指的就是未命名数据流的大小。该属性没有最大最小限制，最小情况是该属性为常驻属性



90H属性

90H属性是索引根属性，该属性是实现NTFS的B+树索引的根节点（仅在根节点中存在），它总是常驻属性。

索引根结构表：

偏移（offset）	长度（字节）	描述
-	-	标准属性头（已经分析过）
0x00	4	属性类型
0x04	4	排序规则
0x08	4	索引项分配大小（字节数）
0x0C	1	每索引记录的簇数
0x0D	3	填充（到8字节）

索引头结构表：

偏移（offset）	长度（字节）	描述
-	-	标准属性头（已经分析过）
0x00	4	第一个索引项的偏移
0x04	4	索引项的总大小
0x08	4	索引项的分配大小
0x0C	1	标志
0x0D	3	填充（到8字节）

索引项结构表：

偏移（offset）	长度（字节）	描述
0x00	8	文件的MFT参考号
0x08	2	索引项大小
0x0A	2	文件名偏移
0x0C	2	索引标志
0x0E	2	填充（到8字节）
0x10	8	父目录的MFT文件参考号
0x18	8	文件创建时间
0x20	8	最后修改时间
0x28	8	文件记录最后修改时间
0x30	8	最后访问时间
0x38	8	文件分配大小
0x40	8	文件实际大小
0x48	8	文件标志
0x50	1	文件名长度（F）
0x51	1	文件名命名空间
0x52	2F	文件名
2F+0x52	P	填充（到8字节）
P+2F+0x52	8	子节点索引缓存的VCN

A0H属性

A0属性是索引分配属性，也是一个索引的基本结构，存储着组成索引的B+树目录索引子节点的定位信息。即，当前文件不是根目录时，用于存储索引项的属性，与90H属性基本一致。

索引项中的首八个字节表示该文件的MFT参考号，据此可以找到该文件的位置。

标准索引头的解释如下：

偏移（offset）	长度（字节）	描述
0x00	4	总是“INDX”
0x04	2	更新序列号的偏移
0x06	2	更新序列号与更新数组以字为单位的大小（S）
0x08	8	日志文件序列号
0x10	8	本索引缓存在索引分配中的VCN
0x18	4	索引项的偏移
0x1C	4	索引项大小
0x20	4	索引项分配大小
0x24	1	如果不是叶节点，置1，表示还有子节点
0x25	3	用0填充
0x28	2	更新序列
0x2A	2S-2	更新序列数组