

中图分类号: O157
密 级: 公开

单位代号: 10280
学 号: 18724781

上海大学



硕士学位论文

SHANGHAI UNIVERSITY
MASTER'S DISSERTATION

题 目	FAT32文件系统的FAT数据结 构
--------	-----------------------

作 者 王鹏

学科专业 计算机视觉

导 师 王宜敏

完成日期 二〇一八年十月

上海大学

本论文经答辩委员会全体委员审查，确认符合上海大学硕士学位论文质量要求。

答辩委员会签名:

主席:

委员:

导 师:

答辩日期:

原创性声明

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签名：_____ 日期：_____

本论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定。即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

(保密的论文在解密后应遵守此规定)

签名：_____ 导师签名：_____ 日期：_____

上海大学工学硕士学位论文

FAT32文件系统的FAT数据结构

作 者：王鹏

导 师：王宜敏

学科专业：计算机视觉

计算机科学与技术学院

上海大学

2018 年 10 月

A Dissertation Submitted to Shanghai University for the
Degree of Master in Science

An Introduction to L^AT_EX Thesis Template of Shanghai University v2.0

Candidate: ahhylau

Supervisor:

Major: Operational Research and Control Theory

**Department of Mathematics
Shanghai University
May, 2017**

摘 要

FAT文件系统是最原始，兼容和简单的文件系统。在这个时代仍然支持数字设备，如迷你MP3播放器，智能手机和数码相机。由于其简单性和传统性，几乎所有操作系统都支持该文件系统。本文综述了FAT32文件系统最重要的构建块数据结构的基本设计技术，约束和公式以及FAT数据结构。

关键词: FAT，文件分配表，FAT32，文件系统，群集，扇区

ABSTRACT

Abstract in English.

Keywords: T_EX; L^AT_EX; Template; Thesis

目 录

第一章 FAT32介绍	1
第二章 结构	2
2.1 FAT32卷的数据结构排序	2
2.2 FAT32文件系统由四个不同的部分组成.....	2
第三章 FAT32文件系统卷的FAT数据结构	3
3.1 基本设计技术	3
3.2 约束	4
3.3 FAT中的前两个条目存储特殊值.....	4
3.4 公式	4
第四章 结论和未来的工作	6
插图索引	7
表格索引	8
参考文献	9
作者在攻读硕士学位期间发表的论文与研究成果	10
致 谢	11
附录 A 经典不等式	12

主要符号对照表

\mathcal{T}	张量
H	超图
$\mathcal{A}(H)$	超图 H 的邻接张量
$\mathcal{L}(H)$	超图 H 的拉普拉斯张量
$\mathcal{Q}(H)$	超图 H 的无符号拉普拉斯张量
ρ	谱半径
G	图
κ	连通度
χ	染色数
Δ	最大度
δ	最小度

第一章 FAT32介绍

FAT（文件分配表）文件系统是在20世纪70年代末和80年代初开发的，是Microsoft®MS-DOS®操作系统1支持的文件系统。它是各种操作系统的主要文件系统，包括DR-DOS，FreeDOS，MS-DOS，OS/2（v1.1）和Microsoft Windows（最多Windows Me）。FAT最初是为小于500K的软盘驱动器开发的。随着存储容量的增加，FAT得到了增强，以支持大型存储介质。因此，我们三个完整记录的FAT文件系统类型：FAT12，FAT16和FAT32。exFAT 2是最近的Microsoft®汇编，而KFAT 3，TFAT 4和FATTY 5是同一研究人员和其他研究人员对实际设计的可靠性增强。

与其他文件系统相比，FAT的性能较差，因为它使用简单的数据结构，在存在许多小文件的情况下使文件操作耗时且低效的磁盘空间利用率。但是对于相同的简单设计和传统，几乎所有现有的个人计算机操作系统都支持它。这使其成为固态存储卡的有用格式，也是在操作系统之间共享数据的便捷方式。

在当今世界，一些便携式数字设备，如迷你MP3播放器，智能手机和数码相机正在成为我们生活的一部分。这些设备经常与台式计算机交换数据。PC将这些设备作为标准USB大容量存储设备发现，并自动将文件系统卷安装在其中。仅当PC的操作系统支持设备中使用的文件系统时，才可以执行此操作。这就是为什么能够解决大型存储介质并且受所有主要桌面操作系统支持的传统FAT32 1文件系统仍然是便携式数字设备中使用最广泛的文件系统的原因。

第二章 结构

2.1 FAT32卷的数据结构排序

The following table shows the order of the data structures that compose a FAT32 disk volume.

表 2.1 模排序结构

Boot Sector	Reserved Sectors	FAT (Copy 1)	FAT (Copy 2)	File& Directory Sectors
----------------	---------------------	-----------------	-----------------	-------------------------------

2.2 FAT32文件系统由四个不同的部分组成

1. 引导扇区位于卷的开头，即第0扇区。它包括一个名为BPB（BIOS参数块）的区域，它从偏移量11开始，包含一些基本的文件系统信息。扇区的其余部分通常包含引导加载程序代码
2. 保留扇区紧跟引导扇区。包括引导扇区的卷的保留扇区数由引导扇区的偏移量14处的BPB指示。通常，保留的扇区包括扇区1的文件系统信息扇区和卷的扇区6的备份引导扇区。
3. 文件分配表是一个32位宽的条目数组，跨越BPB指示的多个扇区，位于引导扇区的偏移量36处。FAT32通常有两个FAT数据结构副本，以便冗余检查磁盘介质，而一个是FLASH介质。引导扇区的BPB偏移40的第7位指示FAT是否镜像。该区域为文件系统和后缀32赋予FAT名称
4. 文件和目录扇区构成文件系统的其余部分，直到存储实际文件和目录数据的卷的末尾。FAT32通常会在第一个集群中占用根目录（一个集群是固定数量的连续扇区，该数量由文件和目录扇区的引导扇区的偏移量13处的BPB指示），并由偏移处的BPB指示引导扇区44。

第三章 FAT32文件系统卷的FAT数据结构

所有FAT文件系统最初都是为IBM PC机器架构开发的，因此FAT对BPB，FAT和文件和目录条目中的条目使用小端格式[1]。FAT数据结构是存储关于使用哪些集群，空闲或可能不可用的信息的表。除此之外，它还存储有关属于特定文件的集群链的信息。根据所使用的FAT文件系统的类型和卷的大小，簇大小会有所不同，每个簇的连续扇区数可能是1,2,4,8,16,32或64 [1]。由于每个容量的内存成本每年都在急剧下降[9]，最大集群数量急剧增加，因此用于标识每个集群的位数也在增长。因此，FAT格式连续主要版本以用于寻址集群的表元素位数来命名：12,16,32和64 [2]。在FAT32中，FAT条目是32位宽，但只有低28位用于寻址 2^{28} 个簇。因此，FAT32体积可以与 $((2^{28}) * 64) / 2KB$ 一样大，等于8TB。

3.1 基本设计技术

FAT32卷上的每个文件和目录（根目录除外）在其父目录中都有一个包含名称，属性，大小等的条目以及分配给它的32位宽的第一个簇号。对应任何簇编号，FAT条目可以具有下面给出的某些允许值：

表 3.1 基本设计技术

FAT32 Cluster Entry Values	Description
0x00000000	Is Free Cluster
0x00000001	Reserved value
0x00000002–0x0FFFFFFF	Is Used Cluster and value points to next cluster in the chain allocated to file/directory
0x0FFFFFFF0–0x0FFFFFFF6	Reserved values
0x0FFFFFFF7	Some Bad sector in Cluster, Unusable
0x0FFFFFFF8–0x0FFFFFFF	Is Last Cluster in file/directory or EOC (End Of Cluster chain) marker

每个文件/目录可以占用一个或多个集群，这取决于其大小和每个集群的扇区数。因此，文件由这些簇的链表示。然而，这些集群不一定在磁盘表面上彼此相邻地存储，而是经常在整个文件和目录扇区中分段。

假设有两个文件，比如MYFILE1.TXT和MYFILE2.TXT目前驻留在FAT32卷上，前者是碎片，3个簇是长的，而后者没有碎片，两个簇长，如图1所示。

MYFILE1.TXT的第一个簇分配为0x00000029，针对该簇的FAT内容显示另一个

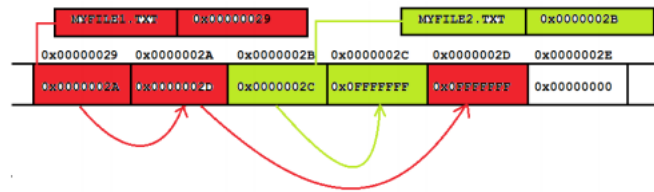


图 3.1 FAT32卷的FAT数据结构快照

簇0x0000002A，然后是0x0000002D，其FAT内容显示该簇是链中的最后一个簇。类似地，对于MYFILE2.TXT，分配的第一个簇是0x0000002B，其FAT内容指向链中的下一个簇0x0000002C，这是其FAT内容指向的链中的最后一个簇。

3.2 约束

FAT32 FAT条目实际上使用低28位来寻址簇。FAT32 FAT条目的高4位是保留的，只有在格式化卷时才会更改，此时整个32位FAT条目应归零，包括高4位。这意味着所有这些32位群集条目值：0xA0000000, 0xB0000000和0x00000000表示空闲群集，因为低28位设置为0。

假设32位空闲簇值当前为0xA0000000，并且我们希望通过在其中存储值0x0FFFFFFF7来将此簇标记为坏。然后32位条目应包含值0xAFFFFFFF7，因为当我们写入0x0FFFFFFF7坏簇标记时，我们必须保留高4位。

因为BPB在引导扇区的偏移11处指示的每个扇区的字节数总是可以被4整除，所以FAT32 FAT条目永远不会跨越扇区边界

3.3 FAT中的前两个条目存储特殊值

- 第一个条目包含引导扇区偏移21处的BPB副本，该副本长度为8位，表示存储介质的类型。此条目的高4和低8之间的剩余20位设置为1。
- 第二个条目存储EOC标记。此条目的高位两位有时用于脏卷管理：如果设置为1则高位指示上次关闭是否干净，否则异常。如果设置为1，则下一个最高位表示在上一次安装期间没有检测到磁盘I/O错误，否则有其他错误。

因为前两个FAT条目存储特殊值，所以没有簇0或1。FAT32 FAT数据结构中的第一个可寻址簇是簇2，这就是为什么引导扇区偏移44处的BPB值表示根目录簇的原因number不能小于2，通常为2，即根目录位于文件/目录区域的开头。

3.4 公式

此处计算的所有扇区号都是相对于FAT32卷的第一个扇区，即引导扇区，并不

一定直接映射到驱动器上，因为由于分区和代码，卷的扇区0不一定是驱动器的扇区0 片段采用CProgramming语言。

文件/目录区域的开始，即集群2的第一个扇区，计算如下：

$$FirstDataSector = BPB_ResvdSecCnt + (BPB_NumFATs * FATSz)$$

*BPB_ResvdSecCnt is the number of reserved
sectors at offset 14 of Boot sector*

*BPB_NumFATs is the count of FAT data
structures at offset 16 of Boot Sector*

*FATSz is the count of sectors occupied by one
FAT copy at offset 36 of Boot Sector*

(3.1)

给定任何有效的数据簇号N，该簇的第一扇区的扇区号计算如下：

$$FirstSectorofCluster = ((N \sim 2) * BPB_SecPerClus) + FirstDataSector$$

*BPB_SecPerClus is the count of sectors per
cluster at offset 13 of Boot Sector*

(3.2)

给定任何有效数据簇号N，包含其条目的FAT扇区号和该FAT扇区中的偏移量计算如下：

$$FATOffset = N * 4$$

$$ThisFATSecNum = BPB_ResvdSecCnt + (FATOffset / BPB_BytsPerSec)$$

$$ThisFATEntOffset = FATOffset \% BPB_BytsPerSec$$

*BPB_BytsPerSec is the count of bytes per sector
at offset 11 of Boot Sector*

(3.3)

上面计算的FAT扇区属于FAT的第一个副本；如果要使用第二个副本，则FAT扇区计算如下：

$$ThisFATSecNum = BPB_ResvdSecCnt + (FATOffset / BPB_BytsPerSec) + FATSz$$
(3.4)

第四章 结论和未来的工作

在本文中，我们讨论了FAT32中FAT数据结构的基本和简单设计技术，并回顾了处理FAT32 FAT数据结构所需的约束和公式。由于其设计简单，它被小型数字设备广泛支持到桌面操作系统。有一些问题，如FAT数据结构的可靠性问题和兼容性问题，优化的算法等。

插图索引

图 3.1	FAT32卷的FAT数据结构快照.....	4
-------	-----------------------	---

表格索引

表 2.1	排序结构	2
表 3.1	基本设计技术	3

参考文献

- [1] Microsoft Corporation, “FAT32 File System Specification” , [http://microsoft.com/whdc/ system/platform/firmware/fatgen.mspix](http://microsoft.com/whdc/system/platform/firmware/fatgen.mspix), 2000
- [2] Microsoft Corporation, "Extended FAT File System", [http://msdn2.microsoft.com/en-us/ library/aa914353.aspx](http://msdn2.microsoft.com/en-us/library/aa914353.aspx), 2007
- [3] M. S. Kwon, S. H. Bae, S. S. Jung, D. Y. Seo, and C. K. Kim, “KFAT: Log-based Transactional FAT File system for Embedded Mobile Systems” , In Proceedings of 2005 US-Korea Conference, ZCTS-142, 2005
- [4] Microsoft MS-DOS Programmer’ s Reference: version 5.0.” , Microsoft press. 1991.
- [5] Liang Alei, Liu Kejia, Li Xiaoyong , “FATTY : A reliable FAT File System” , Proceedings of the 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools, Pages: 390-395, 2007.

作者在攻读硕士学位期间发表的论文与研究成果

发表的学术论文

1. Lele Liu, Liying Kang, Xiying Yuan, On the principal eigenvectors of uniform hypergraphs. *Linear Algebra and its Applications*, 511 (2016) 430-446. (SCI 收录)
2. Wei Zhang, Lele Liu, Liying Kang, Yanqin Bai, Some properties of the spectral radius for general hypergraphs. *Linear Algebra and its Applications*, 513 (2017) 103–119. (SCI 收录)

研究成果

1. Ahhy Lau, 上海大学研究生(硕博)学位论文 \LaTeX 模板 SHUThesis.

致 谢

衷心感谢导师 xxx 教授对本人的精心指导.

附录 A 经典不等式

论文中用到的经典不等式.

(Hölder Inequality) 设 $a_i \geq 0, b_i \geq 0, i = 1, 2, \dots, n$, 且 $p > 1, q > 1$ 满足 $1/p + 1/q = 1$. 则有

$$\sum_{i=1}^n a_i b_i \leq \left(\sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^n b_i^q \right)^{\frac{1}{q}},$$

等号成立当且仅当存在一个常数 c 满足 $a_i^p = c b_i^q$.

(PM Inequality) 设 x_1, x_2, \dots, x_n 是 n 个非负实数. 如果 $0 < p < q$, 那么

$$\left(\frac{x_1^p + x_2^p + \dots + x_n^p}{n} \right)^{\frac{1}{p}} \leq \left(\frac{x_1^q + x_2^q + \dots + x_n^q}{n} \right)^{\frac{1}{q}},$$

等号成立当且仅当 $x_1 = x_2 = \dots = x_n$.

(AM-GM Inequality) 设 x_1, x_2, \dots, x_n 是 n 个非负实数. 则有

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n},$$

等号成立当且仅当 $x_1 = x_2 = \dots = x_n$.