从硬盘到文件系统

这篇文章将要为大家介绍硬盘的组成及其工作原理，然后由硬盘引申到文件系统。

首先来为大家介绍硬盘的分类。

硬盘按接口来分的话一般可以分为：

SATA (Serial Advanced Technology Attachment，串行高级技术附件)接口硬盘。（SATA后面还有SATAⅡ）

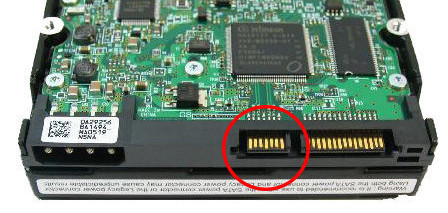


图1

SAS(Serial Attached SCSI)和接口硬盘。



图2

大家有没有觉得SAS接口硬盘和SATA接口硬盘很像？事实是它们俩确实很像，像到了SAS和SATA在物理层上完全兼容的地步，但是SAS和SATA又确确实实是两种不同的接口。

SCSI(Small Computer System Interface（[小型计算机系统接口](http://baike.baidu.com/view/611524.htm)）)接口硬盘。



图3

IDE（Integrated Drive Electronics集成驱动器）接口硬盘。



图4

FC(Fiber Channel光纤通道)接口硬盘



图5

还有大家非常熟悉的USB接口的硬盘。



图6

接下来我们来看看硬盘的组成。

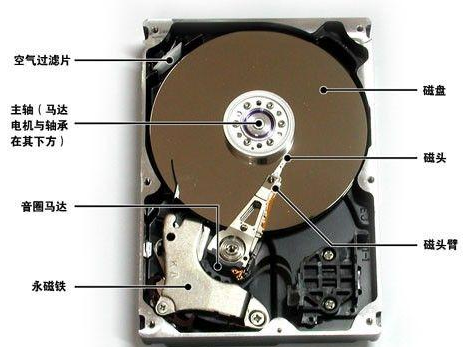


图7

除了固态硬盘之外，硬盘一般都由磁盘、主轴马达、磁头臂、磁头、永磁铁等部分组成。

硬盘的每一个组成部件都有其独特的作用，下面让我们来看看硬盘的工作原理。

硬盘的存储介质是磁性材料，磁头通过电流改变磁盘的磁性来存储数据。

硬盘在逻辑上被划分为磁道、柱面以及扇区。

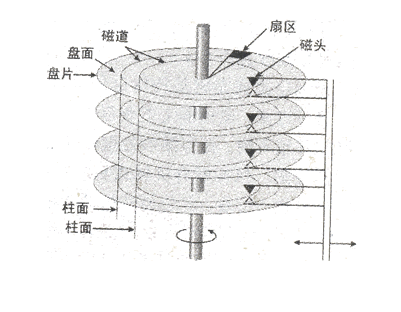


图8

硬盘的每个盘片的每个面都有一个读写磁头，磁盘盘面区域的划分如图所示。

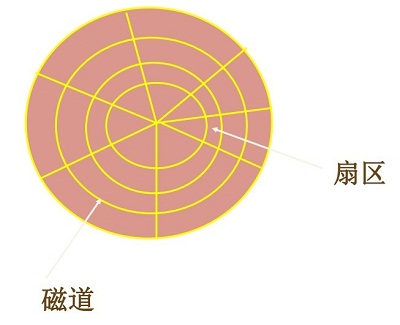


图9

硬盘的读写和扇区有着紧密相连的关系，在说扇区之前，我们来看一下与扇区相关的盘面、磁道、和柱面。盘面、磁道和柱面的位置大家可以参照图8。

盘面：

硬盘的盘面一般用铝合金材料做基片，高速硬盘也用玻璃做基片。硬盘的每一个盘面都有上下两个盘面。一般每个盘面都会利用以存储数据成为有效盘。当然也有极个别的硬盘盘面数为单数。每一个有效盘都有一个盘面好，按顺序从上至下由0开始编号。盘面号又叫磁头号，因为每一个有效盘都有一个对应的读写磁头。

磁道：

磁盘在格式化时被划分为许多同心圆，这些同心圆轨迹叫做磁道。磁道从外向内由0开始编号。硬盘的每一个盘面有300——1024个磁道（新式大容量盘每面的磁道会更多）。信息以脉冲串的形式记录在这些轨迹中，这些同心圆不是连续记录数据，而是被划分为一段段的圆弧，圆弧的角速度这些圆弧的角速度一样。由于径向长度不一样，所以线速度也不一样，外圈的线速度较内圈的线速度大。每段圆弧叫做一个扇区，扇区从1开始编号，每个扇区的数据作为一个单元同时读出或写入。

柱面：

所有盘面上的同一磁道构成一个圆柱，通常称为柱面。每个圆柱上的磁头由上而下从0开始编号。磁头读/写数据时首先在同一柱面内从0磁头开始操作，依次向下在同一柱面的不同盘面上进行操作，只在同一柱面所有磁头全部读/写完毕后磁头才转移到下一柱面。

扇区：

操作系统以扇区形式将信息存储在硬盘上，每个扇区包括512个字节的数据和一些其他信息。一个扇区有两个主要部分：存储数据地点的标识符和存储数据的数据段。扇区的第一个主要部分是标识符。标识符，就是扇区头标，包括组成扇区三维地址的三个数字：

盘面号：扇区所在的磁头（或盘面）

柱面号：磁道

扇面号：扇区在磁道的位置，也叫块号。

好了，我们对磁盘的基础知识就介绍在这了，下面让我们通过磁盘来了解计算机的文件系统。

磁头靠近主轴接触的表面，即线速度最小的地方，有一块很特别的地方。为什么说它特别呢，因为它不存放数据，称为启停区或着陆区，在启停区外就是数据区。在数据区的最外面一圈就是“0”磁道。硬盘的数据存放就是从最外圈开始的。

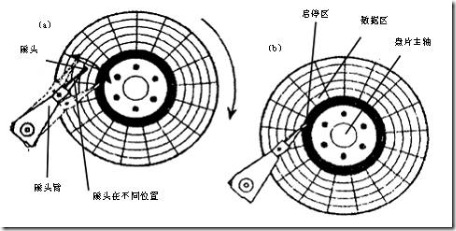


图10

在硬盘启动时，我们的磁头会首先从启停区移动到0磁道，具体是0柱面，0磁头，1扇区。硬盘的0柱面、0磁头、1扇区为主引导扇区，FDISK程序写到该扇区的内容为主引导记录（MBR），该记录占用446个字节，它用于硬盘启动时将系统控制权转移给用户指定的并在分区表中登记了的某个操作系统区。

大家可能会问，什么是主引导扇区和主引导记录。下面让我们来看看这个在硬盘启动时比系统区还要先读取的扇区吧。

主引导记录（MBR，Main Boot Record）是位于磁盘最前边的一段引导代码。它负责[磁盘](http://baike.baidu.com/view/157418.htm)操作系统(DOS)对磁盘进行读写时分区合法性的判别、分区引导信息的定位，它由磁盘操作系统(DOS)在对硬盘进行初始化时产生的。硬盘的主引导记录（MBR）是不属于任何一个操作系统的，它先于所有的操作系统而被调入内存，并发挥作用，然后才将控制权交给[主分区](http://baike.baidu.com/view/1304004.htm)（[活动分区](http://baike.baidu.com/view/493740.htm)）内的操作系统，并用主分区信息表来管理硬盘。

包含主引导记录的扇区叫主引导扇区。

在我们的生活中，许多人都会认为分区就一定要把硬盘划分成好几个部分，其实我们完全可以只创建一个分区使用全部或部分的硬盘空间。不过，不论我们划分了多少个分区，也不论使用的是[SCSI硬盘](http://baike.baidu.com/view/681438.htm)还是[IDE硬盘](http://baike.baidu.com/view/1330089.htm)，都必须把硬盘的[主分区](http://baike.baidu.com/view/1304004.htm)设定为[活动分区](http://baike.baidu.com/view/493740.htm)，这样才能够通过硬盘启动系统。

我们的磁盘最多只能有四个主分区（[硬盘分区](http://baike.baidu.com/view/14995.htm)实质上是对[硬盘](http://baike.baidu.com/view/4480.htm)的一种[格式化](http://baike.baidu.com/view/902.htm)，然后才能使用硬盘保存各种信息。创建分区时，就已经设置好了硬盘的各项物理[参数](http://baike.baidu.com/view/327406.htm)，指定了硬盘[主引导记录](http://baike.baidu.com/view/418401.htm)和引导记录[备份](http://baike.baidu.com/view/235162.htm)的存放位置。）。其中如果要设定扩展分区的话，扩展分区占一个主分区位，扩展分区还可以再分为若干个逻辑分区。（硬盘可以设好多分区，但是能够让你设置为系统盘的就只有那么四个主分区位）

Linux中，IDE硬盘最多只能有59个逻辑分区（5-63号），SATA接口硬盘最多只能有11个逻辑分区（5-15号）

[DOS](http://baike.baidu.com/view/365.htm)和[FAT文件系统](http://baike.baidu.com/view/4757778.htm" \t "_blank)最初都被设计成可以支持在一块[硬盘](http://baike.baidu.com/view/4480.htm)上最多建立24个分区，分别使用从C到Z 24个[驱动器](http://baike.baidu.com/view/15529.htm" \t "_blank)盘符。

为什么一个硬盘只能分四个主分区呢?

这是由个人计算机初期的设计架构决定的.一开始,PC被设计成这样的启动过程,开机时主板BIOS进行自检,当一切OK后,就开始读取硬盘的主引导扇区。通过上面的一些铺垫我们可以知道，主引导扇区的大小是512字节。BIOS把这512字节的内容读出来并执行。这512字节包含了主引导记录MBR，磁盘分区表项，结束标志。由于空间只有区区512字节,而主引导记录要占用446个字节，结束标志要占用2个字节，所以分区表的大小就只有64个字节。他被设计成十分小但是又足够描述硬盘的分区情况，包括分区的开始扇区,结束扇区,引导标志,分区类型等等.每个分区信息用16个字节表示,四个分区就用掉了所有的64个字节的空间。所以我们的硬盘最多只能划分为4个主分区。

图11

我们都知道，在安装操作系统和[软件](http://baike.baidu.com/view/37.htm)之前，首先需要对硬盘进行分区和格式化，然后才能使用硬盘保存各种信息。为什么需要进行格式化呢？这是因为每种操作系统所配置的文件属性/权限并不相同， 为了存放这些文件所需的数据，因此就需要将硬盘分区进行格式化，以成为操作系统能够利用的文件系统格式。

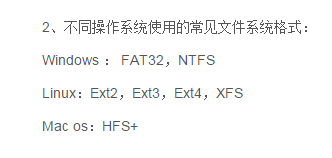
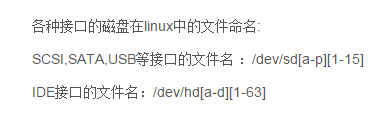
由此我们也能够知道，每种操作系统能够使用的文件系统并不相同。举例来说，windows 98 以前的微软操作系统主要利用的文件系统是 FAT (或 FAT16)，windows 2000 以后的版本有所谓的 NTFS 文件系统，至于 Linux 的正统文件系统则为 Ext2这一个。此外，在默认的情况下，windows 操作系统是不会认识 Linux 的 ext2 的。

图12

那么文件系统是如何运行的呢？这与操作系统的文件数据有关。较新的操作系统的文件数据除了文件实际内容外， 通常含有非常多的属性。

这里以部分文件系统为例向大家简要地介绍文件系统的运行过程。

文件系统是装在硬盘上的。Linux最传统的磁盘文件系统使用的是ext2，所以我们就从ext2介绍文件系统。

Linux 操作系统的文件权限(rwx)与文件属性(拥有者、群组、时间参数等)。 文件系统通常会将这两部份的数据分别存放在不同的区块，权限与属性放置到 inode中（一个文件只有一个inode，同时记录文件数据所存放的data block号码），至于实际数据则放置到 block 区块中（一个文件可有多个block）。另外，还有一个超级区块 (superblock) 会记录整个文件系统的整体信息，包括 inode 与 block 的总量、使用量、剩余量等。

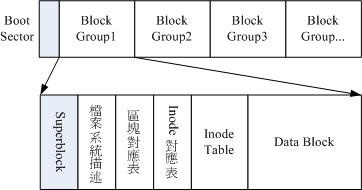
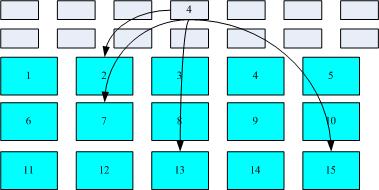
我们ext2文件系统是索引式文件系统。具体的原理如图13、图14所示，灰色的方框代表inode，蓝色的方框代表block。Ext2文件系统的每个文件都由一个inode和一个或多个block组成，当我们读取文件时，首先会访问文件的inode，确认权限属性。如果权限和属性符合要求，那么就由inode里面的记录找到存储数据的block。（这种文件系统产生的磁盘碎片很少，这也就是Linux不需要进行磁盘碎片整理的原因）

图13 图14

与Linux的ext2相比，FAT文件系统就没有inode，只有存储文件数据的block。那么FAT文件系统又是如何解决大文件的存储的呢？当一个文件比较大时，FAT文件系统就会为文件分配多个block，前一个block里面装着下一个block的号码。FAT文件系统通过这样的一种链表的形式存储大文件。FAT文件系统在读文件时只会从文件开始的那个block开始读，读完这个之后才会读取下一个block的数据。FAT文件系统具体如图15所示。（现在这种FAT文件系统主要用在U盘上面）

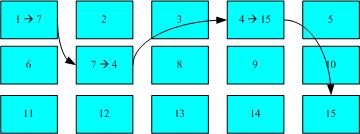


图15

前面提到了Windows的文件系统，我们来看看Windows系统的文件系统的发展吧。

我们最开始用的DOS、Windows 95都是使用的FAT16文件系统，以前流行的Windows 98/2000/XP等系统也均支持FAT16文件系统。它最大可以管理大到2GB的分区，但每个分区最多只能有65525个簇(簇是磁盘空间的配置单位)。随着硬盘或分区容量的增大，每个簇所占的空间将越来越大，从而导致硬盘空间的浪费。

随着大容量硬盘的出现，从Windows 98开始，FAT32开始流行。它是FAT16的增强版本，可以支持大到2TB(2048G的分区。FAT32使用的簇比FAT16小，从而有效地节约了硬盘空间。

在FAT32之后，微软又推出了具有高层次的安全保证的NTFS文件系统和更高级的winFS文件系统。我们现在常用的win7系统支持NTFS和winFS文件系统，而win8加载的则是全新的ReFS系统。

文件系统的产生就是为了能更好地利用硬盘空间，但是硬盘在不断改进，我们的文件系统也在不断进步，一切都在改变。未来的文件系统和硬盘又是个什么样子，谁知道呢？