大容量存储器和文件系统

**江学强，PB16120100**

1. **大容量存储器**

我们知道现代存储器具有明显的层次结构，层次结构主要可以分为CPU缓存、主存以及由大容量存储器组成的次级、三级存储器（也可称为外存）。磁盘是组成现代计算机大容量存储器最常见的物理设备。磁盘片两面涂有磁质材料，通过在磁质材料上进行磁记录可以保存和修改信息。磁盘的物理结构不多赘述，这里只叙述与后面磁盘调度有关的结构。读写头可以在磁盘面上移动用于读写信息，磁盘片的表面被划分为了一个个环形的磁道（track），磁道再进一步划分为扇区(sector)。磁盘上有磁盘控制器。主机控制器将读写命令传送给磁盘控制器，磁盘控制器操纵磁盘驱动器来执行命令。为了加快速率，磁盘控制器中通常置有内置缓存。

为了更有效地使用磁盘硬件，获得更快地访问速度和更宽地磁盘带宽，需要进行磁盘地调度。对磁盘的访问时间主要包括寻道时间和旋转延迟，寻道时间指磁臂将磁头移动到目标扇区的柱面所需的时间，旋转延迟时间指磁盘把目标扇区旋转到磁头下的时间。进程通过向操作系统发送系统调用来请求磁盘的I/O操作，这个系统调用包含一些有用的信息，如果所需的磁盘驱动器和控制器空闲，就会立刻处理请求，如果磁盘驱动器和控制器忙，就会把新的服务请求加入到等待队列中，因为等待队列可以有多个待处理的请求，所以就涉及对请求处理的调度，下面主要叙述几个典型的调度算法。

FCFS调度：这是最简单的调度算法，先请求的先服务；

SSTF调度（shortest-seek-time first）：最短寻道时间优先算法，意思是先处理靠近当前磁头位置的请求；

SCAN调度：SCAN算法是指磁臂从磁盘的一端向另一端移动，磁头移过每个柱面时处理该柱面上的服务请求，当到达另一端时，磁头改变方向，反向扫描。来回如此；

C-SCAN调度：是SCAN调度的变种，C-SCAN将磁头从一端移动到另一端时不反向移动，而是返回磁盘开始再次扫描；

LOOK调度：SCAN和C-SCAN调度都是扫描整个磁盘，但事实上磁头只移动到最远的请求为止而不是到磁盘的尽头，这样的调度算法相应的叫LOOK调度和C-LOOK调度。

操作系统还要负责管理磁盘其他方面的东西，下面叙述磁盘的初始化、磁盘引导。

新的磁盘可以看成是一个白板，只是含有磁性材料的盘子。在使用之前需要把磁盘划分为扇区，这个过程称为低级格式化。每个扇区通常由头、数据区域（通常512B）和尾部组成。头部和尾部包含一些磁盘控制器所需要的信息，如扇区号码和ECC。为了使用磁盘存储文件，操作系统还需要把磁盘分为一个或多个柱面组成的分区，在分区上挂载文件系统。

计算机启动时需要运行bootstrap程序，bootstrap程序初始化系统，一般保存在ROM中，此外在磁盘上还有一个更完整的启动程序，拥有这个启动程序的磁盘分区称为启动盘或者系统盘。例如windows 2000的MBR（主引导记录）,是磁盘上的第一个扇区

磁盘驱动器越来越小也越来越便宜，现在一台计算机系统上装有大量磁盘已经可行了，大量磁盘通过有效的组织可以有效改善读写速度和数据可靠性，这种组织技术称为RAID技术。下面叙述怎样通过RAID技术提高磁盘的读写性能和可靠性。两个主要思想是通过数据的冗余来提高可靠性以及通过数据读写的并行提高性能。

下面简单叙述一下RAID级别：

RAID0：按块分散的磁盘阵列，没有冗余，所以无法保证可靠性也并未提高读写性能；

RAID1：设置一份（不是一个）镜像磁盘，有冗余，保证了数据可靠性；

RAID2：称为内存方式的差错纠正码结构，实现了基于奇偶校验位的检错系统，每个字节的奇偶位记录字节中置为1的bit数是偶数还是奇数，如果字节的一个位损坏，字节的奇偶位将改变。差错纠正方案会储存两个或者多个额外的位，当单个位出错时可用来恢复数据。RAID2相对于RAID1减少了冗余磁盘数，但可靠性buuruRAID1；

RAID3: 基于位交织奇偶结构，相对于RAID2做了改进，与内存系统不同，这里磁盘控制器能够检测到一个扇区是否正确读取，这样单个奇偶位就能用于差错检测和差错纠正，RAID3和RAID2在可靠性上同样好，但是RAID3节省了磁盘，此外由于同一字节的不同位存储在不同的磁盘，因此可以并行读取，相对RAID1也提高了读写性能；

参考文献:

[1] Abraham Silberschatz. Operating System Concepts Essentials Second Edition.Wilay 2014