**设备管理**

1. I/O系统
2. I/O设备
3. 设备与控制器

设备并不是直接与CPU进行通信，而是与设备控制器通信，在设备中的接口中通常有三种类型的信号线

2.设备控制器(硬件驱动)

1. 设备控制器的基本功能: 1)接收和识别命令 2)数据交换 3)标识和报告设备的状态 4)地址识别 5)数据缓冲 6)差错控制

3.I/O通道

(1)I/O通道(I/O Channel)设备

实际上，I/O通道是一种特殊的处理机(CPU)。它具有执行I/O指令的能力，并通过执行通道(I/O)程序来控制I/O操作。但I/O通道又与一般的处理机不同，主要表现在以下两个方面： 一是其指令类型单一, 再就是通道没有自己的内存，通道所执行的通道程序是放在主机的内存中的， 换言之，是通道与CPU共享内存。

(2)通道类型

1) 字节多路通道(Byte Multiplexor Channel)

它通常含有许多非分配子通道，这些子通道按时间片轮换方式共享主通道，按字节为单位进行工作，当第一个子通道完成一个字节的交换后，便转向第二个子通道



1. 数组选择通道(Block Selector Channel)

字节多路通道不适于连接高速设备，这推动了按数组方式进行数据传送的数组选择通道的形成。这种通道虽然可以连接多台高速设备，但由于它只含有一个分配型子通道，在一段时间内只能执行一道通道程序， 控制一台设备进行数据传送， 致使当某台设备占用了该通道后，便一直由它独占，可见，这种通道的利用率很低。

3)数组多路通道(Block Multiplexor Channel)

数组选择通道虽有很高的传输速率，但它却每次只允许一个设备传输数据。数组多路通道是将数组选择通道传输速率高和字节多路通道能使各子通道(设备)分时并行操作的优点相结合而形成的一种新通道。它含有多个非分配型子通道， 因而这种通道既具有很高的数据传输速率，又能获得令人满意的通道利用率。也正因此，才使该通道能被广泛地用于连接多台高、中速的外围设备，其数据传送是按数组方式进行的。

1. I/O控制方式
2. 程序I/O方式

在程序I/O方式中，CPU不断发出请求测试I/O设备

1. 中断驱动I/O控制方式

在I/O设备输入每个数据的过程中，由于无须CPU干预，仅当输完一个字节数据时，I/O设备才向CPU(数据无需进入内存)发出中断处理。从而提高了整个系统的资源利用率及吞吐量。

1. 直接存储器访问DMA I/O控制方式

DMA(Direct Memory Access)控制方式

该方式的特点是：① 数据传输的基本单位是数据块，即在CPU与I/O设备之间，每次传送至少一个数据块；② 所传送的数据是从设备直接送入内存的，或者相反； ③ 仅在传送一个或多个数据块的开始和结束时，才需CPU干预，整块数据的传送是在控制器的控制下完成的。

1. I/O通道控制方式

I/O通道方式是DMA方式的发展，它可进一步减少CPU的干预，即把对一个数据块的读(或写)为单位的干预，减少为对一组数据块的读(或写)及有关的控制和管理为单位的干预。

1. 缓冲管理

1.缓冲的引入

(1)缓和CPU与I/O设备间速度不匹配的矛盾。

(2)减少对CPU的中断频率， 放宽对CPU中断响应时间的限制。

(3)提高CPU和I/O设备之间的并行性。

2.单缓冲(Single Buffer)



3.双缓冲(Double Buffer)

4.循环缓冲

R:用于装入数据的空缓冲区

G:已装满的缓冲区

C:计算进程正在使用的缓冲区

Nextg:指示计算进程下一个处理的缓冲区G

Current:计算进程正在处理的缓冲区

Nexti:指示输入进程下一个可用的空缓冲区R

5.缓冲池(Buffer Pool)

(1)缓冲池的组成

对于既可用于输入又可用于输出的公用缓冲池， 其中至少应含有以下三种类型的缓冲区：① 空(闲)缓冲区； ② 装满输入数据的缓冲区； ③ 装满输出数据的缓冲区。 为了管理上的方便，可将相同类型的缓冲区链成一个队列，于是可形成以下三个队列：

1. 空缓冲队列emq。 (2) 输入队列inq。 (3) 输出队列outq。

(2)缓冲池的工作方式



1. 设备分配
2. 设备控制表DCT

DCT:每个设备都配置一张设备控制表，用于记录本设备的情况



1. 控制器控制表、 通道控制表和系统设备表

COCT:系统为每一个设备控制器都设置一张用于记录本控制器的情况

CHCT:每个通道都配有一张通道控制表

SDT:系统级表，用于记录全部设备的情况，每个设备占一条表目，系统只有一张SDT表

3.设备分配的考虑因素

1. 设备的固有属性:a.独享设备(临界资源)。 b.共享设备。 c.虚拟设备。
2. 设备分配算法:a.先来先服务。 b.优先级高者优先
3. 设备分配中的安全性:a.安全分配方式 b.不安全分配方式
4. 设备独立性: 应用程序独立于具体使用的物理设备。为了实现设备独立性而引入了逻辑设备(一类相同的物理设备总称)和物理设备这两个概念。在应用程序中， 使用逻辑设备名称来请求使用某类设备；而系统在实际执行时， 还必须使用物理设备名称。

逻辑设备名到物理设备名映射的实现

逻辑设备表LUT（a为整个系统设置一张LUT表，则不能同时使用一个相同的逻辑设备名，常用于单用户系统，b为每个用户设置一张LUT表，常用于多用户系统）



4.SPOOLing技术

为了缓和CPU的高速性与I/O设备低速性间的矛盾而引入了脱机输入、 脱机输出技术。 该技术是利用专门的外围控制机， 将低速I/O设备上的数据传送到高速磁盘上；把这种在联机情况下实现的同时外围操作称为SPOOLing(Simultaneaus Periphernal Operating On-Line)，或称为假脱机操作。



输入井和输出井：这是在磁盘上开辟的两大空间，用于暂时设备输入数据，与将要送入输出设备的数据

输入进程和输出进程：用来模拟外围机

1. 磁盘存储器管理
2. 磁盘访问时间

1) 寻道时间Ts（寻找磁道）

这是指把磁臂(磁头)移动到指定磁道上所经历的时间。该时间是启动磁臂的时间s与磁头移动n条磁道所花费的时间之和， 即

Ts=m×n+s

其中，m是一常数，与磁盘驱动器的速度有关，对一般磁盘， m=0.2

2)旋转延迟时间Tτ（寻找扇区）

这是指定扇区移动到磁头下面所经历的时间。对于硬盘，典型的旋转速度大多为5400 r/min，每转需时11.1 ms，平均旋转延迟时间Tτ为5.55 ms；

3)传输时间Tt

这是指把数据从磁盘读出或向磁盘写入数据所经历的时间。 Tt的大小与每次所读/写的字节数b和旋转速度有关：

Tt=b/rN

其中，r为磁盘每秒钟的转数；N为一条磁道上的字节数

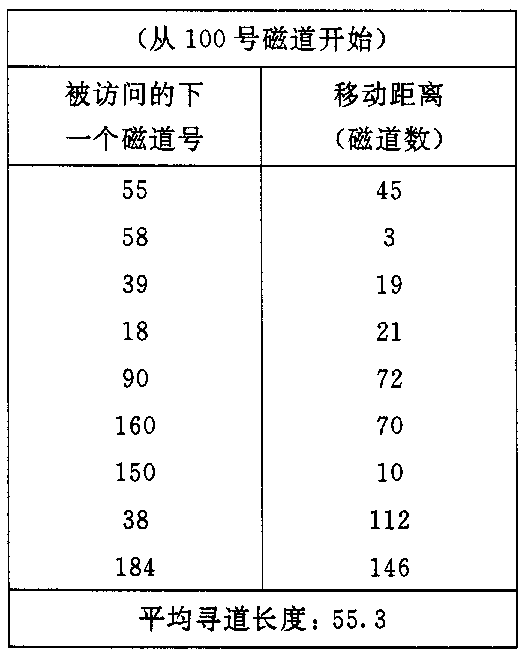
4)总时间Ta

一次读/写的字节数相当于半条磁道上的字节数时，Tt与Tτ相同， 因此， 可将访问时间Ta表示为：

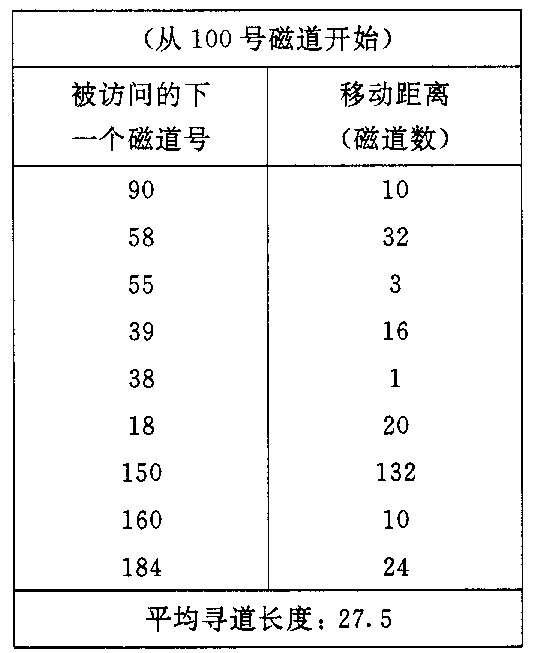
Ta=Ts+1/2r+b/rN

1. 磁盘调度

(1). 先来先服务FCFS(First-Come, First Served)

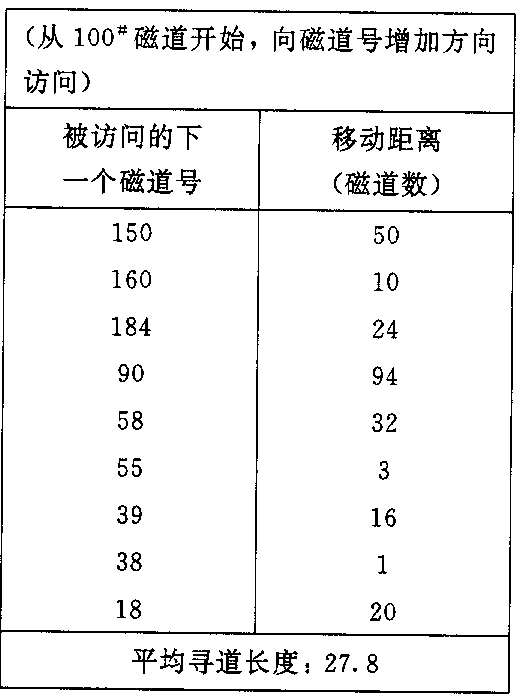


1. .最短寻道时间优先SSTF(Shortest Seek Time First):其要求访问的磁道与当前磁头所在的磁道距离最近，以使寻道时间最短



1. 扫描(SCAN)算法

SSTF算法虽然能获得较好的寻道性能， 但却可能导致某个进程发生“饥饿”(Starvation)现象。因为只要不断有新进程的请求到达， 且其所要访问的磁道与磁头当前所在磁道的距离较近，这种新进程的I/O请求必须优先满足。对SSTF算法略加修改后所形成的SCAN算法， 即可防止老进程出现“饥饿”现象。



该算法不仅考虑欲访问磁道与当前磁道的距离，还考虑磁头当前的移动方向（来回移动）

1. 循环扫描(CSCAN)算法