

- 大容量存储结构概述
- 磁盘特性
- 磁盘调度（本章要掌握的内容）
- 磁盘存储优化分布（这个要了解）
- 磁盘管理
- 磁盘管理

# 大容量存储结构概述

## 大容量存储结构概述

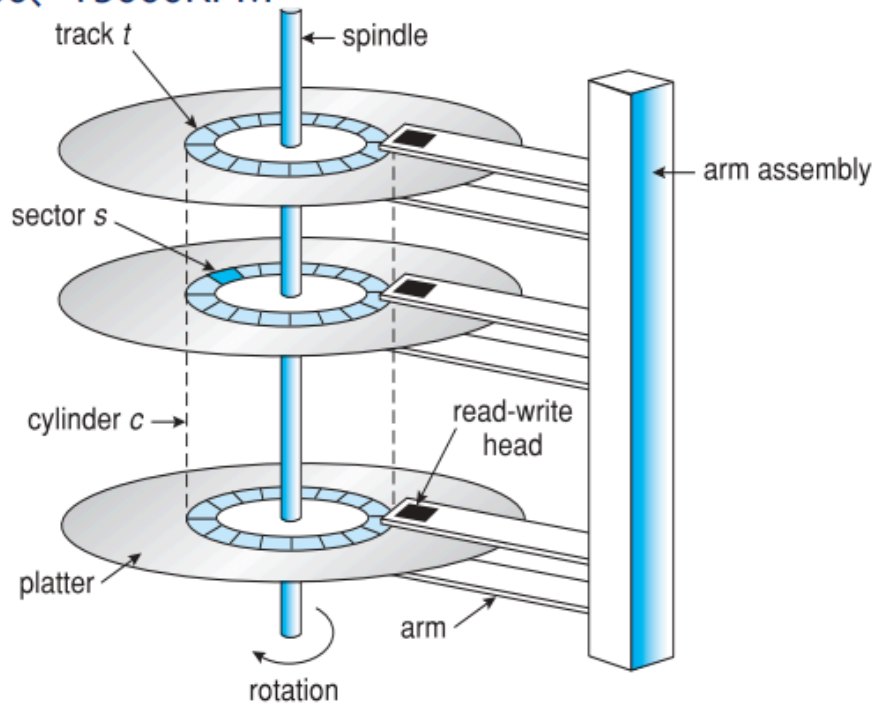
- 文件的存储设备主要有磁带、磁盘、光盘等。
- 存储设备的特性可以决定文件的存取方法。
- 以磁带为代表的顺序存取设备
- 以磁盘为代表的直接存取设备

- 磁盘

# 磁盘机理

## ■ 磁盘

- 为现代计算机提供了大容量的二级存储
- 普通驱动器转速60—250次/秒，按每分钟转速为5400、7200、10000、15000RPM





# 磁盘工作机理

---

- 磁盘访问时间
  - 传输速率(Transfer rate)：MBps级
    - 指从磁盘上读出数据或向磁盘写入数据所经历的时间
  - 定位时间(Positioning time, random-access time):ms级
    - 寻道时间Seek Time：移动磁臂到所需柱面/磁道的时间，由启动磁臂时间和磁头移动多条磁道的时间构成
    - 旋转延迟RPM：所需扇区旋转到磁头下的时间，平均旋转延迟时间是每转所需时间的一半
- 磁头碰撞：磁头和盘面接触的破坏结果
- 驱动器通过I/O总线与计算机相连
  - BUS：EIDE, ATA, SATA, USB, Fiber Channel, SCSI, SAS, Firewire
  - 主机控制器：总线与磁盘控制器对话的部件
  - 磁盘控制器：磁盘驱动器内部的部件

- 磁盘性能

# 磁盘性能

- 访问延迟=平均访问时间=平均寻道时间+平均旋转延迟
  - 如上例子, 最快:  $3\text{ms} + 2\text{ms} = 5\text{ms}$
  - 较慢  $9\text{ms} + 5.56\text{ms} = 14.56\text{ms}$
- 平均I/O时间 = 平均访问时间 + (总传输量 / 传输速率) + 控制器负载时间
- 例: 在7200RPM磁盘上, 传输4KB数据块, 平均寻道时间5ms, 传输速率1Gbps, 控制器负载0.1ms, 则
  - $5\text{ms} + 4.17\text{ms} + 0.1\text{ms} + \text{传输时间}$
  - 传输时间 =  $4\text{KB} / 1\text{Gbps} * 8 = 0.0305\text{ms}$
  - 平均I/O时间为9.3005ms

## 磁盘特性

- 略

# 磁盘调度（本章要掌握的内容）

## 磁盘调度需求

- 操作系统负责高效地使用硬件
  - 对于磁盘，需要更快的访问时间和磁盘带宽
- 目标：最小化寻道时间
- 寻道时间  $\approx$  寻道距离
- 磁盘带宽 = 传输量的总和 / (最后完成传输时间 - 第一个请求服务时间)

### 1. 先来先服务--FCFS

## 1. 先来先服务--FCFS

- 先来先服务算法按进程请求访问磁盘的先后次序进行调度。
- 特点：简单合理，但未对寻道进行优化。

## 先来先服务算法例

下一磁道号	移动距离
55	45
58	3
39	19
18	21
90	72
160	70
150	10
38	112
184	146

从100号磁道开始，磁盘访问请求为：55、58、39、18、90、160、150、38、184

平均寻道长度为：55.3

### 2. 最短寻道时间优先--SSTF

## 2. 最短寻道时间优先--SSTF

- 最短寻道时间优先算法选择与当前磁头所在磁道距离最近的请求作为下一次服务的对象。
- 特点：寻道性能比FCFS好，但不能保证平均寻道时间最短，还可能会使某些请求总也得不到服务。



## 最短寻道时间优先算法例

下一磁道号	移动距离
90	10
58	32
55	3
39	16
38	1
18	20
150	132
160	10
184	24

从100号磁道开始，磁盘访问请求为：55、58、39、18、90、160、150、38、184

平均寻道长度为：27.6

### 3. 扫描算法--SCAN

- 进程饥饿：进程的请求总也得不到服务。
- SCAN算法：在磁头当前移动方向上选择与当前磁头所在磁道距离最近的请求作为下一次服务的对象。**磁头方向可以双向移动。当走到尽头才掉头。**因这种算法中磁头移动规律颇似电梯的运行，故又称为**电梯调度算法(\*)**。
- 特点：具有较好的寻道性能，能避免进程饥饿，但不利于两端磁道的请求。

下一磁道号	移动距离
150	50
160	10
184	24
90	94
58	32
55	3
39	16
38	1
18	20

从100号磁道开始，向磁道号增加方向移动。磁盘访问请求为：55、58、39、18、90、160、150、38、184（假设磁头最远为184）

平均寻道长度为：27.8

#### 4. 循环扫描算法CSCAN

- CSCAN算法是SCAN算法的改良，它规定磁头单向移动。例如，自里向外移动，当磁头移到最外磁道时立即又返回到最里磁道，如此循环进行扫描。
- 特点：该算法消除了对两端磁道请求的不公平。





## 循环扫描算法例

下一磁道号	移动距离
150	50
160	10
184	24
18	166
38	20
39	1
55	16
58	3
90	32

从100号磁道开始，向磁道号增加方向移动。磁盘访问请求为：55、58、39、18、90、160、150、38、184

平均寻道长度为：35.8

5. 一个特别的说明，走到请求的尽头就掉头



## 5. Look & C-Look算法

- SCAN & CSCAN的改良
- 磁头不再走到磁道尽头，而是走到请求的尽头即掉头。
- 注意：在一些国内教材，包括考研时，SCAN & CSCAN是不走到磁道尽头的，且不介绍Look & C-look算法，这时候，就默认SCAN&CSCAN，为只走到请求的尽头即掉头

6. N-Step-SCAN



## 6. N-Step-SCAN

- 若多个进程反复请求对某一磁道的访问，则磁臂可能停留在某处不动，这一现象称为**磁臂粘着**。
- N-Step-SCAN算法：
  - 将磁盘请求队列分成若干个长度为N的子队列
  - 磁盘调度按FCFS算法依次处理这些子队列
  - 处理每个队列时按SCAN算法进行，一个队列处理完后，再处理其他队列。

### 7. FSCAN算法



## 7. FSCAN算法

- FSCAN算法是N-Step-SCAN算法的简化
- 它只将磁盘请求队列分成两个子队列。
  - 一个是当前所有请求磁盘I/O的进程形成的队列，由磁盘调度按SCAN算法进行处理，
  - 另一个队列则是在扫描期间新出现的磁盘请求。

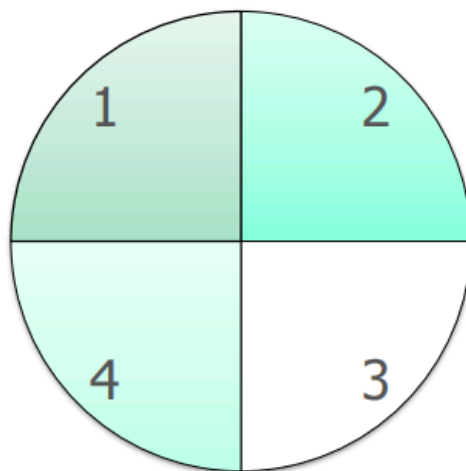
## 磁盘存储优化分布（这个要了解）

- 例子

## 磁盘存储优化分布

- 考虑磁道保存4个记录的旋转型设备，假定收到四个I/O请求。假设每周为20ms。假定平均用1/2周定位，1/4周读出记录。

记录信息



■ 位置1 □ 位置2 ■ 位置3 ■ 位置4

# 多种I/O请求排序方法

## ■ 情形1

- I/O请求次序为读记录4、3、2、1，初始位置未知，考虑一般情况，则平均用 $1/2$ 周定位，再加上 $1/4$ 周读出记录。
- $1/2 + 1/4 + 3 \times 3/4 = 3$ 周，即60毫秒。

## ■ 情形2

- 如果果次序为读记录1、2、3、4。
- 总处理时间等于 $1/2 + 1/4 + 1/4 + 1/4 + 1/4 = 1.5$ 周，即30毫秒。

## ■ 情形3

- 如果知道当前读位置是记录4，可采用次序为读记录4、1、2、3。总处理时间等于1周，即20毫秒

- 优化分布问题，主要是因为处理的时候磁头还在转，因此可以优化

## 优化分布问题

- 考虑10个逻辑记录A, B....., J, 被存于旋转型设备上, 每道存放10个记录, 安排如下:

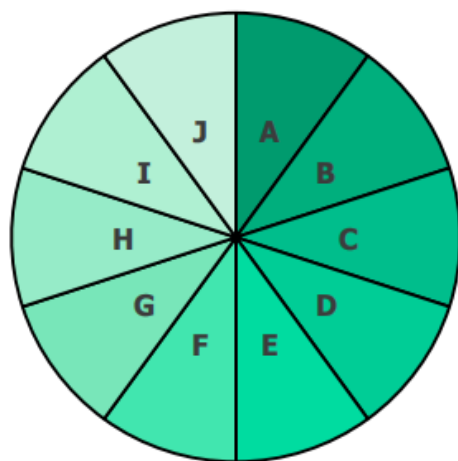
物理块	逻辑纪录
1-10	A-J (A B C D E .....)

- 设磁盘匀速转动, 旋转一周的时间为20毫秒, 读出时间为2毫秒, 处理时间为4毫秒
- 按A B C D E .....J的顺序, 处理10个记录的时间为多少?

## 信息简单顺序分布

- 处理10个记录的总时间:  
10ms(移动到记录A的平均时间)  
+ 2ms(读记录A)  
+ 4ms(处理记录A)  
+  $9 \times [16\text{ms}(\text{访问下一记录}) + 2\text{ms}(\text{读记录}) + 4\text{ms}(\text{处理记录})] = 214\text{ms}$

扇区信息分布





## 如果按照下面方式对信息优化分布

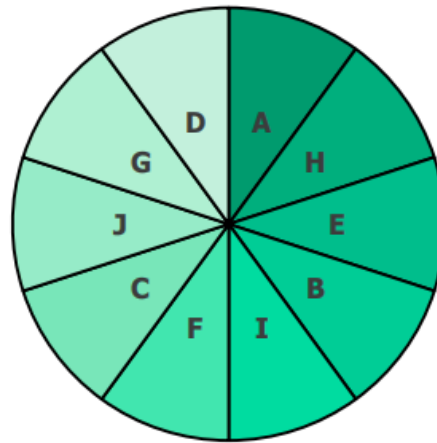
物理块

1  
2  
3  
4  
5  
6  
7  
8  
9  
10

逻辑纪录

A  
H  
E  
B  
I  
F  
C  
J  
G  
D

扇区信息分布



10ms(移动到记录A的平均时间)  
+10× 2ms (读记录)  
+10× 4ms (处理记录) ]  
=70毫秒

## 磁盘管理

- 略

# 磁盘管理

- 容错技术：通过在系统中设置冗余部件来提高系统可靠性的一种技术。
- 磁盘容错技术：通过增加冗余磁盘驱动器、磁盘控制器等方法来提高磁盘系统可靠性的一种技术。也称为系统容错技术。
- 系统容错分为三级：
  - 第一级容错技术：低级磁盘容错技术
  - 第二级容错技术：中级磁盘容错技术
  - 第三级容错技术：高级系统容错技术