

计算机组织结构

9 外部存储器

刘博涵

2023年11月9日



南京大學
NANJING UNIVERSITY

教材对应章节

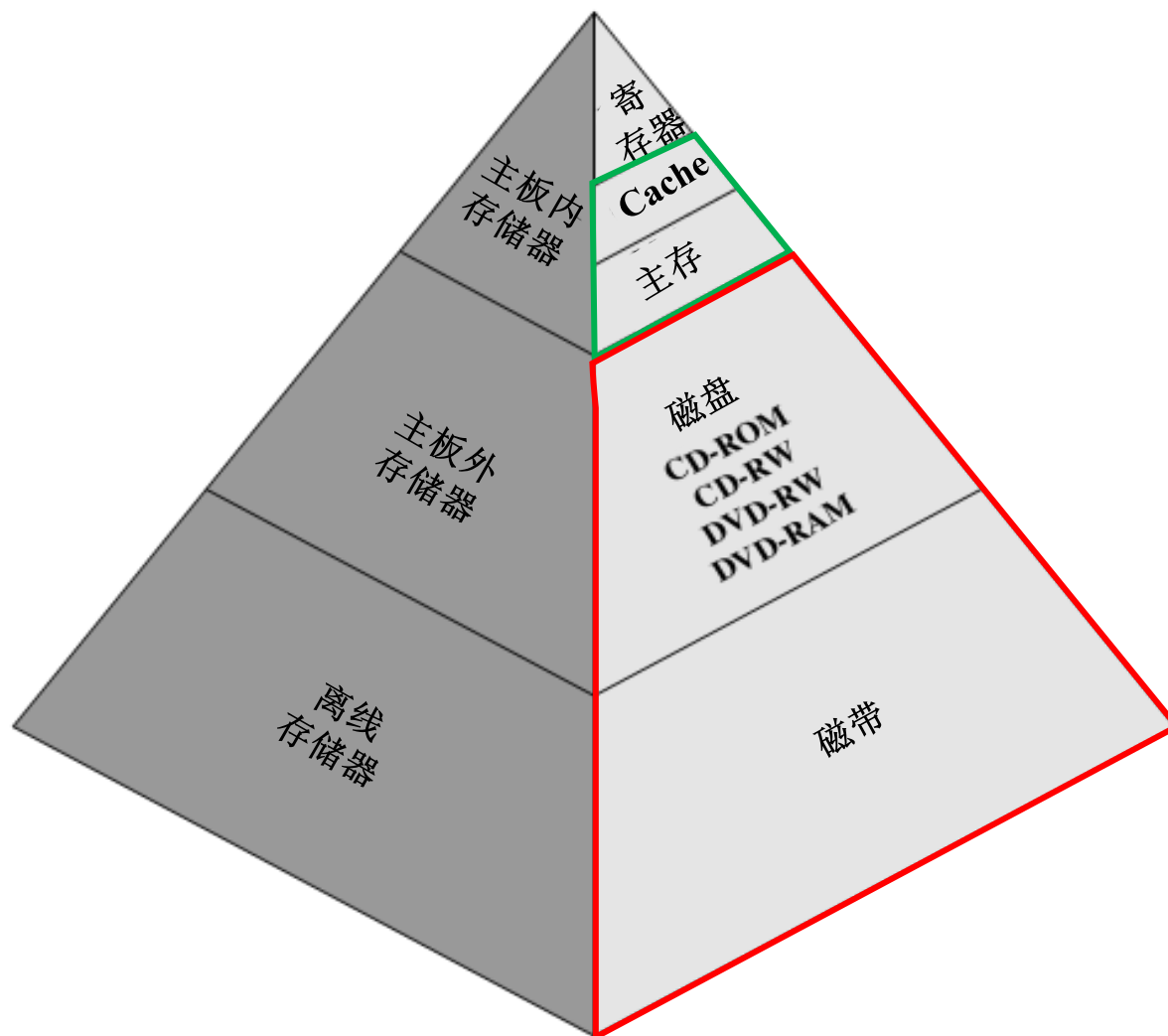


第8章 互连及输入输出组织



第6章 外部存储器

存储器层次结构



外部存储设备发展：以音乐存储为例



盒式磁带（1963年诞生）
1997年第一次购买



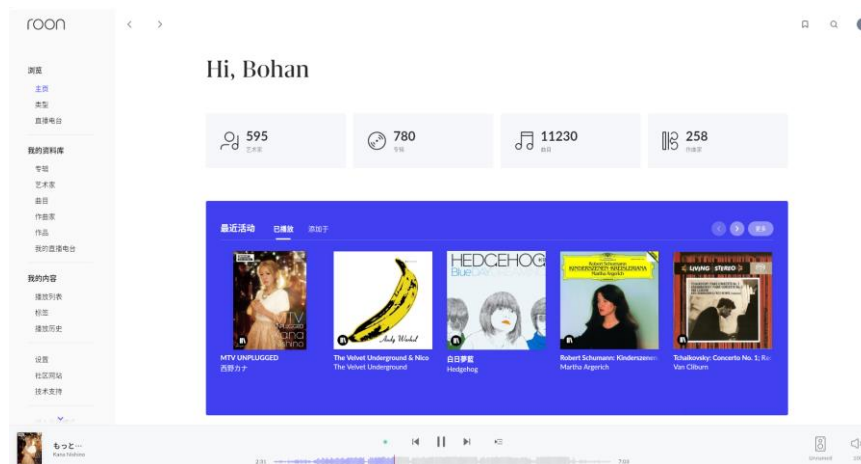
光盘（1965年诞生）
2003第一次购买



黑胶唱片（1948年诞生）
2016第一次购买



机械硬盘（1956年诞生）



2020年至今，我使用NAS存储音乐



外部存储设备发展：一种古老的播放方式



外部存储设备

- **特性**

- 用于存储不经常使用的、数据量较大的信息
- 非易失

- **类型**

- 磁盘存储器 (magnetic disk)
- 光存储器 (optical memory)
- 磁带 (magnetic tape)
- U盘 (USB flash disk) , 固态硬盘 (solid state disk, SSD)

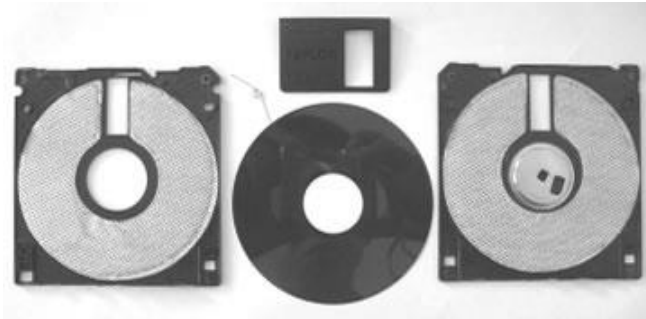


磁盘存储器

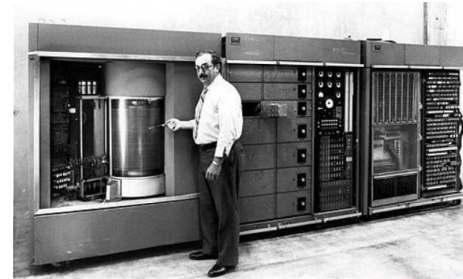
- 磁盘是由涂有**可磁化材料**的**非磁性材料**（基材）构成的圆形盘片
 - 基材：铝、铝合金、玻璃.....
 - **玻璃基材的优势（稳定可靠、为存储更多信息提供基础）**
 - 改善磁膜表面的均匀性，提高磁盘的可靠性
 - 显著减少整体表面瑕疵，以帮助减少读写错误
 - 能够支持（磁头）较低的**飞行高度**
 - 更高的硬度，使磁盘转动时更加稳定
 - 更强的抗冲击和抗损伤能力



磁盘存储器：类型

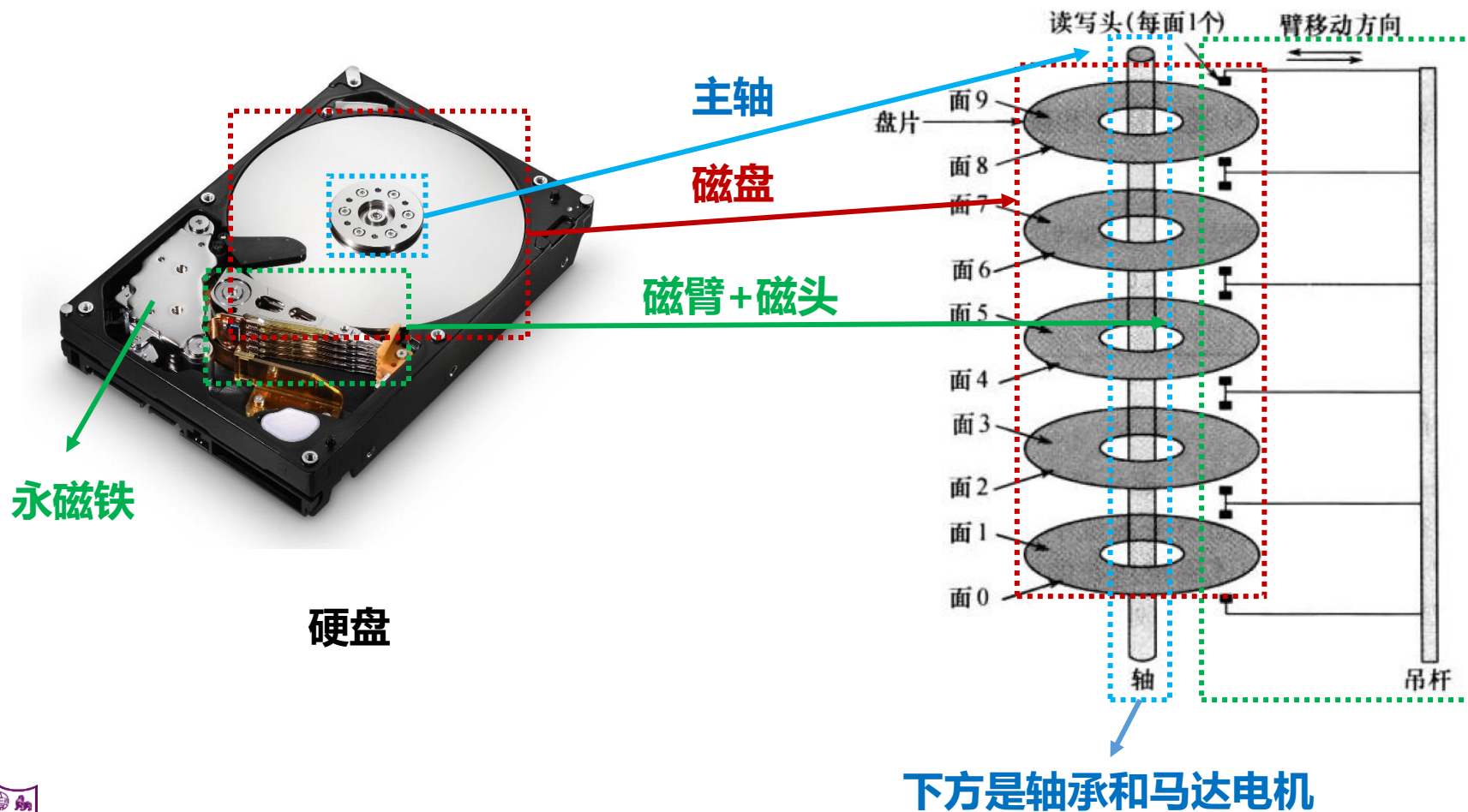


软盘
(目前主流为1.44MB)



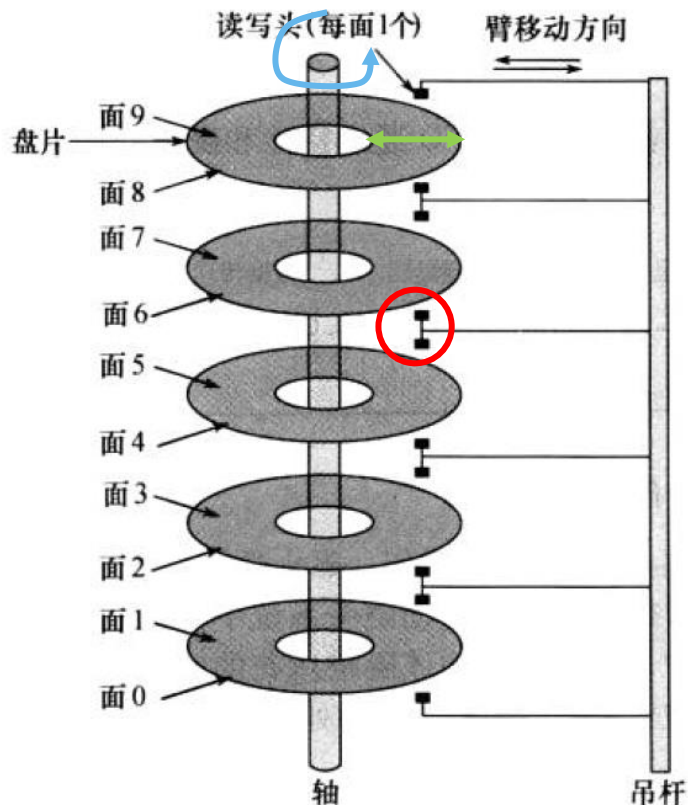
硬盘
(目前主流为TB级)

硬磁盘存储器：结构



硬磁盘存储器：结构（续）

- 磁盘存储器每个盘片表面有一个读写磁头，所有磁头通过机械方式固定在一起，同时移动
- 在任何时候，所有磁头都位于距磁盘中心等距离的磁道上



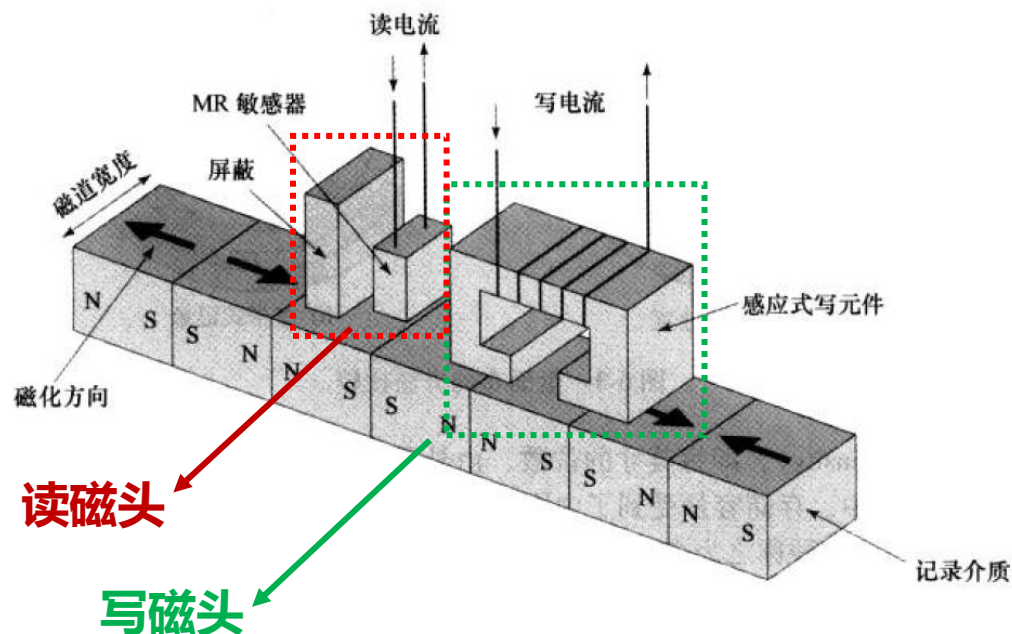
硬磁盘存储器：结构（续）

- 对盘片进行读写操作的装置叫做**磁头 (head)**
 - 磁头必须产生或感应足够大的电磁场，以便正确地读写
 - 磁头越窄，**电磁感应**能力越弱，离盘片的距离就越近
 - 更高的**数据密度**需要更窄的磁头和更窄的磁道，这将导致更高的出错风险
 - 硬盘必须密封
- **温彻斯特磁头 (Winchester head)**
 - 磁头实际上是一个空气动力箔片，当磁盘静止时，它轻轻地停留在盘片的表面上
 - 旋转圆盘时产生的空气压力足以使箔片上升到盘片表面上方



硬磁盘存储器：读写机制

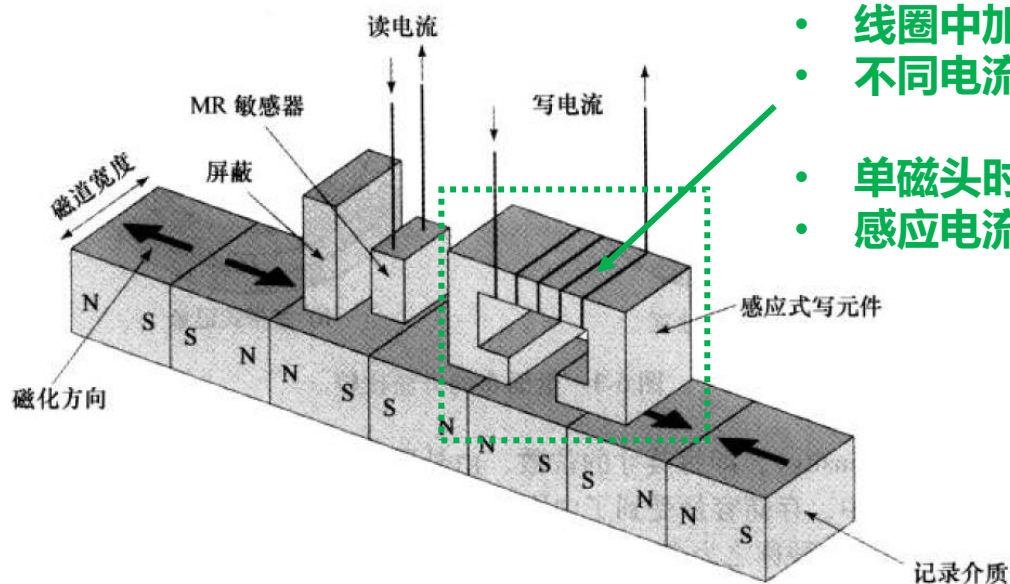
- 在读或写操作期间，**磁头静止**，而**盘片在其下方旋转**
- 磁头的数量
 - **单磁头**：读写公用同一个磁头（软盘、早期硬盘）
 - **双磁头**：使用一个单独的磁头进行读取（当代硬盘）



硬磁盘存储器：读写机制（续）

• 写入机制

- 电流脉冲被发送到写入磁头
- 变化的电流激发出磁场
- 产生的磁性图案被记录在下面的盘片表面上
- 反转电流方向，则记录介质上的磁化方向也会反转

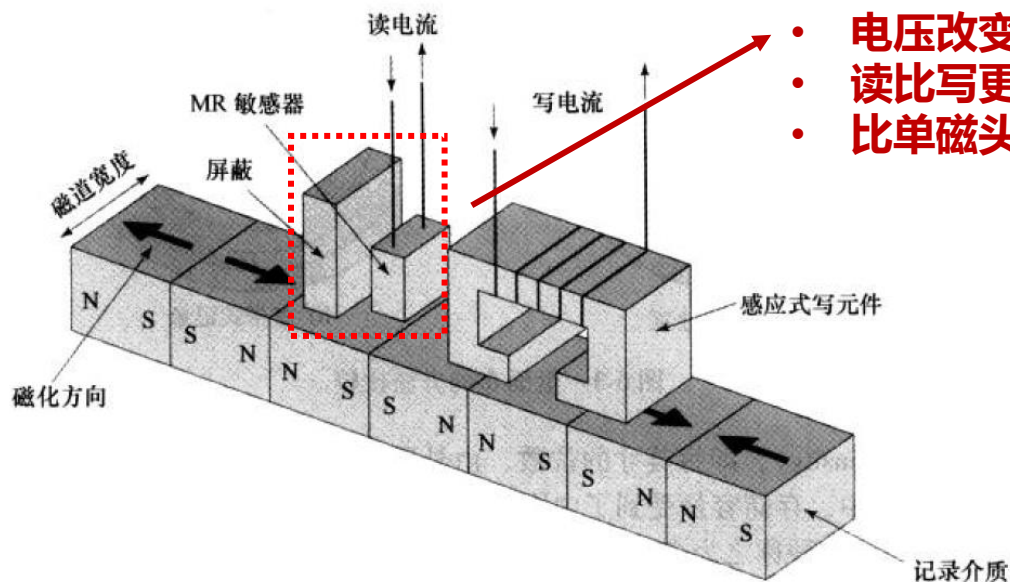


- 线圈中加电流，产生磁场
- 不同电流方向，不同磁场方向
- 单磁头时不加电流可以读
- 感应电流的稳定需要时间

硬磁盘存储器：读写机制（续）

• 读取机制

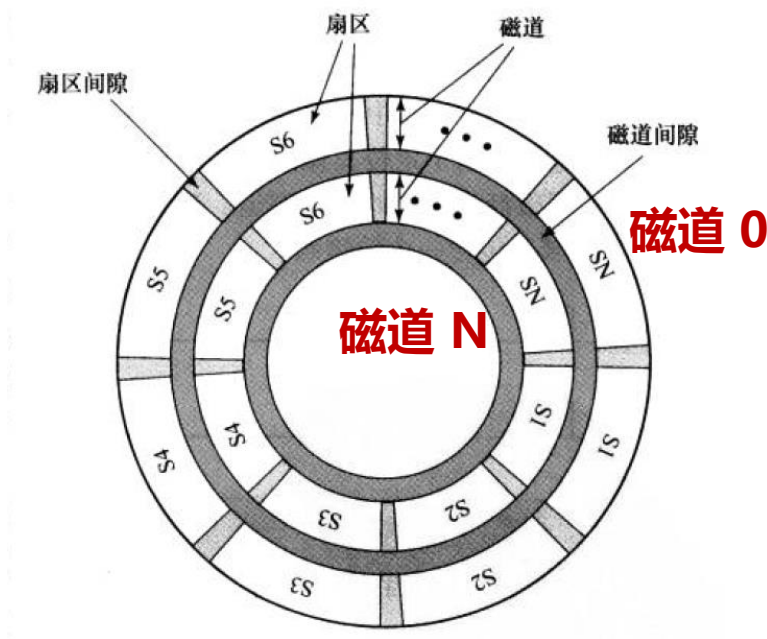
- 读取磁头是由一个部分屏蔽的**磁阻（MR）敏感器**组成，其电阻取决于在其下移动的介质的磁化方向
- **恒定电流**通过MR敏感器时，通过**电压**信号检测其**电阻变化**
- MR敏感器允许更高频率的操作，实现更高的存储密度和更快的操作速度



- **电压改变比电流方向改变更快**
- **读比写更快**
- **比单磁头对磁性介质的要求低**

硬磁盘存储器：数据组织

- 盘片上的数据组织呈现为一组**同心圆环**，称为**磁道 (track)**
- 数据以**扇区 (sector)** 的形式传输到磁盘或从传出磁盘
 - 默认值为**512B**
- 相邻磁道之间有**间隙 (gap)**，相邻的扇区之间也留有间隙，否则磁场容易相互影响
- 磁道编号**从外往里**递增 (0~N)，磁头静止时停在磁道0外侧



硬磁盘存储器：数据组织（续）

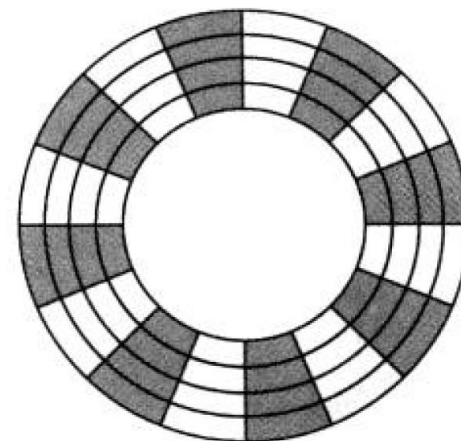
- 扇区划分

- 恒定角速度 (Constant angular velocity, CAV)

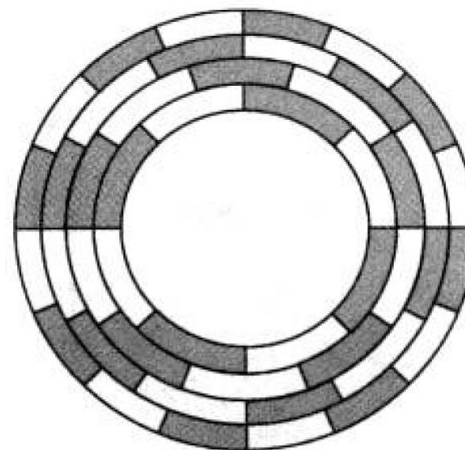
- 增大记录在盘片区域上的信息位的间隔, 使得磁盘能够以恒定的速度扫描信息, 即**恒定的数据传输率**

- **优点：**能以磁道号和扇区号直接寻址各个数据块

- **缺点：**磁盘存储容量受到了最内层磁道所能实现的最大记录密度的限制



a) 恒定角速度



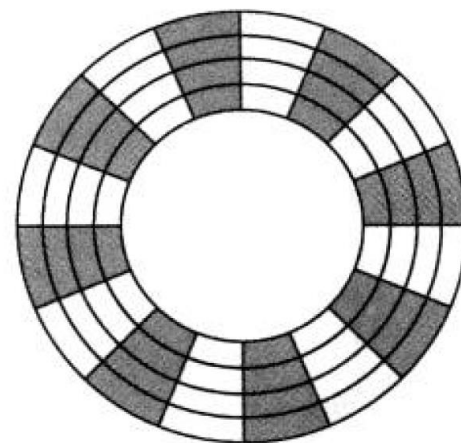
b) 多带式记录

硬磁盘存储器：数据组织（续）

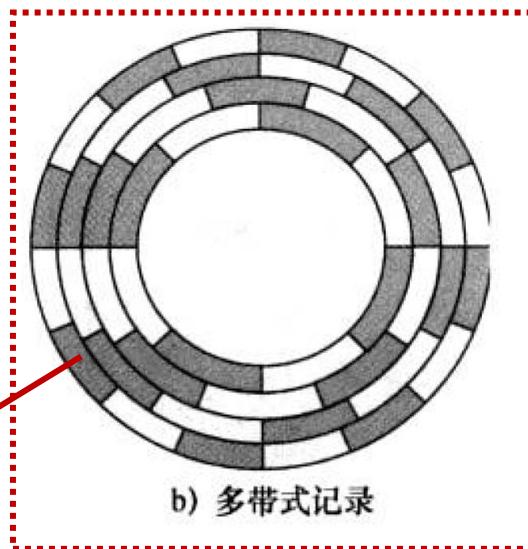
- 扇区划分（续）

- 多带式记录 / 多重区域记录 (Multiple zone recording)

- 将盘面划分为多个同心圆区域，每个区域中各磁道的扇区数量是相同的，距离中心较远的分区包含的扇区数多于距离中心较近的分区
- 优点：提升存储容量
- 缺点：需要更复杂的电路



a) 恒定角速度

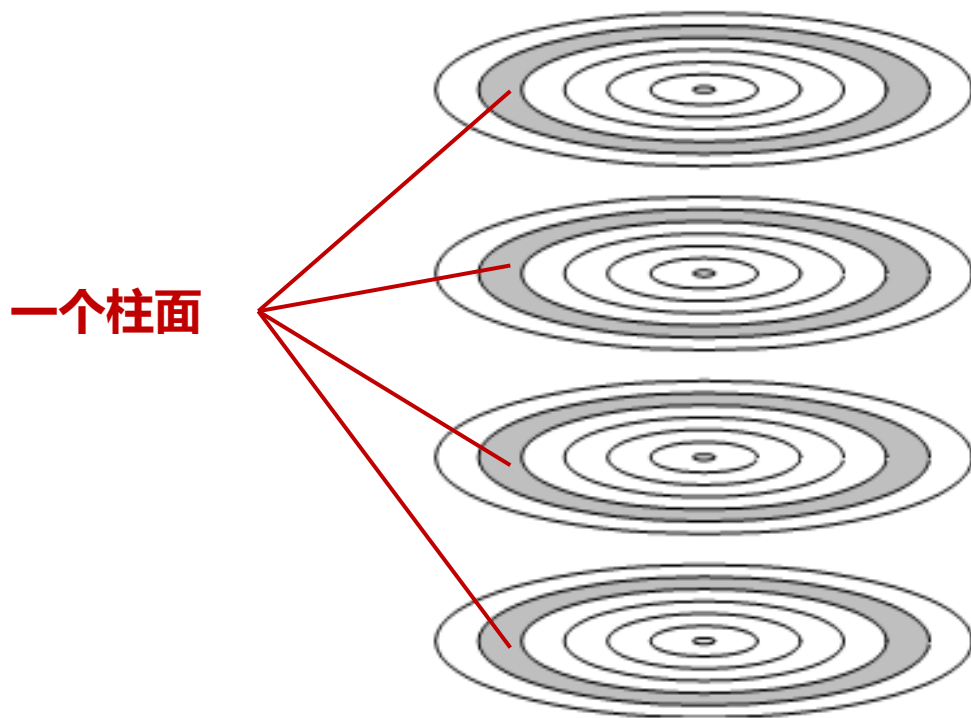


b) 多带式记录

这不是一个磁道，而是包含多个磁道的区域

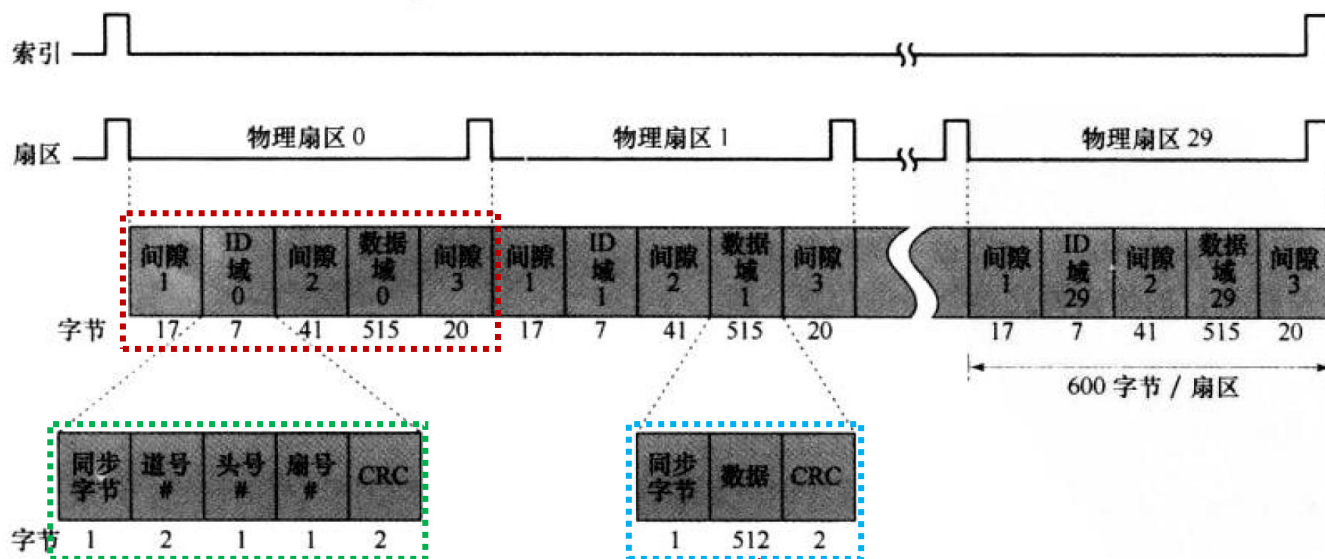
硬磁盘存储器：数据组织（续）

- 所有盘片上处于相同的相对位置的一组磁道被称为**柱面(cylinder)**



硬盘存储器：格式化

- 磁道必须有一些起始点和辨别每个扇区起点及终点的方法
- 格式化时，会附有一些仅被磁盘驱动器使用而不被用户存取的额外数据



ID域开始
特定的01序列

确定一个
唯一的
扇区

循环冗余校验

4096是现在的主流标准



硬盘存储器：格式化类型

- **低级格式化/物理格式化：**创建硬盘扇区（sector）使硬盘具备存储能力的操作
 - 清除数据（如果硬盘分过区，所有区都会被清除）
 - 有损：是一种损伤性操作，它对硬盘寿命有影响
 - 建议场景：硬盘受到外部强磁体、强磁场的影响而受到物理性损伤的情况
- **高级格式化/逻辑格式化：**根据用户选定的文件系统（如FAT12、FAT16、FAT32、exFAT、NTFS、EXT2、EXT3等），在磁盘的特定区域写入特定数据，以达到初始化磁盘或磁盘分区、清除原磁盘或磁盘分区中所有文件的一个操作
 - **快速格式化：**会删除目标磁盘上原有的文件分配表和根目录，不检测坏道，不备份数据，它格式化的速度很快，但不是很稳定。
 - **一般/完全格式化：**会清除目标磁盘上的所有的数据（非物理的）。重新生成引导信息、初始化文件分配表、标注逻辑坏道，一样不备份数据。



硬盘存储器：格式化类型

格式化 娱乐 (F:)

容量(P):

1.63 TB

文件系统(E)

NTFS (默认)

分配单元大小(A)

4096 字节

还原设备的默认值(D)

卷标(L)

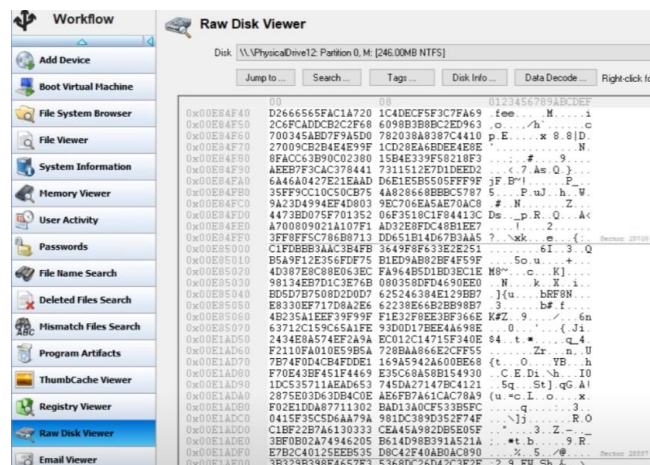
娱乐

格式化选项(O)

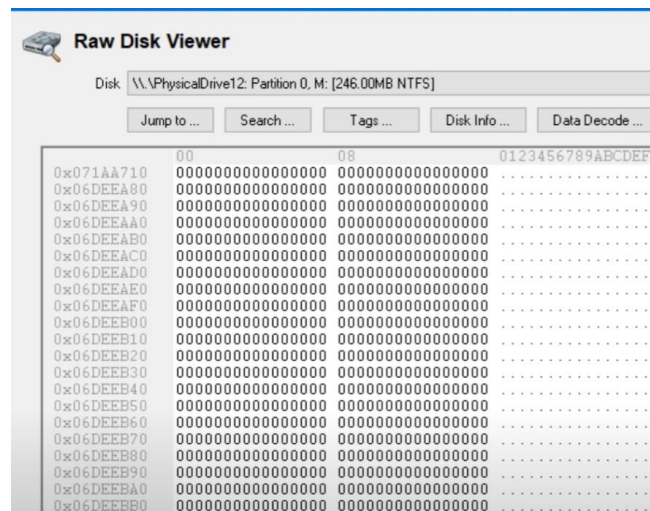
☒ 快速格式化(Q)

开始(S)

关闭(C)



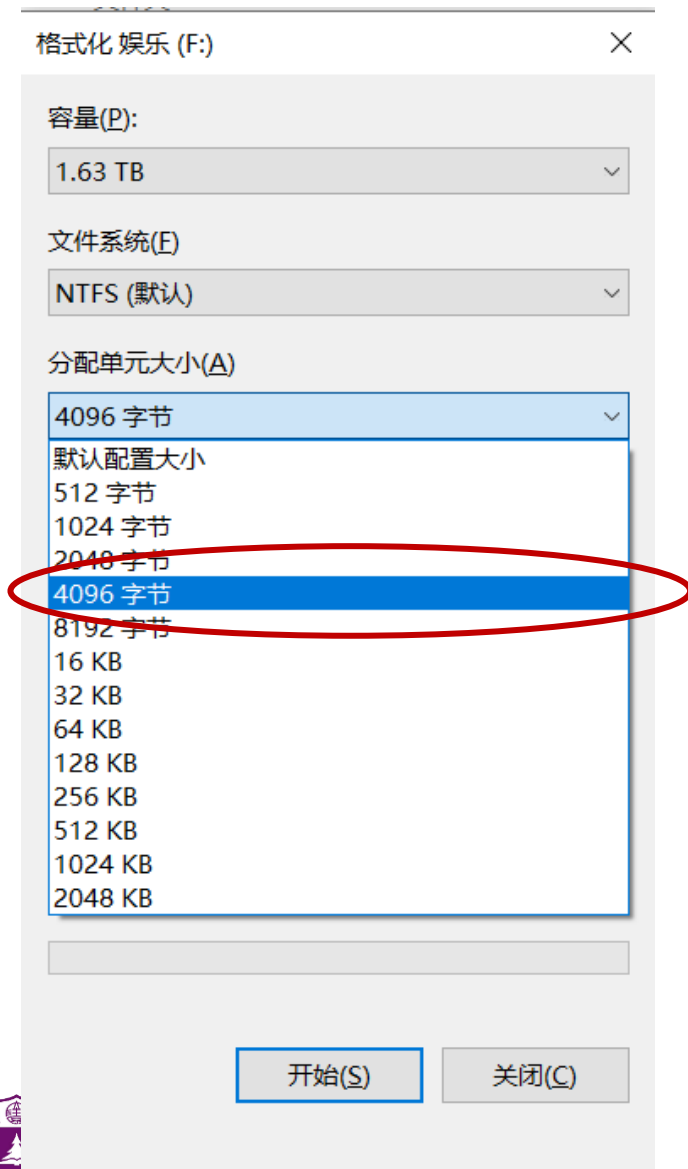
快速



一般

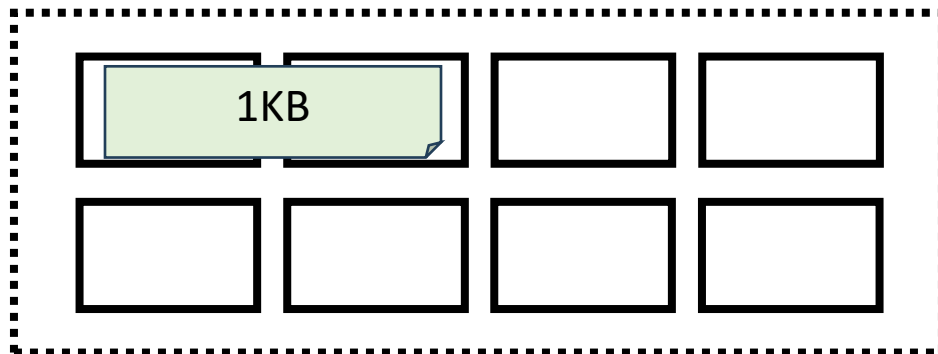


硬盘存储器：格式化类型



硬盘存储数据的**最小**单位是**扇区**，
它的大小是出厂确定的

操作系统存储数据的**最小**单位是**块**，
块的大小可以通过格式化确定



硬盘存储器：I/O访问时间

- **寻道时间 (seek time)**：磁头定位到所需移动到的磁道所花费的时间
 - 初始启动时间，跨越若干磁道所用的时间
- **旋转延迟 (rotational delay)**：等待响应扇区的起始处到达磁头所需的时间
 - 通常是磁道旋转半周所需的时间
- **传送时间 (transfer time)**：数据传输所需的时间

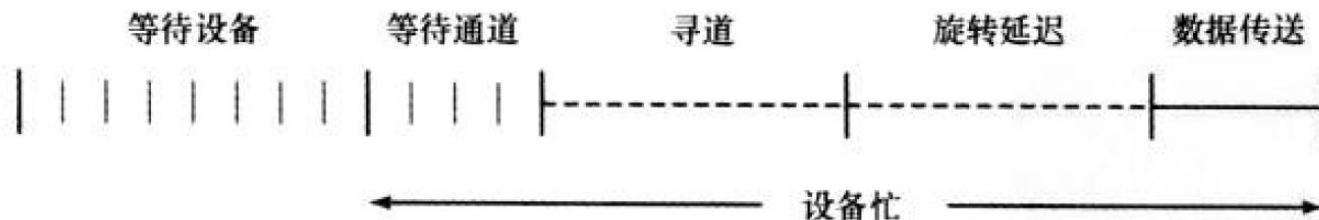
$$T = \frac{b}{rN}$$

T = 传送时间

b = 传送的字节数

N = 每磁道的字节数

r = 旋转速率，单位是转/秒



硬磁盘存储器：I/O访问时间（续）

- 平均访问时间

取平均时间 取实际时间

$$T_a = \boxed{T_s + \frac{1}{2r}} + \boxed{\frac{b}{rN}}$$

T_s 是平均寻道时间

- 当连续访问多个相邻的磁道时，**跨越磁道**：

- 对于**每个**磁道都需要考虑**旋转延迟**
- **通常**只需要考虑**第一个**磁道的**寻道时间**，但在明确知道跨越每个磁道需要的时间时需要考虑

磁道非常多，顺序移到下一个磁道的耗时是非常少的



硬盘存储器：I/O访问时间（示例）

- 假设某个硬盘的平均寻道时间为4ms，转速为15000rpm，每磁道500扇区，**每扇区512B**，现读取一个由2500个扇区组成的文件
- 情况 1： **顺序组织**
 - 该文件占据相邻5个磁道的全部扇区 ($5 \text{ 磁道} \times 500 \text{ 扇区/道} = 2500 \text{ 扇区}$)
- 情况 2： **随机存取**
 - 该文件随机分布在磁盘上的各个扇区



硬磁盘存储器：I/O访问时间（示例）

- 情况 1: 顺序组织

平均寻道	4 ms
平均旋转延迟	2 ms
读500个扇区	4 ms
	<hr/>
	10 ms

访问后续磁道不计寻道时间，因此访问每一个后续磁道的用时为 $2 + 4 = 6$ ms

$$\text{总时间} = 10 + (4 * 6) = 34 \text{ ms} = 0.034 \text{ s}$$

第1个磁道
访问时间

余下4个磁道
访问时间



硬盘存储器：I/O访问时间（示例）

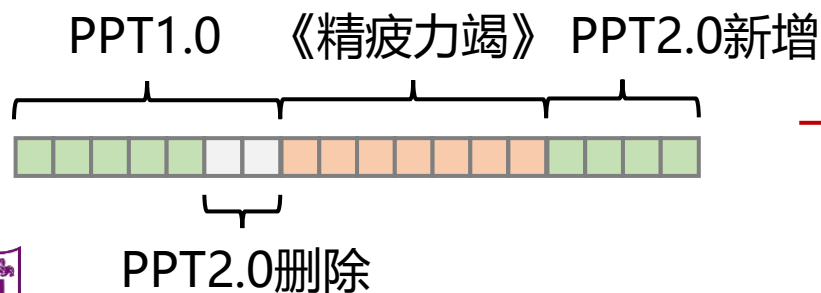
- 情况 2：随机存取

平均寻道	4 ms
平均旋转延迟	2 ms
读1个扇区	0.008 ms
	<hr/>
	6.008 ms

$$\text{总时间} = 2500 * 6.008 = 15020 \text{ ms} = 15.02 \text{ s}$$

需访问
扇区数量

每个扇区
访问时间



磁盘整理：将数据散落在磁盘多个位置的一个文件，读到内存，然后写到一个连续的磁盘区域中



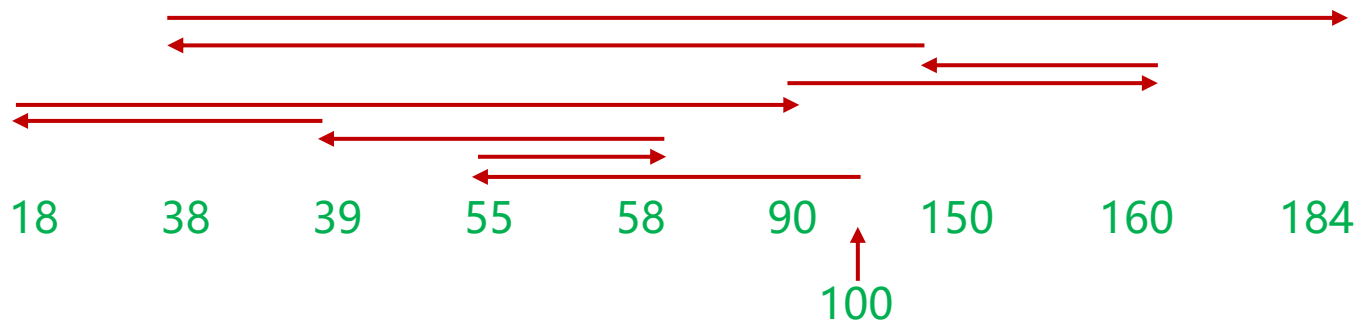
硬磁盘存储器：磁头寻道/磁盘调度

- 目标：当有**多个访问磁盘任务**时，使得**平均寻道时间最小**
- 常见的磁头寻道/磁盘调度算法
 - 先来先服务 (First Come First Service, FCFS)
 - 最短寻道时间优先 (Shortest Seek Time First, SSTF)
 - 扫描/电梯 (SCAN)
 - LOOK
 - 循环扫描 (C-SCAN)
 - C-LOOK



硬磁盘存储器：磁头寻道/磁盘调度（续）

- 先来先服务（FCFS）
 - 按照请求访问磁盘的先后次序进行处理
 - 优点：公平简单
 - 缺点：如果有大量访问磁盘的任务，且请求访问的磁道很分散，则性能上很差，寻道时间长
 - 示例：假设磁头的初始位置是100号磁道，有多个任务先后陆续的请求访问55，58，39，18，90，160，150，38，184号磁道



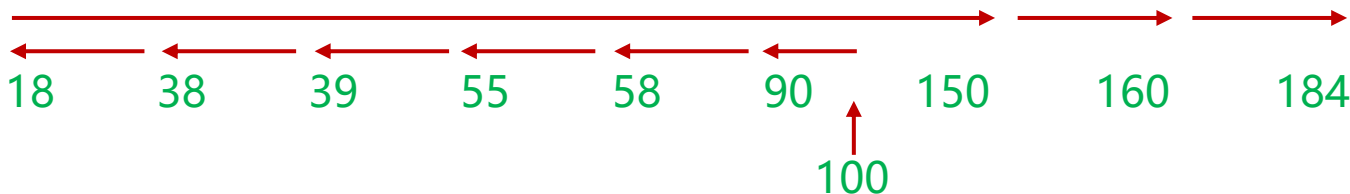
- 磁头总共移动的磁道个数： $45+3+19+21+72+70+10+112+146=498$
- 平均寻道长度： $498/9=55.3$ 个磁道



硬磁盘存储器：磁头寻道/磁盘调度（续）

- 最短寻道时间优先（SSTF）

- 优先处理起始位置与当前磁头位置最接近的读写任务
- 优点：每次的寻道时间最短（局部最优），平均寻道时间缩短
- 缺点：可能产生饥饿现象，尤其是位于两端的磁道请求
- 示例：假设磁头的初始位置是100号磁道，有多个任务先后陆续的请求访问55, 58, 39, 18, 90, 160, 150, 38, 184号磁道



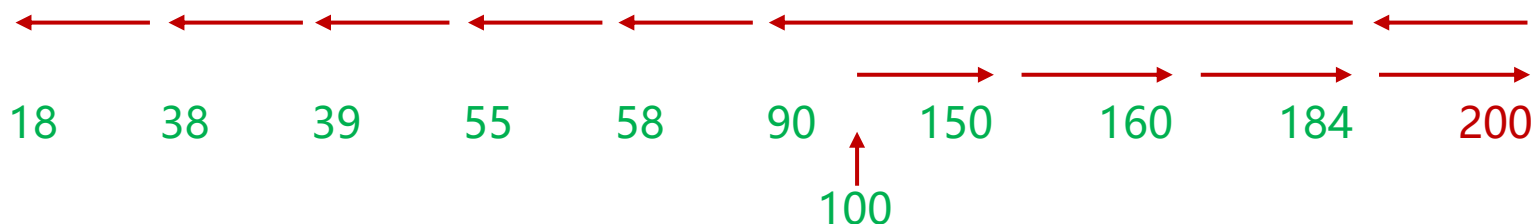
- 磁头总共移动了 $(100-18) + (184-18) = 248$ 个磁道
- 平均寻道长度为 $248/9 = 27.5$ 个磁道



硬磁盘存储器：磁头寻道/磁盘调度（续）

- 扫描/电梯（SCAN）

- 总是按照一个方向进行磁盘调度，直到该方向上的边缘，然后改变方向
- 优点：性能较好，平均寻道时间短，不会产生饥饿现象
- 缺点：只有到最边上的磁道才能改变磁头的移动方向，对于各个位置磁道响应频率不均匀
- 示例：假设磁头的初始位置是100号磁道，有多个任务先后陆续的请求访问55，58，39，18，90，160，150，38，184号磁道



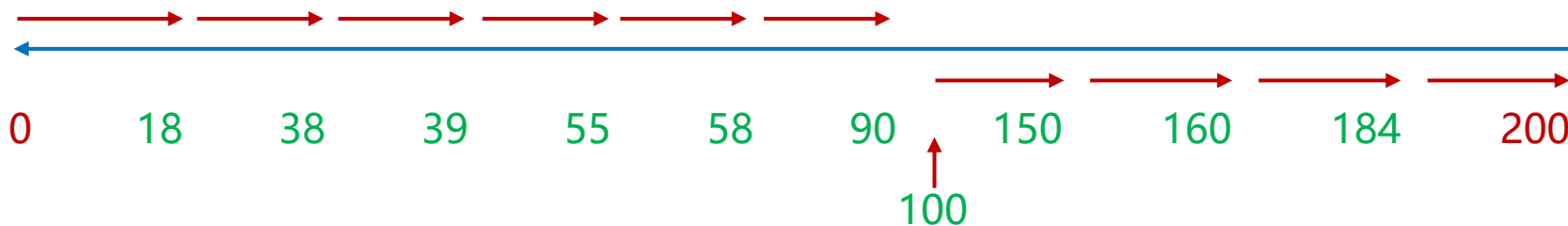
- 磁头总共移动了 $(200-100) + (200-18) = 282$ 个磁道
- 平均寻道长度为 $282/9 = 31.3$ 个磁道



硬磁盘存储器：磁头寻道/磁盘调度（续）

- 循环扫描（C-SCAN）

- 只有磁头朝某个方向移动时才会响应请求，移动到边缘后立即让磁头返回起点，返回途中不做任何处理
- 优点：与SCAN算法相比，对于各位置磁道的响应频率平均
- 缺点：与SCAN算法相比，平均寻道时间更长
- 示例：假设磁头的初始位置是100号磁道，有多个任务先后陆续的请求访问55，58，39，18，90，160，150，38，184号磁道



- 磁头总共移动了 $(200-100)+(200-0)+(90-0)=390$ 个磁道
- 平均寻道长度为 $390/9=43.3$ 个磁道

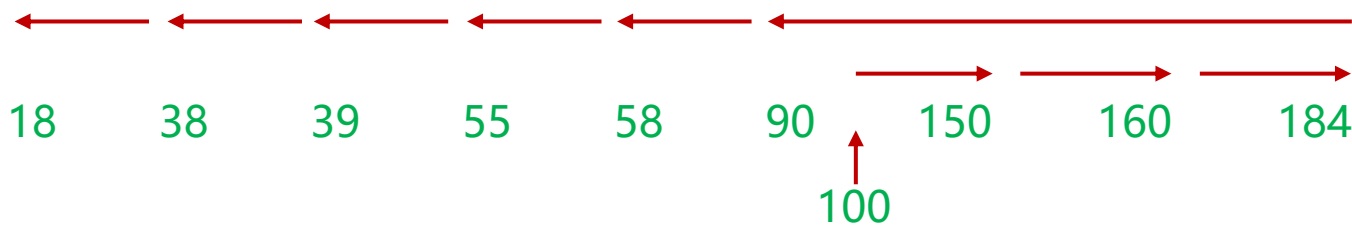
时间短



硬磁盘存储器：磁头寻道/磁盘调度（续）

- LOOK

- **SCAN算法的升级**，只要磁头移动方向上不再有请求就立即改变磁头的方向
- **示例**：假设磁头的初始位置是100号磁道，有多个任务先后陆续的请求访问55, 58, 39, 18, 90, 160, 150, 38, 184号磁道



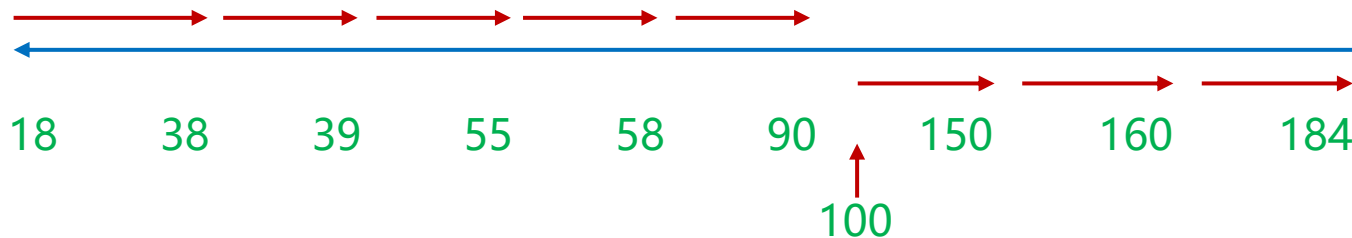
- 磁头总共移动了 $(184-100)+(184-18)=250$ 个磁道
- 平均寻道长度为 $250/9=27.8$ 个磁道



硬磁盘存储器：磁头寻道/磁盘调度（续）

- C-LOOK

- **C-SCAN**算法的改进，只要在磁头移动方向上不再有请求，就立即让磁头返回起点
- **示例：**假设磁头的初始位置是100号磁道，有多个任务先后陆续的请求访问55, 58, 39, 18, 90, 160, 150, 38, 184号磁道



- 磁头总共移动了 $(184-100)+(184-18)+(90-18)=322$ 个磁道
- 平均寻道长度为 $322/9=35.8$ 个磁道



光存储器

• 光存储器产品

- 光盘 (Compact disk, CD)
- 光盘只读存储器 (CD read-only memory, CD-ROM)
- 可刻录光盘 (CD recordable, CD-R)
- 可重写光盘 (CD rewritable, CD-RW)
- 数字多功能光盘 (digital versatile disk, DVD)
- 可刻录DVD (DVD recordable, DVD-R)
- 可重写DVD (DVD rewritable, DVD-RW)
- 高清晰视频光盘 (High definition video disk, Blu-Ray DVD)

名 称	描 述
CD (compact disk, 光盘)	存储数字音频信息的不可擦除盘。标准系统使用 12 厘米的盘, 能够记录可连续播放 60 分钟以上的信息
CD-ROM (compact disk read-only memory, 光盘只读存储器)	用于存储计算机数据的不可擦除盘。标准系统使用 12 厘米的盘, 能够存储 650MB 以上的信息
CD-R (CD recordable, 可刻录光盘)	类似于 CD-ROM, 用户只能向盘写入一次
CD-RW (CD rewritable, 可重写光盘)	类似于 CD-ROM, 用户能多次擦除和重写盘
DVD (digital versatile disk, 数字多功能光盘)	一种制作数字化的压缩的视频信息以及其他大容量数字数据的技术。使用直径为 8 或 12 厘米的盘, 双面容量高达 17GB。基本的 DVD 是只读的 (DVD-ROM)
DVD-R (DVD recordable, 可刻录 DVD)	类似于 DVD-ROM, 用户只能向盘写入一次, 只有一面盘能使用
DVD-RW (DVD rewritable, 可重写 DVD)	类似于 DVD-ROM, 用户能多次擦除和重写盘, 只有一面盘能使用
Blu-Ray DVD (High definition video disk, 高清晰视频光盘)	使用 405nm (蓝-紫色) 的激光, 提供比 DVD 大得多的数据存储密度, 单面单层能存储 25GB 的信息



光存储器 (示例)



CD, **单面**单层, 0.7GB

DVD-5, **单面**单层, 4.7GB

DVD-9, **单面**双层8.5GB

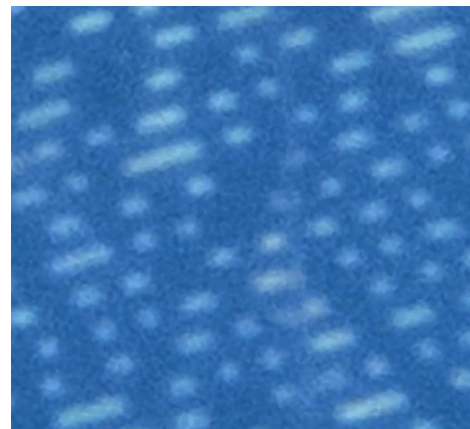
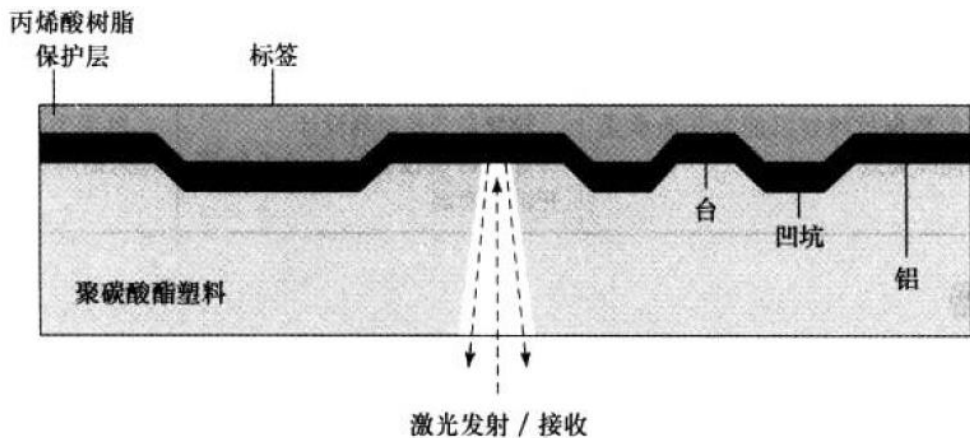
BD25 Blu-ray, **单面**单层, 25GB

BD50 Blu-ray, **单面**双层, 50GB

Ultra HD Blu-ray, **单面**三层, 100GB

CD和CD-ROM

- CD和CD-ROM采用类似的技术，但CD-ROM更加耐用且有纠错功能
- 制造方法
 - 用精密聚焦的高强度激光束制造一个母盘
 - 以母盘作为模板压印出聚碳酸酯的复制品
 - 在凹坑表面上镀上一层高反射材料（一般采用铝，HQCD采用银合金）
 - 使用丙烯酸树脂保护高反射材料
 - 在丙烯酸树脂层上用丝网印刷术印制标签



CD和CD-ROM：读取

- 通过安装在光盘播放器或驱动装置内的低强度激光束从CD或CD-ROM处读取信息
 - 如果激光束照在**凹坑 (pit)** 上，由于凹坑表面有些不平，因此光被散射，反射回**低强度**的激光
 - 如果激光束照在**台/面 (land)** 上，台的表面光滑平坦，反射回来的是**高强度**的激光
 - 光电传感器检测凹坑和台之间的变化，**凹坑的起点或终点**代表 “1” ，没有高度变化则记为 “0” 。
- 盘片上包含**一条单螺旋的轨道**，轨道上的所有扇区长度相同
 - 从内向外扫描，角速度逐渐变小
 - 凹坑被激光以恒定线速度读出

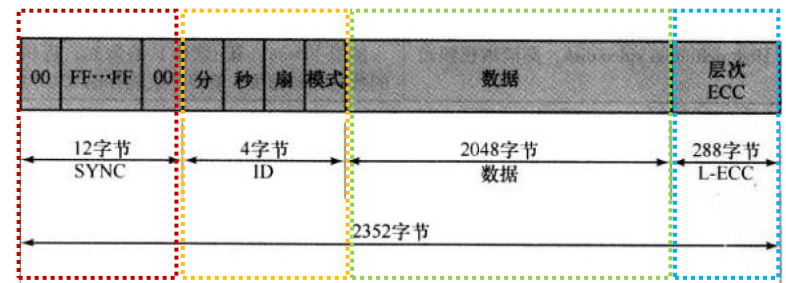
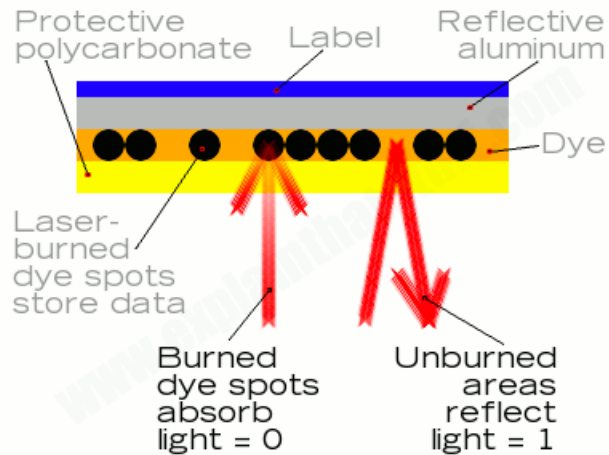


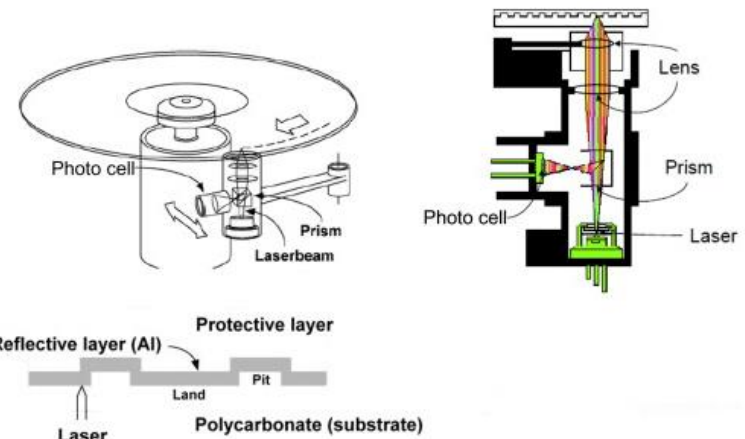
图 6-11 CD-ROM 块格式

CD和CD-ROM: 读取的补充



The information on a CD is encoded in a much more subtle way that uses complex and clever data encoding techniques, including eight-to-fourteen modulation (EFM) and non-return to zero inverted (NRZI) coding.

<https://www.explainthatstuff.com/cdplayers.html>



[Sign in to download full-size image](#)

Figure 1.60. Schematic diagram showing major components of a CD player. The bottom-most figure shows construction of a CD disc media including pits (raised bumps) and lands (flat area between pits) [133].

Optical Media Type	Example
Pre-recorded	CD, CD-ROM, DVD, HD-DVD, BD-ROM
Recordable	CD-R, DVD-R, BD-R
Re-writable	CD-RW, DVD-RW, DVD-RAM, BD-RE

<https://www.sciencedirect.com/topics/engineering/laser-recording>

CD和CD-ROM：优缺点

- **优点**

- 存储有信息的光盘可以廉价地进行大规模复制
- 光盘更换方便
- 寿命长

- **缺点**

- 它是只读的，不能更改
- 其存取时间比磁盘存储器慢得多



CD-R和CD-RW

- **CD-R (写一次)**

- 包含了一个染色层，用于改变反射率，并且由高强度激光激活
- 生成的盘既能在 CD-R 驱动器上也能在 CD-ROM 驱动器上读出

- **CD-RW (写多次)**

- 使用了一种在两种不同相位状态下有两种显著不同反射率的材料，激光束能改变这种材料的相位状态
- 材料老化最终会永久失去相位可变的特性，当前的材料可用于50万次到100万次的擦除

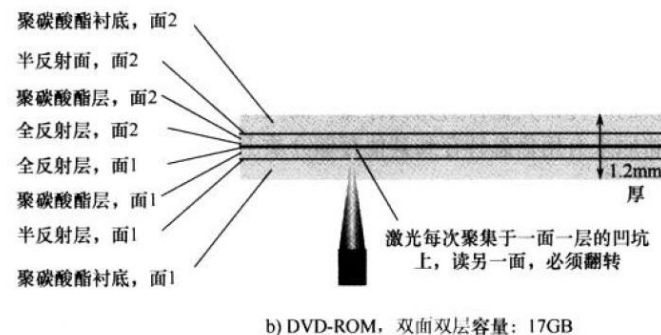
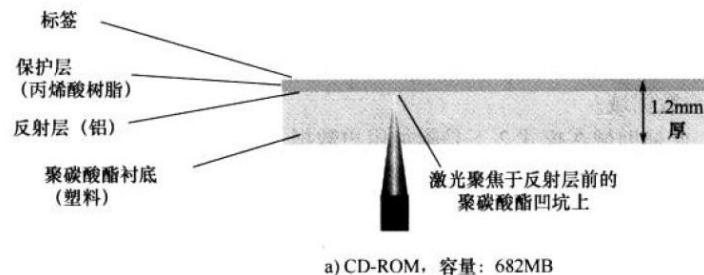


数字多功能光盘 (DVD)

• DVD vs. CD

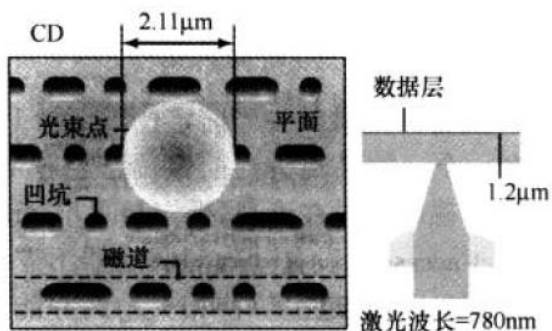
- DVD 上的位组装更紧密：光道间隙，凹坑间距（容量达到4.7GB）
- DVD 采用双层结构：设有半反射层，可以通过调整焦距读取每一层（容量达到8.5GB）
- DVD-ROM 可以用两面记录数据（容量达到17GB）
- DVD的激光束更细，所以能读DVD的设备可以读CD，但反过来不行

• DVD-R和DVD-RW

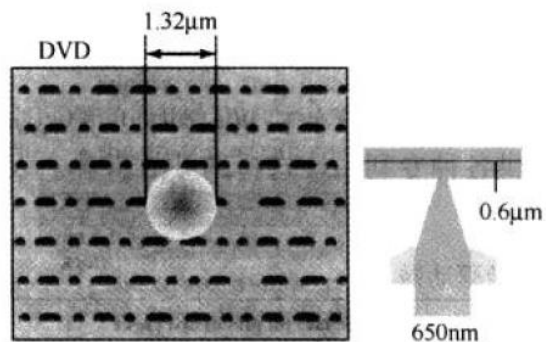


高清晰光盘

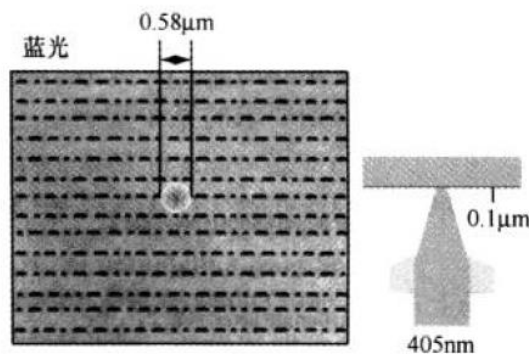
- 通过使用更短波长的激光（在蓝-紫光范围），可以实现更高的位密度（数据凹坑相对更小）



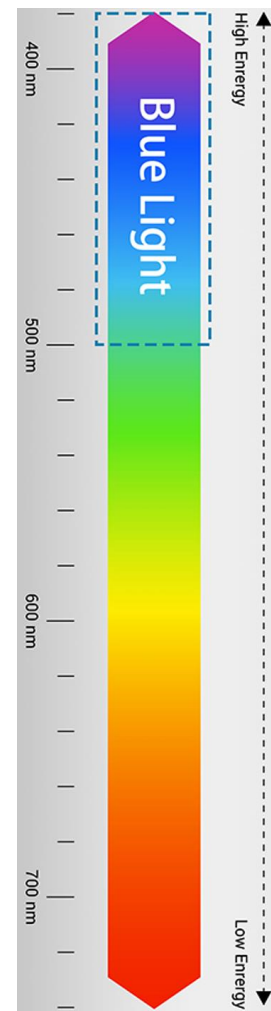
a) CD



b) DVD



c) 蓝光



高清晰光盘

- 格式的战斗 (2003~2008) : HD-DVD vs Blu-Ray

TOSHIBA
405nm, 15GB

VS

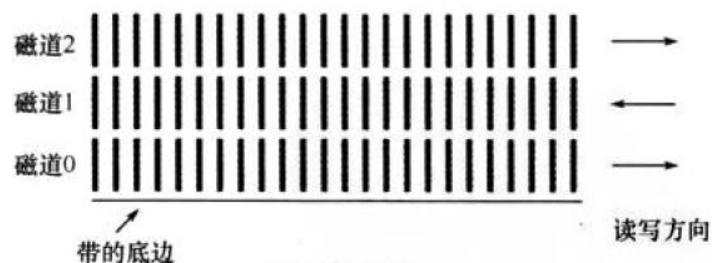
SONY
405nm, 20GB

Walmart
沃尔玛

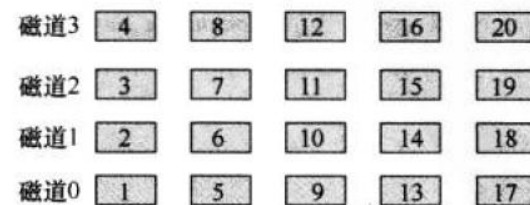


磁带

- 使用与磁盘类似的记录和读取技术
- **记录**
 - 介质是柔韧的聚酯薄膜带，外涂磁性材料
- **读取**
 - 磁带：顺序读取 (sequential-access)
 - 磁盘：直接读取 (direct-access)
 - 并行记录 vs. 串行记录 (蛇形记录)
 - 并行+串行



a) 蛇形读和写



b) 同时读写4磁道的数据块分布

一个大“磁道”包含多个小磁道

磁带：示例



- 便宜
 - 全新4TB硬盘，899元，224.75元/TB
 - Lto6磁带，2.5TB，180元，72元/TB
- 寿命长
 - 硬盘寿命大约十年
 - 磁带寿命数十年
- 适合数据中心冷备份

回顾：快闪存储器

- **Flash Memory**

- 1984年日本电气工程科学家Fujio Masuoka发明了NOR Flash (Flash EEPROM)，可随机寻址，即可用于内部存储。
- 1987年Fujio Masuoka又发明了NAND Flash （用于外部存储器）

- **特性**

- 电可擦除：与EEPROM原理类似，优于EPROM
- 可以在块级擦除，不能在字节级擦除：优于EPROM，不如EEPROM
- 需要先擦除再写入

- **分类**

- **NOR Flash**：数据线和地址线分开，可以读任意字节
- **NAND Flash**：数据线和地址线共用，只能按页读取



U盘和固态硬盘

- **U盘 (USB flash disk)**
 - 采用了快闪存储器，属于非易失性半导体存储器
 - 相比于软盘和光盘：体积小，容量大，携带方便，寿命长达数年
- **固态硬盘 (Solid State Disk/Drive)**
 - 与U盘没有本质区别：容量更大，存储性能更好
 - 与硬磁盘存储器相比：抗振性好，无噪声，能耗低，发热量低



总结

- 磁盘存储器
 - 软盘
 - 硬盘：结构，读写机制，数据组织，格式化，I/O访问时间，磁头寻道
- 光存储器
 - CD, CD-ROM, CD-R, CD-RW
 - DVD, DVD-R, DVD-RW
 - Blu-ray
- 磁带
- U盘和固态硬盘



谢谢

bohanliu@nju.edu.cn



南京大學
NANJING UNIVERSITY