**文件**：**具有名字的一组相关信息的集合，由若干个记录组成(相关数据项集合)**

**数据项（信息项）是数据组织中可以命名的最小逻辑单位**

**文件系统：文件管理程序+它所管理的全部文件**

**文件系统的主要功能：按文件名存取文件信息(文件名→地址)、提供统一的和友好的接口、对文件和文件目录的管理、分配回收存储空间、文件保护和共享**

**模型:用户→接口→管理/操作→对象/属性**

**对象及其属性：文件、目录、存储空间**

**操作/管理(核心):存储空间管理、目录管理、地址转化管理、读写管理、共享/保护**

**接口:命令、图形、程序接口**

**分类目的：为了更好地管理和使用文件、提高存取速度、更好地执行共享与保护**

**按性质和用途分类**

**系统文件:系统软件的文件，用户通过系统调用或系统提供的专用命今来执行，不允许对其进行读写和修改**

**库文件：允许用户读取和执行，不允许修改，主要由各种标准子程序库组成(.lib)**

**用户文件：通过操作系统保存的用户文件，主要有用户的源程序等**

**操作保护分类：只读（r）、可读可写（rw）、可执行(x)**

**使用情况分类：临时文件、永久文件、档案文件**

**按用户观点分类：普通文件(由字符流组成的无结构文件)、目录文件、特殊文件（将设备驱动程序看成文件，便于统一管理）**

**数据形式分类：源文件、目的文件(编译后未链接的目的代码)、可执行文件(链接后的obj)**

**文件操作：**

**建立：新文件→分配外存→建立目录项(文件名、地址等)**

**删除：查找目录→清空(删除？)目的目录项→回收空间**

**打开：访问→将目录信息写入内存活动文件表→可以多次打开→关闭**

**关闭：删除文件表中的目录信息→若修改过则回写辅存**

**读:查找目录→找到目录项→找到外存地址→数据传入内存**

**写：查找目录→找到目录项→找到文件指针→写入信息**

**对文件属性进行操作：重命名、改变拥有者、修改访问权、查看状态**

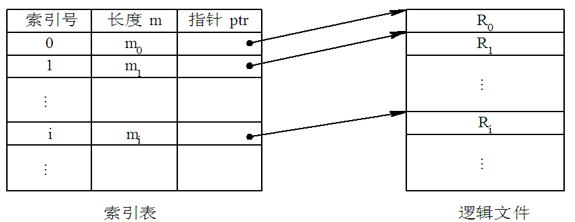
**对目录进行操作：创建目录、删除、修改、调用、**

**有结构文件：定长记录/变长记录、顺序文件/索引文件/顺序索引文件**

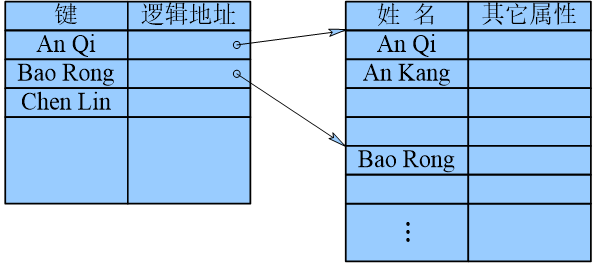
**无结构文件(流式文件)：以字节为单位，系统不对文件进行格式化处理，读写指针**

**顺序文件的优缺点：读写大批量数据时顺序文件最快、查找/修改/增加删除单个记录效率差**

**变长记录→顺序直接存取效率差→索引**



**索引顺序文件**



**直接文件:根据给定的记录键值，直接获得指定记录的物理地址**

**散列文件：hash[记录键值]→地址**

**物理结构：一个文件在外存上的存储组织形式(顺序式、链式、索引式文件结构)**

**顺序结构：将文件信息存放在若干连续的物理块中(顺序读取速度快、寻道少、寻道时间短)**

**缺点：要求有连续的存储空间、必须事先知道文件的长度**

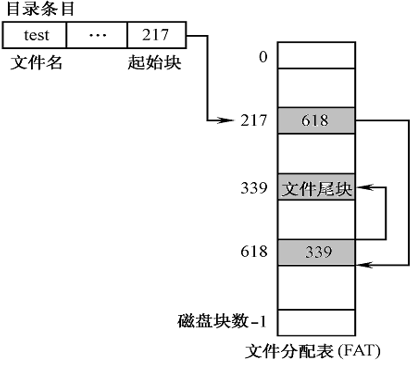
**链式结构：将一个逻辑上连续的文件信息存放在外存的不连续(或连续)物理块中，通过在每个盘块上的链接指针，将同属于一个文件的多个离散的盘块链接成一个链表**

**优点：提高磁盘利用率、无碎片、易扩充、易插入删除**

**缺点：存取速度慢，不适于随机存取、指针出错(可靠性)、寻道数和时间多、指针占空间**

**隐式链接：用文件目录[文件名、始址、末址]**

**显式链接：将盘块中的链接字按盘块号的顺序集中起来，构成文件分配表FAT**



**缺点：占空间、内容够大时FAT可能装不进内存，影响效率**

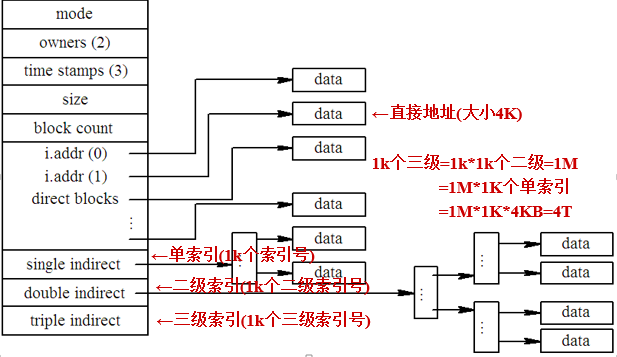
**索引：每个文件建立一个索引表，并将这些块的块号存放在一个索引表[逻辑号、物理号]中**



**单级索引：见上图**

**多级索引:将一个大文件的所有索引表（二级索引)的地址放在另一个索引表（一级索引)中**

**混合索引：多种索引分配方式相结合而形成的一种分配方式，见下图**



**优点：既能顺序存取,又能随机存取，满足了文件动态增长、插入删除的要求，也能充分利用外存空间**

**缺点：较多的寻道次数和寻道时间，索引表本身带来了系统开销**

**文件的存取方法：读写文件存储器上的一个物理块的方法(顺序、直接、按键)**

**顺序定长：读指针rptr=rptr+L，写指针wptr=wptr+L**

**顺序变长：读指针rptr+Li=>rptr写指针wptr+Li=>wptr**

**直接存取(又称随机存取)法：按记录编号来直接存取文件中的任一记录**

**按键存取法实质上是直接存取法，根据文件记录中的关键字(通常称为键)经过某种方法计算处理，转换成相应的物理地址后进行存取，被广泛采用**

**对键或记录的搜索算法：线性搜索法、散列法、二分**

**文件控制块（FCB）：文件控制块就是一个文件目录项，文件存在的标志，FCB=文件名+inode**

**内容：基本信息类(名字、物理位置、逻辑结构、物理结构)、存取控制信息类(文件主的存取权限、核准用户的存取权限、一般用户的存取权限)、使用信息类(因文件操作而产生的信息)**

**索引结点I node：将文件的描述信息（除文件名外）单独形成一个称为索引结点的数据结构**

**内容:标识符、存取权限、类型、物理地址、长度、链接计时、存取时间**

**内存索引结点:文件被打开时结点进入内存，增加内容编号、状态、访问数、设备号、指针**

**目录是文件控制块(FCB)的有序集合，文件系统将若干个文件目录组成一个独立的目录文件**

**目录结构：**

**单级目录：为所有文件建立一个目录表，实现简单、查找困难、不允许重名、不方便共享**

**二级目录：将目录分为2级：主文件目录（MFD） 和用户文件目录（UFD,每个用户有一个）**

**MFD=user+UFD.addr，始于多用户分时系统**

**优点：解决了文件的重名问题和文件共享问题，提高了搜索速度，降低了查找时间**

**缺点：不太适合大量用户和大量文件的大系统，增加了系统开销**

**多级目录结构（树型目录）：目录与文件在一起，目录也做成文件，被广泛采用(始于unix)优点：便于管理和保护、便于分类、解决重命名、提高搜索速率、能够进行存取权限控制**

**缺点：查找一个文件按路径名逐层检查，由于每个文件都放在外存，多次访盘影响速度**

**目录查询算法：线性检索法、Hash方法查找**

**文件存储空间管理：文件新建(连续分配/离散)与删除的管理方法**

**空闲表：为外存上的所有空闲区建立一张空闲表，空闲区对应空闲项(编号、长度、起始、l连续的空闲块集合)，适用于建立连续文件，参照存储管理方案**

**空闲链表法：将所有空闲盘区拉成一条空闲链(空闲盘块链/空闲盘区链)**

**位示图：用一位二进制表以反映整个存储空间的分配情况，0未分配1分配**

**UNIX：略**

**文件共享：指不同的用户可以使用同一个文件**

**实现方法：**

**绕道法：由系统目录实现对文件的共享，用户通过全路径名共享地访问这些文件**

**链接法：在相应目录表之间进行链接**

**基本文件目录表：基本文件目录表(BFD文件信息+内部标识符)+符号文件目录表(SFD符号名+标识符)**

**基于索引节点的共享方式：计数器count表示正在共享的用户/作业数量（详情见前面）**

**利用符号链实现文件共享：共享某文件时，创建一LINK类型的新文件加到用户目录中，该文件仅包含被链接文件F的路径名，称该链接方法为符号链接，该方式中，只有文件所有者才拥有指向其索引结点的指针，其它共享的用户只有该文件的路径名**

**优点：方便地链接任一文件（用路径名），包括网络上文件**

**缺点：访问共享文件时开销大（多次读盘，消耗盘空间）、每一共享文件都要增加一文件名（因路径名各不相同）**

**文件保护：略**