

本节内容

文件的物理结构 (文件分配方式)

王道考研/CSKAOYAN.COM

1

知识总览

操作系统需要对磁盘块进行哪些管理

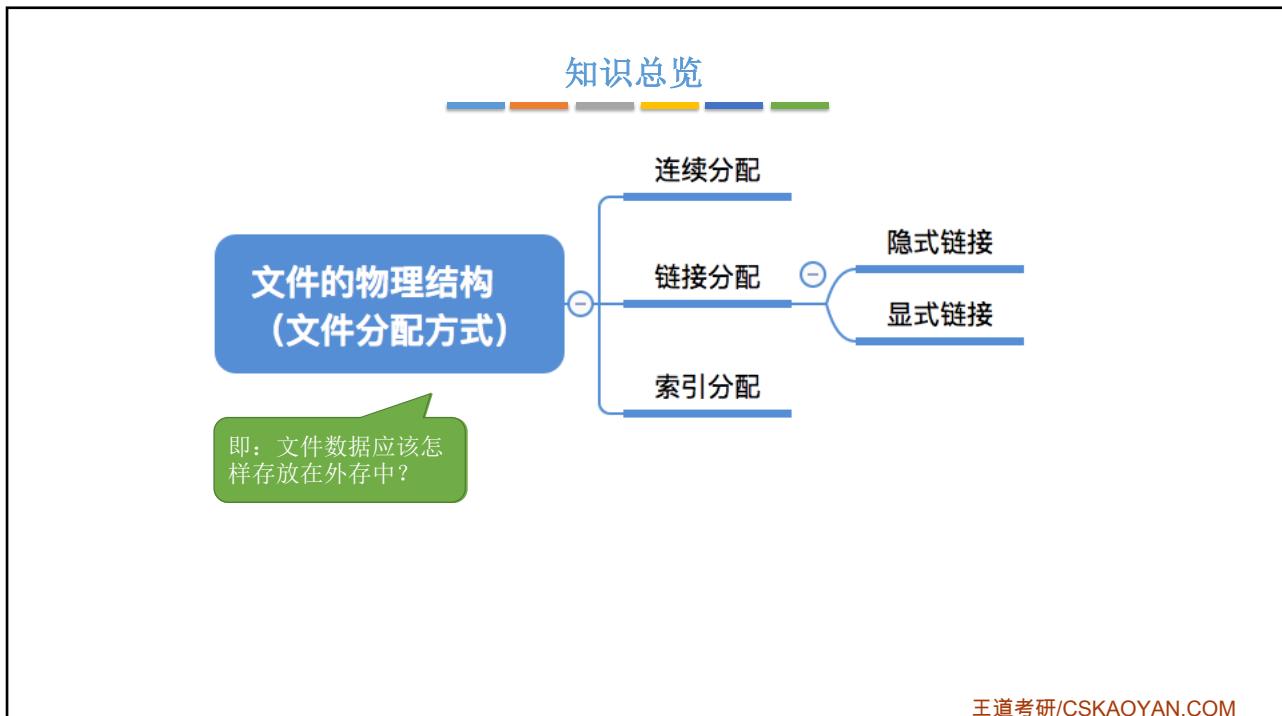
对非空闲磁盘块的管理 (存放了文件数据的磁盘块)

对空闲磁盘块的管理

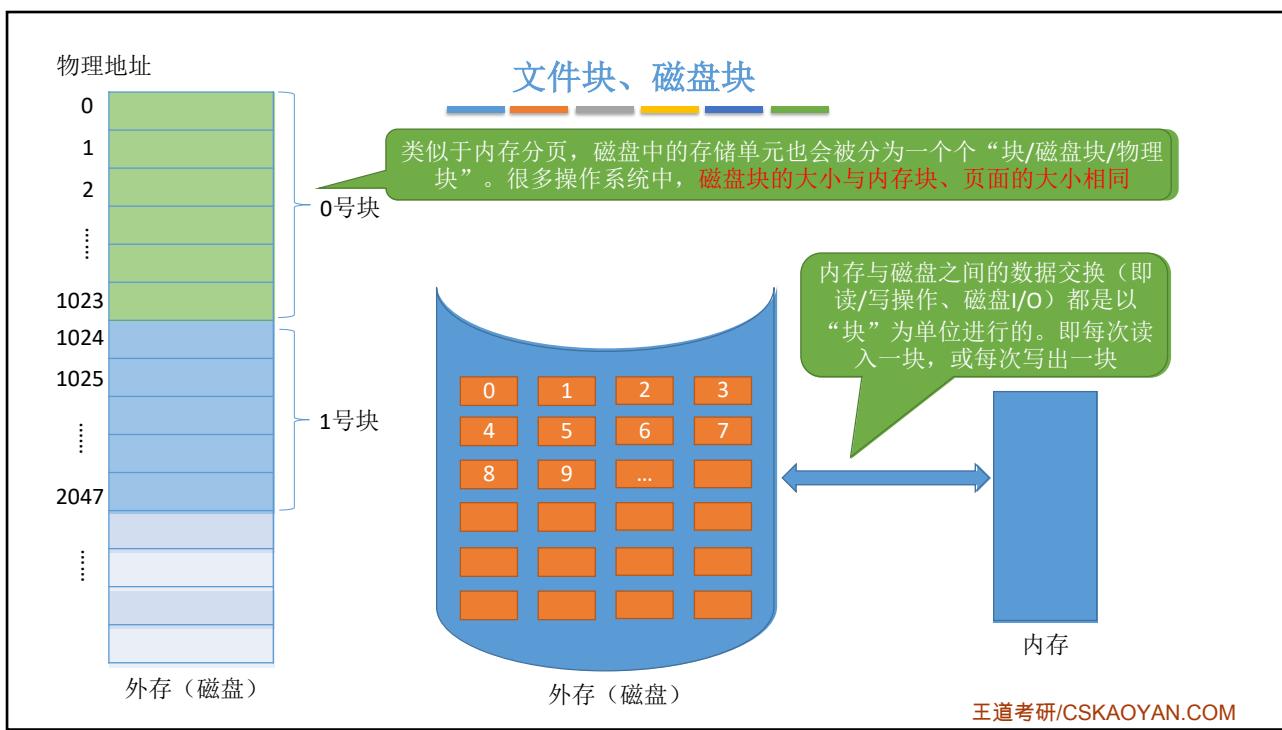
“文件的物理结构/文件分配方式”要探讨的问题

王道考研/CSKAOYAN.COM

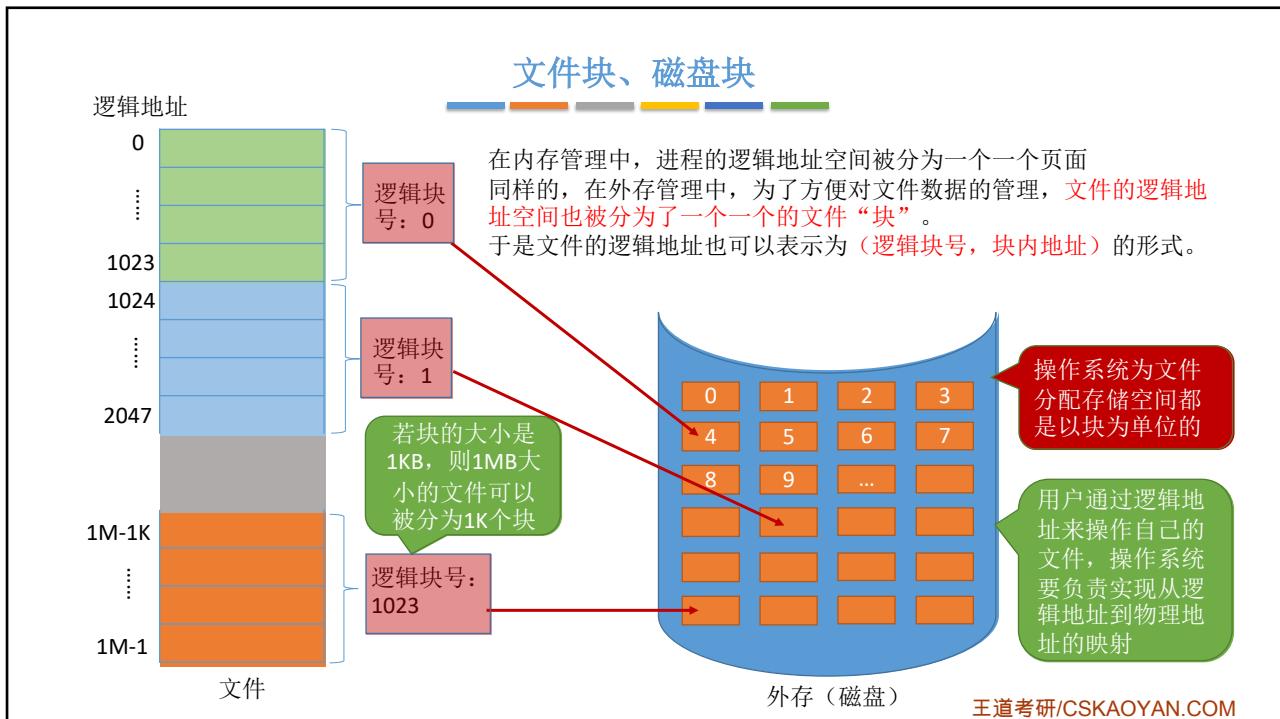
2



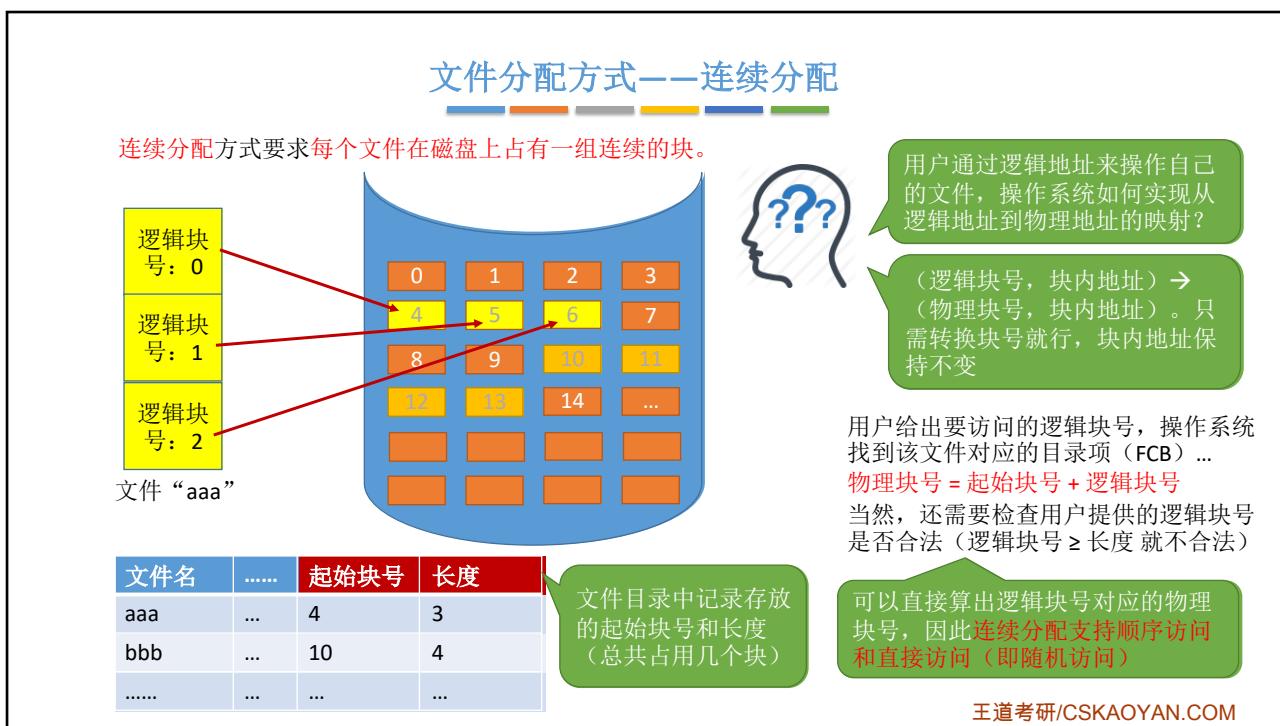
3



4



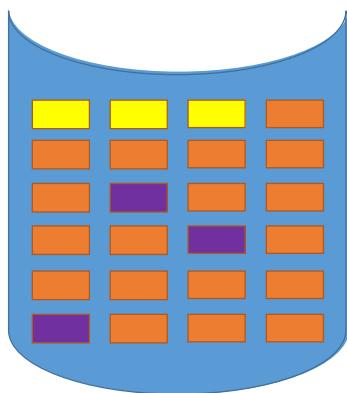
5



6

文件分配方式——连续分配

连续分配方式要求每个文件在磁盘上占有一组连续的块。



读取某个磁盘块时，需要移动磁头。访问的两个磁盘块相隔越远，移动磁头所需时间就越长。

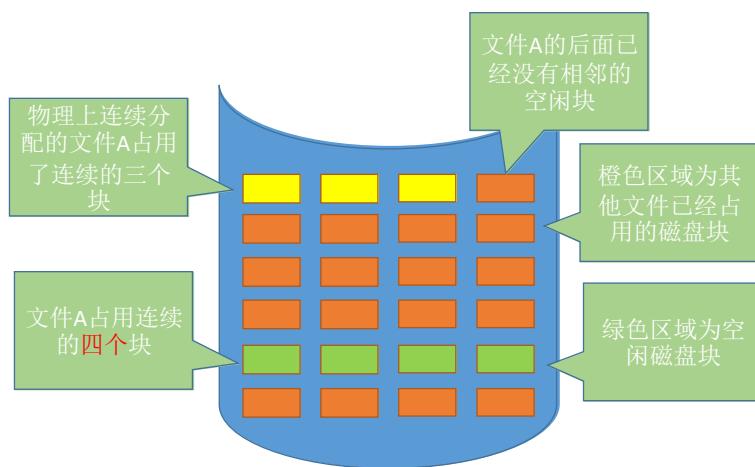
结论：连续分配的文件在顺序读/写时速度最快

王道考研/CSKAOYAN.COM

7

文件分配方式——连续分配

连续分配方式要求每个文件在磁盘上占有一组连续的块。



若此时文件A要拓展，需要再增加一个磁盘块（总共需要连续的4个磁盘块）。由于采用连续结构，因此文件A占用的磁盘块必须是连续的。因此只能将文件A全部“迁移”到绿色区域。

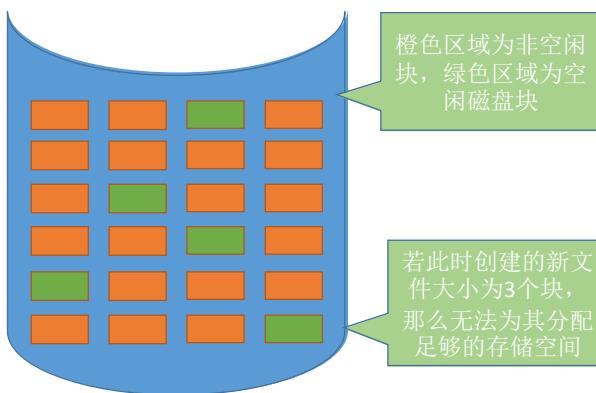
结论：物理上采用连续分配的文件不方便拓展。

王道考研/CSKAOYAN.COM

8

文件分配方式——连续分配

连续分配方式要求每个文件在磁盘上占有一组连续的块。



王道考研/CSKAOYAN.COM

9

连续分配（总结）

连续分配方式要求每个文件在磁盘上占有一组连续的块。

优点：支持顺序访问和直接访问（即随机访问）；连续分配的文件在顺序访问时速度最快
缺点：不方便文件拓展；存储空间利用率低，会产生磁盘碎片

王道考研/CSKAOYAN.COM

10

文件分配方式——链接分配

链接分配采取离散分配的方式，可以为文件分配离散的磁盘块。分为**隐式链接**和**显式链接**两种。

王道考研/CSKAOYAN.COM

11

链接分配——隐式链接



如何实现文件的逻辑块号到物理块号的转变？

用户给出要访问的逻辑块号 i ，操作系统找到该文件对应的目录项 (FCB) ...

从目录项中找到起始块号 (即0号块)，将0号逻辑块读入内存，由此知道1号逻辑块存放的物理块号，于是读入1号逻辑块，再找到2号逻辑块的存放位置.....以此类推。因此，读入 i 号逻辑块，总共需要 $i+1$ 次磁盘 I/O。

结论：采用**链式分配**（**隐式链接**）方式的文件，**只支持顺序访问，不支持随机访问**，查找效率低。另外，指向下一个盘块的指针也需要耗费少量的存储空间。

王道考研/CSKAOYAN.COM

12

链接分配——隐式链接

文件名	起始块号	结束块号
aaa	...	9	8

是否方便拓展文件？

若此时要拓展文件，则可以随便找一个空闲磁盘块，挂到文件的磁盘块链尾，并修改文件的FCB

结论：采用隐式链接的链接分配方式，很方便文件拓展。另外，所有的空闲磁盘块都可以被利用，不会有碎片问题，外存利用率高。

王道考研/CSKAOYAN.COM

13

链接分配——隐式链接

链接分配采取离散分配的方式，可以为文件分配离散的磁盘块。分为**隐式链接**和**显式链接**两种。

隐式链接——除文件的最后一个盘块之外，每个盘块中都存有指向下一个盘块的指针。文件目录包括文件第一块的指针和最后一块的指针。

优点：很方便文件拓展，不会有碎片问题，外存利用率高。

缺点：只支持顺序访问，不支持随机访问，查找效率低，指向下一个盘块的指针也需要耗费少量的存储空间。

王道考研/CSKAOYAN.COM

14

链接分配——显式链接

文件名	起始块号
aaa	...	2
bbb	...	4

目录中只需记录文件的起始块号

物理块号	下一块
0	1
1	-1
2	5
3	-1
4	23
5	0
.....	
22	
23	3

把用于链接文件各物理块的指针显式地存放在一张表中。即文件分配表（FAT, File Allocation Table）

FAT (文件分配表)

假设某个新创建的文件“aaa”依次存放在磁盘块 $2 \rightarrow 5 \rightarrow 0 \rightarrow 1$

假设某个新创建的文件“bbb”依次存放在磁盘块 $4 \rightarrow 23 \rightarrow 3$

注意：一个磁盘仅设置一张FAT。开机时，将FAT读入内存，并常驻内存。FAT的各个表项在物理上连续存储，且每一个表项长度相同，因此“物理块号”字段可以是隐含的。

王道考研/CSKAOYAN.COM

15

链接分配——显式链接

文件名	起始块号
aaa	...	2
bbb	...	4

物理块号	下一块
0	1
1	-1
2	5
3	-1
4	23
5	0
.....	
22	
23	3

FAT (文件分配表)

如何实现文件的逻辑块号到物理块号的转变？

用户给出要访问的逻辑块号 i ，操作系统找到该文件对应的目录项（FCB）...

从目录项中找到起始块号，若 $i > 0$ ，则查询内存中的文件分配表FAT，往后找到 i 号逻辑块对应的物理块号。逻辑块号转换成物理块号的过程不需要读磁盘操作。

结论：采用链式分配（显式链接）方式的文件，支持顺序访问，也支持随机访问（想访问 i 号逻辑块时，并不需要依次访问之前的 $0 \sim i-1$ 号逻辑块），由于块号转换的过程不需要访问磁盘，因此相比于隐式链接来说，访问速度快很多。

显然，显式链接也不会产生外部碎片，也可以很方便地对文件进行拓展。

王道考研/CSKAOYAN.COM

16

链接分配（总结）

链接分配采取离散分配的方式，可以为文件分配离散的磁盘块。分为**隐式链接**和**显式链接**两种。

隐式链接——除文件的最后一个盘块之外，每个盘块中都存有指向下一个盘块的指针。文件目录包括文件第一块的指针和最后一块的指针。

优点：很方便文件拓展，不会有碎片问题，外存利用率高。

缺点：只支持顺序访问，不支持随机访问，查找效率低，指向下一个盘块的指针也需要耗费少量的存储空间。

考试题目中遇到未指明隐式/显式的“链接分配”，默认指的是**隐式链接**的链接分配

显式链接——把用于链接文件各物理块的指针显式地存放在一张表中，即**文件分配表（FAT, File Allocation Table）**。一个磁盘只会建立一张文件分配表。开机时文件分配表放入内存，并**常驻内存**。

优点：很方便文件拓展，不会有碎片问题，外存利用率高，并且**支持随机访问**。相比于隐式链接来说，**地址转换时不需要访问磁盘，因此文件的访问效率更高**。

缺点：文件分配表的需要占用一定的存储空间。

王道考研/CSKAOYAN.COM

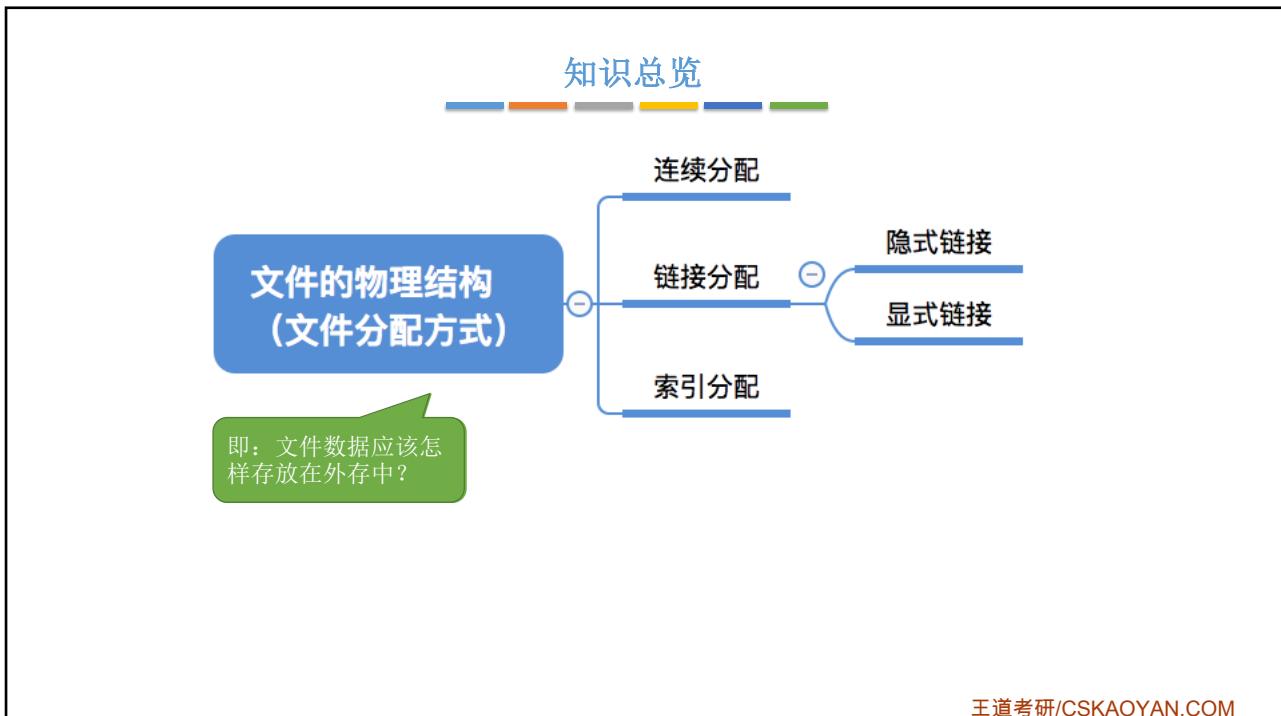
17

本节内容

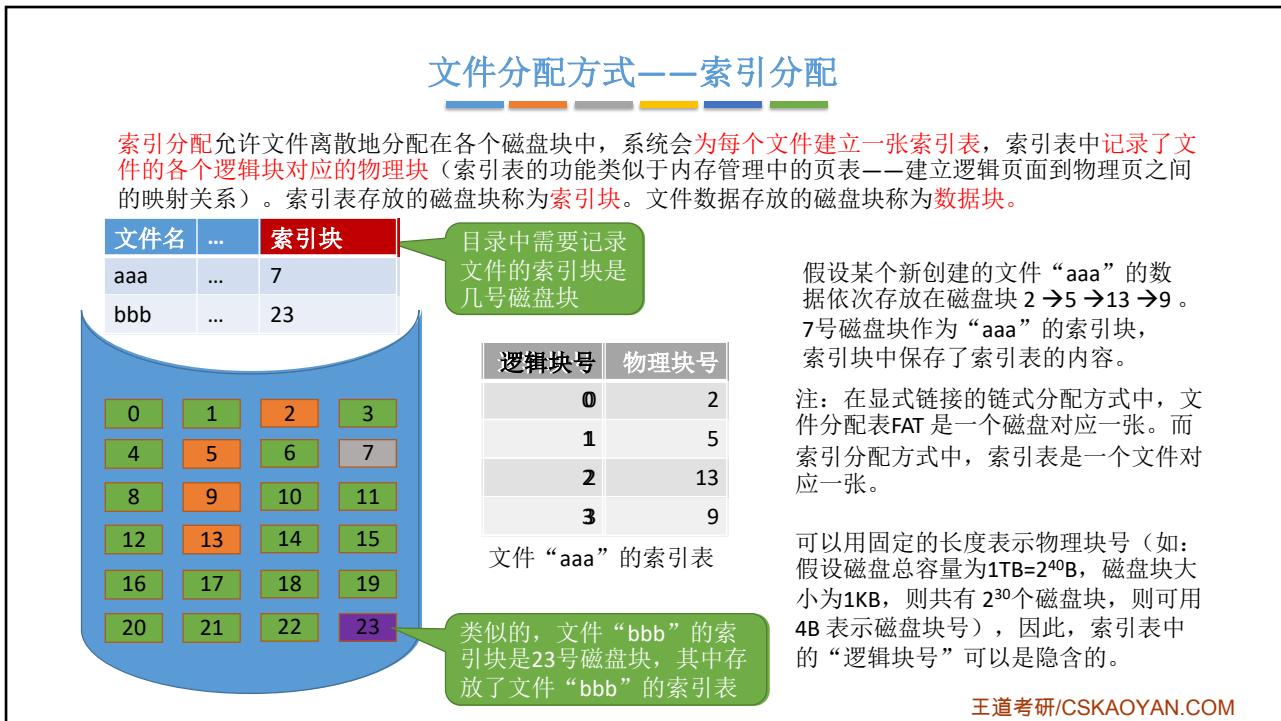
文件的物理
结构
(文件分配方式)

王道考研/CSKAOYAN.COM

18



19



20

文件分配方式——索引分配

索引分配允许文件离散地分配在各个磁盘块中，系统会为每个文件建立一张索引表，索引表中记录了文件的各个逻辑块对应的物理块。索引表存放的磁盘块称为索引块。文件数据存放的磁盘块称为数据块。



目录中需要记录文件的索引块是几号磁盘块



如何实现文件的逻辑块号到物理块号的转换？

逻辑块号	物理块号
0	2
1	5
2	13
3	9
4	19

文件“aaa”的索引表

用户给出要访问的逻辑块号 i ，操作系统找到该文件对应的目录项 (FCB) ...

从目录项中可知索引表存放位置，将索引表从外存读入内存，并查找索引表即可只 i 号逻辑块在外存中的存放位置。

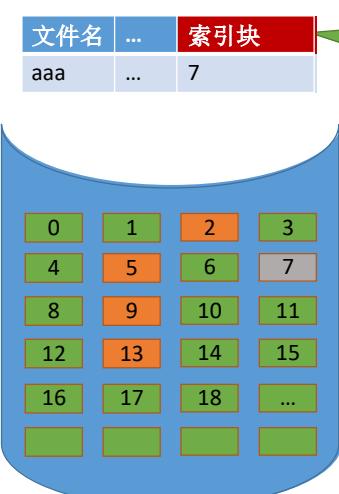
可见，索引分配方式可以支持随机访问。
文件拓展也很容易实现（只需要给文件分配一个空闲块，并增加一个索引表项即可）
但是索引表需要占用一定的存储空间

王道考研/CSKAOYAN.COM

21

文件分配方式——索引分配

索引分配允许文件离散地分配在各个磁盘块中，系统会为每个文件建立一张索引表，索引表中记录了文件的各个逻辑块对应的物理块。索引表存放的磁盘块称为索引块。文件数据存放的磁盘块称为数据块。



目录中需要记录文件的索引块是几号磁盘块



若每个磁盘块1KB，一个索引表项4B，则一个磁盘块只能存放 256 个索引项。

逻辑块号	物理块号
0	2
1	5
2	13
3	9

文件“aaa”的索引表

如果一个文件的大小超过了256块，那么一个磁盘块是装不下文件的整张索引表的，如何解决这个问题？

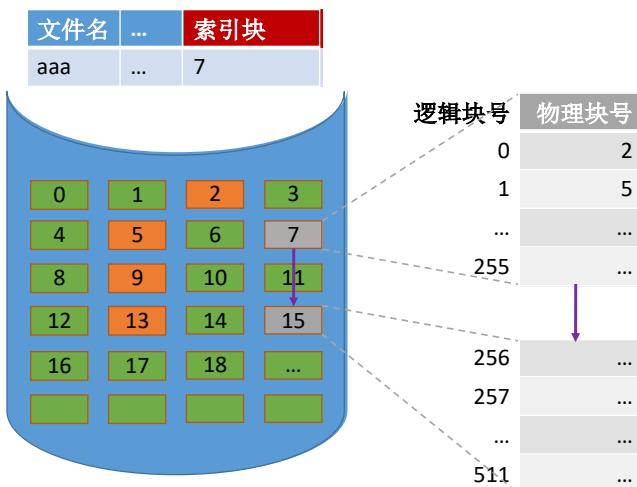
- ①链接方案
- ②多层次索引
- ③混合索引

王道考研/CSKAOYAN.COM

22

文件分配方式——索引分配

①链接方案：如果索引表太大，一个索引块装不下，那么可以将多个索引块链接起来存放。



假设磁盘块大小为1KB，一个索引表项占4B，则一个磁盘块只能存放256个索引项。

若一个文件大小为 $256 \times 256 \text{ KB} = 65,536 \text{ KB} = 64 \text{ MB}$

该文件共有 256×256 个块，也就对应 256×256 个索引项，也就需要 256 个索引块来存储，这些索引块用链接方案连起来。

若想要访问文件的最后一个逻辑块，就必须找到最后一个索引块（第256个索引块），而各个索引块之间是用指针链接起来的，因此必须先顺序地读入前 255 个索引块。

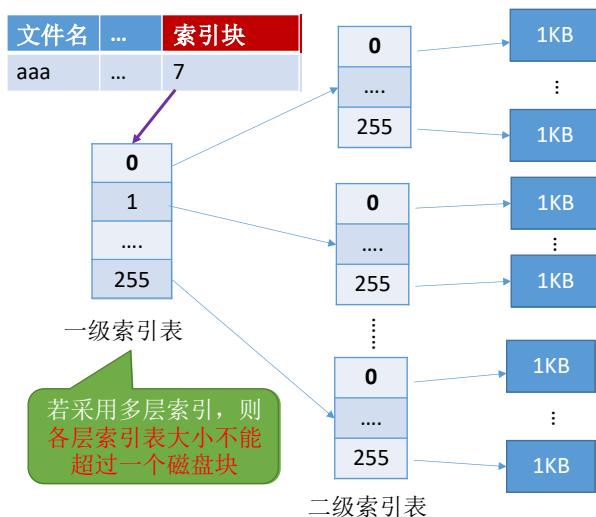
这显然是很低效的。如何解决呢？

王道考研/CSKAOYAN.COM

23

文件分配方式——索引分配

②多层索引：建立多层次索引（原理类似于多级页表）。使第一层索引块指向第二层的索引块。还可根据文件大小的要求再建立第三层、第四层索引块。



假设磁盘块大小为1KB，一个索引表项占4B，则一个磁盘块只能存放256个索引项。

若某文件采用两层索引，则该文件的最大长度可以到 $256 \times 256 \times 1 \text{ KB} = 65,536 \text{ KB} = 64 \text{ MB}$

可根据逻辑块号算出应该查找索引表中的哪个表项。如：要访问 1026 号逻辑块，则

$$1026/256 = 4, 1026 \% 256 = 2$$

因此可以先将一级索引表调入内存，查询 4 号表项，将其对应的二级索引表调入内存，再查询二级索引表的 2 号表项即可知道 1026 号逻辑块存放的磁盘块号了。访问目标数据块，需要 3 次磁盘 I/O。

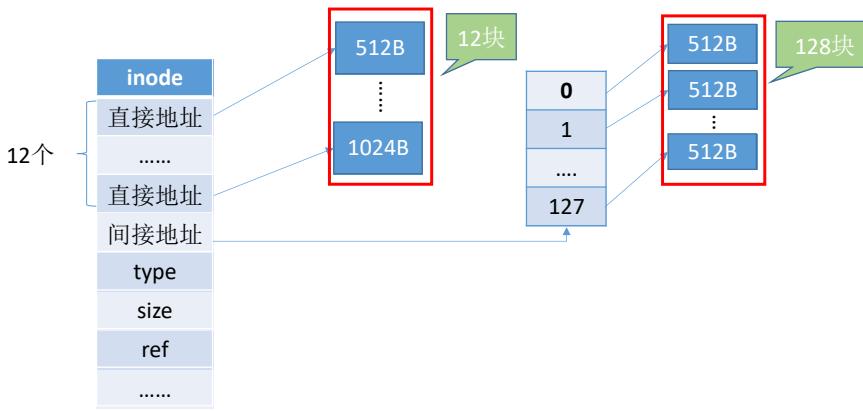
若采用 K 层索引结构，且顶级索引表未调入内存，则访问一个数据块只需要 $K+1$ 次读磁盘操作

王道考研/CSKAOYAN.COM

24

文件分配方式——索引分配

③混合索引：多种索引分配方式的结合。例如，一个文件的顶级索引表中，既包含直接地址索引（直接指向数据块），又包含一级间接索引（指向单层索引表）、还包含两级间接索引（指向两层索引表）。



王道考研/CSKAOYAN.COM

25

索引分配（总结）

索引分配允许文件离散地分配在各个磁盘块中，系统会为每个文件建立一张索引表，索引表中记录了文件的各个逻辑块对应的物理块（索引表的功能类似于内存管理中的页表——建立逻辑页面到物理页之间的映射关系）。索引表存放的磁盘块称为索引块。文件数据存放的磁盘块称为数据块。

若文件太大，索引表项太多，可以采取以下三种方法解决：

①链接方案：如果索引表太大，一个索引块装不下，那么可以将多个索引块链接起来存放。缺点：若文件很大，索引表很长，就需要将很多个索引块链接起来。想要找到 i 号索引块，必须先依次读入 $0 \sim i-1$ 号索引块，这就导致磁盘I/O次数过多，查找效率低下。

②多层次索引：建立多层次索引（原理类似于多级页表）。使第一层索引块指向第二层的索引块。还可根据文件大小的要求再建立第三层、第四层索引块。采用 K 层索引结构，且顶级索引表未调入内存，则访问一个数据块只需要 $K+1$ 次读磁盘操作。缺点：即使是小文件，访问一个数据块依然需要 $K+1$ 次读磁盘。

③混合索引：多种索引分配方式的结合。例如，一个文件的顶级索引表中，既包含直接地址索引（直接指向数据块），又包含一级间接索引（指向单层索引表）、还包含两级间接索引（指向两层索引表）。优点：对于小文件来说，访问一个数据块所需的读磁盘次数更少。

超级超级超级重要考点：①要学会根据多层次索引、混合索引的结构计算出文件的最大长度（Key：各级索引表最大不能超过一个块）；②要能自己分析访问某个数据块所需要的读磁盘次数（Key：FCB中会存有指向顶级索引块的指针，因此可以根据FCB读入顶级索引块。每次读入下一级的索引块都需要一次读磁盘操作。另外，要注意题目条件——顶级索引块是否已调入内存）

王道考研/CSKAOYAN.COM

26

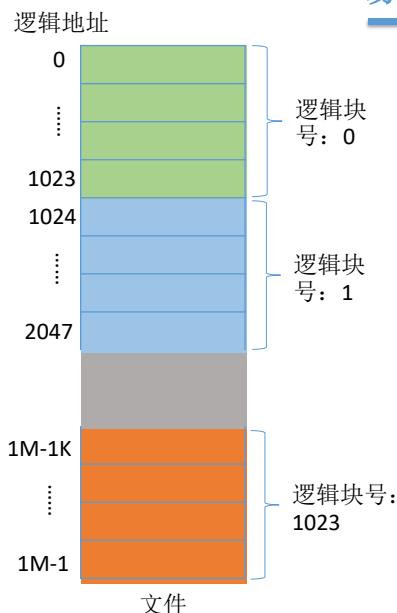
知识点回顾与重要考点

		How?	目录项内容	优点	缺点
顺序分配		为文件分配的必须是连续的磁盘块	起始块号、文件长度	顺序存取速度快，支持随机访问	会产生碎片，不利于文件拓展
链接分配	隐式链接	除文件的最后一个盘块之外，每个盘块中都存有指向下一个盘块的指针	起始块号、结束块号	可解决碎片问题，外存利用率高，文件拓展实现方便	只能顺序访问，不能随机访问。
	显式链接	建立一张文件分配表(FAT)，显式记录盘块的先后关系（开机后FAT常驻内存）	起始块号	除了拥有隐式链接的优点之外，还可通过查询内存中的FAT实现随机访问	FAT需要占用一定的存储空间
索引分配		为文件数据块建立索引表。若文件太大，可采用链接方案、多层次索引、混合索引	链接方案记录的是第一个索引块的块号，多层次/混合索引记录的是顶级索引块的块号	支持随机访问，易于实现文件的拓展	索引表需占用一定的存储空间。访问数据块前需要先读入索引块。若采用链接方案，查搜索引块时可能需要很多次读磁盘操作。

王道考研/CSKAOYAN.COM

27

易混难点：支持随机访问



假设这个文件的逻辑结构是“顺序文件”，并且是定长记录，每个记录长度16B。那么，i号记录的逻辑地址是多少？（从0开始编号）

每块大小为1KB，定长记录长度为16B，因此一块有 $1KB/16B = 64$ 个记录。则...

逻辑块号 $m = i / 64$

块内地址 $n = (i \% 64) * 16$

i号记录的逻辑地址 = (m, n)

“定长记录的顺序文件支持随机访问”——可以直接算出i号记录对应的逻辑地址

若文件记录不定长，则只能先顺序遍历前*i-1*个记录，因此记录不定长时不支持随机访问

“文件的某种逻辑结构支持随机存取/随机访问”是指：采用这种逻辑结构的文件，可以根据记录号直接算出该记录对应的逻辑地址（逻辑块号，块内地址）。

王道考研/CSKAOYAN.COM

28