

本节内容

# 页面分配策略

# 知识总览



## 页面分配、置换策略

**驻留集**：指请求分页存储管理中给进程分配的物理块的集合。

在采用了虚拟存储技术的系统中，驻留集大小一般小于进程的总大小。

若驻留集太小，会导致缺页频繁，系统要花大量的时间来处理缺页，实际运行效率降低；若驻留集太大，又会导致多道程序并发度下降，资源利用率降低。所以应

考虑一个极端情况，若某进程共有100个页面，则该进程的驻留集大小为100时进程可以全部放入内存，运行期间不可能再发生缺页。若驻留集大小为1，则进程运行期间必定会极频繁地缺页

**固定分配**：操作系统为每个进程分配一组固定数目的物理块，在进程运行期间，驻留集大小不变

**可变分配**：先为每个进程分配一定数目的物理块，在进程运行期间，可根据情况做适当的增加或减少。即，驻留集大小可变

**局部置换**：发生缺页时只能选进程自己的物理块进行置换。

**全局置换**：可以将操作系统保留的空闲物理块分配给缺页进程，也可以将别的进程持有的物理块置换到外存，再分配给缺页进程。

	局部置换	全局置换
固定分配	√	—
可变分配	√	√

全局置换意味着一个进程拥有的物理块数量必然会改变，因此不可能是固定分配

## 页面分配、置换策略

**固定分配局部置换：**系统为每个进程分配一定数量的物理块，在整个运行期间都不改变。若进程在运行中发生缺页，则只能从该进程在内存中的页面中选出一页换出，然后再调入需要的页面。这种策略的缺点是：很难在刚开始就确定应为每个进程分配多少个物理块才算合理。（采用这种策略的系统可以根据进程大小、优先级、或是根据程序员给出的参数来确定为一个进程分配的内存块数）

**可变分配全局置换：**刚开始会为每个进程分配一定数量的物理块。操作系统会保持一个空闲物理块队列。当某进程发生缺页时，从空闲物理块中取出一块分配给该进程；若已无空闲物理块，则可选择一个**未锁定**的页面换出外存，再将该物理块分配给缺页的进程。采用这种策略时，**只要某进程发生缺页，都将获得新的物理块**，仅当空闲物理块用完时，系统才选择一个未锁定的页面调出。被选择调出的页可能是任何一个进程中的页，因此这个**被选中的进程拥有的物理块会减少，缺页率会增加**。

系统会锁定一些页面，这些页面中的内容不能置换出外存（如：重要的内核数据可以设为“锁定”）。当某进程发生缺页时，只允许从该进程未锁定的页面中换出。如果进程在运行中频繁地缺页，系统会为该进程多分配几个物理块，直至该进程缺页率趋势适当程度；反之，如果进程在运行中缺页率特别低，则可适当减少分配给该进程的物理块。

可变分配**全局**置换：只要缺页就给分配新物理块

可变分配**局部**置换：要根据发生**缺页的频率**来动态地增加或减少进程的物理块

## 何时调入页面

主要指空间局部性，即：如果当前访问了某个内存单元，在之后很有可能会接着访问与其相邻的那些内存单元。

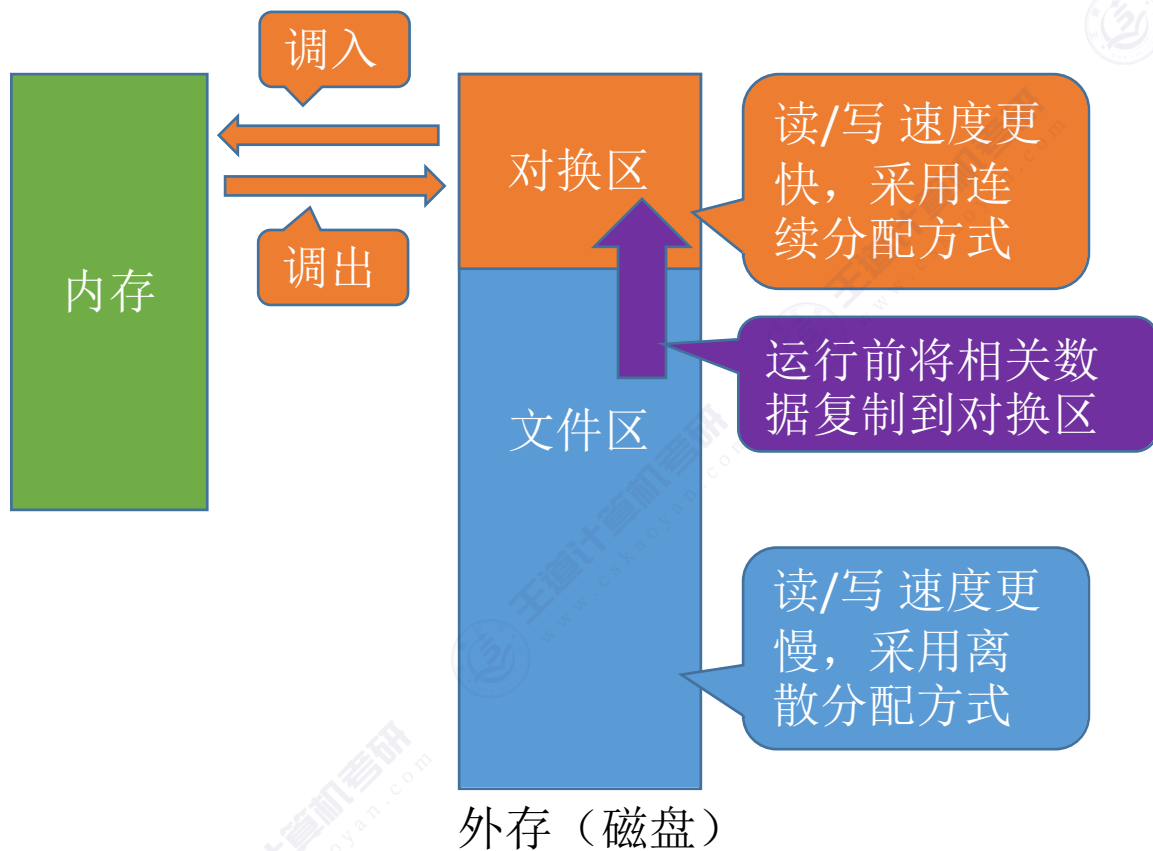
1. 预调页策略：根据局部性原理，一次调入若干个相邻的页面可能比一次调入一个页面更高效。但如果提前调入的页面中大多数都没被访问过，则又是低效的。因此可以预测不久之后可能访问到的页面，将它们预先调入内存，但目前预测成功率只有50%左右。故这种策略**主要用于进程的首次调入**，由程序员指出应该先调入哪些部分。

运行前调入

运行时调入

2. 请求调页策略：进程**在运行期间发现缺页时才将所缺页面调入内存**。由这种策略调入的页面一定会被访问到，但由于每次只能调入一页，而每次调页都要磁盘I/O操作，因此I/O开销较大。

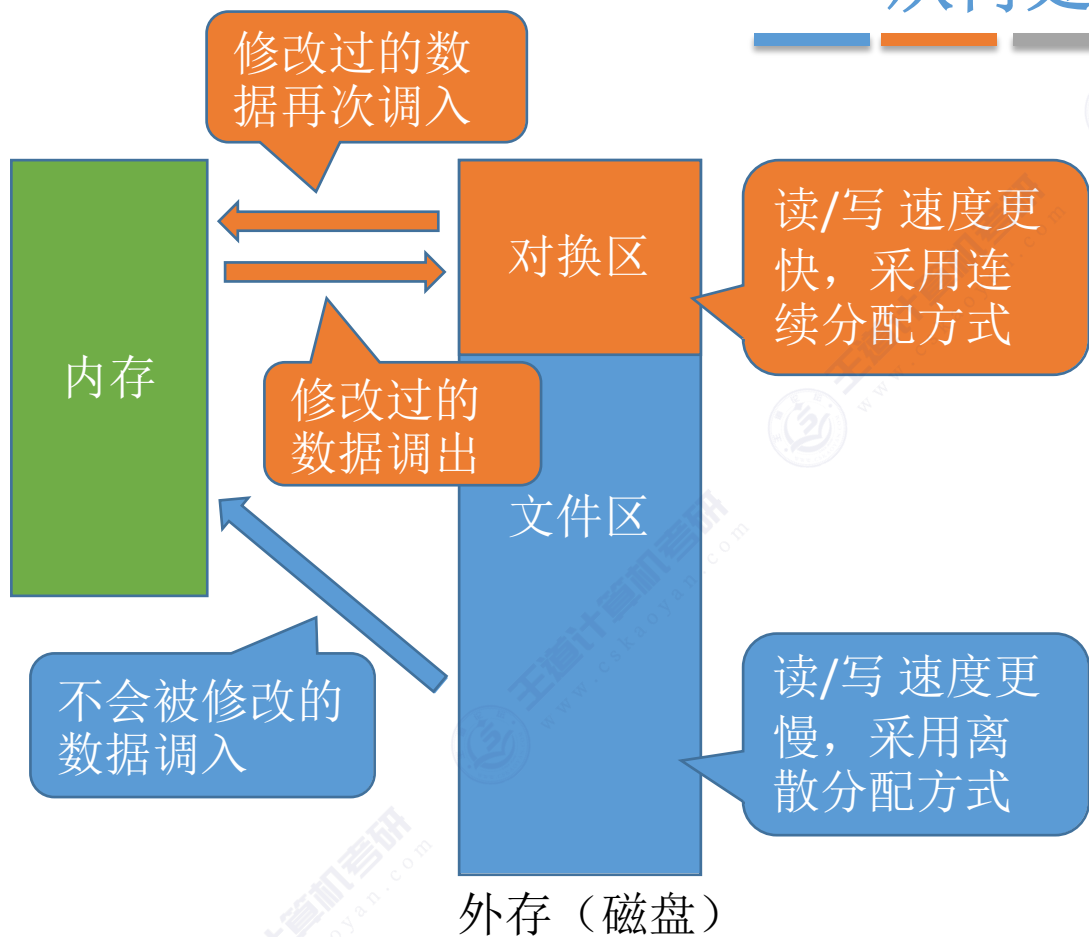
## 从何处调入页面



1. 系统拥有足够的对换区空间：页面的调入、调出都是在内存与对换区之间进行，这样可以保证页面的调入、调出速度很快。在进程运行前，需将进程相关的数据从文件区复制到对换区。

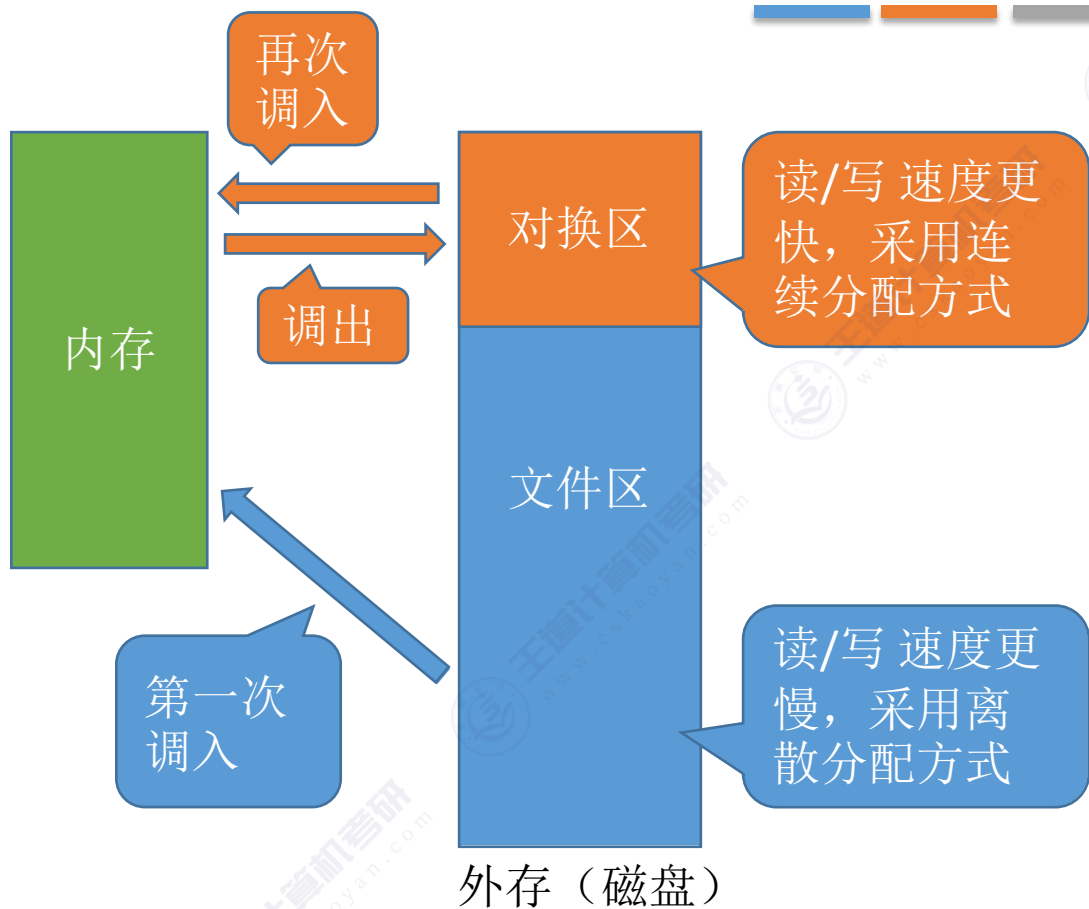


## 从何处调入页面



1. 系统拥有足够的对换区空间：页面的调入、调出都是在内存与对换区之间进行，这样可以保证页面的调入、调出速度很快。在进程运行前，需将进程相关的数据从文件区复制到对换区。
2. 系统缺少足够的对换区空间：凡是不会被修改的数据都直接从文件区调入，由于这些页面不会被修改，因此换出时不必写回磁盘，下次需要时再从文件区调入即可。对于可能被修改的部分，换出时需写回磁盘对换区，下次需要时再从对换区调入。

## 从何处调入页面



1. 系统拥有足够的对换区空间：页面的调入、调出都是在内存与对换区之间进行，这样可以保证页面的调入、调出速度很快。在进程运行前，需将进程相关的数据从文件区复制到对换区。
2. 系统缺少足够的对换区空间：凡是不会被修改的数据都直接从文件区调入，由于这些页面不会被修改，因此换出时不必写回磁盘，下次需要时再从文件区调入即可。对于可能被修改的部分，换出时需写回磁盘对换区，下次需要时再从对换区调入。
3. **UNIX 方式**：运行之前进程有关的数据全部放在文件区，故未使用过的页面，都可从文件区调入。若被使用过的页面需要换出，则写回对换区，下次需要时从对换区调入。



## 抖动（颠簸）现象

刚刚换出的页面马上又要换入内存，刚刚换入的页面马上又要换出外存，这种频繁的页面调度行为称为**抖动**，或**颠簸**。产生抖动的**主要原因**是进程频繁访问的页面数目高于可用的物理块数（**分配给进程的物理块不够**）

为进程分配的物理块太少，会使进程发生抖动现象。为进程分配的物理块太多，又会降低系统整体的并发度，降低某些资源的利用率

为了研究为应该为每个进程分配多少个物理块，Denning 提出了进程“工作集”的概念

## 工作集

**驻留集**：指请求分页存储管理中给进程分配的内存块的集合。

**工作集**：指在某段时间间隔里，进程实际访问页面的集合。

操作系统会根据“窗口尺寸”来算出工作集。例：

某进程的页面访问序列如下，窗口尺寸为4，各时刻的工作集为？

24, 15, 18, 23, 24, 17, 18, 24, 18, 17, 17, 15

工作集：  
24,15,18,23

工作集：  
18,24,17

**工作集大小**可能小于窗口尺寸，实际应用中，操作系统可以统计进程的工作集大小，根据工作集大小给进程分配若干内存块。如：窗口尺寸为5，经过一段时间的监测发现某进程的工作集最大为3，那么说明该进程有很好的局部性，可以给这个进程分配3个以上的内存块即可满足进程的运行需要。

一般来说，**驻留集大小不能小于工作集大小**，否则进程运行过程中将频繁缺页。

拓展：基于局部性原理可知，进程在一段时间内访问的页面与不久之后会访问的页面是有相关性的。因此，可以根据进程近期访问的页面集合（工作集）来设计一种页面置换算法——选择一个不在工作集中的页面进行淘汰。

# 知识回顾与重要考点

## 页面分配策略

### 驻留集

指请求分页存储管理中给进程分配的内存块的集合

固定分配 VS 可变分配：区别在于进程运行期间驻留集大小是否可变

局部置换 VS 全局置换：区别在于发生缺页时是否只能从进程自己的页面中选择一个换出

### 页面分配、置换策略

固定分配局部置换：进程运行前就分配一定数量物理块，缺页时只能换出进程自己的某一页

可变分配全局置换：只要缺页就分配新物理块，可能来自空闲物理块，也可能需换出别的进程页面

可变分配局部置换：频繁缺页的进程，多分配一些物理块；缺页率很低的进程，回收一些物理块。直到缺页率合适

### 何时调入页面

预调页策略：一般用于进程运行前

请求调页策略：进程运行时，发现缺页再调页

### 从何处调页

对换区——采用连续存储方式，速度更快；文件区——采用离散存储方式，速度更慢。

对换区足够大：运行将数据从文件区复制到对换区，之后所有的页面调入、调出都是在内存与对换区之间进行

对换区不够大：不会修改的数据每次都从文件区调入；会修改的数据调出到对换区，需要时再从对换区调入

UNIX方式：第一次使用的页面都从文件区调入；调出的页面都写回对换区，再次使用时从对换区调入

### 抖动（颠簸）现象

页面频繁换入换出的现象。主要原因是分配给进程的物理块不够

### 工作集

在某段时间间隔里，进程实际访问页面的集合。驻留集大小一般不能小于工作集大小



公众号：王道在线



b站：王道计算机教育



抖音：王道计算机考研