

缓存

缓存是用来减少高速设备访问低速设备所需平均时间的组件，文件读写涉及到计算机内存和磁盘，内存操作速度远远大于磁盘，如果每次调用read, write都去直接操作磁盘，一方面速度会被限制，一方面也会降低磁盘使用寿命，因此不管是对磁盘的读操作还是写操作，操作系统都会将数据缓存起来

Page Cache

页缓存（Page Cache）是位于内存和文件之间的缓冲区，它实际上也是一块内存区域，所有的文件IO（包括网络文件）都是直接和页缓存交互，操作系统通过一系列的数据结构，比如inode, address_space, struct page, 实现将一个文件映射到页的级别，这些具体数据结构及之间的关系我们暂且不讨论，只需知道页缓存的存在以及它在文件IO中扮演着重要角色，很大一部分程度上，文件读写的优化就是对页缓存使用的优化

Dirty Page

页缓存对应文件中的一块区域，如果页缓存和对应的文件区域内容不一致，则该页缓存叫做脏页（Dirty Page）。对页缓存进行修改或者新建页缓存，只要没有刷磁盘，都会产生脏页

说明：缓存 本质上是 内存上的区域，内存 是通过页管理的，所以缓存 也叫 page cache，脏页 只是 页缓存的 状态，当写数据时，数据不会立即写入到硬盘，而是写入到页缓存中，这时 页缓存就和硬盘文件内容不一样，该页缓存就叫做脏页。

查看页缓存大小

linux上有两种方式查看页缓存大小，一种是free命令

```
$ free
```

	total	used	free	shared	buffers
cached					
Mem:	20470840	1973416	18497424	164	270208
	1202864				

```
-/+ buffers/cache:      500344    19970496
Swap:                   0          0          0
```

cached那一列就是页缓存大小，单位Byte

另一种是直接查看/proc/meminfo，这里我们只关注两个字段

```
Cached:                1202872 kB
Dirty:                  52 kB
```

Cached是页缓存大小，Dirty是脏页大小

脏页回写参数

Linux有一些参数可以改变操作系统对脏页的回写行为

```
$ sysctl -a 2>/dev/null | grep dirty
vm.dirty_background_ratio = 10
vm.dirty_background_bytes = 0
vm.dirty_ratio = 20
vm.dirty_bytes = 0
vm.dirty_writeback_centisecs = 500
vm.dirty_expire_centisecs = 3000
```

vm.dirty_background_ratio是内存可以填充脏页的百分比，当脏页总大小达到这个比例后，系统后台进程就会开始将脏页刷磁盘

（vm.dirty_background_bytes类似，只不过是字节数来设置）

vm.dirty_ratio是绝对的脏数据限制，内存里的脏数据百分比不能超过这个值。

如果脏数据超过这个数量，新的IO请求将会被阻挡，直到脏数据被写进磁盘

vm.dirty_writeback_centisecs指定多长时间做一次脏数据写回操作，单位为百分之一秒

vm.dirty_expire_centisecs指定脏数据能存活的时间，单位为百分之一秒，比如这里设置为30秒，在操作系统进行写回操作时，如果脏数据在内存中超过30秒时，就会被写回磁盘

这些参数可以通过 `sudo sysctl -w vm.dirty_background_ratio=5` 这样的命令来修改，需要root权限，也可以在root用户下执行 `echo 5 >`

`/proc/sys/vm/dirty_background_ratio` 来修改