

操作系统

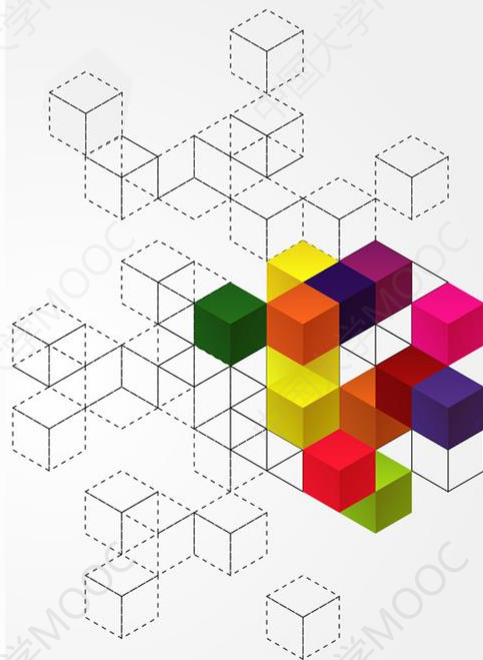
Operating system

胡燕

大连理工大学

零、Linux的IO子系统中的缓存

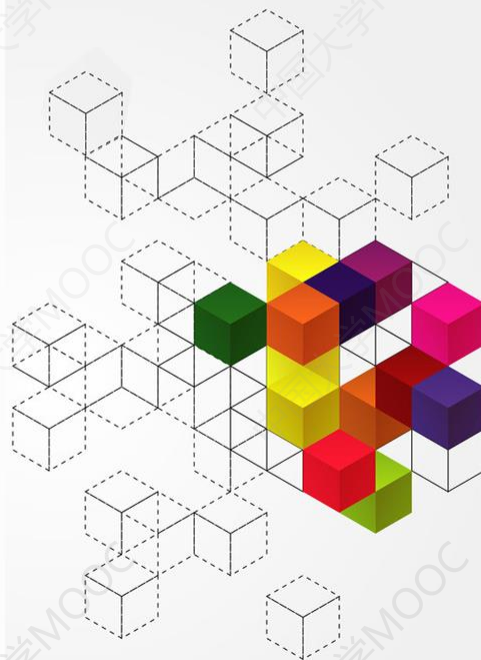
- Linux的IO子系统中，缓冲被重点进行设计，以提升IO子系统性能
- 两类关键缓存：Page Cache和Buffer Cache
 - Page Cache: 以页为单位，缓存文件内容
 - Buffer Cache: 内核为了加速对底层存储介质的访问速度而构建的一层缓存



一、Buffer Cache

二、Page Cache

三、两类Cache演进历史



一、Buffer Cache

- **设立buffer Cache的目的**

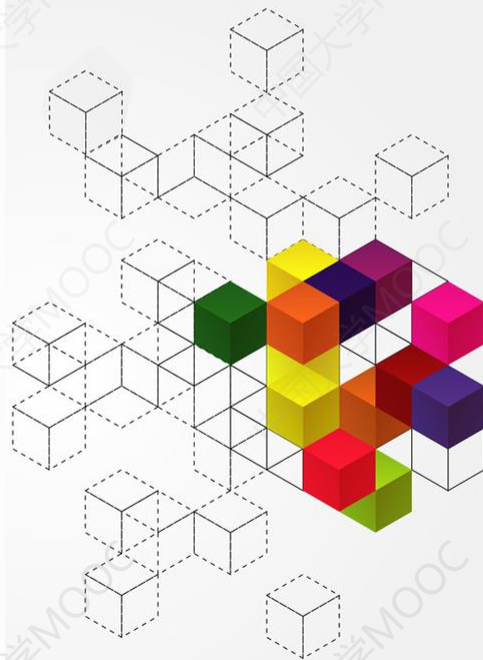
- 在内存中设立磁盘扇区数据的缓存

- **相关背景**

- 磁盘的最小数据单位是扇区，每次读写磁盘都是以扇区为单位进行
 - 如果直接访问磁盘，那么意味着及时用户仅更新某个扇区一个字节的数据，他都必须更新整个扇区数据

- **提升效率的方法**

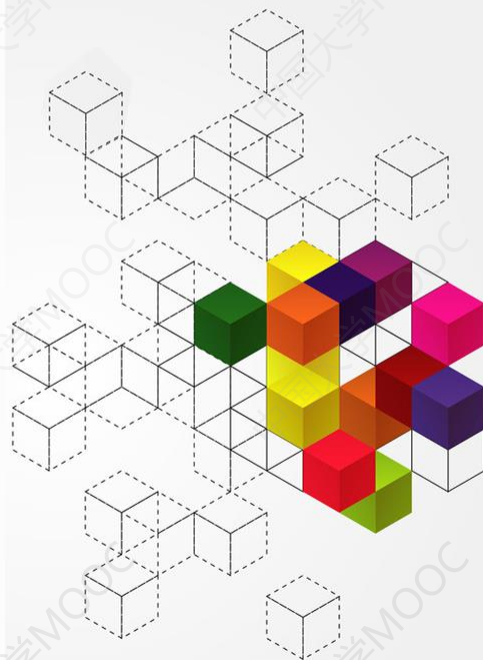
- 为磁盘扇区建立一层缓存，以扇区的整数倍大小构建缓存块



一、Buffer Cache

• 基于Buffer Cache的磁盘扇区数据读写

- 当首次访问某个扇区，在buffer cache中建立新的缓存项
- 此后，对于该扇区的读写请求，直接从内存中读写
- 通过异步方式，将更新后的数据写回对应磁盘扇区

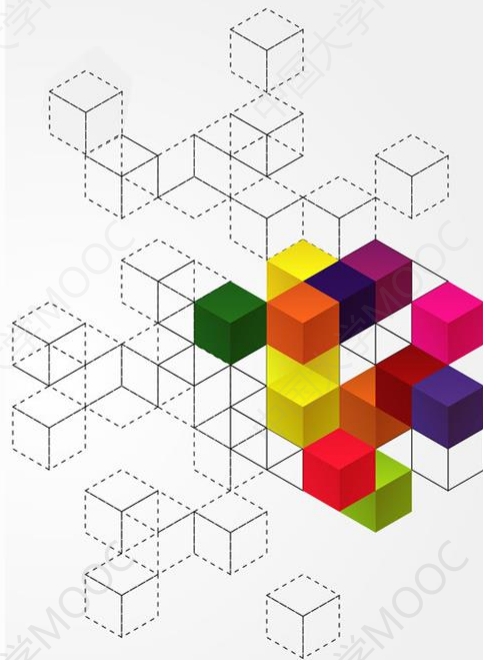


二、Page Cache

• Page Cache

- 缓存在Page Cache中的文件数据，能够更快地被读取
- 进行写入操作时，数据可以在被写入Page Cache后立即返回

Page Cache可以大幅提高上层应用读写文件地整体性能

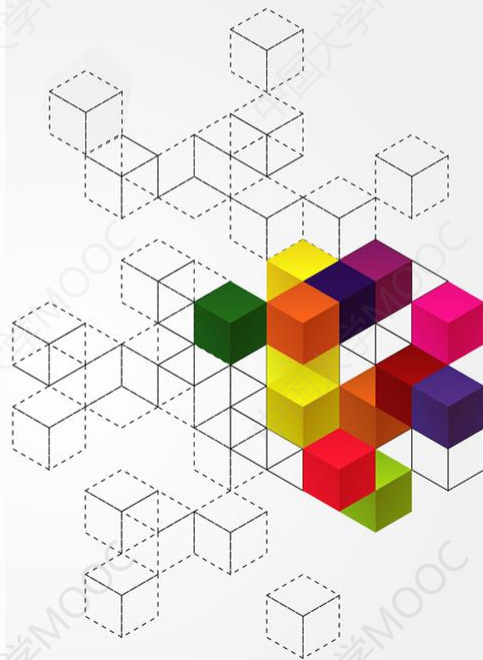


二、Page Cache

- **Page Cache与Buffer Cache的逻辑关系**

- 参考代码：Linux 2.6.18

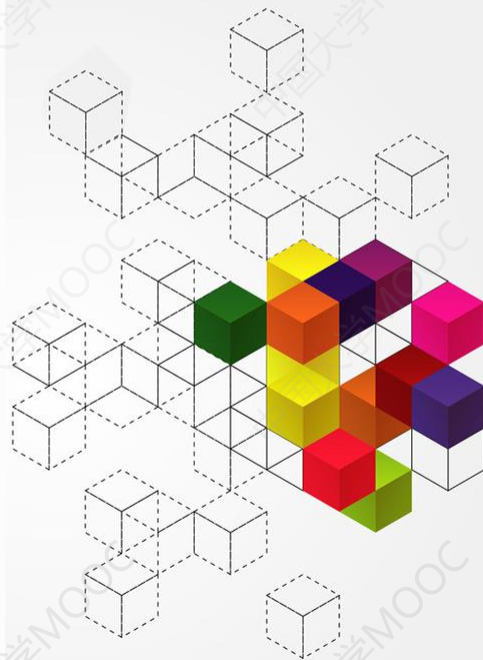
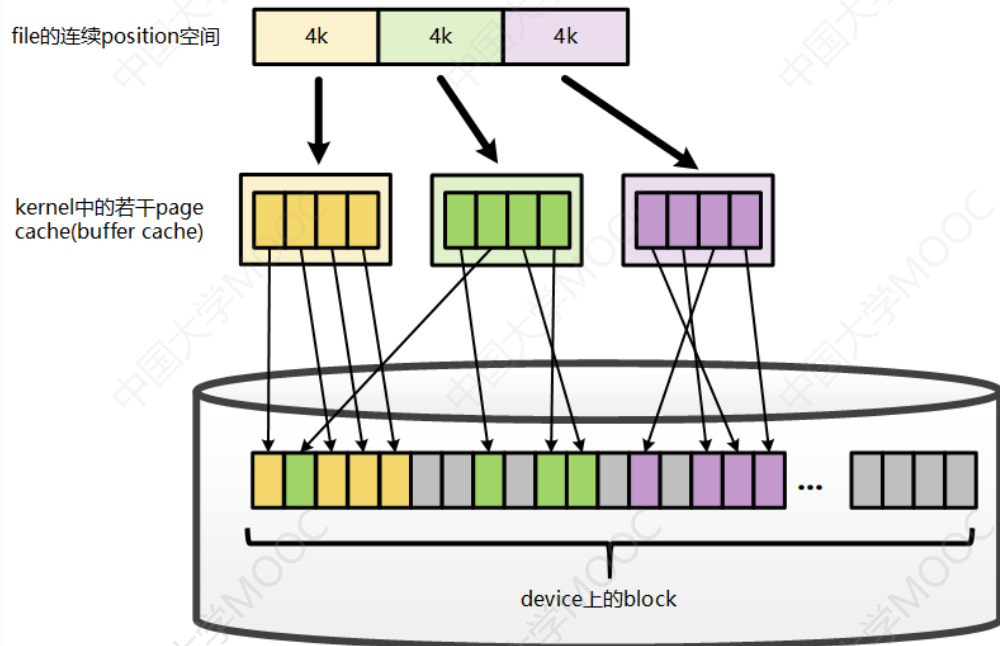
一个作为缓存的页，
对上层，它是某个File的一个Page Cache
对下，它是一个磁盘上的一组Buffer Cache



二、Page Cache

• Page Cache与Buffer Cache的逻辑关系

- 参考代码：Linux 2.6.18



三、两类Cache的演进历史

• 第1阶段：仅有Buffer Cache

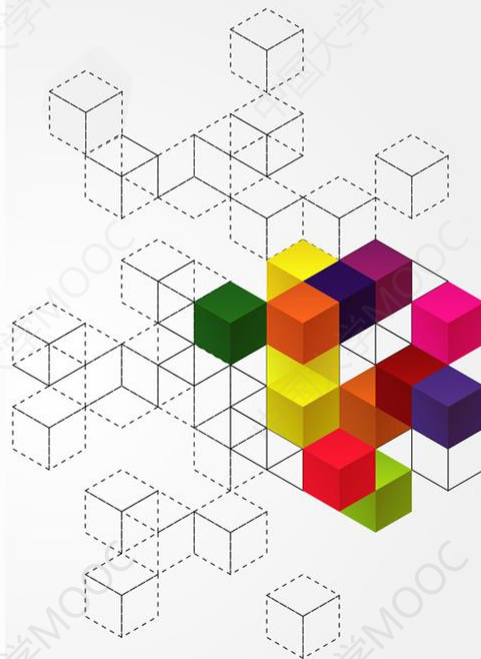
- 参考Linux版本：Linux 0.11

在Linux-0.11的代码中，buffer cache是完全独立的实现

●其中还没有基于page划分内存单元，而是以原始指针的形式出现

●每一个block sector，在kernel内部对应一个独立的buffer cache单元，这个buffer cache单元通过buffer head来描述

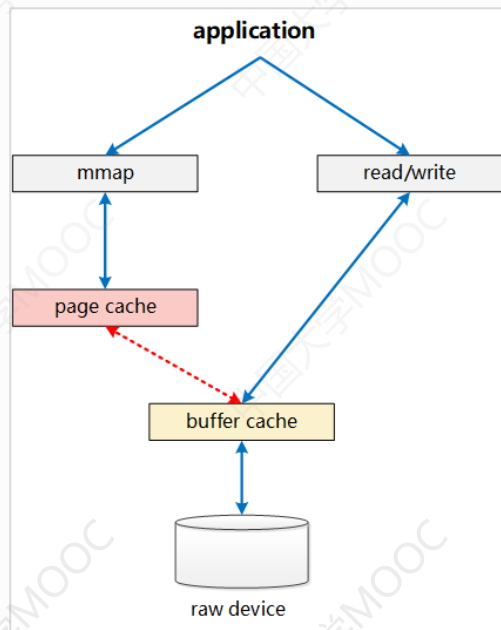
```
68: struct buffer_head {
69:     char * b_data; /* pointer to data block (1024 bytes) */
70:     unsigned long b_blocknr; /* block number */
71:     unsigned short b_dev; /* device (0 = free) */
72:     unsigned char b_uptodate;
73:     unsigned char b_dirt; /* 0-clean,1-dirty */
74:     unsigned char b_count; /* users using this block */
75:     unsigned char b_lock; /* 0 - ok, 1 -locked */
76:     struct task_struct * b_wait;
77:     struct buffer_head * b_prev;
78:     struct buffer_head * b_next;
79:     struct buffer_head * b_prev_free;
80:     struct buffer_head * b_next_free;
81: };
82:
```



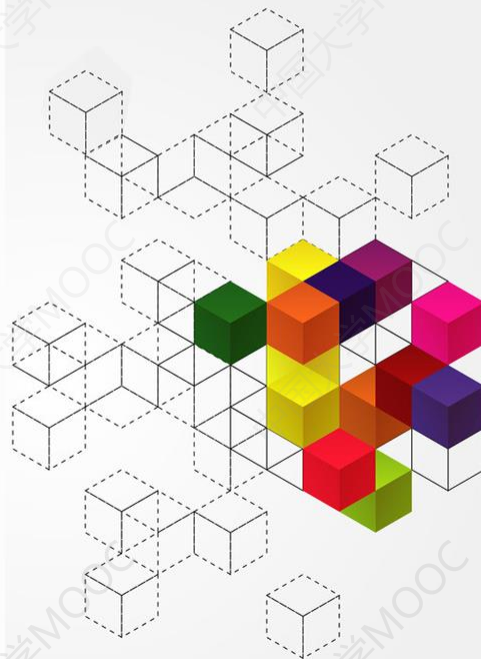
三、两类Cache的演进历史

• 第2阶段: Buffer Cache与Page Cache并存

- 参考Linux版本: Linux 2.2



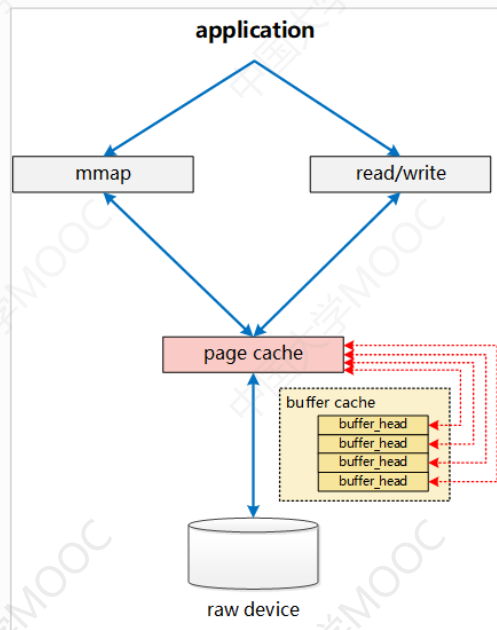
Linux-2.2中, 磁盘读写操作访问的高速缓冲仍然是Buffer Cache。其访问模式与上面Linux-0.11版本的访问逻辑基本类似, 但此时, Buffer Cache已基于page来分配内存, buffer_head内部, 已经有了关于所在page的一些信息



三、两类Cache的演进历史

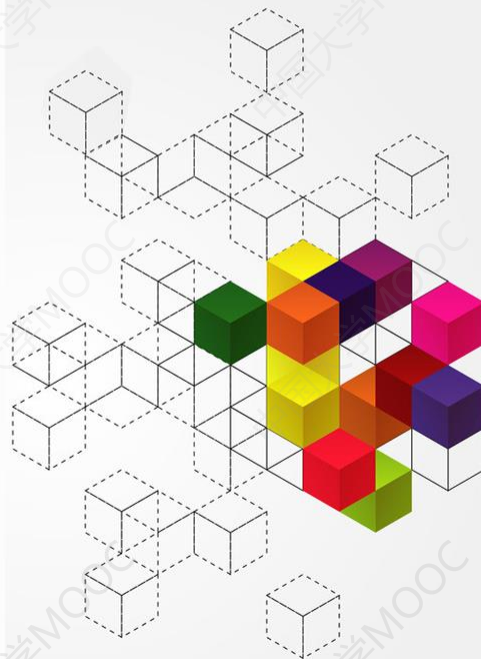
- **第3阶段：Buffer Cache与Page Cache融为一体**

- 参考Linux版本：Linux 2.4.0



Linux-2.4版本中对Page Cache、Buffer Cache的实现进行了融合

- 融合后的Buffer Cache不再以独立的形式存在，Buffer Cache的内容，直接存在于Page Cache中
- 保留了对Buffer Cache的描述符单元：buffer_head



本讲小结

Unix/Linux IO缓冲机制

- ✓ Buffer Cache
- ✓ Page Cache
- ✓ 两类Cache的演进历史

