1.flashcache的ssd分为superblock,metadatablock,cacheblock三大部分

1.1 superblock是创建cachedev后写的相关信息，4KB大小

1.2 metadata是存放描述cacheblock元数据的数据域，4KB大小；其中描述一个cacheblock的元数据有16字节，所以一个metadatablock存放了256个cacheblock的元数据

1.3 cacheblock默认是8个扇区即4KB（此值可以通过参数更改），为了更好管理又设计了一个assoc区域的概念，它包含512个cacheblock，即2MB

以上就是flashcache缓存加速盘的数据结构；

2.内存也要数据结构来管理这些cacheblock，一个cacheblock需要一个18字节的数据结构来描述，因此如果我们的缓存加速盘有500G，则对应需要2.25G的内存来支持

3.flashcache中的cacheblock通过链表来管理使用，一共3个分别是invalid，hotlist，warmlist

3.1 初始状态的cacheblock都是在invalid链表中，表示无数据

3.2 hotlist存放的是最热的数据

3.3 warmlist存放的是降级的数据

优先级 hot MRU > hot LRU > warm MRU > warm LRU

4.当访问数据的请求到达时

4.1 首先在有数据的链表中查找，如果找到：

4.1.1 数据状态是VALID，则将此block状态+1，返回请求

4.1.2 数据是dirty，则将此block状态+1，下刷数据到disk并返回请求

4.2 未找到

4.2.1 invalid有空闲cacheblock，则在invalid链表中找一个cacheblock存放数据，然后将此cacheblock插入到warmlist的LRU

4.2.3 invalid无空闲cacheblock，则执行reclaim策略，有FIFO和LRU两种策略，清除warmlist并且状态为VALID的block，（不选择数据dirty是因为不可能等原数据刷到disk后才使用该cacheblock）存放数据然后将此cacheblock插入到warmlist的LRU

4.3 默认使用的是lru策略，当数据被访问则会提升等级（提升在链表中的位置），删除的时候优先删除warm LRU端

5. 在正常关闭flashcahce之后，有效块和脏块都将在随后的缓存重新加载中保留。在节点崩溃或电源故障后，只有脏缓存块会在随后的缓存重新加载。节点崩溃或电源故障不会导致数据丢失，但会导致缓存丢失有效和非脏缓存块

5.1 脏数据写入磁盘有两个条件

5.1.1 当达到一定时间时，写入磁盘；首先会给脏块一个1标记，当过900s之后标记+1，当达到2之后就写入在此过程中如果有访问读写则要禁用空闲处理机制，将标记重新刷为0

5.1.2 有阈值，当脏块超过脏块阈值时写入磁盘

5.2 当确认要清理某脏块时，会搜索其他脏块合并到一个大IO中一起处理

6. ata trim指令引入，由于ssd不像hdd那样能直接覆盖数据，需要先擦除再写入，所以引入trim指令，它由内核发给ssd主控，告诉ssd哪些地址数据是无效的可以删除，然后执行ssd的GC策略。

7. through 写入时同时写入ssd和hdd，读数据时缓存；删除设备或重启时缓存丢失

Around 写入时不缓存，读数据时缓存到ssd；删除设备或重启时缓存丢失

Back 写入时只写入ssd，通过一些策略将数据刷入hdd，读数据时也缓存；删除设备或重启时缓存保留

Write-only 它是back的一直模式，写入的才缓存，读的数据不缓存；-w参数

8. Ssd在随机读写方面有较大提升，但是顺序读写方面和hdd差别不大，Ssd被顺序写一段时间之后切换为随机写时，带宽会相对较高，IOPS也会很高，然后会慢慢回落到稳定状态；当ssd被随机写一段时间后，当切到顺序后会相对较低，慢慢上升到一个稳定状态