

写了17204字,被8人关注,获得了14个喜欢(/users/1b2925c3117c)

Cache 替换算法之:基本算法

字数857 阅读718 评论0 喜欢1

Cache miss不仅意味着需要从主存获取数据,而且还需要将cache的某一个block替换出去。常用的算法包括FIFO、LRU、RR、Random等

FIFO: First in First out



如上图,不同的色块代表不同的主存数据,按既定的顺序被load到cache中,位于cache中的特定的位置, 当需要被替换出去时,他们也按原来的顺序依次被替换出去。

Round Robin



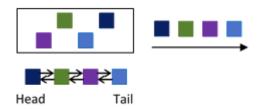
和FIFO相比,RR算法将cache划分成若干个单元,新数据进来时,根据cache单元的位置为顺序,依次将原有数据替换,从结果看,数据被替换出cache 的顺序和进入时的顺序没有必然关系。

Random



真正意义上的随机,你不知道下一次被替换出去的会是哪一个cache block

LRU (Least Recently Used)



按照Cache block被使用的先后顺序组成链表,按最老的数据最先被替换的规则进行替换

MRU(Most Recently Used)和LRU类似,差别在于它是按使用的频度来排序,按最少使用的数据最先被替换出去的规则进行替换。

FIFO、RR和Random算法都没有考虑cache的使用历史信息,而程序的时间和空间局部性都依赖于这些历史信息,因此不少CPU使用了LRU算法。这并不意味着LRU就一定比这些算法强。理论和实验(参考)ligr参考2,参考3)都证明了LRU在某些场景下miss率比其他三种都高,比如访问数组{a, b, c, d, e}命中到同一个组时,Miss的概率非常高,在这种情况下LRU并不比FIFO、RR好多少,而明显弱于Random方式。

LRU算法没有利用访问次数这个重要信息,在处理文件扫描这种空间局限性较弱的场景时就显得有点力不从心,访问的数据量越大,miss率越高,因此LRU出现了改良算法:LRFU和LRU-K

LRFU是LRU和LFU (Least Frequently Used)两者的结合,优先替换访问次数少的数据。LRU-K记录页面访问的次数,K为最大值,实现方法是:先从使用次数为1的页面中根据LRU查找页面进行替换,如果没有1的页面则查找访问次数为2的页面,直到K为止。当K=1时,等效于LRU。现实中LRU-2比较常用。

LUR-K使用多个优先级队列,算法复杂度为O(Log2N),而LRU、FIFO这类算法的复杂度位O(1),因此采用LRU-K算法时需要耗费更多的cycle,同时,多个队列使用互相独立的空间,消耗的空间也较多,因此出现了针对LRU-2优化的2Q算法,其初衷是保证LRU-2效果不变的前提下,减小时间和空间的消耗。

2Q算法有两种实现方式: Simplified 2Q和Full version 2Q,下节详细介绍

参考1Alan Jay Smith [Sep. 1982].Cache Memories. Sep. 1982] ACM Computing Surveys Volume 14 Issue 3

参考 2 : GURURAJ S. RAO [Jul. 1978]. Performance Analysis of Cache Memories. Journal the ACM (JACM) Vol 25, No 3.

参考3: Jan Reineke, Daniel Grund, Christoph Berg, andReinhard Wilhelm[Nov 2007]. Timing Predictabilityof Cache Replacement Policies. Real-Time Systems. Volume 37, Number 2.

◆ 推荐拓展阅读 (/sign_in)

© 著作权归作者所有

如果觉得我的文章对您有用,请随意打赏。您的支持将鼓励我继续创作!

¥打赏支持



豆浆/口及农厅记(/ 31911_111)

被以下专题收入,发现更多相似内容:



Linux **内存管理** (/collection/433654b6371f)

Linux memory management , talking about kernel、userspace and haredware (/collection/433654b6371f) · 30人关注

★添加关注 (/sign_in)