

7.9 假设在 3000 次访存中,第一级 Cache 不命中 110 次,第二级 Cache 不命中 55 次。试问:在这种情况下,该 Cache 系统的局部不命中率和全局不命中率各是多少?

第一级cache不命中率(全局和局部)是110/3000,即3.67%;

第二级cache的局部不命中率是55/110,即50%;

第二级cache的全局不命中率是55/3000,即1.83%。

7.10 给定以下的假设,试计算直接映像 Cache 和两路组相联 Cache 的平均访问时间以及 CPU 的性能。由计算结果能得出什么结论?

(1) 理想 Cache 情况下的 CPI 为 2.0,时钟周期为 2ns,平均每条指令访存 1.2 次;

(2) 两者 Cache 容量均为 64KB,块大小都是 32 字节;

(3) 组相联 Cache 中的多路选择器使 CPU 的时钟周期增加了 10%;

(4) 这两种 Cache 的不命中开销都是 80ns;

(5) 命中时间为 1 个时钟周期;

(6) 64KB 直接映像 Cache 的不命中率为 1.4%,64KB 两路组相联 Cache 的不命中率为 1.0%。

1. 平均访存时间

平均访存时间=命中时间+失效率×失效开销

平均访问时间1-路= $1 \times 2.0 + 1.4\% \times 80 = 3.12\text{ns}$

平均访问时间2-路= $1 \times 2.0 \times (1 + 10\%) + 1.0\% \times 80 = 3.0\text{ns}$

两路组相联的平均访存时间比较低

2. CPU时间

CPU时间 = (CPU执行周期数 + 访存次数 × 不命中率 × 不命中开销) × 时钟周期时间

$$\begin{aligned} \text{CPU时间} &= IC \times \left(CPI_{\text{execution}} + \frac{\text{访存次数}}{\text{指令数}} \times \text{不命中率} \times \text{不命中开销} \right) \times \text{时钟周期时间} \\ &= IC \times (CPI_{\text{execution}} + \text{每条指令的平均访存次数} \times \text{不命中率} \\ &\quad \times \text{不命中开销}) \times \text{时钟周期时间} \end{aligned}$$

$$\text{CPU time 1-way} = \text{IC}(2.0 \times 2 + 1.2 \times 0.014 \times 80) = 5.344\text{IC}$$

$$\text{CPU time 2-way} = \text{IC}(2.2 \times 2 + 1.2 \times 0.01 \times 80) = 5.36\text{IC}$$

$$\text{相对性能比: } 5.36 / 5.344 = 1.003$$

直接映象的访存时间是两路组相联的1.04倍,

两路组相联的平均CPU时间是直接映象的1.003倍。

因此这里选择直接映象。

7.11 在伪相联中,假设在直接映像位置没有发现匹配,而在另一个位置才找到数据(伪命中)时,不对这两个位置的数据进行交换。这时只需要1个额外的周期。假设不命中开销为50个时钟周期,2KB直接映像Cache的不命中率为9.8%,两路组相联的不命中率为7.6%;128KB直接映像Cache的不命中率为1.0%,两路组相联的不命中率为0.7%。

(1) 推导出平均访存时间的公式。

(2) 利用(1)中得到的公式,对于2KB Cache和128KB Cache,计算伪相联的平均访存时间。

(1) 推导平均访存时间公式

- 不管作了何种改进,失效开销相同。不管是否交换内容,在同一“伪相联”组中的两块都是用同一个索引得到的,因此失效率相同,即:
失效率_{伪相联} = 失效率_{2路}。
- 伪相联cache的命中时间等于直接映像cache的命中时间加上伪相联查找过程中的命中时间*该命中所需的额外开销。
- 命中时间_{伪相联} = 命中时间_{1路} + 伪命中率_{伪相联} × 1
- 交换或不交换内容,伪相联的命中率都是由于在第一次失效时,将地址取反,再在第二次查找带来的。
- 因此 伪命中率_{伪相联} = 命中率_{2路} - 命中率_{1路} = (1 - 失效率_{2路}) - (1 - 失效率_{1路}) = 失效率_{1路} - 失效率_{2路}。交换内容需要增加伪相联的额外开销。
- 平均访存时间_{伪相联} = 命中时间_{1路} + (失效率_{1路} - 失效率_{2路}) × 1
- + 失效率_{2路} × 失效开销_{1路}

(2) 计算平均访存时间(假设命中时间_{1路}为1个时钟周期)

- 将题设中的数据带入计算,得到:
- 平均访存时间_{2Kb} = 1 + (0.098 - 0.076) * 1 + (0.076 * 50) = 4.822
- 平均访存时间_{128Kb} = 1 + (0.010 - 0.007) * 1 + (0.007 * 50) = 1.353

7.14 假设一台计算机具有以下特性：

- (1) 95%的访存在 Cache 中命中；
- (2) 块大小为两个字,且不命中时整个块被调入；
- (3) CPU 发出访存请求的速率为 10^9 字/秒；
- (4) 25%的访存为写访问；
- (5) 存储器的最大流量为 10^9 字/秒(包括读和写)；
- (6) 主存每次只能读或写一个字；
- (7) 在任何时候,Cache 中有 30%的块被修改过；
- (8) 写不命中时,Cache 采用按写分配法。

现欲给该计算机增添一台外设,为此首先想知道主存的带宽已用了多少。试对于以下两种情况计算主存带宽的平均使用比例。

- (1) 写直达 Cache；
- (2) 写回法 Cache。

• 采用按写分配

(1) 写直达cache访问命中，有两种情况：

读命中，不访问主存；

写命中，更新cache和主存，访问主存一次。

访问失效，有两种情况：

读失效，将主存中的块调入cache中，访问主存两次；

写失效，将要写的块调入cache，访问主存两次，再将修改的数据写入cache和主存，访问主存一次，共三次。上述分析如下表所示。

访问命中	访问类型	频率	访存次数
Y	读	$95\% \times 75\% = 71.3\%$	0
Y	写	$95\% \times 25\% = 23.8\%$	1
N	读	$5\% \times 75\% = 3.8\%$	2
N	写	$5\% \times 25\% = 1.3\%$	3

一次访存请求最后真正的平均访存次数

$$= (71.3\% \times 0) + (23.8\% \times 1) + (3.8\% \times 2) + (1.3\% \times 3) = 0.35$$

$$\text{已用带宽} = 0.35 \times 10^9 / 10^9 = 35.0\%$$

- (2) 写回法cache访问命中,有两种情况:

读命中, 不访问主存;

写命中, 不访问主存。采用写回法, 只有当修改的cache块被换出时, 才写入主存;

访问失效,有一个块将被换出, 这也有两种情况:

如果被替换的块没有修改过, 将主存中的块调入cache块中, 访问主存两次;

如果被替换的块修改过, 则首先将修改的块写入主存, 需要访问主存两次; 然后将主存中的块调入cache块中, 需要访问主存两次, 共四次访问主存。

访问命中	块为脏	频率	访存次数
Y	N	$95\% \times 70\% = 66.5\%$	0
Y	Y	$95\% \times 30\% = 28.5\%$	0
N	N	$5\% \times 70\% = 3.5\%$	2
N	Y	$5\% \times 30\% = 1.5\%$	4

所以:

一次访存请求最后真正的平均访存次数

$$= 66.5\% \times 0 + 28.5\% \times 0 + 3.5\% \times 2 + 1.5\% \times 4 = 0.13$$

$$\text{已用带宽} = 0.13 \times 10^9 / 10^9 = 13\%$$