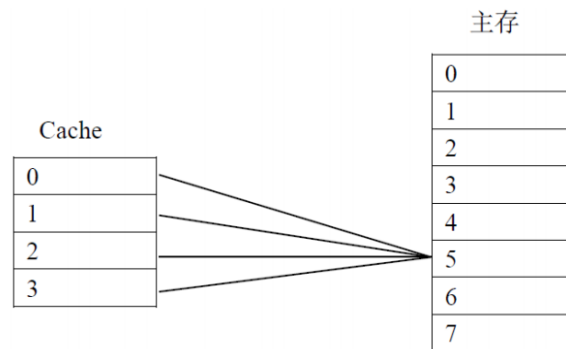


计算机系统结构作业二

7.4 设有一个“Cache-主存”层次。Cache 为 4 块，主存为 8 块，试分别对于以下三种情况，画出其映像关系示意图，并计算访存地址为 5 时的索引。(1)：全相连；(2)：组相连。每组两块；(3)：直接映像。

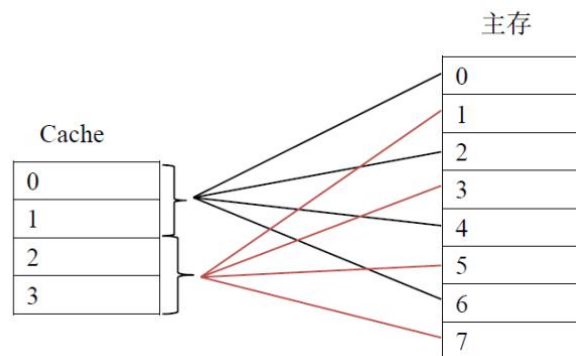
解答：块地址 $(5)_{10} = (101)_2$

(1) 全相连



全相连中没有索引。

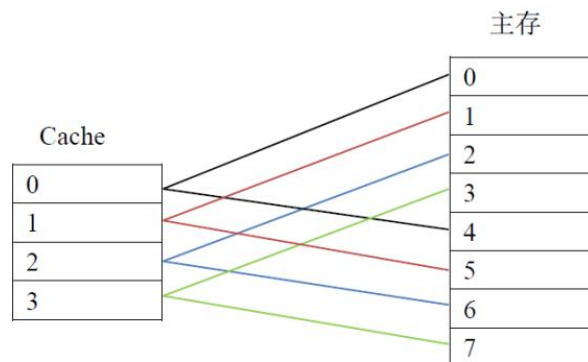
(2) 组相连



$$\text{Index} = (5)_{10} \bmod (2)_{10} = (1)_{10} = (1)_2$$

2 个 cache 块组，只需 1 位作为索引

(3) 直接映像



$$\text{Index} = (5)_{10} \bmod (4)_{10} = (1)_{10} = (01)_2$$

4 个 cache 块，需低 2 位作为索引

PS: 很多同学认为全相连中索引可取 0~3，是错误的!! 实际上全相连因为可以映射到任意位置，所以不需要索引~

7.8 假设对指令 Cache 的访问占全部访问的 75%; 而对数据 Cache 的访问占全部访问的 25%。Cache 的命中时间为 1 个时钟周期，失效开销为 50 个时钟周期，在混合 Cache 中一次 load 或 store 操作访问 Cache 的命中时间都要增加一个时钟周期，32KB 的指令 Cache 的失效率为 0.39%，32KB 的数据 Cache 的失效率为 4.82%，64KB 的混合 Cache 的失效率为 1.35%。又假设采用写直达策略，且有一个写缓冲器，并且忽略写缓冲器引起的等待。试问指令 Cache 和数据 Cache 容量均为 32KB 的分离 Cache 和容量为 64KB 的混合 Cache 相比，哪种 Cache 的失效率更低？两种情况下平均访存时间各是多少？

解答:

(1) 分离 Cache 失效率:

$$75\% \times 0.39\% + 25\% \times 4.82\% = 1.4975\%$$

混合 Cache 失效率:

$$1.35\%$$

所以混合 Cache 的失效率更低

(2) 分离 Cache 的平均访存时间:

$$75\% \times (1 + 0.39\% \times 50) + 25\% \times (1 + 4.82\% \times 50) = 1.74875$$

混合 Cache 的平均访存时间:

$$75\% \times (1 + 1.35\% \times 50) + 25\% \times (1 + 1 + 1.35\% \times 50) = 1.925$$

PS: 平均访存时间 = 命中时间 + 不命中率 × 不命中开销

在计算混合 Cache 的平均访存时间时，因为题目中指出“一次 load 或 store 操作访问 Cache 的命中时间都要增加一个时钟周期”，这里有的同学将其作为对数据 Cache 的访问，因此只在 25% 的指令里增加了命中时间。而有的同学则认为全部指令都增加了命中时间，毕竟如果指令 Cache 不命中时也是需要 load 操作，因此结果如下：

$$2 + 1.35\% \times 50 = 2.675$$

助教想说毕竟题目也没有完全说清楚，所以两者都算对，重要的是大家要理解如何计算平均访存时间的公式!!

7.10 给定以下的假设，试计算直接映象 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 的失效开销都是 80ns；

(5) 命中时间为 1 个时钟周期；

(6) 64KB 直接映象 Cache 的失效率为 1.4%，64KB 两路组相联 Cache 的失效率为 1.0%。

解答：

(1) Cache 的平均访问时间

直接映像 $T_1 = 1 \times 2.0\text{ns} + 1.4\% \times 80\text{ns} = 3.12\text{ ns}$;

两路组相连 $T_1 = 1.1 \times 2.0\text{ns} + 1.0\% \times 80\text{ns} = 3.0\text{ ns}$;

因此，两路组相连的平均访问时间比较低

(2) CPU 的性能

直接映像 CPU 时间 = $IC \times (2.0 \times 2 + 1.2 \times 1.4\% \times 80) = 5.344 \times IC$;

两路组相连 CPU 时间 = $IC \times (2.0 \times 1.1 \times 2 + 1.2 \times 1.0\% \times 80) = 5.36 \times IC$;

直接映射的 CPU 性能比较高。这是因为在两路组相连的情况下，虽然不命中次数减少了，但所有指令的时钟周期时间都增加了 10%。由于 CPU 时间是进行评价的基准，而且直接映像的 Cache 的实现更为简单，所以本题中直接映射 Cache 是较好的选择

PS：本题与教材 204 页的例 7.2 几乎完全相同。主要就是两个公式的运用：

平均访存时间 = 命中时间 + 不命中率 \times 不命中开销

CPU 时间 = $IC \times (CPI \times \text{时钟周期时间} + \text{每条指令的平均访存次数} \times \text{不命中率} \times \text{不命中开销} \times \text{时钟周期时间})$

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，且失效时整个块被调入；
- (3) CPU 发出访存请求的速率为 10^9 字/s；
- (4) 25% 的访存为写访问；
- (5) 存储器的最大流量为 10^9 字/s（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) 写失效时，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频带已用了多少。试对于以下两种情况计算主存频带的平均使用比例。

- (1) 写直达 Cache；
- (2) 写回法 Cache。

解答：

(1) 写直达 Cache

操作	事件频率	每次字数	流量	说明
读命中	75%*95%	0	0	
读失效, 写回	75%*5%*30%	0	0	写直达法不需要写回
读失效, 读入	75%*5%	2	$10^9 * 75\% * 5\% * 2$	每次读 1 块
写命中	25%*95%	1	$10^9 * 25\% * 95\%$	每次写 1 字
写失效, 按写分配, 写回	25%*5%*30%	0	0	写直达法不需要写回
写失效, 按写分配, 读入	25%*5%	2	$10^9 * 25\% * 5\% * 2$	每次读 1 块
写失效, 按写分配读入后再写	25%*5%	1	$10^9 * 25\% * 5\% * 1$	每次写 1 字

因此, 主存频带的平均使用比例:

$$10^9 \times [75\% \times 5\% \times 2 + 95\% \times 5\% \times 1 + 25\% \times 5\% \times (2+1)] \div 10^9 = 35\%$$

(2) 写回法 Cache

操作	事件频率	每次字数	流量	说明
读命中	75%*95%	0	0	
读失效, 写回	75%*5%*30%	2	$10^9 * 75\% * 5\% * 30\% * 2$	整块写回
读失效, 读入	75%*5%	2	$10^9 * 75\% * 5\% * 2$	每次读 1 块
写命中	25%*95%	0	0	写回法不访问主存
写失效, 按写分配, 写回	25%*5%*30%	2	$10^9 * 25\% * 5\% * 30\% * 2$	整块写回
写失效, 按写分配, 读入	25%*5%	2	$10^9 * 25\% * 5\% * 2$	每次读 1 块
写失效, 按写分配读入后再写	25%*5%	0	0	写回法不访问主存

因此, 主存频带的平均使用比例:

$$10^9 \times [75\% \times 5\% \times (2+2 \times 30\%) + 25\% \times 5\% \times (2+2 \times 30\%)] \div 10^9 = 13\%$$

PS: 本题的关键在于对写直达、写回和写分配的理解。写直达是在每次写操作的时候直接对内存进行更改。因为知道更改的是那个字, 所以按照题目假设, 只进行 1 次写即可。但是对于读或写不命中时, 将内存中的数据读入 Cache 则需要将整个块读入, 即 2 次访存操作。

写回则是当脏块替换出去时才对内存进行写操作, 且需要进行整个块的操作, 因为不知道脏块中的哪个字发生了变动。