

# Homework 06

## 1. (改编自 COD\_CH, P337, 5.2)

现对一个64位存储器进行访问，表 1\_1 的第一列是以字地址形式给出的访问地址顺序。（假设字的大小为64位）

- 1) 假设存储器的 *Cache* 共有 16 个基本块，每个块大小为 1 个字。请参考表 1\_1 的形式，给出这些访问地址在 *Cache* 中对应的：二进制字地址、索引、标签和访问的命中情况。
- 2) 假设存储器的 *Cache* 共有 8 个基本块，每个块大小为 2 个字。请参考表 1\_1 的形式，给出这些访问地址在 *Cache* 中对应的：二进制字地址、索引、标签和访问的命中情况。
- (*Cache* 最初为空，替换策略采用 *LRU*，索引、标签给出二进制形式即可)

Word Address	Binary Address	index		Tag		Hit/Miss	
		(1)	(2)	(1)	(2)	(1)	(2)
0x03	00000011	0011	001	0000	0000	M	M
0xb4	10110100	0100	010	1011	1011	M	M
0x2b	00101011	1011	101	0010	0010	M	M
0x02	00000010	0010	001	0000	0000	M	H
0xbf	10111111	1111	111	1011	1011	M	M
0x58	01011000	1000	100	0101	0101	M	M
0xbe	10111110	1110	111	1011	1011	M	H
0x0e	00001110	1110	111	0000	0000	M	M
0xb5	10110101	0101	010	1011	1011	M	H
0x2c	00101100	1100	110	0010	0010	M	M
0xba	10111010	1010	101	1011	1011	M	M
0xfd	11111101	1101	110	1111	1111	M	M

表 1\_1

## 2. (改编自 COD\_CH, P339, 5.7)

考虑以下的程序和 *Cache* 行为：

每 1000 条指令的 数据读次数	每 1000 条指令的 数据写次数	指令 <i>Cache</i> 失效率	数据 <i>Cache</i> 失效率	块大小 (字节)
250	100	0.30%	2%	64

\*存储器位宽为 32-bit，带宽单位采用: bytes/cycle

- 1) 假设一个带有写直达、写分配 *Cache* 的 *CPU* 实现了 2 的 *CPI*，那么 *RAM* 和 *Cache* 之间的读写带宽（用每个周期的字节数进行测量）是多少？（假设每个失效会生成一个块的请求）
- 2) 对于一个写回、写分配 *Cache* 来说，假设替换出的数据 *Cache* 块中有 30% 是脏块，那么为了实现 *CPI* 为 2，读写带宽需要达到多少？
- 1) 读带宽： $0.5 \times 64 \times 0.003 + \frac{1}{4} \times 0.5 \times 0.02 \times 64 + \frac{1}{16} \times 0.5 \times 0.02 \times 64 = 0.32 \text{ bytes/cycle}$   
写带宽： $0.5 \times 0.1 \times 4 = 0.2 \text{ bytes/cycle}$
- 2) 读带宽：同上 0.32 bytes/cycle  
写带宽： $0.5 \times (0.25 + 0.1) \times 0.02 \times 64 \times 0.3 = 0.0672 \text{ byte/cycle}$

3. 假设 *CPU* 执行某段程序时，共访问 *Cache* 命中 2000 次，访问主存 50 次。已知 *Cache* 存取周期为 50ns，主存存取周期为 200ns。求 *Cache*—主存系统的命中率和平均访问时间。

命中率： $\frac{2000}{2000+50} \approx 0.976$

平均访问时间： $\frac{50 \times 2000 + 250 \times 50}{2050} \text{ ns} \approx 54.8 \text{ ns}$

4. (改编自 COD\_CH, P341, 5.12)

多级 Cache 是一种重要的技术，可以克服在一级 Cache 提供的有限空间的同时仍然保持速度。考虑具有一下参数的处理器：

无内存停顿 的基本CPI	处理器速度	主存访问 时间	每条指令的 L1 Cache 失效率	L2 直接映射 Cache 速度	L2 直接映射 Cache 全局失效率	L2 八路组 相联 Cache 速度	L2 八路组 相联 Cache 全局失效率
1.5	2GHz	100ns	7%	12 cycles	3.5%	28 cycles	1.5%

\*L1 Cache 失效率是针对每条指令而言的。假设 L1 Cache 的总失效数量（包括指令和数据）为总指令数的 7%

1) 分别计算在下列情况下的处理器 CPI:

- 仅有 L1 Cache
- 使用 L2 直接映射 Cache
- 使用 L2 八路组相联 Cache

2) 假设处理器采用 L2 直接映射 Cache，设计人员希望添加一个 L3 Cache，该 Cache 访问时间为 50 个时钟周期，并且具有 13% 的失效率，请计算此处理器的 CPI。

3) 在较老的处理器中，例如 Intel Pentium 或 Alpha 21264，L2 Cache 在主处理器和 L1 Cache 的外部（位于不同芯片上）。虽然这种做法使得大型 L2 Cache 成为可能，但是访问 L2 Cache 的延迟也变得很高，并且因为 L2 Cache 以较低的频率运行，所以带宽通常也很低。假设 512KiB 的片外 L2 Cache 的失效率为 4%，如果每增加一个额外的 512KiB 片外 L2 Cache 能够降低 0.7% 的失效率，并且片外 L2 Cache 的总访问时间为 50 个时钟周期，那么片外 L2 Cache 容量必须多大才能与第一问列出的 L2 直接映射 Cache 的性能相匹配？

1).  $T = \frac{1}{2GHz} = 0.5PS$       主存访问所需周期:  $\frac{100ns}{0.5PS} = 200.$

仅有 L1 Cache:  $1.5 + 0.07 \times 200 = 15.5.$

、L2 直接映射:  $1.5 + 0.07 \times (12 + 0.035 \times 200) = 2.83$

、L2 八路组联:  $1.5 + 0.07 \times (12 + 0.35 \times (28 + 0.015 \times 200)) = 3.67$

2)  $1.5 + 0.07 \times (12 + 0.035 \times (50 + 0.13 \times 200)) = 2.5262$

3)  $1.5 + 0.07 \times (50 + 200x) < 2.83$        $x < 0.155$

∴ 没有满足条件的容量.