

5.8

解:

(1) 64B统一Cache的失效率: 1.35%

32KB + 32KB分离Cache的总体失效率:

取指比例 \times 指令失效率 + 数据比例 \times 数据失效率

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

$$= 1.4975\%$$

所以, 统一Cache结构的失效率低于分离Cache结构

(2) 平均访存时间 = 取指比例 \times (指令命中时间 + 指令失效率 \times 失效开销) + 数据比例 \times (数据命中时间 + 数据失效率 \times 失效开销)

$$\text{分离Cache的平均访存时间} = 75\% \times (1 + 0.39\% \times 50) + 25\% \times (1 + 4.82\% \times 50) =$$

$$75\% \times 1.195 + 25\% \times 3.41 = 0.89625 + 0.8525 = 1.74875$$

$$\text{统一Cache的平均访存时间} = 75\% \times (1 + 0.39\% \times 50) + 25\% \times (1 + 1 + 4.82\% \times 50) =$$

$$75\% \times 1.195 + 25\% \times 4.41 = 0.89625 + 1.1025 = 1.99875$$

5.9

解:

第一级cache的局部 (和全局) 不命中率是 $110/3000 \times 100\% = 3.67\%$

第二级cache的局部不命中率是 $55/110 \times 100\% = 50\%$

第二级cache的全局不命中率是 $55/3000 \times 100\% = 1.83\%$

5.10

解:

平均访存时间 = 命中时间 + 不命中率 \times 不命中开销

$$\text{直接映像平均访存时间} = 1 \times 2\text{ns} + 1.4\% \times 80\text{ns} = 3.12\text{ns}$$

$$\text{两路组相联平均访存时间} = 1 \times 2\text{ns} \times (1 + 10\%) + 1.0\% \times 80 = 3.00\text{ns}$$

$$\begin{aligned} \text{CPU时间} &= \text{IC} \times (\text{CPI执行时间} + \text{访存次数} / \text{指令数} \times \text{不命中率} \times \text{不命中开销}) \times \text{时钟周期时间} \\ &= \text{IC} \times (\text{CPI执行时间} + \text{每条指令平均访存次数} \times \text{不命中率} \times \text{不命中开销}) \times \text{时钟周期时间} \end{aligned}$$

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

$$\text{两路组相联平均访寸时间} = \text{IC} \times (2.0 \times 2\text{ns} \times 1.10 + 1.2 \times 1.0\% \times 80\text{ns}) = 5.361\text{IC}$$

结论: 两路组相联的平均访存时间比较低; 而直接映像的Cache性能好一点, 综合来说, 选择直接映像会更好

5.11

解:

(1) 两种方法的失效开销是相同的, 不命中率也是相同的。

平均访存时间 = 直接映像命中时间 + (直接相联不命中率 - 两路组相联不命中率) \times 1 + 两路组相联不命中率 \times 直接映像不命中开销

$$(2) \quad 2\text{KB平均访存时间} = 1 + (9.8\% - 7.6\%) \times 1 + (7.6\% \times 50) = 4.822$$

$$128\text{KB平均访存时间} = 1 + (1.0\% - 0.7\%) \times 1 + (0.7\% \times 50) = 1.353$$

因此, 128KB的伪相联Cache更快。

5.15

解:

实际带宽 = 0.6 × 最大带宽 = 4MB / s

又 最大带宽 = 字长 × m / 存储周期

所以 m = 最大带宽 × 存储周期 / 字长 = (4MB/s / 0.6) × 2us / 4B = 3.495

因为 m 应该是个整数, 所以 m 为 4

5.16 设主存由 8 个存储体按低位交叉编址方式组成, 主存容量 1MB, Cache 容量 4KB, 要求一个主存周期从主存取得一个块。采用全相联地址映像, 用相联目录表实现地址变换, 求出相联目录表的行数, 比较位数, 宽度和总位数。

解:

一个主存周期可取出一个块, 所以块长可视为 1

$$1MB = 1 \times 2^{20} B = 8 \times 2^{20} = 2^{23} b$$

$$4KB = 2^{12} B = 2^{15} b$$

相联目录表行数为 Cache 块数, 即:

$$2^{15} \div m = \frac{2^{15}}{m}$$

相联目录表比较位数为主存块号长度, 即:

$$\log_2 2^{23} = 23$$

相联目录表的宽度 = 主存块号的长度 + Cache 块号长度 + 有效位之和 = 15 + 23 + 1 = 39

总位数 = 目录表行数 × 目录表宽度

$$= \frac{2^{15} \times 39}{m}$$