

第一章 计算机抽象及相关技术

1.3 [2]<1.3> 请描述高级语言 (例如 C) 编写的程序转化为能够直接在计算机处理器上执行的表示的具体步骤。

高级语言通过编译器 (编译型的高级语言, 如 C、C++) 或解释器 (解释型的高级语言, 如 Python、JavaScript) 转化成指令, 再转化成二进制的机器码, 或者直接转化成二进制的机器码 (不同的硬件对应的机器码可能不同), 机器码就是能够直接在计算机处理器上执行的表示。

1.5 [4]<1.6> 有 3 种不同的处理器 P1、P2 和 P3 执行同样的指令系统。P1 的时钟频率为 3GHz,CPI 为 1.5;P2 的时钟频率为 2.5GHz, CPI 为 1.0;P3 的时钟频率为 4GHz,CPI 为 2.2。

a. 以每秒执行的指令数为标准, 哪个处理器性能最高?

$$\begin{aligned} P1: & \frac{3\text{GHz}}{1.5\text{时钟周期数/指令}} = 2 \times 10^9 \\ P2: & \frac{2.5\text{GHz}}{1.0\text{时钟周期数/指令}} = 2.5 \times 10^9 \\ P3: & \frac{4\text{GHz}}{2.2\text{时钟周期数/指令}} \approx 1.818 \times 10^9 \end{aligned}$$

所以处理器 P2 性能最高。

b. 如果每个处理器执行一个程序都花费 10 秒时间, 求它们的时钟周期数和指令数。

	时钟周期数	指令数
P1	$3\text{GHz} \times 10\text{s} = 30 \times 10^9$	$\frac{30 \times 10^9}{1.5} = 20 \times 10^9$
P2	$2.5\text{GHz} \times 10\text{s} = 25 \times 10^9$	$\frac{25 \times 10^9}{1.0} = 25 \times 10^9$
P3	$4\text{GHz} \times 10\text{s} = 40 \times 10^9$	$\frac{40 \times 10^9}{2.2} \approx 18.1818 \times 10^9$

c. 我们试图把执行时间减少 30%, 但这会引起 CPI 增大 20%。请问为达到时间减少 30% 的目标, 时钟频率应达到多少?

$$\begin{aligned} \text{执行时间} \times 70\% &= \text{指令数} \times (\text{CPI} \times 120\%) / \text{时钟频率} \\ \text{时钟频率} &= \frac{\text{指令数} \times \text{CPI}}{\text{执行时间}} \times \frac{1.2}{0.7} \approx 1.714 \times \text{原始时钟频率} \end{aligned}$$

所以时钟频率要达到原始时钟频率的大约 1.714 倍。

1.6 [20]<1.6> 同一个指令系统体系结构有两种不同的实现方式。根据 CPI 的不同将指令分成四类 (A、B、C 和 D)。P1 的时钟频率为 2.5GHz，CPI 分别为 1、2、3 和 3;P2 时钟频率为 3GHz，CPI 分别为 2、2、2 和 2。

给定一个程序，有 1.0×10^6 条动态指令，按如下比例分为 4 类:A, 10%;B, 20%;C, 50%;D, 20%。

a. 每种实现方式下的整体 CPI 是多少?

$$P1: 1 \times 10\% + 2 \times 20\% + 3 \times 50\% + 3 \times 20\% = 2.6$$

$$P2: 2$$

b. 计算两种情况下的时钟周期总数。

$$P1: 1.0 \times 10^6 \times 2.6 = 2.6 \times 10^6$$

$$P2: 1.0 \times 10^6 \times 2 = 2 \times 10^6$$

1.7 [15]<1.6> 编译器对应用程序性能有极深的影响。假定对于一个程序，如果采用编译器 A，则动态指令数为 1.0×10^9 ，执行时间为 1.1s；如果采用编译器 B，则动态指令数为 1.2×10^9 ，执行时间为 1.5s。

a. 在给定处理器时钟周期长度为 1ns 时，求每个程序的平均 CPI。

$$\text{编译器 A: } \frac{1.1s}{1 \times 10^{-9}s} / 1.0 \times 10^9 = 1.1$$

$$\text{编译器 B: } \frac{1.5s}{1 \times 10^{-9}s} / 1.2 \times 10^9 = 1.25$$

b. 假定被编译的程序分别在两个不同的处理器上运行。如果这两个处理器的执行时间相同，求运行编译器 A 生成之代码的处理器时钟比运行编译器 B 生成之代码的处理器时钟快多少。

$$\begin{aligned} \text{执行时间}_A &= \text{执行时间}_B \\ \frac{\text{指令数}_A \times CPI_A}{\text{时钟频率}_A} &= \frac{\text{指令数}_B \times CPI_B}{\text{时钟频率}_B} \\ \frac{\text{时钟频率}_A}{\text{时钟频率}_B} &= \frac{\text{指令数}_A \times CPI_A}{\text{指令数}_B \times CPI_B} = \frac{1.0 \times 10^9 \times 1.1}{1.2 \times 10^9 \times 1.25} = \frac{11}{15} = 0.733 \end{aligned}$$

所以题意应该是错了，运行编译器 A 生成之代码的处理器时钟比运行编译器 B 生成之代码的处理器时钟慢，约 0.733 倍。

c. 假设开发了一种新的编译器，只需 6.0×10^8 条指令，程序平均 CPI 为 1.1。求这种新的编译器在原处理器环境下相对于原编译器 A 和 B 的加速比。

记这个新的编译器为 C。

$$\frac{\text{性能}C}{\text{性能}A} = \frac{\text{执行时间}A}{\text{执行时间}C} = \frac{\frac{\text{指令数}_A \times CPI_A}{\text{时钟频率}}}{\frac{\text{指令数}_C \times CPI_C}{\text{时钟频率}}} = \frac{\text{指令数}_A \times CPI_A}{\text{指令数}_C \times CPI_C} = \frac{1.0 \times 10^9 \times 1.1}{6.0 \times 10^8 \times 1.1} = \frac{5}{3} \approx 1.667$$

同理,

$$\frac{\text{性能}C}{\text{性能}B} = \frac{\text{指令数}_B \times CPI_B}{\text{指令数}_C \times CPI_C} = \frac{1.2 \times 10^9 \times 1.25}{6.0 \times 10^8 \times 1.1} = \frac{25}{11} \approx 2.273$$

所以这种新的编译器在原处理器环境下相对于原编译器 A 和 B 的加速比分别约为 1.667 和 2.273。

1.9 在某处理器中, 假定算术指令、load/store 指令和分支指令的 CPI 分别是 1、12 和 5。同时假定某程序在单个处理器核上运行时需要执行 2.56×10^9 条算术指令、 1.28×10^9 条 load/store 指令和 2.56×10^8 条分支指令, 并假定处理器的时钟频率为 2GHz。现假定程序并行运行在多核上, 分配到每个处理器核上运行的算术指令和 load/store 指令数目为单核情况下相应指令数目除以 $0.7 \times p$ (p 为处理器核数), 而每个处理器的分支指令的数量保持不变。

1.9.1 [5] <1.7> 求出当该程序分别运行在 1、2、4 和 8 个处理器核上的执行时间, 并求出其他情况下相对于单核处理器的加速比。

$$\begin{aligned} \text{多核执行时间} &= \frac{\text{指令数} \times CPI}{\text{时钟周期}} = \frac{\frac{2.56 \times 10^9}{0.7 \times p} \times 1 + \frac{1.28 \times 10^9}{0.7 \times p} \times 12 + 2.56 \times 10^8 \times 5}{2 \times 10^9} = \frac{16(p+20)}{25p} s \\ \text{单核执行时间} &= \frac{\text{指令数} \times CPI}{\text{时钟周期}} = \frac{2.56 \times 10^9 \times 1 + 1.28 \times 10^9 \times 12 + 2.56 \times 10^8 \times 5}{2 \times 10^9} = \frac{48}{5} = 9.6s \\ \text{多核与单核的加速比} &= \frac{\text{单核执行时间}}{\text{多核执行时间}} = \frac{\frac{48}{5}}{\frac{16(p+20)}{25p}} = \frac{15p}{p+20} \end{aligned}$$

处理器核数	执行时间	加速比
1	9.600	1.000
2	7.040	1.364
4	3.840	2.500
8	2.240	4.286

1.9.2 [10] <1.6,1.8> 如果算术指令的 CPI 加倍, 对分别运行在 1、2、4 和 8 个处理器核上的执行时间有何影响?

$$\begin{aligned} \text{多核执行时间} &= \frac{\frac{2.56 \times 10^9}{0.7 \times p} \times 2 + \frac{1.28 \times 10^9}{0.7 \times p} \times 12 + 2.56 \times 10^8 \times 5}{2 \times 10^9} = \frac{16(7p+160)}{175p} s \\ \text{单核执行时间} &= \frac{2.56 \times 10^9 \times 2 + 1.28 \times 10^9 \times 12 + 2.56 \times 10^8 \times 5}{2 \times 10^9} = \frac{272}{25} = 10.88s \end{aligned}$$

处理器核数	之前的执行时间	现在的执行时间
1	9.600	10.880
2	7.040	7.954
4	3.840	4.297
8	2.240	2.469

1.9.3 [10]<1.6,1.8> 如果要使单核处理器的性能与四核处理器相当，单处理器中 load/store 指令的 CPI 应该降低多少？此处假定四核处理器的 CPI 保持原数值不变。

设单处理器中 load/store 指令的 CPI 应该降低到原 CPI 的 x 倍。

$$\begin{aligned} \text{单核处理器 (CPI降低) 的执行时间} &= \text{四核处理器 (原CPI) 的执行时间} \\ \frac{2.56 \times 10^9 \times 1 + 1.28 \times 10^9 \times 12x + 2.56 \times 10^8 \times 5}{2 \times 10^9} &= \frac{16 \times (4 + 20)}{25 \times 4} \\ \frac{192x}{25} + \frac{48}{25} &= \frac{96}{25} \\ x = \frac{96 - 48}{192} &= \frac{1}{4} = 0.25 \end{aligned}$$

所以单处理器中 load/store 指令的 CPI 应该降低到原 CPI 的 0.25 倍，即 $12 \times 0.25 = 3$ 。