

第五章课后习题

- 第二题解：

1. 单周期处理器的**CPI = 1**。单周期 CPU 的时钟周期必须长到足以容纳**最慢的指令**所需时间。性能差是因为：所有指令都被迫使用**同一个超长周期**。
2. 设计思想：将一条指令分成多个步骤，每个步骤用一个时钟周期执行。每条指令在多周期处理器中的**CPI不相同**，因为不同指令执行阶段数量不同，因此占用的周期数也不同。
3. 单条指令的执行时间**并没有变短**，甚至通常会变长。程序的执行时间则会大大**缩短**。
4. 流水线的时钟周期由所有流水级中**延迟最长的那个阶段**，再加上流水寄存器延迟共同决定。理想情况下，单条流水线处理器的**CPI=1**。
5. 在各流水段之间**加寄存器**的原因是：保存每一阶段的计算结果；实现阶段间的数据隔离；保持时序正确，形成同步电路。各流水段寄存器的宽度**不相同**。这是因为：每个阶段需要**传递的数据和控制信号不同**，位宽自然不同。

- 第六题解：

1. RegWr 恒为 0：所有需要写回寄存器堆的指令都会失败。
2. ALUASrc 恒为 0：所有需要以 PC 为 ALU 操作数来计算跳转和分支目标的指令。
3. Branch 恒为 0：所有有条件分支指令。
4. Jump 恒为 0：无条件跳转或者函数调用类指令。
5. MemWr 恒为 0：所有 store 指令。
6. MemtoReg 恒为 0：load 指令。

- 第十四题解：

流水线时钟周期 = 所有阶段中最大的延迟 = **max(200, 50, 150, 200, 50) = 200 ps**。

1. 已知ALU 原延迟 = 150 ps，减少 20%后得到 新延迟 = $150 \times 0.8 = 120 \text{ ps}$ 。然而流水线时钟周期仍然由 **200 ps**所决定，所以**无法加快流水线速度**。这是因为时钟周期由最长的所决定，这里也就是存储器，而非 ALU。
2. 与上问类似，因为时钟周期由最长的所决定，这里也就是存储器，而非 ALU。所以最大延迟仍然是 **200 ps**，对流水线性能没有影响。
3. 当 ALU 延迟增加 40% 时，ALU 新延迟 = $150 \times 1.4 = 210 \text{ ps}$ ，此时 ALU 延迟成为新的流水线**时钟周期**。性能下降约为：

$$\left(1 - \frac{200}{210}\right) \times 100\% \approx 4.76\%$$

- 第十五题解：

1. 我们设程序指令数 $N = 1000000 = 10^6$ ，每条指令在非流水线处理器上的执行时间 $t = 100 \text{ ps} = 10^{-10} \text{ s}$ 。

非流水线下，每条指令串行完成，程序总时间为：

$$T = Nt = 10^6 \times 10^{-10} = 10^{-4}s$$

答案为约0.0001秒。

2. 当新CPU采用20级流水线，理想状态下，每条指令被均匀分成 $k = 20$ 个阶段，管道填充和排空开销相对于 N 可忽略，即 $N \gg k$ 。

每个阶段的理想时钟周期为 $t' = \frac{t}{k} = 5ps$ ，流水线总周期数大约为 $N + (k - 1)$ ，所以总时间为：

$$T = (N + k - 1) \times t' = 5.000095 \times 10^{-6}s \approx 5 \times 10^{-6}s$$

因此，理想加速比为：

$$Speedup \approx \frac{10^{-4}}{5 \times 10^{-6}} \approx 20$$

所以，在理想情况下，流水线大约比非流水线快**20**倍。

这些开销会影响吞吐率，并可能使时钟周期变长；对单条指令的执行时间影响较小。

- 第十七题解：

首先给出的程序指令序列为：

```
I1: add s3, s1, s0
I2: add t2, s3, s3
I3: lw t1, 0(t2)
I4: add t3, t1, t2
```

发生数据相关指令对为：(I1, I2), (I2, I3), (I2, I4), (I3, I4)。

在不采用转发技术的情况下，在每个发生数据相关的“后面指令前”各插入**2**个nop，也就是一共插入**6**个nop。

在采用转发的情况下，该序列仅需在I3(lw)和I4(add)之间插入**1**个nop来避免冒险，其它相关都由转发处理。

- 第十八题解：

由题意知，程序指令序列为：

```
I1: add s3, s1, s0
I2: sub t2, s3, s3
I3: lw t1, 0(t2)
I4: add t3, t1, t2
I5: add t1, s4, s5
```

1. 寄存器的转发以及转发的目的地如下表：

指令序号	寄存器	指令序号
I_1	s3	I_2
I_2	t2	I_3
I_2	t2	I_4
I_3	t1	I_4

2. 是，存在这种数据冒险： $I_3 (\text{lw } t1, 0(t2)) \rightarrow I_4 (\text{add } t3, t1, t2)$ 。

3. 第五周期结束时，各指令的执行状态如下表：

指令序号	执行状态
I_1	WB
I_2	MEM
I_3	EX
I_4	ID
I_5	IF

t1 和 **t2** 两个寄存器的数据正在被读出，寄存器 **s3** 将被写入。

- 第十九题解：

1. 带转发的 **CPI = 1**

2. 不带转发的 **CPI = 1**

- 第二十一题解：

1. 因为分支指令在 EX 阶段确定跳转，所以必须清空 ID 阶段的指令，进而损失 **1 个时钟周期**。因此，分支延迟槽数量 = **1**。

2. 对于指令 **4 → 指令 5** 分析： $\text{lw } t0, 0(t1) \rightarrow \text{bne } t0, s5, \text{ Exit}$ 。这是 **load-use** 冒险，进一步分析知：指令 5 需要阻塞 **1 个周期**。

- 第二十三题解：

1. 静态预测：总是预测 Taken

分析后可知：总正确次数 = $3 + 0 + 3 + 4 + 5 = 15$ ，总执行次数 = $3 + 4 + 6 + 5 + 7 = 25$ 。

所以，准确率为：

$$\eta = \frac{15}{25} = 0.60$$

2. 静态预测：总是预测 Not Taken

分析后可知：总正确次数 = 0 + 4 + 3 + 1 + 2 = 10， 总执行次数 = 3 + 4 + 6 + 5 + 7 = 25。

所以，准确率为：

$$\eta = \frac{10}{25} = 0.40$$

3. 1-bit 动态预测，初始预测 Taken

分析后可知：总正确次数 = 3 + 3 + 1 + 3 + 3 = 13， 总执行次数 = 3 + 4 + 6 + 5 + 7 = 25。

所以，准确率为：

$$\eta = \frac{13}{25} = 0.52$$

4. 2-bit 动态预测，初始预测弱转移 (WT)

分析后可知：总正确次数 = 3 + 3 + 3 + 4 + 5 = 18， 总执行次数 = 3 + 4 + 6 + 5 + 7 = 25。

所以，准确率为：

$$\eta = \frac{18}{25} = 0.72$$