

高等计算机体系结构，2021 年春季

作业 4：流水线 2

主讲教师：栾钟治

助讲教师：杨海龙；助教：孙庆骁

作业下发时间：2021 年 4 月 23 日

作业回收时间：2021 年 5 月 7 日

1 分支预测 35 分

考察以下高级语言代码段：

```
int array[1000] = { /* random values */ };
int sum1 = 0, sum2 = 0, sum3 = 0, sum4 = 0;
for (i = 0; i < 1000; i++) // 分支 1: 循环分支
{
    if (i % 4 == 0) // 分支 2: If 条件分支 1
        sum1 += array[i]; // 发生分支的路径
    else
        sum2 += array[i]; // 不发生分支的路径
    if (i % 2 == 0) // 分支 3: If 条件分支 2
        sum3 += array[i]; // 发生分支的路径
    else
        sum4 += array[i]; // 不发生分支的路径
}
```

(a) 当使用 last-time 预测器时三个分支的预测准确率分别是多少？(假设每一个分支的 last-time 计数器起始状态是‘不发生’) 请写出你的计算过程和依据。

(b) 当使用基于 2-bit 饱和计数器的预测器时三个分支的预测准确率分别是多少？(假设每一个分支的 2-bit 计数器起始状态是‘强不发生’) 请写出你的计算过程和依据。

(c) 当分支 2 和分支 3 的 2-bit 计数器起始状态分别是 (i)‘弱不发生’；(ii)‘弱发生’ 时，预测准确率分别是多少？请写出你的计算过程和依据。

(d) 当使用两层全局历史分支预测器（2bit 全局历史寄存器+每分支一张独立的模式历史表，模式历史表的每个表项是一个 2-bit 饱和计数器）时，三个分支的预测准确率分别是多少？假设全局历史寄存器每一位的初始状态都是‘不发生’，模式历史表中的 2-bit 饱和计数器初始状态都是‘强不发生’，计算预测准确率时，忽略前 500 次循环迭代。

2 两层分支预测器 25 分

假设一个两层全局预测器由全局历史寄存器和所有分支共享的一张模式历史表组成(称为预测器 A)

1) 我们把分支预测器中不同的分支映射到相同位置的情况称为"分支干扰"。预测器 A 的结构中，不同分支会在哪里发生这种干扰？

2) 另一个两层全局预测器由全局历史寄存器和每个分支一张模式历史表组成 (称为预测器 B"),

(a) 什么情况下预测器 A 的预测准确率低于预测器 B? 请解释理由，并举例说明。可以通过代码来说明。

(b) 预测器 A 能获得比预测器 B 更高的预测准确率吗? 请解释理由，并举例说明。可以通过代码来说明。

(c) 分支干扰是否总是影响预测器的预测准确率? 请解释理由，并举例说明。可以通过代码来说明。

3 分支预测和推断 30 分

考察两台具有 15 段流水线的机器 A 和 B，流水段分布如下：

取指 (1 个阶段)

译码 (8 个阶段)

执行 (5 个阶段)

写回 (1 个阶段)

两台机器都会在发生流相关时采用数据转发。在译码的最后一个阶段检测是否有流相关，指令会在停顿在这个阶段等待检测结果。

机器 A 有一个预测准确率为 P% 的分支预测器，分支方向和目标在执行的最后一个阶段产生。

机器 B 采用推断执行，类似咱们在课堂上讲的方式。

1) 考察以下在机器 A 上执行的代码段:

```
Add r3 r1, r2
sub r5 r6, r7
beq r3, r5, X
addi r10 r1, 5
add r12 r7, r2
add r1 r11, r9
X: addi r15 r2, 10
....
```

我们把这段代码转化为在机器 B 上执行的推断代码，大概是下面这个样子:

```
add r3 r1, r2
sub r5 r6, r7
cmp r3, r5
addi.ne r10 r1, 5
add.ne r12 r7, r2
add.ne r14 r11, r9
addi r15 r2, 10
....
```

(假设条件结果由‘cmp’指令计算，推断由‘ne’指令根据条件结果执行。条件结果在执行的最后一个阶段计算并且可以像其他值一样被转发)

这一段代码会重复执行上百次，分支 40% 可能性发生，60% 可能性不发生，平均而言，P 取什么样的值会使机器 A 比机器 B 具有更高的指令吞吐量?

2) 考察在机器 A 上执行的另一段代码:

```
add r3 r1, r2
sub r5 r6, r7
beq r3, r5, X
```

```
addi r10 r1, 5
add r12 r10, r2
add r14 r12, r9
X: addi r15 r14, 10
.....
```

我们把这段代码转化为在机器 B 上执行的推断代码，大概是下面这个样子：

```
add r3 r1, r2
sub r5 r6, r7
cmp r3, r5
addi.ne r10 r1, 5
add.ne r12 r10, r2
add.ne r14 r12, r9
addi r15 r14, 10
.....
```

(假设条件结果由‘cmp’指令计算，推断由‘.ne’指令根据条件结果执行。条件结果在执行的最后一个阶段计算并且可以像其他值一样被转发)

这一段代码会重复执行上百次，分支 40%可能性发生，60%可能性不发生，平均而言，P 取什么样的值会使机器 A 比机器 B 具有更高的指令吞吐量？