

作业 7 多线程

下面是一个链表结点结构，包含一个 **key**，指向下一个结点的指针，指向数据的指针。

```
struct node {
    int key;
    struct node *next;
    struct data*ptr;
}
```

下面代码搜索链表，找到某个特定 **key** 值的结点。

```
; R1: a pointer to the linked list
; R2: the key to find
loop:    LW      R3, 0(R1) ;    load a key
         LW      R4, 4(R1) ;    load the next pointer
         SEQ     R3, R3, R2 ;    set R3 if R3 == R2
         BNEZ    R3, END      found the entry
         ADD     R1, R0, R4
         BNEZ    R1, loop ;    check the next node
End:
         ; R1 contains a pointer to the matching entry or zero
         ; if not found
```

假设在一个**单发射、顺序处理器**上运行上面程序。该处理器每个周期可以取 1 条指令、发射 1 条指令。如果指令存在相关性，需要停顿。**整数运算需要 1 个周期完成**，计算结果在下一个周期可用。例如，如果 **SEQ** 在第 1 个周期执行，那么 **BNEZ** 可以在第 2 个周期执行。假设处理器**所有分支预测都 100%正确**。

(1) 假设该处理器没有 **cache**。**每个访存操作需要 100 个周期完成**。**load/store** 功能单元基于流水线实现，并且非阻塞。即，处理器提交访存请求后，不需要等待结果可以继续执行后面指令，直到遇到有相关性的指令。那么，在稳定状态下，完成 1 次循环需要多少个周期？

稳定状态下，每个周期执行 1 条指令，除非遇到相关性。所以指令执行时间如下表所示，一共需要 104 个周期。

指令	开始周期	结束周期
LW R3, 0(R1)	1	100
LW R4, 4(R1)	2	101
SEQ R3, R3, R2	101	101
BNEZ R3, End	102	102
ADD R1, R0, R4	103	103
BNEZ R1, Loop	104	104

(2) 假设采用**细粒度多线程**技术，每个线程执行相同的代码，但是搜索不同的 **key** 值。硬件

负责线程调度，采用简单的轮循方式，每个周期调度 1 个线程来执行，所以每个线程每 N 个周期执行 1 条指令。那么，如果让该处理器利用率达到 100%（每个周期都有指令执行），最少要有多少个线程？

假设有 N 个线程，第 1 个线程的 LW 指令在第 1,2 个周期开始，因为每个线程有两个 load 指令，所以第 1 个线程的第 3 条指令（SEQ）要在 $2N$ 个周期后才能执行。如果要充分利用处理器，直到 SEQ 指令执行之前都不能有空闲，所以要填补 100 个周期，即需要 $N=50$ 。

(3) 假设采用细粒度多线程技术，并且处理器的利用率达到了 100%。将如何影响 throughput（单位时间内完成的搜索次数）和 latency（单次搜索用的周期数）呢？（只回答变好或者变坏即可）

throughput 变好了，但是 latency 变坏了，因为每个线程的后面几条指令要排队执行。

(4) 假设采用粗粒度多线程技术，当一个线程因为数据相关性停顿的时候再切换到其他线程。那么，最少需要多少个线程才能让处理器的利用率达到 100%？

稳定状态下，每个线程最多可以连续执行 6 条指令，顺序为 SEQ, BNEZ, ADD, BNEZ, LW, LW。6 条指令后被切换出去，等到下次切换进来至少要等 99 个周期（第 2 个 LW 和 SEQ 之前差 99 个周期），这个期间需要 $\lceil 99/6 \rceil + 1 = 18$ 个线程才能填满。