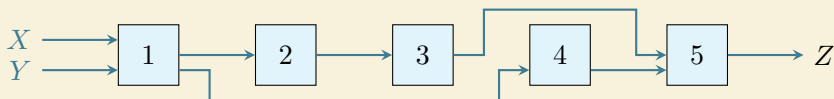


设有两个向量  $A, B$ ，各有 4 个元素，在静态双功能流水线上，计算向量点积  $A \cdot B = \sum_{i=1}^4 a_i b_i$ 。其中， $1 \rightarrow 2 \rightarrow 3 \rightarrow 5$  组成加法流水线， $1 \rightarrow 4 \rightarrow 5$  组成乘法流水线。又设每个流水线所经过的时间为  $\Delta t$ ，而且流水线的输出结果可以直接返回到输入或暂存于相应的缓冲寄存器中，其延迟时间和功能切换所需的时间都可以忽略不计。请使用合理的算法，完成向量点积  $A \cdot B$  所用的时间最短，并求出流水线在此期间实际的吞吐率  $TP$  和效率  $\eta$ 。



明确静态流水线的特征是 **加法和乘法运算不能同时进行**。

对于此题，我们可以连续计算  $a_1 b_1, a_2 b_2, a_3 b_3, a_4 b_4$  共 4 次乘法运算（将 4 个乘积分别记为  $c_1, c_2, c_3, c_4$ ），随后进行  $(a_1 b_1 + a_2 b_2) + (a_3 b_3 + a_4 b_4)$  共 3 次加法运算，相当于总共 7 条指令：

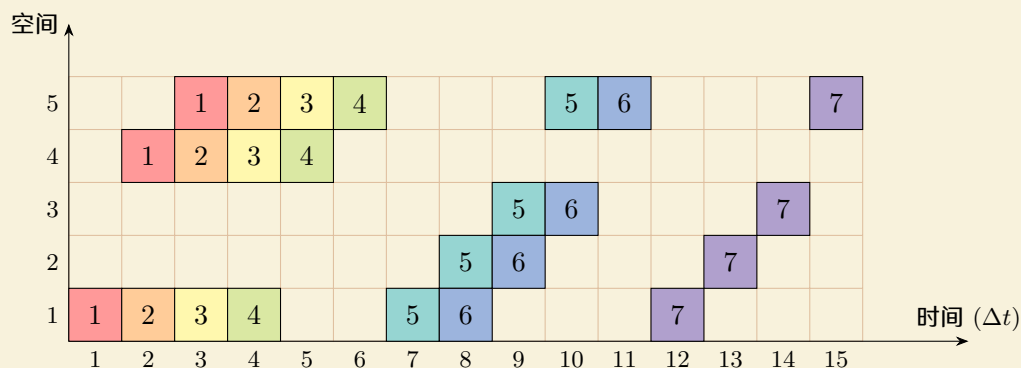
1:	mul	$c_1$	$a_1$	$b_1$
2:	mul	$c_2$	$a_2$	$b_2$
3:	mul	$c_3$	$a_3$	$b_3$
4:	mul	$c_4$	$a_4$	$b_4$
5:	add	$S_{12}$	$c_1$	$c_2$
6:	add	$S_{34}$	$c_3$	$c_4$
7:	add	$S$	$S_{12}$	$S_{34}$

根据本题题设，可以提炼出如下规则：

- 第 4 条指令执行完毕以后，第 5 条指令才能开始（因为静态流水线）
- 第 1、2 条指令执行完毕以后，第 5 条指令才能开始（因为需要  $c_1, c_2$  值，「写后读」相关）
- 第 3、4 条指令执行完毕以后，第 6 条指令才能开始（因为需要  $c_3, c_4$  值，「写后读」相关）

- 第 5、6 条指令执行完毕以后，第 7 条指令才能开始（因为需要  $S_{12}, S_{34}$  值，「写后读」相关）

如此可以做出时空图：



从上面的「彩虹时空图」可以看出，整个流水线经历了 15 个  $\Delta t$ 。

- 实际吞吐率  $TP = \frac{\text{指令数}}{\text{流水线用时}} = \frac{7}{15\Delta t}$
- 加速比  $S_p = \frac{\text{非流水线用时}}{\text{流水线用时}} = \frac{3 \times 4\Delta t + 4 \times 3\Delta t}{15\Delta t} = 1.6$
- 效率  $\eta = \frac{\text{有效时空区}}{\text{总时空区}} = \frac{3 \times 4 + 4 \times 3}{15 \times 5} = \frac{24}{75} = 32\%$

【结论】 $TP = \frac{7}{15\Delta t}$ ， $\eta = 32\%$ 。

【点评】这是一道指令流水线的分析题。其中，本题的「静态流水线」是一个坑点，因为它规定了乘法和加法指令不能同时运行，这一点对时空图的影响比较关键。此外，同学们还需要掌握流水线的几个性能指标，如最大吞吐率、实际吞吐率、加速比、效率等。