

一、概念题

解释下列名词：

1. 纵横处理方式
把长度为 N 的向量分为若干组，每组长度为 n ，组内按纵向方式处理，依次处理各组，组数为 $\lceil N/n \rceil$ ，适合流水处理。
2. 向量流水线链接
具有先写后读相关的两条指令，在不出现功能部件冲突和 V_i 冲突的情况下，可以把功能部件链接起来进行流水处理，以达到加快执行的目的。
3. V_i 冲突
指并行工作的各向量指令的源向量或结果向量要使用相同的 V_i
4. 半性能向量长度
向量处理机的性能达到其最大性能 R_∞ 的一半时所需的向量长度。
5. 流水线冲突
指对于具体的流水线来说，由于相关的存在，使得指令流中的一条指令不能在指定的时钟周期执行。
6. 指令的静态调度
指依靠编译器对代码进行静态调度，以减少相关和冲突。它不是在程序执行的过程中，而是在编译期间进行代码调度和优化的。

二、选择题

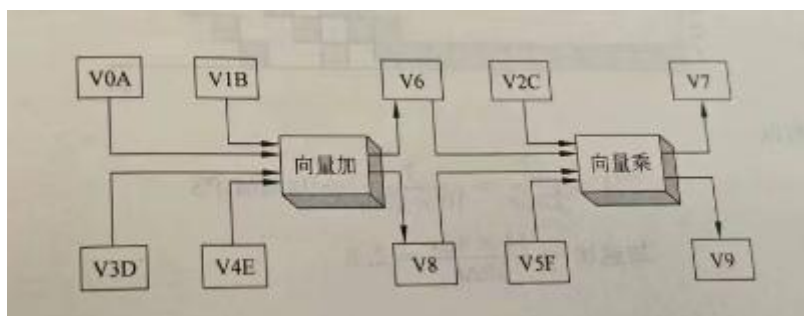
1. Cray-1 的流水线是 (A)
A. 多条单功能流水线
B. 一条单功能流水线
C. 多条多功能流水线
D. 一条多功能流水线
2. Cray-1 向量处理机要实现指令间的链接，必须满足下列条件中的 (C)
A. 源向量相同，功能部件不冲突，有指令相关
B. 源向量不同，功能部件相同，无指令相关
C. 源向量、功能部件都不相同，指令有写后读冲突
D. 源向量、功能部件都不相同，指令有读后写冲突
3. Cray-1 的两条向量指令：
 $V_1 \leftarrow V_2 + V_3$
 $V_4 \leftarrow V_1 \times V_5$
属于 (B)

- A. 没有功能部件冲突和源向量冲突，可以并行
 B. 没有功能部件冲突和源向量冲突，可以链接
 C. 没有源向量冲突，可以交换执行顺序
 D. 有向量冲突，只能串行
4. 假设每种向量的功能部件只有一个，而且不考虑向量链接，那么下面的一组向量指令能分成 (D) 个编队。
- | | | |
|--------|------------|--------------------|
| LV | V1, Rx | //取向量 x |
| LV | V3, Ry | //取向量 y |
| MULTSV | V2, R0, V1 | //向量 x 和标量 (R0) 相乘 |
| ADDV | V4, V2, V3 | //相加，结果保存到 V4 中 |
| SV | V4, Ry | //存结果 |
- A, 1
 B, 2
 C, 3
 D, 4

三、问答题

- 采用链接技术时，向量指令能够链接执行必须满足哪些条件？
 - 向量指令之间要求无 V_i 冲突和无功能部件使用冲突。
 - 只有在前一条指令的第一个结果元素送入结果向量寄存器的那一个时钟周期才可以进行链接。如果错过这个时刻，就无法进行链接了。
 - 当一条向量指令的两个源操作数分别来自前面紧邻的两条指令的结果时，要求这两条指令产生运算结果的时间必须相等，即要求有关功能部件的通过时间相等。
 - 链接执行的向量指令的向量长度必须相等，否则无法进行链接。
- 某向量处理机有 16 个向量寄存器，其中 $V_0 \sim V_5$ 中分别放有向量 A、B、C、D、E、F，向量长度均为 8，向量各元素均为浮点数；处理部件采用两条单功能流水线，加法功能部件时间为 2 拍，乘法功能部件时间为 3 拍。采用类似于 Cray-1 的链接技术，先计算 $(A+B) \times C$ ，在流水线不停流的情况下，接着计算 $(D+E) \times F$ 。
 - 求此链接流水线的通过时间（设寄存器入、出各需 1 拍）？
 - 假如每拍时间为 50ns，完成这些计算并把结果存进相应寄存器，此处理部件的实际吞吐率多少 MFLOPS？

答案：（1）在这里假设 $A+B$ 的中间结果放在 V_6 中， $(A+B) \times C$ 的最后结果放在 V_7 中， $D+E$ 的中间结果放在 V_8 中， $(D+E) \times F$ 的最后结果放在 V_9 中。具体实现如下所示。



通过时间应该为前者 $(A+B) \times C$ 通过的时间：

$$T_{\text{通过}} = (1+2+1) + (1+3+1) = 9 \text{ (拍)}$$

(2) 在做完 $(A+B) \times C$ 之后，做 $(C+D) \times E$ 就不需要通过时间了。

$$V6 \leftarrow A+B$$

$$V7 \leftarrow V6 \times C$$

$$V8 \leftarrow D+E$$

$$V9 \leftarrow V8 \times F$$

$$T = T_{\text{通过}} + (8-1) + 8 = 24 \text{ (拍)} = 1200 \text{ (ns)}$$

$$TP = 32/T = 26.67 \text{ MFLOPS}$$

3. 在一向量处理机上实现 $A=B \times S$ 计算，其中 A 和 B 是长度为 $N=200$ 的向量， S 是一个标量，向量寄存器长度为 $MVL=64$ ，各功能部件的启动时间为：取数和存数部件为 12 个时钟周期、乘法部件为 7 个时钟周期，执行标量代码的开销 $T_{\text{loop}} = 15$ 个时钟周期，对一个向量元素执行一次操作的时间 $T_g = 1$ 个时钟周期。求 A 的总执行时间。

解：假设向量 A 和 B 存放在向量寄存器 Ra 和 Rb 中，标量 S 存放在标量寄存器 $R0$ 中，由下面 3 条指令完成计算：

LV V1, Rb

MULTSV V2, R0, V1

SV V2, Ra

这 3 条指令之间存在相关，需分为 3 个编队 $m=3$ 。向量需要分为 $\lceil 200 / 64 \rceil = 4$ 组进行计算。

由题目得：

$$T_{\text{start}} = 12 + 12 + 7 = 31, T_{\text{loop}} = 15$$

$$\begin{aligned} T &= 4 \times (T_{\text{loop}} + T_{\text{start}}) + 3 \times 200 \times 1 \\ &= 4 \times (15 + 31) + 3 \times 200 \times 1 = 784 \text{ (时钟周期)} \end{aligned}$$

4. 为了保证程序执行的正确性，必须保持哪两个最关键的属性？简述其含义。

最关键的属性是：数据流和异常行为。

保持异常行为是指：无论怎么改变指令的执行顺序，都不能改变程序中异常的发生情况。即原来程序中是怎么发生的，改变执行顺序后还是怎么发生。这个条件经常被弱化为：指令执行顺序的改变不能导致程序中新发生的异常。

数据流是指数据值从其产生者指令到其消费者指令的实际流动。

5. 假设浮点流水线中各部件的延迟如下。

加法需 2 个时钟周期；

乘法需 10 个时钟周期；

除法需 40 个时钟周期。

代码段和记分牌信息的起始点状态如下图所示。给出 SUB.D 准备写结果之前的记分牌状态。

指令	指令状态表			
	流出	读操作数	执行	写结果
L.D F6,34(R2)	✓	✓	✓	✓
L.D F2, 45(R3)	✓	✓	✓	
MULT.D F0, F2, F4	✓			
SUB.D F8, F6, F2	✓			
DIV.D F10, F0, F6	✓			
ADD.D F6, F8, F2				

部件名称	功能部件状态表								
	Busy	Op	Fi	Fj	Fk	Qj	Qk	Rj	Rk
Integer	yes	L.D	F2	R3				no	
Mult1	yes	MULT.D	F0	F2	F4	Integer		no	yes
Mult2	no								
Add	yes	SUB.D	F8	F6	F2		Integer	yes	no
Divide	yes	DIV.D	F10	F0	F6	Mult1		no	yes

部件名称	结果寄存器状态表							
	F0	F2	F4	F6	F8	F10	...	F30
	Mult1	Integer			Add	Divide		

MIPS 记分牌中保存的信息

指令	指令状态表			
	流出	该操作数	执行	写结果
L.D F6,34(R2)	✓	✓	✓	✓
L.D F2,45(R3)	✓	✓	✓	✓
MULT.D F0,F2,F4	✓	✓	✓	
SUB.D F8,F6,F2	✓	✓	✓	
DIV.D F10,F0,F6	✓			
ADD.D F6,F8,F2				

部件名称	功能部件状态表								
	Busy	Op	Fi	Fj	Fk	Qj	Qk	Rj	Rk
Integer	no								
Mult1	yes	MULT.D	F0	F2	F4			no	no
Mult2	no								
Add	yes	SUB.D	F8	F6	F2			no	no
Divide	yes	DIV.D	F10	F0	F6	Mult1		no	yes

部件名称	结果寄存器状态表							
	F0	F2	F4	F6	F8	F10	...	F30

部件名称	Mult1				Add	Divide		
------	-------	--	--	--	-----	--------	--	--