

本科生期末试卷 (1)

一. 填空题。(每小题 2 分, 共 14 分)

1. A 是促使计算机系统结构发展最重要的因素, B 是促使计算机系统结构发展最根本的动力, 而 C 是促使计算机系统结构发展最活跃的因素。

2. 流水线技术是一种 A 并行技术, 它把一个 B 的过程分解为若干个子过程, 每个子过程与其他子过程并行进行。

3. 在一个过程中进行指令一级或操作一级的 A 处理, 称为 B 并行性。

4. 标量是指 A , 而向量是指 B , 而对一组数的运算叫做 C 处理。

5. 为了反映不同互连网络的连接特性, 每种互连网络可用一组 A 来描述, 它表示相互连接的 B 端号和 C 端号之间的一一对应关系。

6. 阵列处理机又称 A , 主要技术手段是硬件上采用 B 的方法来实现并行性。

7. 多处理机结构由若干台独立的计算机组成, 每台计算机能够独立执行自己的 A , Flynn 称这种结构为 B 结构。

二. 解释下列术语。(每个 3 分, 共 21 分)

1. 计算机系统结构
2. 结构相关
3. 细粒度并行性
4. 向量流水处理
5. 互连网络
6. 并行处理机
7. 共享存储多处理机

三. (13 分) 假定要将某一执行部件改进后速度提高 10 倍, 改进后被改进部件执行时间占系统总时间的 50%。问改进后, 获得的加速比 S_p 是多少?

四. (13 分) 如果一条指令的执行过程分为取指令、指令分析、指令执行三个子过程, 且这三个子过程延迟时间相等, 即每个子过程时间都为 t 。假设某过程共有 $n=1000$ 条指令, 请写出如下三种情况下机器执行该程序所需时间, 并加以比较。

- (1) 指令顺序执行方式;
- (2) 一次重叠执行方式;
- (3) 二次重叠执行方式。

五. (13 分) 有 A、B、C、D 四个存储器操作数, 要求完成 $(A \times B) + (C + D)$

的运算, 原来使用的程序如下:

```
I1  LOAD  R1, M(A)           ; R1 ← M(A)
I2  LOAD  R2, M(B)           ; R2 ← M(B)
I3  MUL   R5, R1, R2         ; R5 ← (R1) * (R2)
I4  LOAD  R3, M(C)           ; R3 ← M(C)
I5  LOAD  R4, M(D)           ; R4 ← M(D)
I6  ADD   R2, R3, R4         ; R2 ← (R3) + (R4)
I7  ADD   R2, R2, R5         ; R2 ← (R2) + (R5)
```

现采用静态指令调度方法, 请写出该程序调度后的指令序列。

六. (13 分) 设网络结点数 $N=8$, 通过立方体置换网络进行连接。试写出互连函数表达式, 并画出他们的连接图。

七. (13 分) 一个具有 32 台处理机的系统, 对远程存储器的访问时间是 2000ns。除了通信以外, 假设计算中的访问均命中局部存储器。当发出一个远程请求时, 本处理机挂起。处理机的时钟周期时间是 10ns, 假设指令基本的 CPI 为 1.0 (设所有访存均命中 Cache)。

求在没有远程访问的状态下, 与有 0.5% 的指令需要远程访问的状态下, 前者比后者快多少?

本科生期末试卷(1) 答案

一. 填空题。

1. A. 软件 B. 应用 C. 器件
2. A. 时间 B. 重复
3. A. 并行 B. 细粒度
4. A. 单个量 B. 一组标量 C. 向量
5. A. 互连函数 B. 输入 C. 输出
6. A. 并行处理机 B. 资源重复
7. A. 程序 B. MIMD

二. 解释术语。

1. 计算机系统结构: 机器语言程序员所看到的传统机器级所具有的属性, 它包括概念性结构和功能特性两个方面。
2. 结构相关: 某些指令组合在流水线重叠执行过程中, 如果硬件资源满足不了指令重叠执行的要求, 便会产生资源冲突, 则称流水线有结构相关。
3. 细粒度并行性: 是指在一个进程中进行指令一级或操作一级的并行处理。
4. 向量流水处理: 一条向量指令可以处理 N 个或 N 对操作数。我们把这 N 个互相独立的数叫做向量, 对这样一组数的运算叫做向量处理。
5. 互连网络: 互连网络是一种由高速开关元件按照一定的拓扑结构和控制方式构成的网络, 用来实现计算机系统内部多个处理机或多个功能部件之间的相互连接。
6. 并行处理机: 主要技术手段是硬件上采用资源重复的方法来实现并行性。因为是单指令流多数据流结构, 也称为 SIMD 计算机。
7. 共享存储多处理机: 共享存储型多处理机 SMP, 也称为对称型多处理机。它又分为 UMA、NUMA、COMA 三种结构模型, 彼此的区别在于存储器和外围资源如何共享或分布。

三. 解:

假设系统在改进前后的执行时间分别为 T_0 和 T_n , 则

$$S_p = T_0 / T_n = \frac{1}{(1 - F_e) + F_e / S_e} \quad (1)$$

由(1)式得

$$T_n = [(1 - F_e) + F_e / S_e] \times T_0 \quad (2)$$

根据题意:

$$50\% T_n = (F_e \times T_0) / 10 \quad (3)$$

由(2)式 - (1)式得

$$50\% T_n = (1 - F_e) \quad (4)$$

$$F_e = \frac{10 \times T_n}{2 \times T_0} \quad (5)$$

由(4)、(5)式可得

$$S_p = T_0 / T_n = 5.5$$

四. 解:

(1) 顺序执行方式

$$T_1 = 3t \times n = 3nt = 3 \times 1000 \times t = 3000t$$

(2) 一次重叠执行方式

$$T_2 = 2t \times n + t = (2n+1)t = (2 \times 1000 + 1)t = 2001t$$

(3) 二次重叠执行方式

$$T_3 = t \times n + 2t = (n+2)t = (1000+2)t = 1002t$$

(4) $T_1 > T_2 > T_3$

五. 解:

```

I1  LOAD  R1, M(A)
I2  LOAD  R2, M(B)
I3  LOAD  R3, M(C)
I4  LOAD  R4, M(D)
I5  MUL   R5, R1, R2
I6  ADD   R2, R3, R4
I7  ADD   R2, R2, R5
    
```

六. 解:

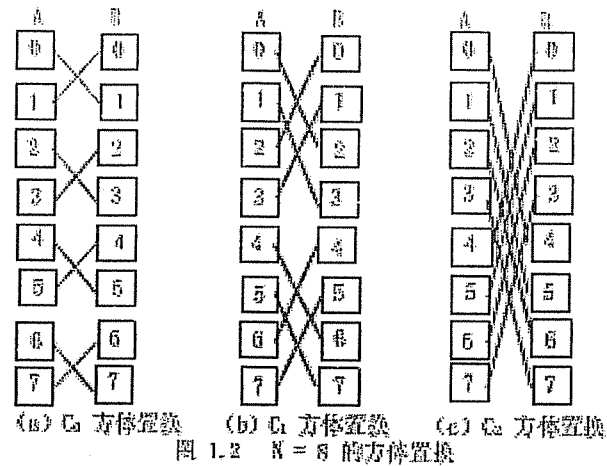
互连函数表达式为

$$C_0(x_2 x_1 x_0) = x_2 x_1 \bar{x}_0$$

$$C_1(x_2 x_1 x_0) = x_2 \bar{x}_1 x_0$$

$$C_2(x_2 x_1 x_0) = \bar{x}_2 x_1 x_0$$

本科生期末试卷 (2)



七. 解:

已知远程访问率 $p=0.5\%$, 远程访问时间 $t=2000\text{ns}$, 时钟周期 $T=10\text{ns}$

远程访问开销 $C=t/T=2000\text{ns}/10\text{ns}=200$ (时钟周期数)

有 0.5% 远程访问的机器的实际 CPI_2 为:

$$\text{CPI}_2 = \text{CPI}_1 + p \times C = 1.0 + 0.5\% \times 200 = 2.0$$

只有局部访问的机器的基本 $\text{CPI}_1 = 1.0$

$$\text{CPI}_2 / \text{CPI}_1 = 2.0 / 1.0 = 2 \text{ (倍)}$$

因此,没有远程访问状态下的机器速度是有 0.5% 远程访问的机器速度的 2

倍。

一. 填空题。(每小题 2 分, 共 14 分)

1. 衡量计算机性能的评价标准是程序执行的 A 或 B 。
2. 表示流水线的常用方法是 A 法和 B 法。
3. 指令级并行技术是指 A 并行性, 它不仅包括 B 并行技术, 还包括 C 并行技术。
4. 对向量的运算可以采用 A 方法, B 方法, C 方法, 但向量流水处理机中不适合采用 B 种方法。
5. 互连网络用来实现计算机系统内部多个处理机或多个功能部件之间的相互连接, 它是一种由开关元件按照一定的 A 结构和 B 方式构成的网络。
6. 并行处理机又称 A, 因为是单指令流多数数据流结构, 也称为 B 计算机。
7. 多处理机属于 A 计算机, 处理机之间按某种形式互连, 从而实现 B 之间的数据交换和同步。

二. 解释下列术语 (每个 3 分, 共 21 分)

1. 透明性
2. 重叠执行
3. 粗粒度并行性
4. 向量处理方式
5. 互连函数
6. 阵列处理机
7. 分布存储多处理机

三. (13 分) 假设系统某一部件的处理速度加快 9 倍, 但该部件的原处理时间仅为整个运行时间的 45%, 则采用加快措施后能使整个系统的性能提高多少? 如果部件改进后获得的加速比变为 $S_p=5.5$, 问改进前的执行时间占改进后的百分比是多少?

四. (13 分) 图 2.1 所示的静态加、乘双功能流水线中, 由段 S_1 、 S_2 、 S_3 、 S_4 、 S_6 组成乘法流水线, 由段 S_1 、 S_5 、 S_6 组成加法流水线。设向量 $a = (a_1, a_2, a_3, a_4)$, 向量 $b = (b_1, b_2, b_3, b_4)$, 计算 $\prod_{i=1}^4 (a_i + b_i)$ 。画出流水线时空图, 求 P、S、E。

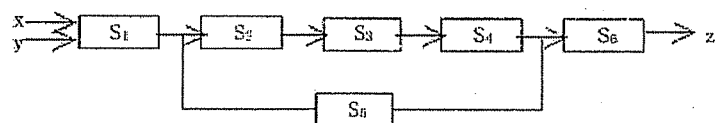


图2.1 连接图

五. (13分) 假定有多个加法器, 不存在加法器的资源冲突。有三条连续指令组成的程序代码如下:

I₁ ADD R1, R2, R4 ; R₁ ← (R₂)+(R₄)

I₂ ADD R2, R1, 1 ; R₂ ← (R₁)+1

I₃ SUB R1, R4, R5 ; R₁ ← (R₄)-(R₅)

(1) 分析程序代码段中的数据相关;

(2) 采用何种硬件技术可解决这些数据相关? 请加以说明。

六. (13分) 设网络结点数 N=8, 分别通过均匀洗牌置换网络和移数置换网络进行连接。试写出互连函数表达式, 并画出他们的连接图。

七. (13分) 何谓多处理机的 Cache 一致性问题? 出现这种问题的原因何在?

本科生期末试卷(2) 答案

一. 填空题。

1. A. 时间 B. 速率
2. A. 连接图 B. 时空图
3. A. 细粒度 B. 时间 C. 空间
4. A. 横向处理 B. 纵向处理 C. 纵横处理
5. A. 拓扑 B. 控制
6. A. 阵列处理机 B. SIMD
7. A. MIMD B. 程序

二. 解释术语。

1. 透明性: 在计算机技术中, 本来存在的事物或属性, 但从某种角度看又好像不存在的概念称为透明性。
2. 重叠执行: 例如一次重叠执行方式: 把执行第 k 条指令与取第 k+1 条指令同时进行。
3. 粗粒度并行性: 在多处理机上分别运行多个进程, 由多台处理机合作完成一个程序。
4. 向量处理方式: 对数组的运算称为向量处理。绝大多数向量处理机采用流水线结构, 从并行性考虑, 向量处理机采用纵向处理方式或纵横相结合的处理方式。
5. 互连函数: 在互连网络中, 用输入变量 x 表示输入, 用函数 f(x) 表示输出, 通过数学表达式建立输入输出端的一一对应关系。
6. 阵列处理机: 又称并行处理机, 主要技术手段是硬件上采用资源重复的方法来实现并行性。
7. 分布存储多处理机: 属于紧耦合系统, 它的共享存储器分布在每台处理机中, 每台处理机都带有自己的本地存储器, 组成一个“处理机-存储器”单元。但是这些分布在每台处理机中的实际存储器又合在一起统一编址, 在逻辑上组成一个共享存储器。这些处理机存储器单元通过互连网络连接在一起, 每台处理机除了能访问本地存储器外, 还能通过互连网络直接访问在其他处理机存储器单元中的“远程存储器”。

三. 解:

(1) 由题意可知, $F_c=0.45$, $S_c=9$, 代入公式:

$$S_p = \frac{T_0}{T_n} = \frac{1}{(1-F_c) + F_c/S_c}$$

$$= \frac{1}{0.55 + 0.45/9} = \frac{1}{0.64} = 1.56$$

(2) 根据题意有

$$S_p = \frac{1}{(1-F_c)+F_c/10} = 5.5$$

由此可得 $F_c = 91\%$

四. 解:

由于此流水线是双功能流水线, 计算要求先做加, 后做乘, 故应先设置加法功能, 连续计算出 (a_1+b_1) 、 (a_2+b_2) 、 (a_3+b_3) 、 (a_4+b_4) 四个加法, 然后切换为乘法功能, 按 $[(a_1+b_1) \times (a_2+b_2)] \times [(a_3+b_3) \times (a_4+b_4)]$ 做三次乘法。

图 2.2 中画出了流水线时空图, 其中 $A=a_1+b_1$, $B=a_2+b_2$, $C=a_3+b_3$, $D=a_4+b_4$ 。

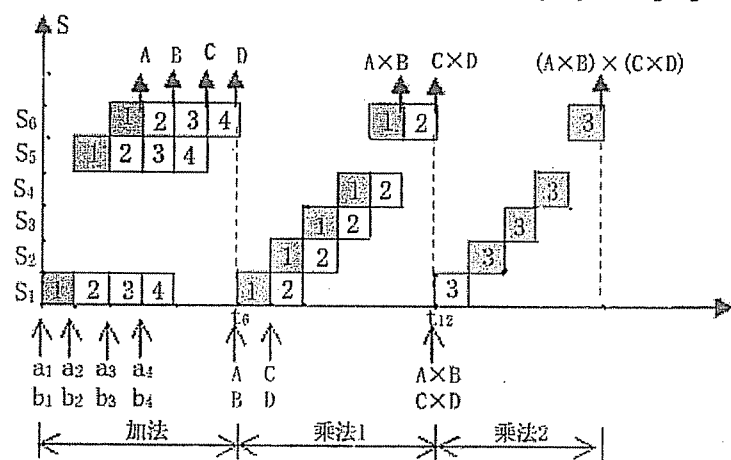


图2.2 时空图

由时空图看出, 在总共 17 个 Δt 时间内输出 7 个结果, 因此有

$$P = \frac{n}{T_k} = \frac{7}{17\Delta t}$$

当用串行方法完成上述操作时, 需 4 次加法和 3 次乘法, 求一次加法需 $3\Delta t$, 求一次乘法需 $5\Delta t$, 总共所需时间为

$$T_0 = 4 \times 3\Delta t + 3 \times 5\Delta t = 27\Delta t$$

$$\therefore S = \frac{T_0}{T_k} = \frac{27\Delta t}{17\Delta t} = 1.88$$

$$E = \frac{\text{有效时空区面积}}{\text{全部时空区总面积}} = \frac{3 \times 4\Delta t + 5 \times 3\Delta t}{6 \times 17\Delta t} = \frac{27}{102} = 0.264$$

五. 解:

(1) 指令 I_1 和 I_2 之间有 RW 相关, I_2 和 I_3 之间有 RW 相关, I_1 和 I_3 之

间有 WW 相关, I_1 和 I_2 之间还有 WR 相关。

(2) 对 I_1 和 I_2 之间的 WR 相关, 可用定向传送解决。根据寄存器重命名技术, 对引起 RW 相关的 I_2 中的 R_2 , 对引起 WW 相关的 I_3 中的 R_1 , 可分别换成备用寄存器 R_2' 、 R_1' 。经寄存器重命名后, 程序代码段实际执行时变为:

```

I1      ADD  R1, R2, R4
I2      ADD  R2', R1, 1
I3      SUB  R1', R4, R5

```

六. 解:

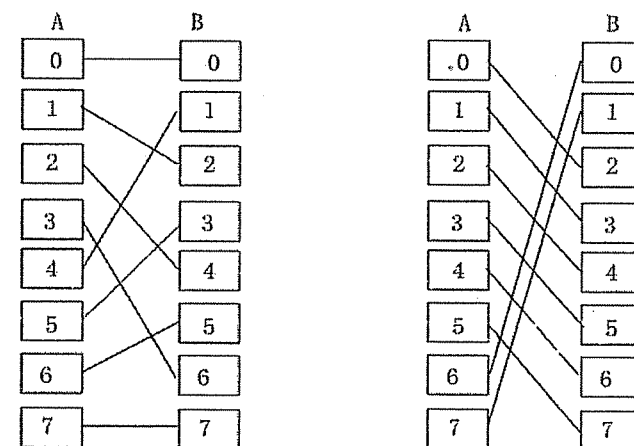
(1) 均匀洗牌置换互连函数表达式为:

$$\sigma(X_{n-1}X_{n-2} \cdots X_1X_0) = X_{n-2}X_{n-3} \cdots X_1X_0X_{n-1}$$

$$\sigma(X_7X_6X_5X_4X_3X_2X_1X_0) = X_6X_5X_4X_3X_2X_1X_0X_7$$

(2) 移数置换互连函数表达式为:

$$a(x) = (x+k) \bmod n \quad 0 \leq x \leq n$$



(1) N=8 的均匀洗牌置换

(2) N=8 的移数置换 $k=2$

图2.4

七. 解:

在多台处理机系统中的私有 Cache 会引起 Cache 中的内容相互之间以及共享存储器之间互不相同的问题, 称为多处理机的 Cache 一致性问题。出现 Cache 一致性问题原因主要有如下三点:

- (1) 共享可写数据引起的不一致性;
- (2) 进程迁移引起的数据不一致性;
- (3) I/O 传输所造成的数据不一致性。

本科生期末试卷 (3)

一. 填空题。(每小题 2 分, 共 14 分)

1. 一个理想的计算机系统结构应在 A、B 和 C 之间取得平衡。
2. 解决流水线瓶颈问题的方法有两种: 一是将瓶颈部分 A, 二是 B 瓶颈流水段, 让多个瓶颈流水段并行工作。
3. 机器并行性是指处理机获取 A 并行性好处的 B 大小。
4. 为了提高机器通用性, 向量处理机应同时具有处理 A 和处理 B 的功能。
5. 互连网络中结点数称为 A, 与结点相连的边数称为 B。
6. 阵列处理机操作模型用五元组表示: $SIMD = (N, C, I, M, R)$, 其中 N 表示机器的 A 数, I 表示由控制部件 CU 广播至所有 PE 进行并行执行的 B。
7. 多处理机与并行处理机的本质差别在于并行性级别的不同: 多处理机实现 A 的并行, 而并行处理机则实现同一指令多数据流的 B 的并行。

二. 解释下列术语(每个 3 分, 共 21 分)

1. 计算机实现
2. 流水线吞吐率
3. 指令级并行性
4. 向量和标量的平衡点
5. 静态网络
6. SIMD 计算机
7. SMP

三. (13 分) 有两台机器, CPU_A 实现一次条件转移要执行“比较”和“测试”两条指令, CPU_B 实现一次条件转移只需一条指令(比较、测试两种功能合一)。假设两台机器的指令系统中, 执行条件转移指令需两个时钟周期, 而其它指令只需一个时钟周期。又假设 CPU_A 中条件转移指令占总执行指令的 20% (比较指令也占 20%), CPU_B 执行条件转移指令时, 其时钟周期比 CPU_A 慢 25%。

问:

- (1) CPU_A 和 CPU_B 哪个速度更快?
- (2) 如果 CPU_B 的时钟周期比 CPU_A 慢 10%, 哪个 CPU 工作速度更快?

四. (13 分) 图 3.1 所示为一个 4 段的流水线连接图。

- (1) 设 $n=4$, 画出它的流水线时空图。
- (2) 求流水线的 P 、 P_{max} 、 S 、 E 值。

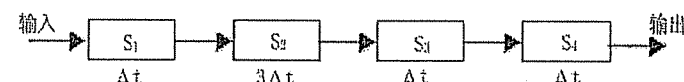


图 3.1 流水线连接图

五. (13 分) 用一台每个时钟周期发射两条指令的超标量处理机运行下面一段程序。所有指令都要进行取指 (IF)、译码 (ID)、执行、写结果 (WB) 四个阶段。其中 IF、ID、WB 三个阶段各为一个流水段, 其延迟时间都为 10ns。在执行阶段, LOAD 操作、AND 操作各延迟 10ns, ADD 操作延迟 20ns, MUL 操作延迟 30ns。这四种功能部件各设置一个, 它们可以并行工作。ADD 部件和 MUL 部件都采用流水结构, 每一级流水线的延迟时间都是 10ns。

```

I1  LOAD  R0,  M(A)           ; R0 ← M(A)
I2  ADD   R1,  R0             ; R1 ← (R1) + (R0)
I3  LOAD  R2,  M(B)           ; R2 ← M(B)
I4  MUL   R3,  R4             ; R3 ← (R3) × (R4)
I5  AND   R4,  R5             ; R4 ← (R4) ∧ (R5)
I6  ADD   R2,  R5             ; R2 ← (R2) + (R5)
    
```

请列出程序代码中所有的数据相关及其相关类型。

六. (13 分) 设网络结点数 $N=8$, 通过加 2^i 置换网络进行连接, 试写出互连函数表达式, 并画出它们的连接图。

七. (13 分) 设 a 为一个计算机系统中 n 台处理机可以同时执行的程序代码的百分比, 其余代码必须用单台处理机顺序执行。每台处理机的执行速率为 x MIPS, 并假设所有处理机的处理能力相同。

(1) 试用参数 n 、 a 、 x 推导出系统专门执行该程序时的有效 MIPS 速率表达式。

(2) 假设 $n=32$, $x=8$ MIPS, 要求得到的系统性能为 64 MIPS, 试求 a 值。

本科生期末试卷(3) 答案

一. 填空题。

1. A. 处理速度 B. 存储容量 C. I/O 吞吐量
2. A. 再细分 B. 重复设置
3. A. 指令级 B. 能力
4. A. 向量 B. 标量
5. A. 网络规模 B. 结点度
6. A. 处理单元 B. 指令集
7. A. 任务一级 B. 操作一级

二. 解释术语。

1. 计算机实现: 指的是计算机组成的物理实现, 包括处理机、主存等部件的物理结构, 器件的集成度和速度, 模块、插件、主板的划分与连接, 信号传输, 电源、冷却及整机装配技术等。
2. 流水线吞吐率: 是指在单位时间内流水线所完成的任务数量, 或是输出结果的数量。
3. 指令级并行性: 程序中的指令是顺序安排的, 当这些指令间不存在相关而能在流水线中通过时间重叠方法来并行执行时, 则存在指令级并行性。
4. 向量和标量的平衡点: 为了使向量硬件设备和标量硬件设备的利用率相等, 通常使用术语向量和标量的平衡点, 它定义为一个程序中向量代码所占的百分比。
5. 静态网络: 是指处理单元间有着固定连接的一类网络, 在程序执行期间, 这种点到点的链接保持不变。
6. SIMD 计算机: 阵列处理机又称并行处理机, 主要技术手段是硬件上采用资源重复的方法来实现并行性。因为是单指令流多数据流结构, 也称为 SIMD 计算机。
7. SMP: 共享存储型多处理机 SMP(shared memory multiprocessors), 也称为对称型多处理机。它又分为 UMA、NUMA、COMA 三种结构模型, 彼此的区别在于存储器和外围资源如何共享或分布。

三. 解:

(1) 设 CPU_A 的时钟周期长度为 t_A , CPU_B 的时钟周期长度为 t_B

$$CPI_A = 0.2 \times 2 + 0.8 \times 1 = 1.2$$

$$T_{CPUA} = I_{NA} \times 1.2 \times t_A$$

CPU_B 中由于没有比较指令, 使转移指令由原来占 20% 上升为 20% ÷ 80% = 25%, 它需要两个时钟周期, 而其余的 75% 指令只需 1 个时钟周期, 所以

$$CPI_B = 0.25 \times 2 + 0.75 \times 1 = 1.25$$

CPU_B 中由于没有比较指令, 因此 $I_{NB} = 0.8 \times I_{NA}$ 。又因 $t_B = 1.25t_A$, 所以

$$T_{CPUB} = I_{NB} \times CPI_B \times t_B = 0.8 I_{NA} \times 1.25 \times 1.25 t_A = 1.25 I_{NA} \times t_A$$

可见 $T_{CPUA} < T_{CPUB}$, 故 CPU_A 比 CPU_B 运行得更快些。

$$(2) T_{CPUA} = 1.2 I_{NA} \times t_A$$

因为 $t_B = 1.1 t_A$, 故

$$T_{CPUB} = 0.8 I_{NA} \times 1.25 \times 1.1 t_A = 1.1 I_{NA} \times t_A$$

由于 $T_{CPUB} < T_{CPUA}$, 故 CPU_B 比 CPU_A 运行得更快些。

四. 解:

(1)

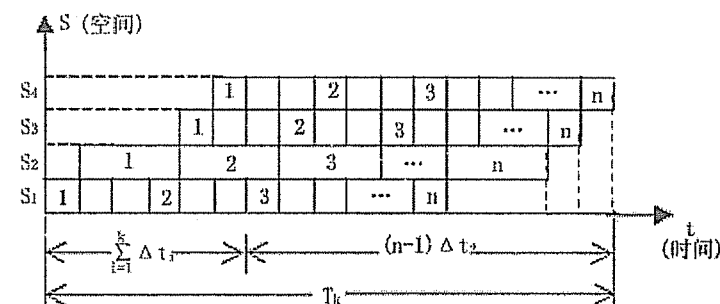


图3.2 流水线时空图

(2) 由时空图看出, $n = 4$, $T_k = t_{15} = 15 \Delta t$

$$P = \frac{n}{T_k} = \frac{4}{15 \Delta t} \quad P_{\max} = \frac{1}{3 \Delta t}$$

$$T_0 = n \cdot k \Delta t = 4 \cdot (1+1+3+1) \Delta t = 24 \Delta t$$

$$S = \frac{T_0}{T_k} = \frac{24 \Delta t}{15 \Delta t} = 1.6$$

$$\text{时空图有效面积} = 3 \times (1 \times \Delta t \times 4) + (1 \times 3 \Delta t) \times 4 = 24 \Delta t$$

$$4 \text{ 个流水段所围总面积} = 4 \times t_{15} = 4 \times 15 \Delta t = 60 \Delta t$$

$$E = \frac{24 \Delta t}{60 \Delta t} = \frac{2}{5} = 0.4$$

五. 解:

指令 I_1 、 I_2 间有寄存器 R_0 的 WR 相关;

指令 I_3 、 I_6 间有寄存器 R_2 的 WR 相关;

指令 I_4 、 I_5 间有寄存器 R_4 的 RW 相关;

指令 I_3 、 I_6 间有寄存器 R_2 的 WW 相关。

六. 解:

$+2^i$ 互连函数表达式为

$$PM_{2+i}(x) = x + 2^i \bmod N$$

其中 $0 \leq x \leq N-1$ $0 \leq i \leq n-1$ $n = \log_2 N$

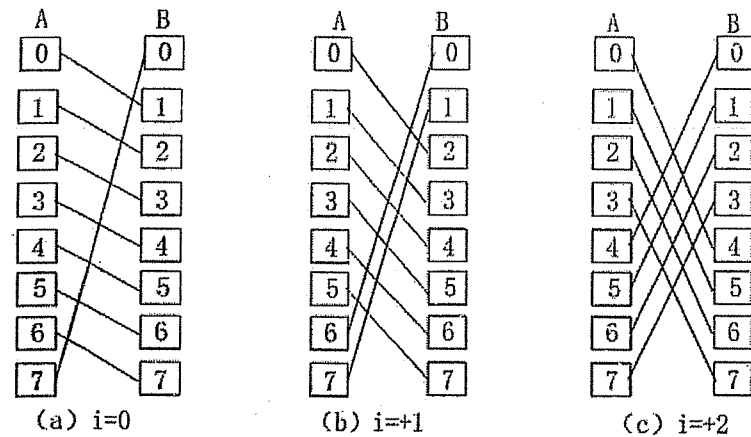


图3.3 N=8 的 PM2I 置换

七. 解:

(1) 设总指令数为 m , 并行代码数为 $m(p)$, 顺序代码数为 $m(s)$, 则总执行时间 T 为:

$$T = \frac{m(p)}{nx} + \frac{m(s)}{x} = \frac{ma}{nx} + \frac{m(1-a)}{x}$$

有效 MIPS 表达式为:

$$\frac{m}{T} = \frac{nx}{n(1-a) + a} \quad (3.1)$$

(2) 由表达式 (3.1), 代入已知条件, 求得 $a = 0.96$

本科生期末试卷 (4)

一. 填空题。(每小题 2 分, 共 14 分)

- 通过 A 技术提高计算机系统的性能, 正是计算机 B 研究的主要内容。
- 衡量流水线性能的主要技术指标是 A、B 和 C。
- 指令级并行度是指在一个 A 内流出的 B 数目。
- 向量流水处理机的基本思想是把两个向量的 A 进行计算, 产生一个 B 向量。一条向量指令可以处理 C 操作数。
- 互连网络中, 两结点间相连的最少边数称为 A。网络中任意两个结点间距离的最大值称为 B。
- 从处理单元看, 阵列处理机可看成是一个 A 并行机, 但实际的阵列处理机系统是一个 B 多处理机系统。
- 多处理机实现 A 的并行性, 并行性存在于 B 外部。

二. 解释下列术语 (每个 3 分, 共 21 分)

- Amdahl 定律
- 流水线加速比
- 机器并行性
- 链接技术
- 动态网络
- 标量处理机
- MPP

三. (13 分) 已知四个程序在 A、B、C 三台计算机上的执行时间如下表所示:

程序	执行时间		
	计算机A	计算机B	计算机C
程序1	1	10	20
程序2	1000	100	20
程序3	500	1000	50
程序4	100	800	100

- 假设四个程序中每一个都有 100000000 条指令要执行, 计算这三台计算机中每台机器上每个程序的 MIPS 速率;
- 用算术平均 (A_m) 和调和平均 (H_m) 法来比较三台计算机的相对性能。

四. (13 分) 图 4.1 所示为一个四段的流水线连接图。

- 设 $n=12$, 画出瓶颈段再细分的流水线连接图和时空图;
- 求改进流水线的 P 、 P_{max} 、 S 、 E 值。

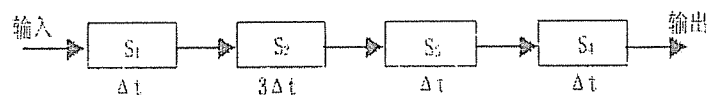


图4.1 流水线连接图

- 五. (13 分) 用一台每个时钟周期发射两条指令的超标量处理机运行下面一段程序。所有指令都要进行取指 (IF)、译码 (ID)、执行、写结果 (WB) 四个阶段。其中 IF、ID、WB 三个阶段各为一个流水段，其延迟时间都为 10ns。在执行阶段，LOAD 操作、AND 操作各延迟 10ns，ADD 操作延迟 20ns，MUL 操作延迟 30ns。这四种功能部件各设置一个，它们可以并行工作。ADD 部件和 MUL 部件都采用流水结构，每一级流水线的延迟时间都是 10ns。

I_1 LOAD R0, M(A) ; $R_0 \leftarrow M(A)$
 I_2 ADD R1, R0 ; $R_1 \leftarrow (R_1) + (R_0)$
 I_3 LOAD R2, M(B) ; $R_2 \leftarrow M(B)$
 I_4 MUL R3, R4 ; $R_3 \leftarrow (R_3) \times (R_4)$
 I_5 AND R4, R5 ; $R_4 \leftarrow (R_4) \wedge (R_5)$
 I_6 ADD R2, R5 ; $R_2 \leftarrow (R_2) + (R_5)$

假设所有运算型指令都在译码(ID)流水段读寄存器，在写结果(WB)流水段写寄存器，采用顺序发射顺序完成的调度方法。

- (1) 画出流水线的时空图；
- (2) 计算执行这个程序所用的时间。

- 六. (13 分) 如图所示，输入端为 8 个处理机，输出端为 8 个存储器，通过三级立方体互连网络连接，采用级控方式。其中所有交换开关均为二功能(控制信号为“0”时直通，为“1”时交换)。若级控信号为：① $K_0K_1K_2=100$ ；② $K_0K_1K_2=110$ ；③ $K_0K_1K_2=111$ ；请在表中填写出对应于 8 个处理机而实际连通的 8 个存储器的排列顺序。

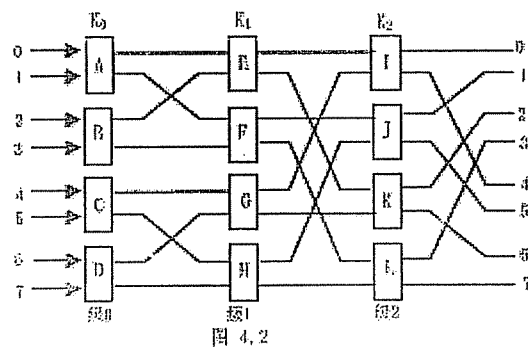


图 4.2

输入排列 (处理机)	输出排列 (存储器)		
	$K_0K_1K_2=100$	$K_0K_1K_2=110$	$K_0K_1K_2=111$
0			
1			
2			
3			
4			
5			
6			
7			

- 七. (13 分) 请用块结构语言 Cobegin-Coend 写出图 4.3 中所示嵌套并行算法优先关系图的程序。

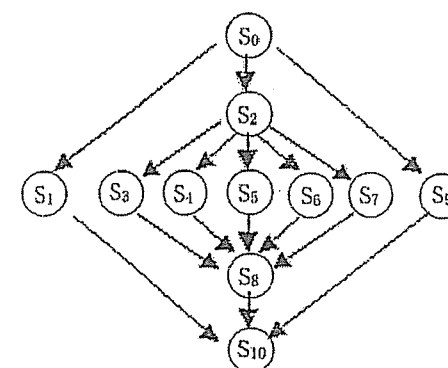


图 4.3

本科生期末试卷(4) 答案

一. 填空题。

- 1. A. 并行性 B. 系统结构
- 2. A. 吞吐量 B. 加速比 C. 效率
- 3. A. 时钟周期 B. 指令
- 4. A. 对应分量 B. 结果 C. N 个或 N 对
- 5. A. 距离 B. 网络直径
- 6. A. 同构型 B. 异构型
- 7. A. 任务一级 B. 指令

二. 解释术语。

- 1. Amdahl 定律: 加快某部件执行速度所获得的系统性能加速比, 受限于该部件在系统中的重要性。
- 2. 流水线加速比: 完成同样一批任务, 不使用流水线所用的时间与使用流水线所用的时间之比称为流水线加速比。
- 3. 机器并行性: 是指处理机获取指令级并行性好处的能力大小。
- 4. 链接技术: 利用向量指令间存在的先写后读的数据相关性来加快向量指令序列执行速度的技术称为链接技术。
- 5. 动态网络: 是用交换开关构成的, 可按运行程序的要求动态地改变连接状态的网路。
- 6. 标量处理机: 处理机在处理数据时以标量为基本处理对象。
- 7. MPP: 即大规模并行处理, 按照当前的标准, 具有几百台~几千台处理机的任何机器都是大规模并行处理系统。

三. 解:

(1) 因为 $MIPS = I_N / (T \times 10^6)$, 所以每台计算机每个程序的 MIPS 速率见下表所示:

不同程序在三台计算机上的 MIPS 速率

程序	MIPS 速率 (百万指令/秒)		
	计算机A	计算机B	计算机C
程序 1	100	10	5
程序 2	0.1	1	5
程序 3	0.2	0.1	2
程序 4	1	0.125	1

(2) A_m 、 H_m 法求得的数据见下表所示:

不同计算机 MIPS 速率与平均执行时间

平均执行时间	MIPS 速率 (百万指令/秒)		
	计算机A	计算机B	计算机C
算术平均 (A_m)	25.3	2.81	3.25
调和平均 (H_m)	0.25	0.20	2.1

从表看出, 按 A_m 比较性能, A 机最快, B 机最慢。按 H_m 比较性能, 则 C 机最快, B 机最慢。

四. 解:

(1)

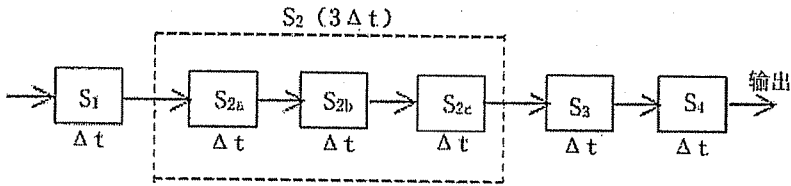


图4.3 流水线连接图

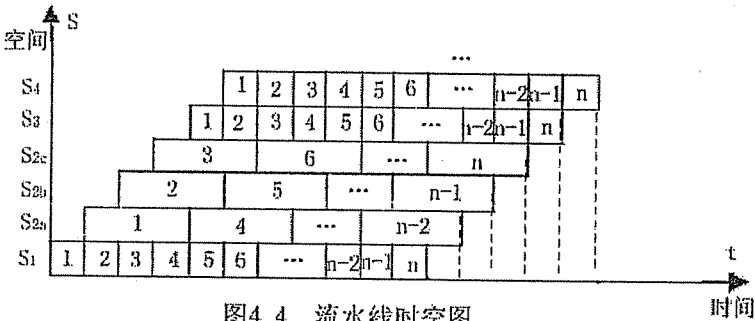


图4.4 流水线时空图

(2) $n = 12$, $T_k = t_{17} = 17 \Delta t$, $k = 6$

$$P = \frac{n}{T_k} = \frac{12}{17 \Delta t} \quad P_{\max} = \frac{1}{\Delta t}$$

$$S = \frac{nk}{k+n-1} = \frac{12 \times 6}{6+11} = \frac{72}{17} = 4.24$$

$$E = \frac{n}{k+n-1} = \frac{12}{17} = 0.7$$

$$E = \frac{S}{k} = 0.7$$

五. 解:

(1) 采用顺序发射顺序完成调度方法的流水线时空图如下图所示:

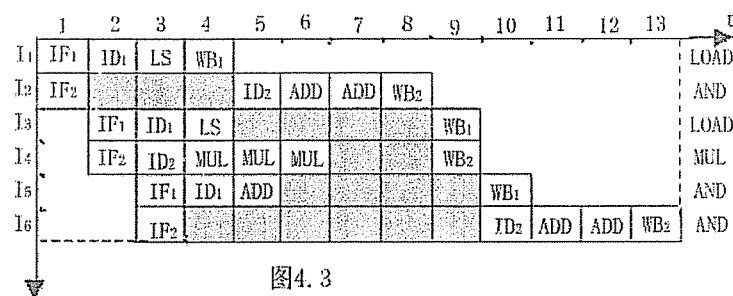


图4.3

(2) 从时空图看到, 执行这个程序共用 130ns。

六. 解:

输入排列 (处理机)	输出排列 (存贮器)		
	$K_0K_1K_2=100$	$K_0K_1K_2=110$	$K_0K_1K_2=111$
0	4	6	7
1	5	7	6
2	6	4	5
3	7	5	4
4	0	2	3
5	1	3	2
6	2	0	1
7	3	1	0

七. 解:

```

begin
  s0;
  Cobegin
    s1;
    begin
      s2;
      Cobegin s3; s4; s5; s6; s7; Coend
    end
    s8;
  end
  s9;
Coend
s10;
end

```

本科生期末试卷 (5)

一. 填空题。(每小题 2 分, 共 14 分)

- 并行性技术包括 ____A____ 并行、 ____B____ 并行、 ____C____ 并行和资源共享四种方法。
- 指令流水线中, 如果硬件资源满足不了指令重叠执行的要求, 便会发生 ____A____ 冲突, 则称该流水线有 ____B____ 相关。
- 衡量指令级并行性的一个指标是 ____A____, 要提高指令级并行性, 人们追求的目标是 A ____B____ 1。
- 向量流水处理机采用 ____A____ 结构或 ____B____ 结构。
- 互连网络中网络被切成相等的两半时, 沿切口的最小边数称为 ____A____, 任两个结点间线的长度称为 ____B____。
- 阵列处理机的基本结构可分为 ____A____ 的阵列机和 ____B____ 的阵列机两大类。
- 多处理机实现 ____A____ 的并行性。一般来说, 在同一时刻, 不同的处理机执行着不同的 ____B____。

二. 解释下列术语。(每个 3 分, 共 21 分)

- MIPS
- 流水线效率
- 指令级并行度
- 分段开采
- 自适应寻径
- 阵列控制器
- 机群

三. (13 分) 用一台 40 MHz 处理机执行标准测试程序, 它包含的混合指令数和相应所需的平均时钟周期数如下表所示:

指令类型	指令数	平均时钟周期数
整数运算	45000	1
数据传送	32000	2
浮点运算	15000	2
控制传送	8000	2

求: 有效 CPI、MIPS 速率和程序执行时间 T。

四. (13 分) 图 5.1 所示为一个四段的流水线连接图。

- 设 $n=12$, 画出瓶颈段重复设置的流水线连接图和时空图;
- 改进流水线的 P、 P_{max} 、S、E 值。

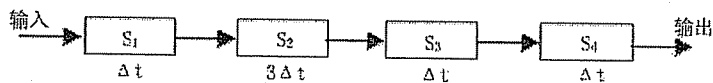


图 5.1 流水线连接图

- 五. (13 分) 用一台每个时钟周期发射两条指令的超标量处理机运行下面一段程序。所有指令都要进行取指 (IF)、译码 (ID)、执行、写结果 (WB) 四个阶段。其中 IF、ID、WB 三个阶段各为一个流水段，其延迟时间都为 10ns。在执行阶段，LOAD 操作、AND 操作各延迟 10ns，ADD 操作延迟 20ns，MUL 操作延迟 30ns。这四种功能部件各设置一个，它们可以并行工作。ADD 部件和 MUL 部件都采用流水结构，每一级流水线的延迟时间都是 10ns。

I_1 LOAD R0, M(A) ; $R_0 \leftarrow M(A)$
 I_2 ADD R1, R0 ; $R_1 \leftarrow (R_1) + (R_0)$
 I_3 LOAD R2, M(B) ; $R_2 \leftarrow M(B)$
 I_4 MUL R3, R4 ; $R_3 \leftarrow (R_3) \times (R_4)$
 I_5 AND R4, R5 ; $R_4 \leftarrow (R_4) \wedge (R_5)$
 I_6 ADD R2, R5 ; $R_2 \leftarrow (R_2) + (R_5)$

假设所有运算型指令都在译码 ID 流水段读寄存器，在写结果 WB 流水段写寄存器，采用顺序发射乱序完成的调度方法：

- (1) 画出流水线的时空图和各条指令完成的时间图；
- (2) 计算执行这个程序所用的时间。

- 六. (13 分) 要求用直径最小的三维环网、六维二元超立方体和带环立方体(CCC)设计一台由 64 个结点组成的多计算机直接连接网络。令 d 为网络的结点度， D 为直径， l 为链路总数。假设网络的性能可用 $(d \times D \times l)^{-1}$ 来衡量，试根据其性能列表排出上述 3 种网络的名次。

- 七. (13 分) 一台单处理机采用如下串行程序实现 $A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8$ 各矩阵 (均为 $n \times n$) 的累加求和运算：

$A_1 = A_1 + A_2$
 $A_3 = A_3 + A_4$
 $A_5 = A_5 + A_6$
 $A_7 = A_7 + A_8$
 $A_1 = A_1 + A_3$
 $A_5 = A_5 + A_7$
 $A_1 = A_1 + A_5$

现用 4 台多处理机系统实现上述矩阵的求和运算。请画出任务优先关系图，并用块结构语言 Cobegin-Coend 写出并行程序。

本科生期末试卷 (5) 答案

一. 填空题。

1. A. 时间 B. 空间 C. 时间+空间并行
2. A. 资源 B. 结构
3. A. CPI B. 小于等于
4. A. 存储器-存储器 B. 寄存器-寄存器
5. A. 等分宽度 B. 结点间线长
6. A. 分布式存储器 B. 共享式存储器
7. A. 粗粒度 B. 指令

二. 解释术语。

1. MIPS : 即 million instructions per second , 表示每秒百万条指令数。
2. 流水线效率: 流水线的设备利用率称为流水线的效率。
3. 指令级并行度: 在一个时钟周期内流水线上流出的指令数。
4. 分段开采: 当向量的长度大于向量寄存器的长度时, 必须把长向量分成长度固定的段, 分段开采技术一次处理一个向量段。
5. 自适应寻径: 通信的通路每次都要根据通信资源或者网络的情况来选择, 这样就可以避开拥挤的或有故障的结点, 从而使网络的利用效率可以得到改进。
6. 阵列控制器: 阵列控制器实际上是一台计算机, 它除了对阵列的处理单元实行控制以外, 还能利用本身的内部资源执行一整套指令, 用以完成标量操作, 且在时间上与各 PE 的数组操作并行进行。
7. 机群: 是并行或分布计算机系统的一种类型, 它是由一组完整的计算机 (结点) 通过高性能的网络或局域网互连而成的系统, 它作为一个单独的统一计算资源来使用。

三. 解:

$$CPI = \sum_{i=1}^n (CPI_i \times \frac{I_i}{I_N})$$

$$= \frac{45000 \times 1 + 32000 \times 2 + 15000 \times 2 + 8000 \times 2}{45000 + 32000 + 15000 + 8000} = 1.55 \text{ (周期/指令)}$$

$$MIPS = \frac{f}{CPI \times 10^6} = \frac{40 \times 10^6}{1.55 \times 10^6} = 25.81 \text{ (百万指令/秒)}$$

$$T = \frac{45000 \times 1 + 32000 \times 2 + 15000 \times 2 + 8000 \times 2}{40 \times 10^6} = 3.875 \times 10^{-3} \text{ (秒)}$$

四. 解:

(1)

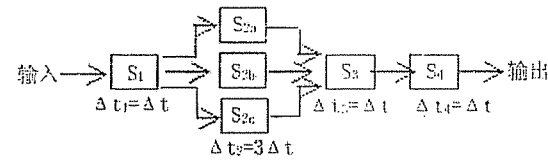


图5.2 流水线连接图

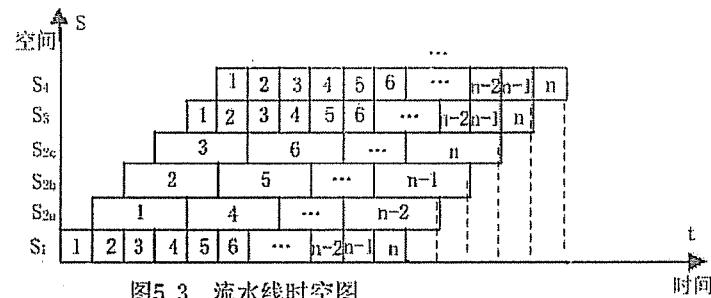


图5.3 流水线时空图

(2) $n = 12, k = 6$

$$P = \frac{n}{T_k} = \frac{12}{17\Delta t} \quad P_{\max} = \frac{1}{\Delta t}$$

$$S = \frac{nk}{k+n-1} = \frac{12 \times 6}{6+11} = 4.24$$

$$E = \frac{n}{k+n-1} = \frac{12}{17} = 0.7$$

五. 解:

(1) 采用顺序发射乱序完成调度方法的流水线时空图, 如图 5.4 (a) 所示, 各条指令完成的时间图如图 5.4 (b) 所示。

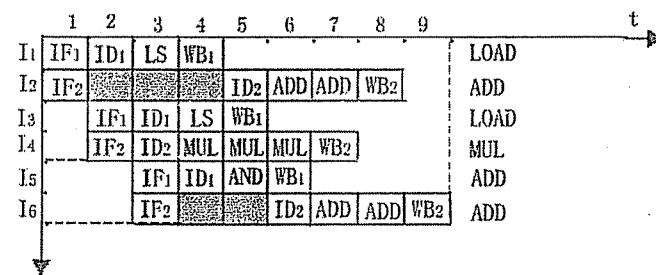


图5.4 (a) 流水线时空图

时钟周期	4	5	6	7	8	9	10
流水线 1	I1	I3	I5	空	空	空	空
流水线 2	忙	忙	忙	I4	I2	I6	空

图 5.4 (b) 各条指令完成的时间图

(2) 执行这个程序共用 90ns。

六. 解:

三种网络的性能如下表所示:

网络	d	D	l	$(d \times D \times l)^{-1}$	名次
3 维网络	6	6	192	1/6912	2
6 维超立方体	6	6	192	1/6912	2
带环立方体	3	9	96	1/2592	1

七. 解:

根据所给条件, 采用 4 台处理机时串程序可做如下改造:

- S0 传送数据
- S1 $A_1 = A_1 + A_2$
- S2 $A_3 = A_3 + A_4$
- S3 $A_5 = A_5 + A_6$
- S4 $A_7 = A_7 + A_8$
- S5 传送数据
- S6 $A_1 = A_1 + A_3$
- S7 $A_5 = A_5 + A_7$
- S8 传送数据
- S9 $A_1 = A_1 + A_5$

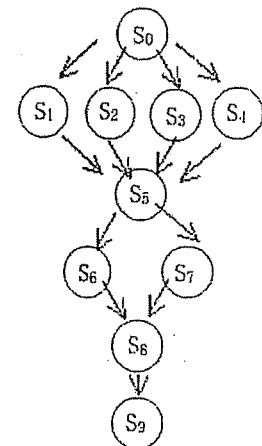


图 5.5

由此可画出并行优先关系图如图 5.5 所示。

并行程序如下:

```

begin
  S0;
  Cobegin  S1; S2; S3; S4; Coend
  S5;
  Cobegin  S6; S7; Coend
  S8;
  S9;
end
  
```

本科生期末试卷 (6)

一. 填空题。(每小题 2 分, 共 14 分)

1. 由软件实现的机器称为 A。在一个计算机系统中, 低层机器的属性对高层机器的程序员往往是 B 的。
2. 当指令在流水线中重叠执行时, 流水线有可能改变指令读/写操作数的顺序, 从而导致 A 供求关系上的冲突, 称为 B 相关。
3. Intel 公司的 Pentium, Motorola 公司的 MC88110, IBM 公司的 Power6000, SUN 公司的 Super SPARC 是 A 处理机。SGI 公司的 MIPS R4000 是 B 处理机。DEC 公司的 Alpha 是 C 处理机。
4. 为了消除存储器争用冲突, 存储器—存储器结构的向量机采用 A 存储器, 并在运算器输入和输出端增加 B。
5. 互连网络传输消息的最大速率称为 A, 采用的单位是 B。
6. Illiac IV 阵列机是一个由以下三种类型处理机联合组成的多机系统: (1) 处理部件阵列 PU 专门用于 A 运算; (2) 阵列控制器 CU 是一台相对独立的小型 B 处理机; (3) B6700 计算机担负 I/O 系统和操作系统管理功能。
7. 在多台处理机中, 为使并发进程之间保持 A 所需要的正确顺序, 必须采取特殊的 B 措施。

二. 解释下列术语。(每个 3 分, 共 21 分)

- | | | |
|---------------|------------|-----------|
| 1. 计算机组成 | 2. 结构相关 | 3. 寄存器重命名 |
| 4. 向量循环 | 5. 网络直径 | |
| 6. 分布式存储器的阵列机 | 7. 虚拟共享存储器 | |

三. (13 分) 设有一个 20000 条指令的程序在一台时钟速率为 40MHz 的流水线处理机上执行。假设该指令流水线有 5 段, 并且每个时钟周期发射一条指令, 忽略转移指令造成的损失。

- (1) 用该流水线执行此程序, 画出流水线的时空图。
- (2) 计算流水处理机的加速比、吞吐率和效率。

四. (13 分) 用一台每个时钟周期发射两条指令的超标量处理机运行下面一段程序。所有指令都要进行取指 (IF)、译码 (ID)、执行、写结果 (WB) 四个阶段。其中 IF、ID、WB 三个阶段各为一个流水段, 其延迟时间都为 10ns。在执行阶段, LOAD 操作、AND 操作各延迟 10ns, ADD 操作延迟 20ns, MUL 操作延迟 30ns。这四种功能部件各设置一个, 它们可以并行工作。ADD 部件和 MUL 部件都采用流水结构, 每一级流水线的延迟时间都是 10ns。

I ₁	LOAD	R0,	M(A)		; R ₀ ← M(A)
I ₂	ADD	R1,	R0		; R ₁ ← (R ₁) + (R ₀)
I ₃	LOAD	R2,	M(B)		; R ₂ ← M(B)
I ₄	MUL	R3,	R4		; R ₃ ← (R ₃) × (R ₄)
I ₅	AND	R4,	R5		; R ₄ ← (R ₄) ∧ (R ₅)
I ₆	ADD	R2,	R5		; R ₂ ← (R ₂) + (R ₅)

假设每个操作部件的输出端都有直接数据通路输入端相连, 采用顺序发射乱序完成调度方法:

- (1) 画出流水线的时空图和各条指令完成的时间图;
- (2) 计算执行这个程序所用的时间。

五. (13 分) 有一个循环程序代码段:

```
DO 10 i=1, n
10 A(i)=5*B(i)+C
```

其中 n 和 C 为常数, 设向量寄存器长度为 64 元素, 请用分段开采技术改造成向量循环形式。

六. (13 分) 设 E 为交换函数, S 为均匀洗牌函数, B 为蝶式函数, PM2I 为移数函数, 函数的自变量是十进制数表示的处理机编号。现有 32 台处理机, 其编号为 0、1、2、……、31。

(1) 分别计算下列互连函数:

$E_2(12)$ $S(8)$ $B(9)$ $PM2I_{+3}(28)$ $E_0(S(4))$ $S(E_0(18))$

(2) 用 E_0 和 S 构成均匀洗牌交换网 (每步只能使用 E_0 和 S 一次), 网络直径是多少? 从 5 号处理机发送数据到 7 号处理机, 最短路径需要经过几步? 列出经过的处理机编号。

(3) 采用移数函数构成互连网, 网络直径是多少? 结点度是多少? 与 2 号处理机距离最远的是几号处理机?

七. (13 分) 有三个程序段 P_1 、 P_2 、 P_3 , 分别求 3 个矩阵算术表达式:

$P_1: X = (A+B) \times (A-B)$

$P_2: Y = (C+1) \times (C+D)$

$P_3: Z = X \times Y$

其中 A, B, C, D, X, Y, Z 均为 $n \times n$ 矩阵。请判别 P_1 , P_2 , P_3 程序段的并行性。

本科生期末试卷 (6) 答案

一. 填空题。

1. A. 虚拟机器 B. 透明
2. A. 数据 B. 数据
3. A. 超标量 B. 超流水线 C. 超标量超流水线
4. A. 多模块 B. 延迟缓冲器
5. A. 频带 B. 位/秒 或 兆位/秒
6. A. 数组 B. 标量
7. A. 程序 B. 同步

二. 解释术语。

1. 计算机组成：指的是计算机系统结构的逻辑实现，包括物理机器级内的数据流和控制流的组成以及逻辑设计等。
2. 结构相关：某些指令组合在流水线重叠执行过程中，如果硬件资源满足不了指令重叠执行的要求，便会产生资源冲突，则称该流水线有结构相关。
3. 寄存器重命名：RW 和 WW 相关，从根本上说是存储位置的资源冲突。凡是资源冲突都可以设法使资源增加成多套的方法来减少冲突的发生。既然 RW 和 WW 相关都是由于写入寄存器引起的，因此可以将寄存器增加成多套的方法来解决，这就是寄存器重命名技术。
4. 向量循环：处理长向量的程序结构称为向量循环。
5. 网络直径：网络中任意两个结点之间距离的最大值，用结点之间的连接边数表示。
6. 分布式存储器的阵列机：它含有多个同样结构的处理单元 PE，通过数据寻径网络以一定方式互相连接。每个 PE 有各自的本地存储器 LM。在阵列控制部件的统一指挥下，实现并行操作。
7. 虚拟共享存储器：是在基于分布存储器的多处理机上，实现物理上分布但逻辑上共享的存储系统。其基本思想是：将物理上分散的各个处理机所用的局部存储器，在逻辑上加以统一编址，形成一个统一的虚拟地址空间来实现存储器的共享。每个处理机可以访问全局存储器的任一位置，用户可以把它当成全局共享存储系统。

三. 解：

(1) 流水线时空图如下图所示：

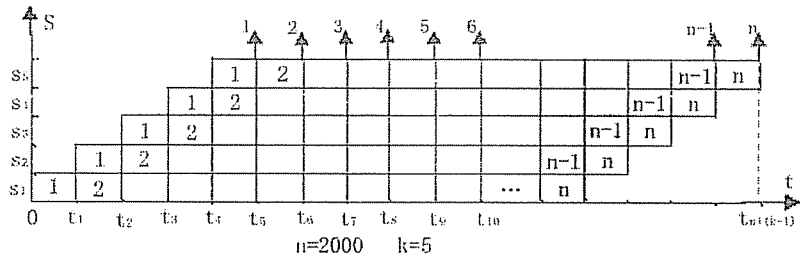


图 6.1

(2) 按题意知 $n=20000$ ， $k=5$ ， $f=40 \times 10^6/s$

$$S = \frac{nk}{k + (n-1)} = \frac{20000 \times 5}{5 + (20000 - 1)} = 5.00$$

$$E = \frac{nk \Delta t}{k[k + (n-1)] \Delta t} = \frac{20000}{20004} = 1.00$$

$$P = \frac{nf}{k + (n-1)} = \frac{20000 \times 40 \times 10^6}{5 + (20000 - 1)} = 4.00 \times 10^7 \text{ (MIPS)}$$

四. 解：

(1) 采用顺序发射乱序完成调度方法的流水线时空图见图 6.2 (a) 所示，各条指令完成的时间图见图 6.2 (b) 所示。

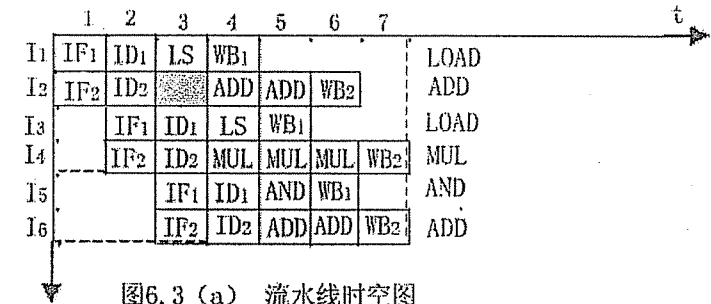


图6.3 (a) 流水线时空图

时钟周期	4	5	6	7	8
流水线 1	I ₁	I ₃	I ₅	I ₆	空
流水线 2	忙	忙	I ₂	I ₄	空

图6.3 (b) 各条指令完成的时间图

(2) 执行这个程序共用 70ns。

五. 解：

LOW=1
 $VL = (n \text{ mod } 64)$: 找出余数长度值
 DO 20 j=1, (n/64) : 外循环

本科生期末试卷 (7)

```

DO 10 i=LOW, LOW+VL-1 ; 以长度VL操作
    A(i)=5*B(i)+C      ; 主要操作
10 continue
LOW=LOW+VL             ; 下一向量的开始
VL=64                  ; 将向量长度值恢复成64
20 continue

```

如果用指令实现, 当 $n \leq 64$ 时, 产生 A 数组的 7 条指令序列完成的操作是:

```

LD    S1, M(5)          ; 常数 5 装入标量寄存器 S1
LD    S2, M(C)           ; 常数 c 装入标量寄存器 S2
LD    V, M(n)            ; 在 VL 寄存器内设置向量长度
LV    V0, B              ; 将 B 向量读入向量寄存器 V0
MULV  V1, S1, V0          ; B 数组的每个分量和常数 c 相乘
ADDV  V2, S2, V1          ; C 和 5*B(X) 相加得结果向量
SV    A, V2              ; 结果向量存入 A 数组

```

六. 解:

$$(1) E_2(12) = 8 \quad S(8) = 16 \quad B(9) = 24 \quad PM2I_{+3}(28) = 4$$

$$E_0(S(4)) = E_0(8) = 9 \quad S(E_0(18)) = S(19) = 7$$

(2) 网络直径是 9, 最短路径需要经过 6 步。

00101 E₀00101 S01000 E₀01001 S10010 E₀10011 S00111

(3) 网络直径是 3, 结点度是 9, 与 2 号处理机距离最远的是 13、15、21、23 号处理机。

七. 解:

根据伯恩斯坦准则, 输入输出变量集分别为:

$$I_1 = \{A, B\}, \quad I_2 = \{C, D\}, \quad I_3 = \{X, Y\}$$

$$O_1 = \{X\}, \quad O_2 = \{Y\}, \quad O_3 = \{Z\}$$

① 由于 $I_1 \cap O_2 = \emptyset$, $I_2 \cap O_1 = \emptyset$, $O_1 \cap O_2 = \emptyset$, 故 P_1 、 P_2 两个程序段可以并行执行;

② 由于 $I_3 \cap O_1 \neq \emptyset$, $I_3 \cap O_2 \neq \emptyset$, 故 P_3 程序段必须在 P_1 、 P_2 程序段执行完毕后方可执行。

一. 填空题。(每小题 2 分, 共 14 分)

- 按指令流和数据流的不同组织方式, 计算机系统结构分为 A、B、C、D 四类。
- 指令流水线中, 由于程序执行 A 类指令而引起的相关, 称为 B 相关。它会引起程序执行方向的改变, 所以是 C 相关。
- 数据相关分三类, 它们是 A 相关、B 相关、C 相关。
- CRAY-1 是一台典型的 A 结构的向量处理机, 其特色是主存与流水结构的运算器之间有一级或两级的 B, B 的速度要比 Cache C。
- 消息通过互连网络的时间称为 A, 它等于 B 除以频宽。
- Illiac IV 阵列 PU 由 64 个 A、64 个 B 和存储器逻辑部件 (MLU) 组成。
- 设 R 代表程序的执行时间, C 代表用于通讯的开销。在 A 并行情况下, R/C 比值比较 B, 每个单位计算只需少量的通信。

二. 解释下列术语。(每个 3 分, 共 21 分)

- 系列机
- 数据相关
- 动态指令调度
- 向量指令执行时间
- 等分宽度
- 屏蔽方案
- Cache 一致性

三. (13 分) 浮点加法器流水线如图所示。用此线性流水线计算 8 个浮点数的和, 即 $Z = a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8$ 。

(1) 画出求 8 个数之和的流水线时空图;

(2) 计算流水线的吞吐率、加速比和效率。



图 7.1

四. (13 分) 画出基准标量机、超标量机、超流水线机、超标量超流水线机、超长指令字机五种不同类型处理机的时空图, 并分别标明其并行度 ILP。

五. (13 分) A、B 两个向量存放于存储器, 其向量长度为 64。设流水加法器有 4 级, 流水线时钟周期为 10ns, 读出 A、B 向量第一对元素到流水线始端所需的时钟周期数为 2, 求执行向量加法指令 ADDV 所需的时间。

六. (13 分) 假设循环移数网络有 64 个结点, 记为 N_0, N_1, \dots, N_{63} , 且网络链路能双向工作。

(1) 列出该网络从结点 N_0 出发, 正好以 3 步可到达的所有结点。

(2) 指出数据从任意一个结点 N_i 传送至另一个结点 N_j 所需的最少寻径步的最小上界。

七. (13 分) 假设使用 100 台多处理机系统获得加速比 80, 求原计算程序中串行部分所占的比例是多少?

本科生期末试卷(7) 答案

一. 填空题。

1. A. SISD B. SIMD C. MISD D. MIMD
2. A. 转移 B. 控制 C. 全局
3. A. 写后读 (WR) B. 读后写 (RW) C. 写后写 (RR)
4. A. 寄存器-寄存器 B. 中间寄存器 C. 快
5. A. 传输时间 B. 消息长度
6. A. 处理单元 (PE) B. 局部存储器 (PEM)
7. A. 粗粒度 B. 大

二. 解释术语。

1. 系列机: 是指在一个厂家内生产的具有相同的系统结构, 但具有不同组成和实现的一系列不同型号的机器。
2. 数据相关: 程序原有先后顺序的两条指令要对共享变量进行读、写操作, 如果在指令流水线中实际完成的顺序与原有顺序不一致, 那么流水线就可能输出错误的结果, 这称为数据相关。
3. 动态指令调度: 是由硬件在程序实际运行时实施的, 其基本思想是对指令流水线互锁控制进一步改进, 能实时的判断出是否有 WR、RW、WW 相关存在, 利用硬件绕过或防止这些相关的出错, 并允许多条指令在具有多功能部件的执行段中并行操作, 从而提高流水线的利用率且减少停顿现象。
4. 向量指令执行时间: 在向量处理机的流水功能部件中, 执行一条向量长度为 n 的向量指令的执行时间 T_{vp} 可表示为: $T_{vp} = T_s + T_e + (n-1)T_c$ 。式中, T_s 为流水线建立时间, T_e 为第一对向量元素通过流水线的时间, T_c 为流水线的失重周期。
5. 等分宽度: 当某一网络被切成相等的两半时, 沿切口的最小边数(通道)称为通道等分宽度, 用 b 表示。而线等分宽度就是 $B = b \times w$ 。其中 w 为通道宽度(用位表示)。该参数主要反映了网络最大流量。
6. 屏蔽方案: 每种屏蔽将所有 PE 划分成允许操作和禁止操作两种工作模式。
7. Cache 一致性: 在多处理机系统中的私有 Cache 会引起 Cache 中的内容相互之间以及共享存储器之间互不相同的问题, 称为多处理机的 Cache 一致性问题。

三. 解:

为了避免数据相关, 令 $A = a_1 + a_2$, $B = a_3 + a_4$, $C = a_5 + a_6$, $D = a_7 + a_8$, 将表达式做如下变换:

$$Z = [(a_1 + a_2) + (a_3 + a_4)] + [(a_5 + a_6) + (a_7 + a_8)] = [A + B] + [C + D]$$

小括号内的 4 个加法操作之间没有数据相关, 可以连续输入到流水线中。只要

前两个加法的结果出来之后，第一个中括号的加法就可以开始进行。因此 8 个浮点数求和的流水线时空图如下图所示。

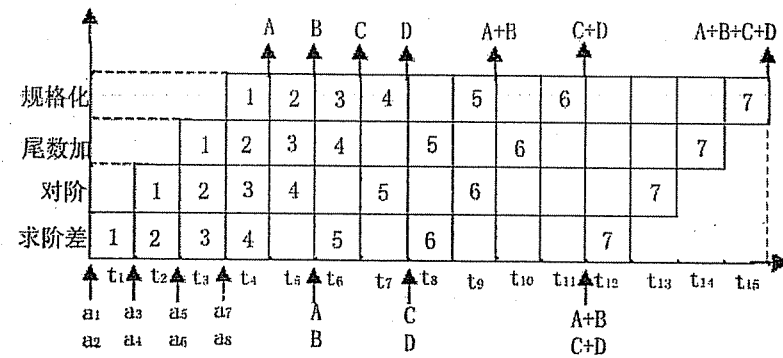


图 7.2 浮点加法器流水线时空图

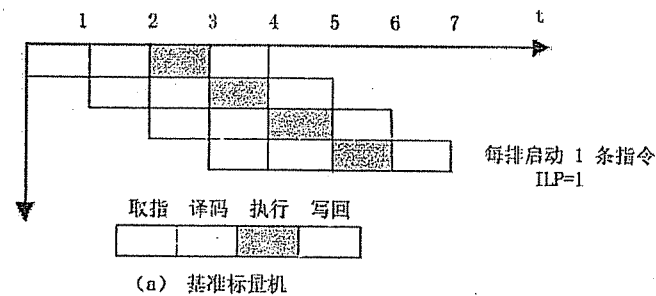
(2) 从时空图看到，7 次浮点加法操作共用了 15 个时钟周期，流水线共有 7 个输出。

流水线吞吐率： $P = 7 / (15 \times \Delta t) = 0.47 / \Delta t$

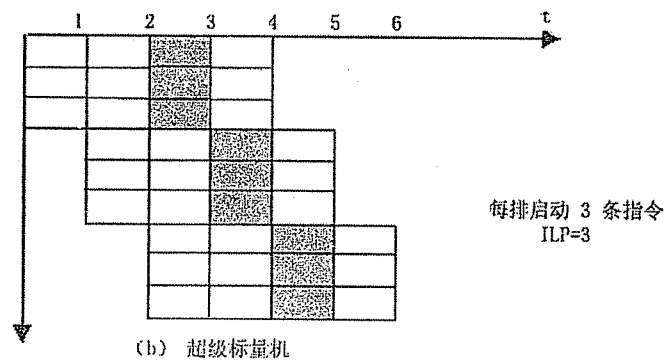
流水线加速比： $S = T_0 / T_k = (4 \times 7 \times \Delta t) / (15 \times \Delta t) = 1.87$

流水线的效率： $E = T_0 / (k \times T_k) = (4 \times 7 \times \Delta t) / (4 \times 15 \times \Delta t) = 0.47$

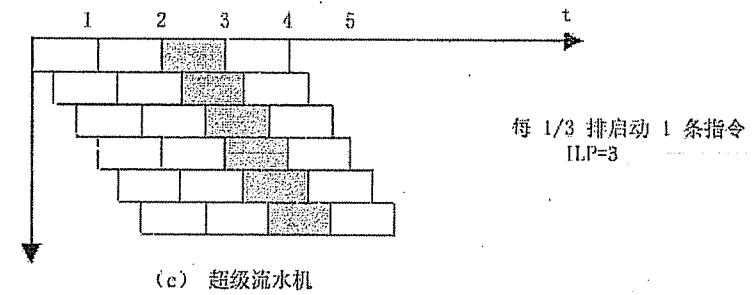
四. 解：



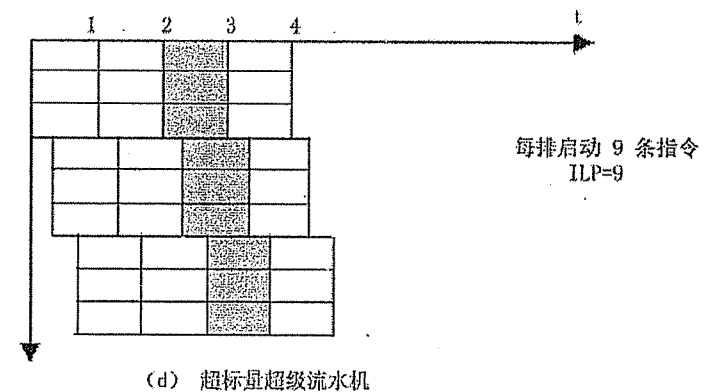
(a) 基准标量机



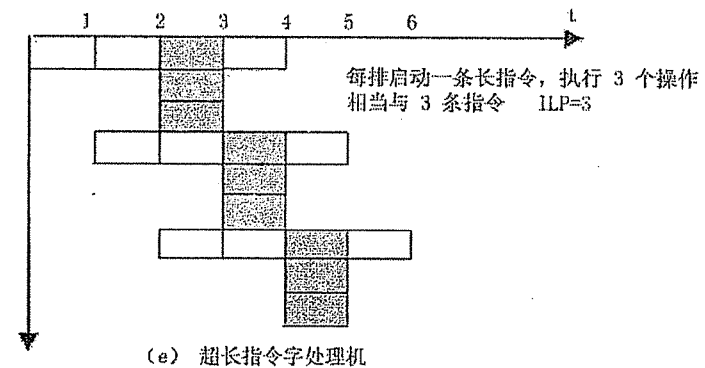
(b) 超级标量机



(c) 超级流水线



(d) 超标量超级流水线



(e) 超长指令字处理机

图 7.3

五. 解：

由题意知， $n = 64$ ， $l = 4$ ， $s = 2$ ， $T_c = 10 \text{ ns}$

$T_{vp} = (s + l + n - 1) T_c = (2 + 4 + 64 - 1) 10 \text{ ns} = 690 \text{ ns}$

六. 解：

(1) 循环移数网络中，与任意结点 N_i ($x_0 x_1 x_2 x_3 x_4 x_5 x_6 x_7$) 相连的结点是与其距离为 2 整数幂的结点。由此可得：

第一步： $N_1, N_2, N_4, N_8, N_{16}, N_{32}$

第二步： $N_3, N_5, N_6, N_7, N_9, N_{10}, N_{12}, N_{14}, N_{15}, N_{17}, N_{18}, N_{20}, N_{24},$

$N_{28}, N_{30}, N_{31}, N_{33}, N_{34}, N_{36}, N_{40}, N_{48}$

第三步: $N_{11}, N_{15}, N_{19}, N_{21}, N_{22}, N_{23}, N_{25}, N_{26}, N_{27}, N_{29}, N_{35}, N_{37}, N_{38},$
 $N_{39}, N_{41}, N_{42}, N_{44}, N_{46}, N_{47}, N_{49}, N_{50}, N_{52}, N_{56}, N_{60}, N_{62}, N_{63}$

(2) 所需的最少寻径步的最小上界即为网络的直径, 故为:

$$\log_2 64 / 2 = 3$$

七. 解:

设加速比为 k , 可加速部分比例为 F_e , 理论加速比为 S_e , 根据 Amdahl 定律:

$$k = \frac{1}{(1 - F_e) + F_e / S_e}$$

为了简单化, 假设程序只在两种模式下运作: (1) 使用所有处理机的并行模式;
 (2) 只用一个处理机的串行模式。假设并行模式下的理论加速比 S_e 即为多处理机的台数, 加速部分的比例 F_e 即并行部分所占的比例, 代入上式有:

$$80 = \frac{1}{(1 - F_e) + F_e / 100}$$

求得并行比例 $F_e = 0.9975 = 99.75\%$

本科生期末试卷 (8)

一. 填空题。(每小题 2 分, 共 14 分)

- 为了在不同系统结构的机器之间实现软件移植, 可采用 A 或 B 方法。
- 延迟转移技术是一种 A 方法, 它由 B 程序重排指令序列来实现。
- RW 和 WW 相关都是由于 A 寄存器引起的, 因此可将寄存器增加成多套的办法来解决相关, 这种技术称为 B 技术。
- 改善向量处理机性能的常用方法有: 采用 A 操作部件; 采用 B 技术; 采用 C 技术加速向量循环; 采用 D 系统。
- 互连网络中消息的第 1 位信息到达接收方所花费的时间称为 A, 它包括由于网络中转发或者其他硬件所引起的 B。
- 91 年问世的 MP-1 阵列机是一台 SIMD 计算机, 它由四个子系统组成: A 阵列、B 部件 ACU、具有 I/O 标准的 UNIX 子系统、高速 C 子系统。
- 多处理机系统中为了解决 Cache 不一致性问题, 提出了 A 协议和 B 的协议, 它们适合于不同的系统结构。

二. 解释下列术语。(每个 3 分, 共 21 分)

- 并行性
- 控制相关
- 静态指令调度
- 向量流水线最大性能
- 存储转发
- 寻径功能
- 监听协议

三. (13 分) 下列五条指令组成的汇编代码在一台三段流水线处理机上执行, 每一流水段上都有相关检测和分解。这三个流水段是: 取指令 IF、取操作数 OF (一个或多个)、执行 EX (包括写回)。请画出时空图说明:

- 在代码执行中所有可能的相关情况;
- 避免相关发生的调度方法。

```

INC    R0                ; R0 ← (R0) + 1
MUL    ACC, R0            ; ACC ← (ACC) × (R0)
STR    R1, ACC            ; R1 ← (ACC)
ADD    ACC, R0            ; ACC ← (ACC) + (R0)
STA    M, ACC             ; M ← (ACC)
    
```

四. (13 分) 有 A、B、C、D 四个存储器操作数, 要求完成 $(A \times B) + (C + D)$

本科生期末试卷(8) 答案

的运算, 原来使用的程序如下:

```

I1  LOAD  R1, M(A)      ; R1 ← M(A)
I2  LOAD  R2, M(B)      ; R2 ← M(B)
I3  MUL   R5, R1, R2    ; R5 ← (R1) × (R2)
I4  LOAD  R3, M(C)      ; R3 ← M(C)
I5  LOAD  R4, M(D)      ; R4 ← M(D)
I6  ADD   R2, R3, R4    ; R2 ← (R3) + (R4)
I7  ADD   R2, R2, R5    ; R2 ← (R2) + (R5)
    
```

现采用下图所示的七段指令流水线: ①执行段划分成三个流水段(共三个时钟周期); ②LOAD/STORE 部件、加法器、乘法器都是执行段的多功能部件; ③存储器也采用流水方式, 存/取操作需三个时钟周期, 故相邻两次存/取操作只要延迟一个时钟就无资源冲突; ④每个功能部件都有自己的“写回”部件; ⑤当发射段(I)判定一条指令所需的功能部件可用时, 读取寄存器操作数并进入E段(发射段I允许与前面指令W段重叠)。请画出静态调度后的指令序列流水的时空图。

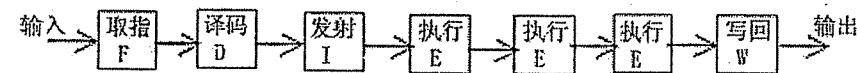


图 8.1

五. (13 分) 在一台向量处理机上实现 $A=s \times B$ 操作, 其中 A、B 是长度为 200 的向量, s 是一标量。向量寄存器长度为 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。T_c=15, 求总的执行时间及一个元素的平均执行时间。

六. (13 分) (1) 画出 2×2 开关构成的 16 个输入端的 Omega 网络。

(2) 结点 11 传送信息给结点 5, 同时结点 7 传送信息给结点 9, 画出完成这一寻径的开关设置。这种情况会出现阻塞吗?

(3) 这个网络实现任意个置换, 最多的通过次数是多少?

七. (13 分) 简述多处理机系统的主要特点。

一. 填空题。

1. A. 模拟 B. 仿真
2. A. 软件 B. 编译
3. A. 写入 B. 寄存器重命名
4. A. 多功能 B. 链接 C. 分段开采 D. 多处理机
5. A. “飞行”时间 B. 时延
6. A. 数组 B. 标量
7. A. 监听 B. 目录表

二. 解释术语。

1. 并行性: 是指计算机系统具有可以同时进行运算或操作的特性, 它包括同时性与并发性两种含义。
2. 控制相关: 流水线的控制相关是因为程序执行转移类指令而引起的相关。转移类指令如无条件转移、条件转移、子程序调用、中断等, 它们属于分支指令, 执行中可能改变程序的方向, 从而造成流水线断流。
3. 静态指令调度: 静态调度是由优化的编译程序来完成, 其基本思想是重排指令序列, 拉开具有数据相关的有关指令间的距离。由于是用便宜程序判断潜在的数据相关, 并在程序运行之前完成调度, 故称静态调度。
4. 向量流水最大性能: 当向量长度为无穷大时的向量流水线的性能。
5. 存储转发: 是最简单的分组交换方式, 用这种方式进行报文传递, 每经过一个中间结点时, 要由这个中间结点先把报文的一个信包全部存储起来, 然后再传递给下一个结点。
6. 寻径功能: 互连网络中 PE 间通信所需要的各种设置模式。
7. 监听协议: 它利用总线具有的广播能力, 用分散控制的办法解决 Cache 一致性问题。

三. 解:

(1) 图 8.2 画出了指令执行过程的时空图。

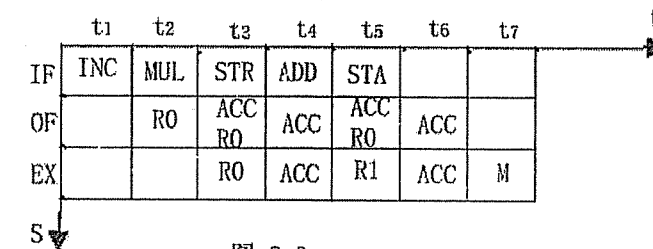


图 8.2

从图中看出:

在 t_3 时刻, INC 指令写 R_0 寄存器, MUL 指令读 R_0 寄存器, 会发生 WR 相关。

在 t_4 时刻, MUL 指令写 ACC 累加器, STR 指令读 ACC 累加器, 会发生 WR 相关。

在 t_6 时刻, ADD 指令写 ACC, STR 指令读 ACC, 会发生 WR 相关。

(2) 图 8.3 画出了避免相关发生的调度方法。

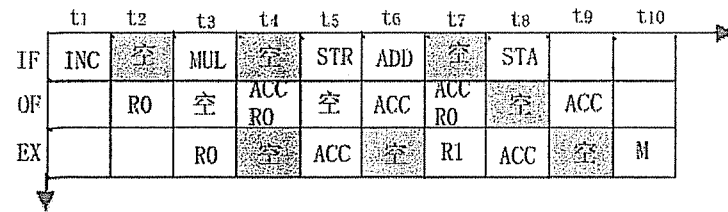


图 8.3

四. 解:

静态调度后的指令序列流水线时空图见图 8.4 所示。

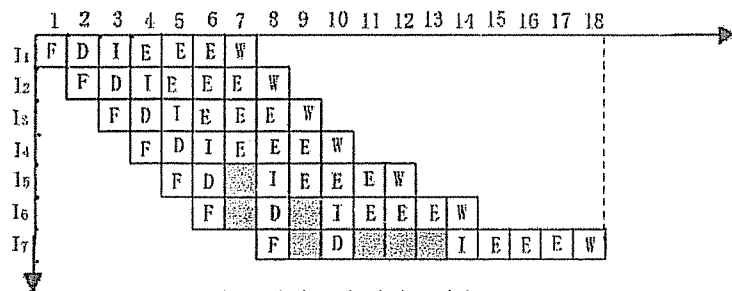


图8.4 重排指令序列的流水时空图

五. 解:

因为向量长度 $n=200$ 超过了向量寄存器长度 $MVL=64$, 所以采取分段开采方法。每次循环主要由下面 3 条指令组成:

LV V1, M(B) ; 取向量 B
 MULV V2, F0, V1 ; 向量 B 与标量 S 相乘
 SV M(A), V2 ; 存向量 A

这三条指令存在相关性它们必须分属于 3 个编队, 因此 $T_{ch}=3$ 。

$$T_n = \left\lceil \frac{n}{MVL} \right\rceil \times (T_e + T_s) + n \times T_{ch}$$

$$= \left\lceil \frac{200}{64} \right\rceil \times (15 + T_s) + 200 \times 3$$

$$= 4 \times (15 + T_s) + 600 = 660 + 4T_s$$

其中 T_s 是向量取、向量乘、向量存三条指令的启动时间 (分别为 12、7、12 个时钟周期) 的和, 即

$$T_s = 12 + 7 + 12 = 31 \quad \text{代入上式有}$$

$$T_{200} = 660 + 4 \times 31 = 784 \text{ (时钟周期)}$$

$$\text{一个元素的平均执行时间为 } \frac{T_n}{n} = \frac{784}{200} = 3.9 \text{ (时钟周期)}$$

六. 解:

(1) 16 个输入端的 Omega 网络如图 8.5 所示。

(2) 寻径开关设置见图 8.5 所示, 因此不会出现阻塞。

(3) 最多通过 $\log_2 16=4$ 次, 就可以实现 1 个置换。

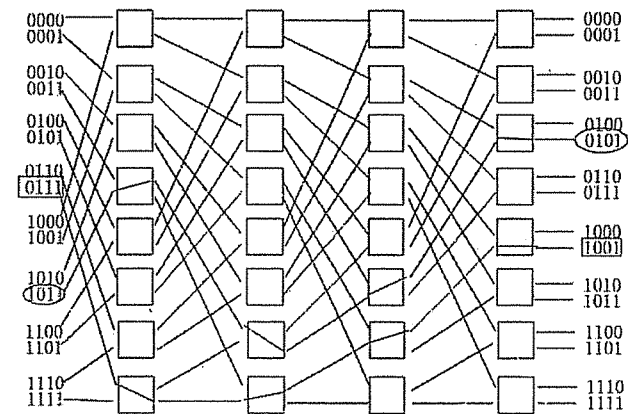


图 8.5 有 16 个输入端的 Omega 网络

七. 解:

(1) 结构的灵活性。多处理机的结构具有较强的通用性, 并能适应多样的算法, 因而具有灵活多变的系统结构。

(2) 程序并行性。多处理机的并行性体现在指令外部, 即表现在多个任务之间。

(3) 并行任务的派生。多处理机是多指令操作方式, 一个程序中就存在并发的程序段, 需要专门的程序段来表示它们的并发关系以控制它们的并发执行。

(4) 进程同步。多处理机实现指令、任务、程序级并行, 在同一时刻不同的处理机执行着不同的指令, 进程之间的数据相关和控制依赖关系, 决定了要采取一定的进程同步策略。

(5) 资源分配和进程调度。多处理机执行并发任务, 需用处理机的数目没有固定要求, 个处理机进入或退出任务的时刻互不相同, 所需共享资源的品种、数量又随时变化。因此存在资源分配和进程调度的问题。

本科生期末试卷 (9)

一. 填空题。(每小题 2 分, 共 14 分)

1. 系列机是指在一个厂家内生产的具有相同的 A, 但具有不同的 B 的一系列不同型号的机器。
2. 控制相关会造成流水线性能的重大损失, 必须采取技术措施来减少这种影响。通常采用的两类分支技术是 A 分支技术和 B 分支技术。
3. 静态指令调度技术是优化的 A 来完成, 其基本思想是重排指令序列, 拉开具有 B 的有关指令间的距离。
4. 存储器-存储器结构的向量机利用 N 个独立的 A 模块来支持相对独立的数据 B 访问, 从而达到所要求的存储器 C。
5. 传输时延是消息在互连网络上所花费的时间, 它等于 A 时间和 B 时间之和。
6. 在 MP-1 阵列机中, 每块处理器板上有 1024 个 A 和相关的局部存储器, 排列成 64 个 B 阵列。
7. 多处理机系统有四大类, 它们是: A 系统、B 系统、C 系统和 D 系统。

二. 解释下列术语。(每个 3 分, 共 21 分)

1. 兼容机
2. 定向技术
3. 超标量机
4. 一组向量操作执行时间
5. 线路交换
6. 阵列控制器
7. 目录表协议

三. (13 分) 有一个 4 段的指令流水线, IF、ID、EX、WB 分别代表取指、译码、执行、写回。分支指令在第 2 个时钟周期末决定是否条件失败分支, 第 3 个时钟周期末决定是否条件失败分支。又假定第一个时钟周期的操作和和条件分支无关, 并忽略其它流水线停顿。要求:

- (1) 画出无条件分支、发生的条件分支、不发生的条件分支三种情况下指令执行过程的时空图。
- (2) 假设条件分支指令占所有指令的 20% (其中条件成功分支有 60% 可能执行), 无条件分支占 5%, 问实际的流水线加速比是多少?

四. (13 分) 使用的程序包含如下指令序列:

```

I1    LOAD  R1, M(A)      ; R1 ← M(A)
I2    ADD   R2, R1        ; R2 ← (R2) + (R1)
I3    SUB   R3, R4        ; R3 ← (R3) - (R4)

```

```

I4    MUL   R4, R5        ; R4 ← (R4) × (R5)
I5    LOAD  R6, M(B)      ; R6 ← M(B)
I6    MUL   R6, R7        ; R6 ← (R6) × (R7)

```

(1) 采用图 9.1 所示的超标量流水线结构模型, 用优化编译程序方法重排上述指令序列。

(2) 画出无序发射无序完成的超标量流水线时空图。

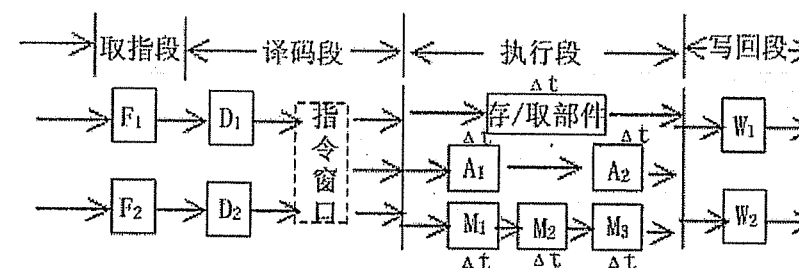


图9.1 流水线结构模型

五. (13 分) 计算下列两种向量超级计算机的峰值性能 (GFLOPS 表示):

- (1) CARY Y-MP C-90 (系统时钟周期长度为 4.2ns);
- (2) NEG SX-X (系统时钟周期长度为 2.9ns);
- (3) 解释这两种机器为什么能提供最大 64 路并行度的向量操作。

六. (13 分) 通过推理或分析证明下列命题正确:

- (1) 采用虫孔方式的超立方体多计算机系统, 相邻结点之间有 1 对方向相反的单向通道, 试证明在该系统上实现 E 立方体寻径不会死锁。
- (2) 试证明在二维网格上实现 X-Y 寻径不会死锁。

七. (13 分) 按多处理机系统的组成结构来分, 现有的 MIMD 计算机分哪几种类型? 简要说明其特点。

本科生期末试卷（9）答案

一. 填空题。

1. A. 系统结构 B. 组成和实现
2. A. 静态 B. 动态
3. A. 编译程序 B. 数据相关
4. A. 存储器 B. 并发 C. 带宽
5. A. “飞行” B. 传输
6. A. PE B. PE 群
7. A. 多向量处理机 B. SMP C. MPP D. 机群

二. 解释术语。

1. 兼容机：不同厂家生产的具有相同系统结构的计算机称为兼容机。
2. 定向技术：又称为旁路技术或专用通路技术，其基本思想是：如果后续指令要使用前面指令的运算结果值，则通过硬件专门电路将该运算结果值提前传送到有关缓冲寄存器，使后续指令得以不停顿地进入流水线，并及时得到所需要的操作数。
3. 超标量机：主要是借助对硬件资源重复来实现空间的并行操作。
4. 一组向量操作执行时间：一组向量操作的执行时间主要取决于下面 3 个因素：向量的长度，向量操作之间是否存在流水功能部件的使用冲突和数据相关性。
5. 线路交换：在传递一个消息之前，先建立一条从源结点到目的结点的物理通路，然后再传递消息。
6. 阵列控制器：阵列控制器实际上是一台计算机。它除了对阵列的处理单元实行控制以外，还能利用本身的内部资源执行一整套指令，用以完成标量操作，且在时间上与各的数组操作并行进行。
7. 目录表协议：是在主存储器中另外单独设置一个中央目录表来存放主存中每个数据块(行)的状态。

三. 解：

(1) 图 9.2 画出了无条件分支、发生的条件分支、不发生的条件分支三种情况下指令指令过程的时空图，说明产生的流水线停顿。

指令	t ₁	t ₂	t ₃	t ₄	t ₅	t ₆	说明
无条件分支 I + 1 I + 2 I + 3	IF	ID IF	EX IF 空	WB ID IF 空	EX ID IF	...	t ₂ 末无条件分支指令的目的地址已确定，t ₃ 时再次执行 IF 以取出转移目标处的指令 有 1 个时钟周期的停顿
发生的条件分支 I + 1 I + 2 I + 3	IF	ID IF	EX 空 空	WB IF 空	ID IF 空	...	条件分支的目的地址在 t ₃ 时刻才能决定 有 2 个时钟周期的停顿
不发生的条件分支 I + 1 I + 2 I + 3	IF	ID IF	EX 空 空	WB ID IF	EX ID IF	...	t ₂ 取到的是下一条顺序指令，条件分支的目的地址在 t ₃ 末确定 有 1 个时钟周期的停顿

图9.2 指令执行过程的时空图

(2) 从上图看出，无条件分支有 1 个时钟周期停顿，发生的条件分支有 2 个时钟周期停顿，不发生的条件分支有 1 个时钟周期停顿。

流水线的平均停顿数 m_c 为

$$m_c = (1 \times 5\%) + (2 \times 20\% \times 60\%) + (1 \times 20\% \times 40\%) = 0.37$$

由于控制相关引起了流水线的效率下降，故实际流水线加速比公式可用下式计算：

$$S = \frac{k}{1 + m_c} = \frac{4}{1 + 0.37} = 2.92$$

其中 k 为流水线段数。该公式可看成是流水线加速比公式的极限情况，即理想情况下，如果没有控制相关，也没有流水线停顿 ($m_c=0$)，则 $S=k$ 。

四. 解：

(1) 将原程序 I₁、I₂ 间的距离以及 I₃、I₄ 间的距离拉开，将 I₅、I₆ 安排在最前面。重排指令序列如下：

I₁ LOAD R6, M(B)
I₂ MUL R6, R7
I₃ LOAD R1, M(A)
I₄ SUB R3, R4
I₅ ADD R2, R1
I₆ MUL R4, R5

(2) 无序发射无序完成的超标量流水线时空图如图 9.3 所示。

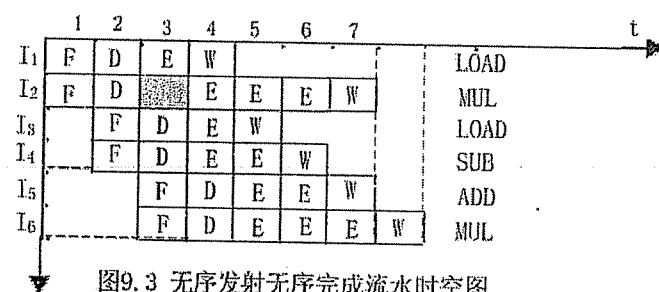


图9.3 无序发射无序完成流水时空图

五. 解:

(1) C-90 由 16 个向量处理器组成, 每个处理器有 2 条向量流水线。每个流水线都有一个可以并行操作的浮点乘法和加法部件。因此, 一条向量流水线每个时钟周期可以执行 2 个浮点操作。每个时钟周期总共可以执行 $16 \times 2 \times 2 = 64$ 个浮点操作。由于系统时钟周期长度为 4.2ns , 所以峰值性能 $= 64/4.2 = 15.2$ GFLOPS。

(2) SX-X 有 4 个处理器, 每个处理器有 4 组向量流水线。每组流水线都有 2 个 add/shift 和 2 个 multiply/logical 流水线。因此, 每个时钟周期总共可以执行 $4 \times 4 \times 2 \times 2 = 64$ 个操作。由于系统时钟周期为 2.9ns , 所以峰值性能 $= 64/2.9 = 22$ GFLOPS。

(3) C-90 每台处理器每个时钟周期可以产生 4 个向量计算结果, 相当于 4 路并行, 故 16 台处理器相当于 64 路并行。同样, SX-X 也可以在一个时钟周期执行 64 个浮点操作。

六. 解:

(1) 在 n 维超立方体中, 把一个 n 维二进制数 k 的结点定义为 n_k , 它有 n 个输出通道, 每维 1 个通道, 表示为 $C_{0k}, \dots, C_{(n-1)k}$, E 立方体按维的增序来寻径路由。 n_1 预定的消息到达结点 n_l 被送到通道 C_{ik} (其中 i 是确定 k 与 l 不同的最低位), 消息按维的增序路由, 增加通道下标不会在通道附图中产生回路, 故 E 立方体不会死锁。

(2) 与东-北, 西-北, 东-南, 西-南的路由方向相对应, X-Y 寻径有 4 种模式。在 Y 方向上有 2 对虚拟通道, 4 种寻径模式中每种都不会形成回路, 即 X-Y 寻径不会死锁。

七. 解:

按多处理机系统的组成结构来分, 现有的 MIMD 计算机分为如下 5 种类型: 并行向量处理机 (PVP)、对称多处理机 (SMP)、大规模并行处理机 (MPP)、分布共享存储器多处理机 (DSM)、工作站机群 (COW)。

(1) 并行向量处理机是由少数几台巨型向量处理机采用共享存储器方式互连而成。(2) 对称多处理机由一组处理机和一组存储器模块经过互连网络连接而成。有多个处理机且是对称的, 每台处理机的能力都完全相同。每次访问存储器时, 数据在处理机和存储器模块间的传送都要经过互连网络。由于是紧耦合系统, 不管访问的数据在哪一个存储器模块中, 访问存储器所需的延迟时间都是一样的。(3) 大规

模并行处理机, 每个计算机模块称为一个结点。每个结点有一台处理机和它的局部存储器 (LM)、结点接口 (NIC), 有的还有本身的 I/O 设备, 这几部分通过结点内的总线连在一起。计算机模块又通过结点接口连接到互连网络上。(4) 分布共享存储器多处理机共享存储器分布在各台处理机中, 每台处理机都带有自己的本地存储器, 组成一个处理机存储器单元。但是这些分布在各台处理机中的实际存储器又合在一起统一编址, 在逻辑上组成一个共享存储器。(5) 工作站机群, 通常是采用商品化的互连网络。

本科生期末试卷 (10)

一. 填空题。(每小题 2 分, 共 14 分)

1. 不同厂家生产的具有相同 A 的计算机称为兼容机, 它一般指 B 兼容。
2. 动态分支预测的依据是从 A 指令过去的行为来预测它将来的行为, 即根据近期转移是否成功的 B 记录, 来预测下一次转移的 C。
3. 动态指令调度技术是由 A 在程序实际运行中来完成。著名的动态调度算法有 B 法和 C 法。
4. 寄存器-寄存器结构的向量机构造一个满足 A 要求的高速 B, 并能实现 B 与主存储器之间的快速 C。
5. 各结点间有专用连接通路且运行中不能改变的网络称为 A 互连网络。设置有源开关, 可以根据需要借助控制信号对连接通路加以重新组合以实现所需要的通信模式的网络叫做 B 互连网络。
6. 91 年问世的 MP-1 阵列处理机的处理器有各种配置, 最少为 A 个, 最多为 B 个。
7. 当前的 MPP 系统具有 A 台处理机, 其最重要的特点是进行 B。

二. 解释下列术语。(每个 3 分, 共 21 分)

1. 模拟
2. 静态分支技术
3. 超流水线机
4. 半性能向量长度
5. 虚拟通道
6. 闭合螺线阵列
7. 伯恩斯坦准则

三. (13 分) 在一台单流水线处理机上执行下面程序。每条指令都要经过“取指”、“译码”、“执行”、“写结果”四个流水段, 每个流水段延迟时间都是 5ns。但在“执行”流水段 LS 部件和 ALU 部件只能其中一个工作, 即 LS 部件完成 LOAD 和 STORE 操作, ALU 部件完成其它操作。这两个操作部件的输出端和输入端有直接输出通路相互切换连接, 且 ALU 部件产生的条件码也能互接送入控制器。

I₁ SUB R0, R0 ; R₀ ← 0
I₂ LOAD R1, #8 ; R₁ ← 向量长度 8
I₃ LOOP: LOAD R2, A(R1) ; R₂ ← A 向量的一个元素
I₄ MUL R2, R1 ; R₂ ← (R₂) × (R₁)
I₅ ADD R0, R2 ; R₀ ← (R₀) + (R₂)
I₆ DNE R1, LOOP ; R₁ ← R₁ - 1, 若 (R₁) ≠ 0 转向 LOOP

I₇ STORE R0, S ; 保存结果

假定采用静态分支预测技术, 每次都预测转移不成功。要求:

- (1) 画出指令流水线的时空图 (中间部分可省略, 图中可用指令序号表示)。
- (2) 计算流水线的吞吐率和加速比, 并分别计算译码部件和 ALU 部件的使用效率。

四. (13 分) 比较并行度为 (m,n) 的超标量超流水线处理机与并行度为 (1, 1) 的基准标量处理机的性能。在如下限制条件下, 试分析下面公式的加速比表达式:

$$S(m,n) = \frac{T(1,1)}{T(m,n)} = \frac{k+N-1}{k+(N-m)/mn} = \frac{mn(k+N-1)}{mnk+N-m}$$

- (1) 在 $1 \leq m \leq 4$ 和 $1 \leq n \leq 6$ 的范围内, 对加速比 $S(m,n)$ 最大化后的最佳流水线段数是多少?
- (2) 阻碍超标量度 m 增长的实际限制是什么?
- (3) 阻碍超标量度 n 增长的实际限制是什么?

五. (13 分) 在某向量处理机上执行 $Y = a \times X + Y$ 运算, 其中 X 、 Y 是向量, 最初存放在内存, a 是一个标量。它们的向量指令如下:

LV V1, M(X) ; 取向量 X
MULV V2, F0, V1 ; 标量 a 和向量 X 相乘
LV V3, M(Y) ; 取向量 Y
ADDV V4, V2, V3 ; 完成 $Y = aX + Y$
SV M(Y), V4 ; 存结果 Y

求完成上述向量操作的总执行时间。

六. (13 分) 列出互连网络中的 4 种信息传递方式, 并分别给出其传输时延公式。

七. (13 分) 简述多处理机的特点。

本科生期末试卷(10) 答案

一. 填空题。

1. A. 系统结构 B. 向后
2. A. 转移 B. 历史 C. 方向
3. A. 硬件 B. Tomasulo 令牌 C. CDC 记分牌
4. A. 带宽 B. 中间存储器 C. 数据交换
5. A. 静态 B. 动态
6. A. 1024 B. 16384
7. A. 几百或几千台 B. 大规模并行处理

二. 解释术语。

1. 模拟: 是指用软件方法在一台现有的计算机上实现另一台计算机的指令系统。
2. 静态分支技术: 为减少转移指令对流水线的影响可采用的静态处理分支的方法主要包括: 尽早判别转移是否发生, 尽早生成转移目标地址、提前形成条件码、硬件上设置两个指令缓冲栈、延迟转移技术等。
3. 超流水线机: 在一个时钟周期内能够分时发射多条指令的处理机称为超流水线处理机, 另外, 也把指令流水线的段数大于等于 8 的流水线处理机称为超流水线处理机。
4. 半性能向量长度: 是为达到一半最大性能值所需的向量长度。
5. 虚拟通道: 虚拟通道是两个结点间的逻辑链, 它由源结点的片缓冲区、结点间的物理通道以及接受结点的片缓冲区组成。
6. 闭合螺线阵列: Illiac IV 阵列 PU 由 64 个处理单元(PE)、64 个局部存储器(PEM)和存储器逻辑部件(MLU) 组成。这个阵列的 64 个处理部件 $PU_0 \sim PU_{63}$ 排列成 8×8 方阵。每一个 PU_i 只和其上、下、左、右 4 个近邻 $PU_{i-8} \pmod{64}$ 、 $PU_{i+8} \pmod{64}$ 、 $PU_{i-1} \pmod{64}$ 、 $PU_{i+1} \pmod{64}$ 有直接连接。按此规则, 上下方向上同一列的 PU 两端相连成一个环, 左右方向上每一行的右端 PU 与下一行的左端 PU 相连, 最下面一行的右端 PU 与最上面一行的左端 PU 相连, 从而构成了一个闭合的螺线形状。因此, Illiac IV 的阵列结构又称为闭合螺线阵列。
7. 伯恩斯坦准则: 每个程序段在执行过程中通常要使用输入和输出这两个分离的变量集。若用 I_i 表示 P_i 程序段中操作所要读取的存储单元集, 用 O_i 表示要写入的存储单元集, 则 P_1 和 P_2 两个程序段能并行执行的伯恩斯坦准则是:

- (1) $I_1 \cap O_2 = \Phi$, 即 P_1 的输入变量集与 P_2 的输出变量集不相交;
- (2) $I_2 \cap O_1 = \Phi$, 即 P_2 的输入变量集与 P_1 的输出变量集不相交;
- (3) $O_1 \cap O_2 = \Phi$, 即 P_1 和 P_2 的输出变量集不相交。

三. 解:

(1) 采用静态分支预测技术, 每次都预测转移不成功, 指令流时空图如图 10.1 所示。

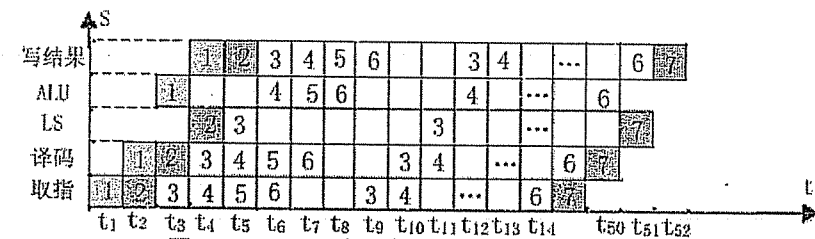


图 10.1 流水线时空图

$$(2) \text{ 吞吐率: } P = \frac{4 \times 8 + 3}{5ns \times (3 + 4 \times 8 + 2 \times 7 + 3)} = 135 \text{ (MIPS)}$$

$$\text{加速比: } S = \frac{(4 \times 8 + 3) \times 4}{52} = 2.69$$

$$\text{译码部件效率: } E_b = 35/52 = 0.673$$

$$\text{ALU 部件效率: } E_a = 25/52 = 0.481$$

四. 解:

(1) 对于给定的 m 、 n 值范围, 有不等式

$$mn(N-1) \geq N-1 \geq N-m$$

对给定加速比公式改写如下:

$$S(m, n) = \frac{mn(N-1) + mnk}{(N-m) + mnk} \geq \frac{(N-m) + mnk}{(N-m) + mnk} \geq 1$$

当 mnk 取最小值时, $S(m, n)$ 有最大值。因此, 为了使加速比最大, 流水线的段数应该是 1。

(2) 指令级的并行度限制了超标量度 m 的增长。

(3) 时钟多相(各部件间的同步)技术存在困难, 限制了超流水度 n 的增长。

五. 解:

将上述 5 条指令分成 3 个编队, 并且前两个编队中每两条指令组成一个链接:

LV V1, M(X) MULV V2, F0, V1 编队1: 取数与乘法指令链接

LV V3, M(Y) ADDV V4, V2, V3 编队2: 取数与加法指令链接

SV M(Y), V4 编队3: 把结果存入

可知 $T_{ch}=3$, $T_e=15$, $T_s=12+7+12+6+12=49$, $MVL=64$, 代入下式,

$$T = \left\lceil \frac{n}{MVL} \right\rceil \times (T_e + T_s) + n \times T_{ch}$$

本科生期末试卷 (11)

$$= \left\lceil \frac{n}{64} \right\rceil \times (15+49) + 3n$$

$$= (n+64) + 3n = 4n+64 \quad (\text{时钟周期})$$

六. 解:

(1) 线路交换寻径, 传输时延公式为:

$$T = (L_i/B) \times D + L/B$$

式中 L_i 为建立路径所需的小信息包长, L 为信息包长, D 为经过的结点数, B 为带宽 (以下同)。

(2) 存储转发寻径, 传输时延公式为:

$$T = (L/B) \times D + L/B = (D+1) \times L/B$$

(3) 虚拟直通, 传输时延公式为:

$$T = (L_h/B) \times 1 + L/B = (L_h \times D + L) \times L/B \approx L/B$$

式中 L_h 是消息的寻径头部的长度, 一般 $L \geq L_h \times D$ 。

(4) 虫孔方式, 传输时延公式为:

$$T = T_r \times D + L/B = (L_r/B) \times D + L/B = (L_r \times D + L) / B \approx L/B$$

式中 L_r 是片的长度, T_r 是片经过一个结点所需的时间, 一般 $L \gg L_r \times D$ 。

七. 解:

(1) 结构的灵活性。多处理机的结构具有较强的通用性, 并能适应多样的算法, 因而具有灵活多变的系统结构。

(2) 程序并行性。多处理机的并行性体现在指令外部, 即表现在多个任务之间。

(3) 并行任务的派生。多处理机是多指令操作方式, 一个程序中就存在并发的程序段, 需要专门的程序段来表示它们的并发关系以控制它们的并发执行。

(4) 进程同步。多处理机实现指令、任务、程序级的并行, 在同一时刻, 不同的处理机执行着不同的指令, 进程之间的数据相关和控制依赖关系, 决定了要采取一定的进程同步策略。

(5) 资源分配和进程调度。多处理机执行并发任务, 需用处理机的数目没有固定要求, 各处理机进入或退出任务的时刻互不相同, 所需共享资源的品种、数量又随时变化。因此存在资源分配和进程调度问题。

一. 填空题。(每小题 2 分, 共 14 分)

1. 紧耦合系统又称 A 系统, 一般是通过 B 或 C 实现计算机间的互连, 可以共享主存。

2. 消除结构相关的有效方法是采用 A 方法。消除数据相关的有效方法是设置 B 通路。

3. 指令发射策略是指指令发射所用的 A。当指令按程序的次序发射时, 称之为 B, 不按程序原有次序发射指令, 称之为 C。

4. 链接技术是标量流水 A 方法在 B 中的应用。

5. 信息传递方式有 A、B、C、D 几种。

6. 91 年问世的 MP-1 阵列处理机中, 处理器之间通过以下三种机制进行通信:
(1) A 通信; (2) B 通信; (3) 全局交叉开关寻径器通信。

7. SMP 称为 A, 也称为 B。

二. 解释下列术语。(每个 3 分, 共 21 分)

1. 仿真
2. 动态分支技术
3. 发射策略
4. CPI
5. 缓冲器死锁
6. 并行算法
7. 并行程序设计语言

三. (13 分) 在一台单流水线处理机上执行下面程序。每条指令都要经过“取指”、“译码”、“执行”、“写结果”四个流水段, 每个流水段延迟时间都是 5ns。但在“执行”流水段 LS 部件和 ALU 部件只能其中一个工作, 即 LS 部件完成 LOAD 和 STORE 操作, ALU 部件完成其它操作。这两个操作部件的输出端和输入端有直接输出通路相互切换连接, 且 ALU 部件产生的条件码也能互接入控制器。

I ₁	SUB	R0, R0	; R ₀ ← 0
I ₂	LOAD	R1, #8	; R ₁ ← 向量长度 8
I ₃	LOOP: LOAD	R2, A(R1)	; R ₂ ← A 向量的一个元素
I ₄	MUL	R2, R1	; R ₂ ← (R ₂) × (R ₁)
I ₅	ADD	R0, R2	; R ₀ ← (R ₀) + (R ₂)
I ₆	DNE	R1, LOOP	; R ₁ ← R ₁ - 1, 若 (R ₁) ≠ 0 转向 LOOP
I ₇	STORE	R0, S	; 保存结果

假定采用静态分支预测技术, 每次都预测转移成功。要求:

- (1) 画出指令流水线的时空图 (中间部分可省略, 图中可用指令序号表示)。
- (2) 计算流水线的吞吐率和加速比, 并分别计算译码部件和 ALU 部件的使

用效率。

四. (13分) 超标量机、超流水线机、VLIW 机、超标量超流水线机都能开发指令级的并行性。假设这四种机器的流水线都为 4 段, 每段均有 1 个时钟周期; 又假设超标量机每个时钟周期可以同时启动 3 条指令; 超流水线机每隔 1/3 时钟周期启动 1 条指令; VLIW 机每个时钟周期启动 1 条指令, 执行 3 个操作; 超标量超流水线机每个时钟周期可以启动 9 条指令。

现要执行 9 条指令的代码序列, 问这四种机器各需多少个时钟周期才能完成? 并画出时空示意图。

五. (13分) 向量处理机上进行 $D = A \times (B \times C)$ 的向量计算, 假设向量长度 $N \leq 64$, B 和 C 已由寄存器取至 V0 和 V1, 则下面三条指令可以完成上述运算:

LD V3, A ; $V_3 \leftarrow A$

ADDV V2, V0, V1 ; $V_2 \leftarrow V_0 + V_1$

MULV V4, V2, V3 ; $V_4 \leftarrow V_2 \times V_3$

(1) 分析上面的向量指令执行过程, 画出向量指令链接图;

(2) 采用链接技术时, 计算执行时间;

(3) 若三条指令全部用串行方法, 计算执行时间。

(4) 若前两条指令并行执行, 第三条指令串行执行, 计算执行时间。

六. (13分) 通过推理或分析证明下列命题正确:

(1) 采用虫孔方式的超立方体多计算机系统, 相邻结点之间有一对方向相反的单向通道, 证明在该系统上实现 E 立方体寻径不会死锁。

(2) 试证明在 2 维网络上实现 X-Y 寻径不会死锁。

七. (13分) 通常采用何种协议机制来解决多处理机系统中的 Cache 不一致性问题? 说明协议机制的适用环境。

本科生期末试卷(11) 答案

一. 填空题。

1. A. 直接耦合 B. 总线 C. 高速开关
2. A. 资源重复 B. 专用
3. A. 协议或规则 B. 按序发射 C. 无序发射
4. A. 定量传送 B. 向量寄存器
5. A. 线路交换 B. 存储转发寻径 C. 虚拟直通 D. 虫蚀寻径
6. A. ACU-PE 阵列 B. X-NET 近邻
7. A. 共享存储型多处理机 B. 对称型多处理机

二. 解释术语。

1. 防真: 用微程序直接解释另一种机器指令系统的方法称为防真。
2. 动态分支技术: 预测的依据是从转移指令过去的行为来预测它将来的行为, 即根据近期转移是否成功的历史记录, 来预测下一次转移的方向。
3. 发射策略: 是指指令发射所用的协议或规则。
4. CPI: 大多数的时钟速度是固定的, 它的运行周期称为时钟周期, 而每条指令的平均时钟周期数即为 CPI。
5. 缓冲区死锁: 由于资源的共享导致需求大于现有资源, 因此缓冲区的循环等待会引起死锁。
6. 并行算法: 阵列机的研究必须与并行算法的研究密切结合, 以使它的求解算法具有更强的适应性和更广的应用面。
7. 并行程序设计语言: 并行程序设计语言通常是在高级语言基础上扩展的, 常见的 3 种扩展并行语言是: (1) FORK JOIN 表示法; (2) 块结构语言 Cobegin Coend 表示法; (3) 块结构语言 Parbegin Parend 表示法。

三. 解:

(1) 每次都预测转移成功时, 指令流水线时空图如下图。

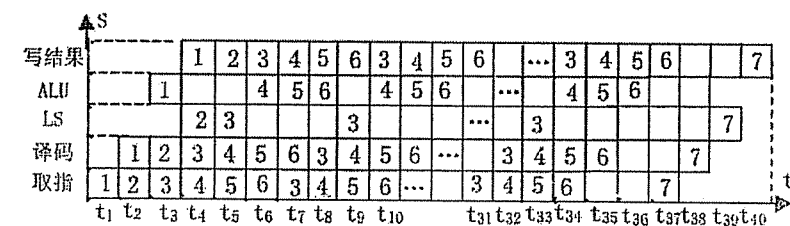


图11.1 流水线时空图

(2) 吞吐率: $P = 135$ (MIPS)

加速比: $S = 2.69$

译码部件效率: $E_b = 35/40 = 0.875$

ALU 部件效率: $E_A = 25/40 = 0.625$

四. 解:

(1) 画出四种机器执行指令的时空图如下:

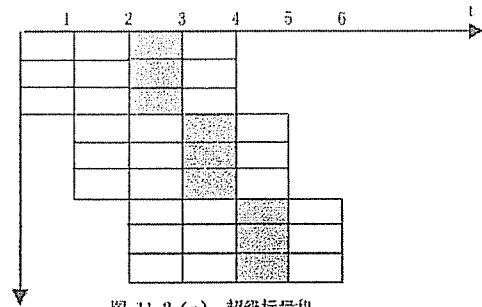


图 11.2 (a) 超级标量机

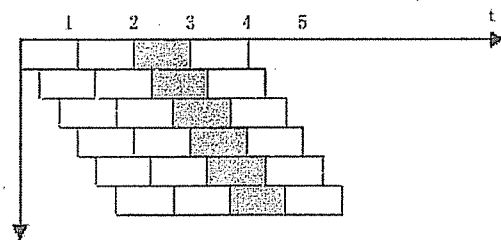


图 11.2 (b) 超级流水线

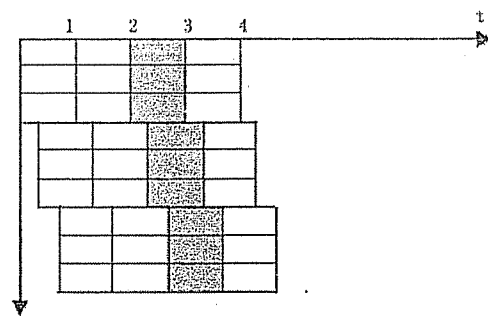


图 11.2 (c) 超标量超级流水线

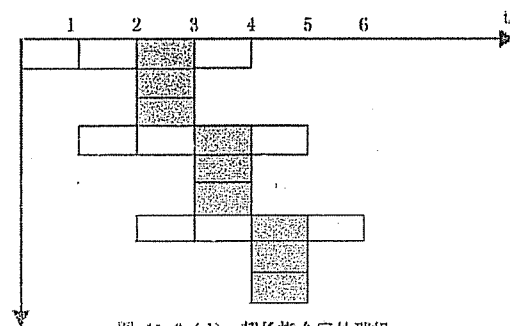


图 11.2 (d) 超长指令字处理机

(2) 超标量机需 6 个时钟周期;

超流水线机需 $20/3$ 个时钟周期;

VLIW 机需 6 个时钟周期;

超标量超流水线机需 $14/3$ 个时钟周期。

五. 解:

(1) 第一、二条指令无向量寄存器使用冲突, 也无功能部件使用冲突, 所以这两条指令可以并行执行。第三条指令与第一、二条指令均存在先写后读的相关冲突, 因而可以将第三条指令与第一、二条指令链接执行, 如下图所示:

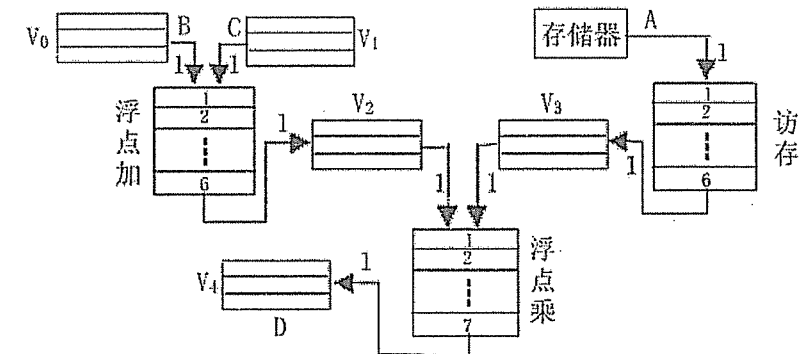


图 11.3 向量连接图

(2) 由于同步的要求, 数据进入和流出每个功能部件, 包括存储器读取, 需要一个时钟周期时间。浮点加和访存操作六个时钟周期 (并行), 浮点乘操作七个时钟周期。故采用链接技术时执行时间为:

$$T = (1+6+1) + (1+7+1) + (N-1) = N+16 \quad (\text{时钟周期})$$

(3) 三条指令全部串行执行时执行时间为:

$$T = [(1+6+1) + N-1] + [(1+6+1) + N-1] + [(1+7+1) + N-1] = 3N+22 \quad (\text{时钟周期})$$

(4) 前两条指令并行执行, 第三条指令串行执行, 执行时间为:

$$T = [(1+6+1) + N-1] + [(1+7+1) + N-1] = 2N+15 \quad (\text{时钟周期})$$

六. 解:

(1) 在 n 维超立方体中, 把一个 n 位二进制数 k 的结点定义为 n_k , 它有 n 个输出通道, 每维 1 个通道, 表示为 $C_{0k}, \dots, C_{(n-1)k}$, E 立方体按维的增序来寻径路由。 n_i 预定的消息到达结点 n_k 被送到通道 C_{ik} (其中 i 是确定 k 与 1 不同的最低位), 消息按维的增序路由, 增加通道下标不会在通道附图中产生回路, 故 E 立方体不会死锁。

(2) 与东-北, 西-北, 东-南, 西-南的路由方向相对应, X - Y 寻径有 4 种模式。在 Y 方向上有 2 对虚拟通道, 4 种寻径模式中每种都不会形成回路, 即 X - Y 寻径不会死锁。

七. 解:

本科生期末试卷 (12)

(1) 监听协议: 在基于总线互连结构的系统中, 由于系统中每个处理机都能察觉到存储器系统正在进行的活动的, 在每个活动破坏了 Cache 的一致性时, Cache 控制器将采取相应的方法使有关的拷贝无效或更新。使用监听协议时, 有两种保持 Cache 一致性的方法: 写无效策略和写更新策略。

(2) 目录表协议: 当采用非总线方式构造大型系统或当在多级网络上实现广播功能代价很大时, 监听协议无法使用, 采用目录表的协议。该协议的核心是把其他 Cache 数据块无效的一致性命令只发给存放相应数据块的 Cache。在多级网络中, 用 Cache 目录存放有关 Cache 拷贝驻留在哪里信息, 从而支持 Cache 一致性。

一. 填空题。(每小题 2 分, 共 14 分)

1. 松耦合系统又称 A 系统, 一般是通过 B 或 C 实现计算机间的互连, 可以共享某些外存设备。

2. 超标量处理机主要开发 A 并行性, 超流水线处理机主要开发 B 并行性。

3. 在一个时钟周期内能够并行发射 A 条以上指令的处理机称为 B 处理机。在一个时钟周期能够分时发射多条指令的处理机称为 C 处理机。

4. 同一个编队中的向量指令一定不存在功能部件的 A 和 B 相关性。

5. 缓冲区或通道上的 A 可能产生死锁, 利用 B 可以解决死锁。

6. 阵列机的研究必须与 A 的研究密切结合, 以使它的求解算法具有更强的适应性和更广的 B。

7. 机群系统是利用高速 A 将一组高性能 B, 按某种结构连接起来, 并在并行程序设计环境支持下实现高效 C 的系统。

二. 解释下列术语。(每个 3 分, 共 21 分)

- | | | |
|------------|---------|-------------|
| 1. 程序局部性原理 | 2. 精确断点 | 3. 超标量超流水线机 |
| 4. 紧耦合系统 | 5. 虫孔方式 | |
| 6. 递归折叠求和 | 7. MPP | |

三. (13 分) 图 12.1 所示为一个 4 段的流水线连接图。

(1) 设 $n=4$, 画出它的流水线时空图, 求其 P , P_{\max} , S , E 值。

(2) 设 $n=12$, 画出瓶颈段再细分的流水线连接图和时空图, 并求 P , P_{\max} , S 和 E 值

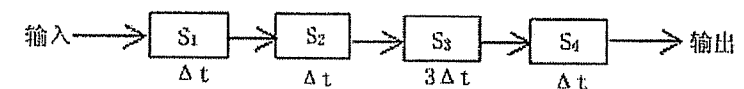


图 12.1

四. (13 分) 假定有多个加法器, 不存在加法器的资源冲突。有三条连续指令组成的程序代码如下:

I_1	ADD	R_1, R_2, R_4	$; R_1 \leftarrow (R_2) + (R_4)$
I_2	ADD	$R_2, R_1, 1$	$; R_2 \leftarrow (R_1) + 1$

- I₃ SUB R1, R4, R5 ; R₁ ← (R₄)-(R₅)
- (1) 分析程序代码段中的数据相关;
- (2) 采用何种硬件技术可解决这些数据相关? 请加以说明。

五. (13 分) 在某向量处理机上执行 $B=c \times A+B$ 运算, 其中 A、B 是向量, 最初存放在内存, c 是一个标量。它们的向量指令如下:

LV V1, M(A) ; 取向量 A

MULV V2, F0, V1 ; 标量 c 和向量 A 相乘

LV V3, M(B) ; 取向量 B

ADDV V4, V2, V3 ; 完成 $B=c \times A+B$

SV M(Y), V4 ; 存结果 B

求完成上述操作的总执行时间。

六. (13 分) 在一个 8×8 网格上, 根据下面条件确定 1 条优化的选播路径, 源结点是 (3, 5), 10 个目标结点是 (1, 1), (1, 2), (1, 6), (2, 1), (4, 1), (5, 5), (5, 7), (6, 1), (7, 1), (7, 5)。试确定网格计算机中的最优寻径路径, 要求选播路径应使从源结点到每个目的结点的距离最短。

七. (13 分) 说明阵列处理机和向量处理机的相同特点和不同特点。

本科生期末试卷 (12) 答案

一. 填空题。

1. A. 简洁耦合 B. 通道 C. 通信线路
2. A. 空间 B. 时间
3. A. 2 B. 超标量 C. 超流水线
4. A. 使用冲突 B. 数据
5. A. 循环等待 B. 虚拟通道
6. A. 并行算法 B. 应用面
7. A. 通用网络 B. 工作站或高档 PC 机 C. 并行处理

二. 解释术语。

1. 程序局部性原理: 实验统计表明: 一个程序用 90% 的执行时间去执行仅占 10% 的程序代码。这一规律称为程序访存的局部性原理。
2. 精确断点: 流水线处理机的中断处理中, 精确断点法需要把断点处的指令的执行结果都保存下来, 为此要设置一定数量的后援寄存器。
3. 超标量超流水线机: 为了进一步提高指令级并行度, 可以把超标量技术与超流水线技术结合在一起, 这就是超标量超流水线处理机。
4. 紧耦合系统: 又称直接耦合系统, 指计算机间物理连接的频带较高, 一般是通过总线或高速开关实现计算机间的互连, 可以共享主存。由于具有较高的信息传输率, 因而可以快速并行处理作业或任务。
5. 虫孔方式: 使用这种方式时, 把信息包再进一步等分成更小的“片”(flit), “片”是队列或通路所能处理的最小信息单位。
6. 递归折叠求和: 对求累加和这样的操作, 为了加快并行计算, 常采用递归折叠求和算法。一些看起来只能串行计算的问题可以用并行算法得到解决, 并有较好的加速比, 在求累加和的过程中, 并非每个处理器始终参加操作。连续模型算法采用近邻互连方式, 这种互连方式每一步只能将信息传播固定有限的距离。
7. MPP: 大规模并行处理 MPP(massively parallel processing)系统的定义按照当前的标准, 具有几百台~几千台处理机的任何机器都是大规模并行处理系统。MPP 系统最重要的特点是进行大规模并行处理。

三. 解:

- (1) 段 3 为瓶颈的流水线时空图如图 12.1 所示。

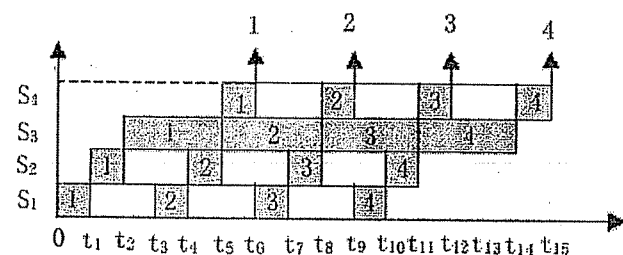


图 12.1 段 3 为瓶颈时的流水线时空图

由时空图看出, $n=4$, $T_k=t_{15}=15\Delta t$

$$P = \frac{n}{T_k} = \frac{4}{15\Delta t} \quad P_{\max} = \frac{1}{3\Delta t}$$

$$T_0 = n \cdot k \Delta t = 4 \cdot (1+1+3+1) \Delta t = 24\Delta t$$

$$S = \frac{T_0}{T_k} = \frac{24\Delta t}{15\Delta t} = 1.6$$

时空图有效面积 $= 3 \times (1 \times \Delta t \times 4) + (1 \times 3 \Delta t) \times 4 = 24\Delta t$

4 个流水段所围总面积 $= 4 \times t_{15} = 4 \times 15\Delta t = 60\Delta t$

$$E = \frac{24\Delta t}{60\Delta t} = \frac{2}{5} = 0.4$$

(2) 瓶颈段再细分的流水线连接图和时空图如图 12.2、图 12.3 所示。

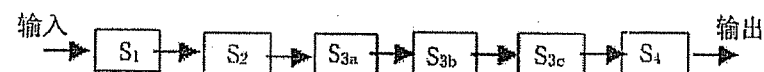


图12.2 段 3 再细分的流水线连接图

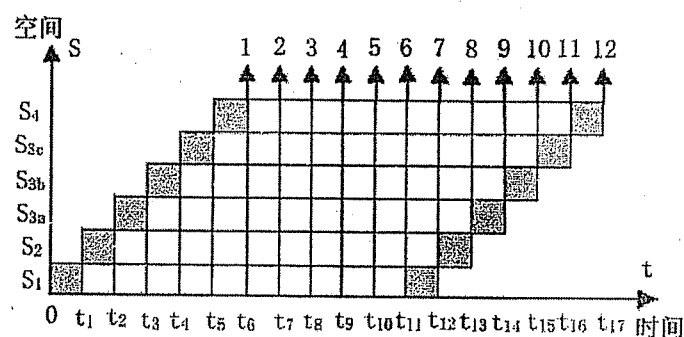


图12.3 段 3 再细分的流水线时空图

$n=12$, $T_k=t_{17}=17\Delta t$, $k=6$

$$P = \frac{n}{T_k} = \frac{12}{17\Delta t} \quad P_{\max} = \frac{1}{\Delta t}$$

$$S = \frac{nk}{k+n-1} = \frac{12 \times 6}{6+11} = \frac{72}{17} = 4.24$$

$$E = \frac{n}{k+n-1} = \frac{12}{17} = 0.7$$

$$\text{或} \quad E = \frac{S}{k} = 0.7$$

四. 解:

(1) 指令 I_1 和 I_2 之间有 RW 相关, I_2 和 I_3 之间有 RW 相关, I_1 和 I_3 之间有 WW 相关, I_1 和 I_2 之间还有 WR 相关。

(2) 对 I_1 和 I_2 之间的 WR 相关, 可用定向传送解决。根据寄存器重命名技术, 对引起 RW 相关的 I_2 中的 R_2 , 对引起 WW 相关的 I_3 中的 R_1 , 可分别换成备用寄存器 R_2' 、 R_1' 。经寄存器重命名后, 程序代码段实际执行时变为:

I_1 ADD R_1 , R_2 , R_4

I_2 ADD R_2' , R_1 , 1

I_3 SUB R_1' , R_4 , R_5

五. 解:

将上述五条指令分成三个编队, 并且前两个编队中每两条指令组成一个链接:

LV V_1 , M(A) MULV V_2 , F0 , V_1 编队 1: 取数与乘法指令链接

LV V_3 , M(B) ADDV V_4 , V_2 , V_3 编队 2: 取数与加法指令链接

SV M(B) , V_4 编队 3: 把结果存入

已知 $T_{ch} = 3$, $T_e = 15$, $T_s = 12+7+12+6+12 = 49$, $MVL = 64$

$$T = \left\lceil \frac{n}{MVL} \right\rceil \times (T_e + T_s) + n \times T_{ch}$$

$$= \left\lceil \frac{n}{64} \right\rceil \times (15 + 49) + 3n$$

$$= 4n + 64 \quad (\text{时钟周期})$$

六. 解:

从源结点到每个目标结点的距离最短的选播路径如下图所示。其中: 通道数为 22, 距离为 8。

