[向量处理机 2](#_Toc507317123)

## 向量处理机

向量处理机专用于哪些可以转化为相联个运算的问题

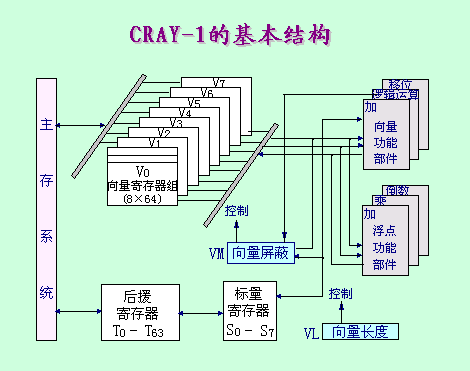
设置向量数据表示和相应的向量指令

向量：**A**=(,,…,)，**B**=(,,…,)，**C**=(,,…,)，**D**=(,,…,)

计算：**D**=**A**×(**B**+**C**)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 向量  处理  方式 | 横向  水平 | 执行n次循环  不利于流水线 | 1.=×(+)  2.=×(+)  ……  =×(+) | |
| 纵向  垂直 | 将整个向量按相同的运算处理完毕后，在去执行其他运算  适用流水线 | =+  =+  ……  =+ | =×  =×  ……  =× |
| 纵横  分组 | 把一个向量分成若干组，每个组有k个元素，  组内纵向处理，结果横向处理 |  | |

## 数据链路



|  |  |
| --- | --- |
| 12条可并行工作的单功能流水线 | 流水运算：地址、向量、标量  提供6个单功能流水部件：  整数加(3T)  逻辑运算(2T)  移位(4T)  浮点加(6T)  浮点乘(7T)  浮点迭代求倒数(14T) |
| 6个单功能流水部件  进行向量运算 | 整数加（3拍）  逻辑运算（2拍）  移位（4拍）  浮点加（6拍）  浮点乘（7拍）  浮点迭代求倒数（14拍）  括号中的数字为其流水经过的时间，每拍为一个时钟周期，即12.5 ns。 |
| 向量寄存器组 | 512个64位寄存器组成，分成8块，编号V0~V7，  每一块称为一个向量寄存器，可存放一个长度不超过64的向量(元素个数)  超过的需要用向量循环进行分段处理  每个向量寄存器可以每拍向功能部件提供一个数据元素，或者每拍接受一个从功能部件来的结果元素 |
| 标量寄存器 | S0~S7 |
| 快速暂存器 | 用于在标量寄存器和存储器之间提供缓冲 |
| 向量屏蔽寄存器  VM | 64bit  每一个bit对应向量寄存器的一个单元，  作用：用于向量的归并、压缩、还原、测试、某些元素的单独运算 |

|  |  |
| --- | --- |
| 特点 | 每个向量寄存器Vi都有连到6个向量功能部件的单独总线  每个向量功能部件也都有把运算结果送回向量寄存器组的总线。  只要不出现Vi冲突和功能部件冲突，各Vi之间和各功能部件之间都能并行工作，大大加快了向量指令的处理。  Vi冲突：并行工作的各向量指令的源向量或结果向量使用了相同的Vi。  例如：源向量相同  V3←V1＋V2  V5←V4∧V1  功能部件冲突：并行工作的各向量指令要使用同一个功能部件。  例如：都需使用乘法功能部件  V3←V1×V2  V5←V4×V6 |
| 向量指令类型 | Vk ← Vi op Vj  Vk ← Si op Vj  Vk ← 主存  主存 ← Vi |

## 性能提高

|  |  |
| --- | --- |
| 设置多个功能部件  使它们并行工作 | 如4组12个单功能流水部件  向量部件：向量加、移位、逻辑运算  浮点部件：浮点加、浮点乘、浮点求倒数  标量部件：标量加、移位、逻辑运算、计数  地址运算部件：整数加、整数乘  独立的功能部件，可并行工作 |
| 链接技术 | 加快一串向量指令的执行，类似于定向技术  链接特征：具有先写后读相关的两条指令，在不出现功能部件冲突和源向量冲突的情况下，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。  链接特性的实质：把流水线定向的思想引入到向量执行过程的结果。  进行向量链接的要求  1.保证：无向量寄存器使用冲突和无功能部件使用冲突  2.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。  3.当一条向量指令的两个源操作数分别是两条先行指令的结果寄存器时，要求先行的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。  4.要进行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。 |
| 分段开采技术 | 加快循环的处理  如果向量的长度大于向量寄存器的长度，该如何处理呢？  1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，然后循环分段处理，每一次循环只处理一个向量段。  2.这种技术称为分段开采技术。  由系统硬件和软件控制完成，对程序员是透明的。 |
| 多处理机系统 | 进一步提高性能 |

## 性能

|  |  |
| --- | --- |
| 向量指令处理时间Tvp |  |
| 峰值性能R |  |
| 半性能向量长度n | 性能达到R/2时所需的向量长度 |
| 向量长度临界值n | 对于某一计算任务  向量方式的处理速度优于标量串行方式处理速度所需的最小向量长度 |