

大型向量寄存器的设计思想及其实现

张民选 刘毅之 刘进学

摘要 本文提出了一种设计大型向量寄存器为三维动态结构的设计思想,描述了其工作原理,并给出了用这种思想设计和扩充 CRAY-1 类巨型机向量寄存器的两个实例。

一、向量寄存器概述

程序可编址的大型向量寄存器,是第二代巨型机进行向量处理的核心部件。它是解决功能部件与主存间频带不匹配的一个有效措施。用它满足功能部件向量操作时所需要的大量数据,可简化机器结构,提高系统速度。

美国巨型机 CRAY-1S,它设置了8个64单元的向量寄存器。其向量寄存器的作用有:

1. 为向量指令的并行执行提供了物质基础。它允许许多个不同的功能部件、对不同的向量数据同时加工,最多可允许7条向量指令同时(并行)执行。

2. 使功能部件之间自动链接执行处理功能成为可能。这样便提高了向量操作并行度,减少了数据相关的影响,因而对提高系统性能起了重要作用。

3. 它允许全部向量型算术/逻辑运算指令是寄存器型的,因而向量处理的“建立”时间较短,避免存储器型向量指令“起步时间”长的弊病。即使向量长度为2~4的短向量的处理,也能显示出其优越性。

4. 在标量计算为主的程序中,可以把向量寄存器作为其后援寄存器以提高处理速度。

5. 降低对主存频带宽度的要求。由它作为源寄存器、结果寄存器和中间寄存器,最后处理结果再送回主存,这样便缓解了多个功能部件操作需要的数据量大而主存能提供的数据流量受限的矛盾。向量寄存器在 CRAY-1 中的作用,在流水线向量机中带有普遍性。

CRAY-1 机中向量寄存器的8个输入/输出端口,与6个功能部件相连。就其平均并行度来说,频带是匹配的,这时可使功能部件的平均并行度为2~3,与经验公式得到的平均并行度相符。参考文献[5]对此已进行了详细讨论。但在大型科学计算(尤其是多维大型科学计算)、数据处理、信号分析中,发现有的算题需要8个(64单元)以上的向量寄存器。譬如,有些本来还要反复使用的数据,也不得不送往主存保存,以便释放向量寄存器为后继向量指令流出提供必要的条件,有的一次处理或变换中,须提供

9个以上的寄存器,为此不得不分开处理而影响效率,使系统的处理能力得不到充分发挥。所以,增加程序可编此向量寄存器的个数和容量和增加向量寄存器与主存之间的数据通路,成为最近推出的流水线向量机的的措施之一。表1列出了它们的向量寄存器设置情况。

表1 近期提供市场的巨型机的向量寄存器设置比较

型 机	厂 家	机器CP (ns)	V寄存器容量 (字)	V寄存器个数	每个V单元数
CRAY X-MP	(美)CRAY研究 公司	9.5	1 K	16 (8×2)	64
CRAY-2	"	4.1	2 K	32 (8×4)	64
FACOM VP/100	(日)富士通	7.5	4 K	128 (128×1)	32
FACOM VP/200	"	7.5	8 K	256 (128×2)	32
FACOM VP/400	"	7.5	16K	512 (128×4)	32
HITAC S810/10	(日)日立	15	4 K	16 (16×1)	256
HITAC S810/20	"	15	8 K	32 (16×2)	256
NEC Sx-1	(日)日电	7	5 K	20 (10×2)	256
NEC Sx-2	"	6	10K	40 (10×4)	265

设置程序可编址大容量寄存器给系统带来的益处很大,但其代价也是很高的。如国外一巨型机向量寄存器的插件数量约为CPU的插件总量的22%,如果将向量寄存器的个数增加一倍,则其插件数占CPU的30%,这个比例是惊人的。我们对程序可编址向量寄存器的设置,进行了初步研究。根据大规模RAM组件和文件寄存器的速度、容量和读/写控制特点,以及机器对向量寄存器容量和个数的需要,进行了综合分析,提出了立体动态结构的向量寄存器和I/O端口独立使用的设计思想,以便既可以增加程序可编址向量寄存器的个数,又使其成本和体积变得可以接受。

二、三维动态结构向量寄存器

图1是单阵列部件的三维结构向量寄存器。向量寄存器分为 n 组,每组 m 个,每个含 l 个单元。我们用 (n, m, l) 表示其基本参数。同一组的 m 个寄存器,公用一套输入端口和输出端口。在同一时刻, m 个寄存器中仅允许一个寄存器作为操作数寄存器,一个作为结果数寄存器。换言之,同一组中的各个寄存器只能分时占用输入和输出端口。指令可以寻址到某组某个寄存器,而对每个寄存器内单元的访问,一般由0号单元开始,直到向量长度寄存器VL内容所指定的单元为止,如需要,也可设计成从任意单元开始连续访问(VL)个单元。对于某些传送指令,也可直接访问某组、某个寄存器的某个单元。

这种立体结构的设计思想,非常适合使用大容量RAM组件来设计大型向量寄存器,使得设计的向量寄存器具有总容量大、程序可编址的特点,与常规的“平面”结构的向量寄存器相比,具有器材省、体积小、成本低、效率高的优点。

对于指令中用以指定向量寄存器号的地址码,相应分为两个部分,一部分为组号地

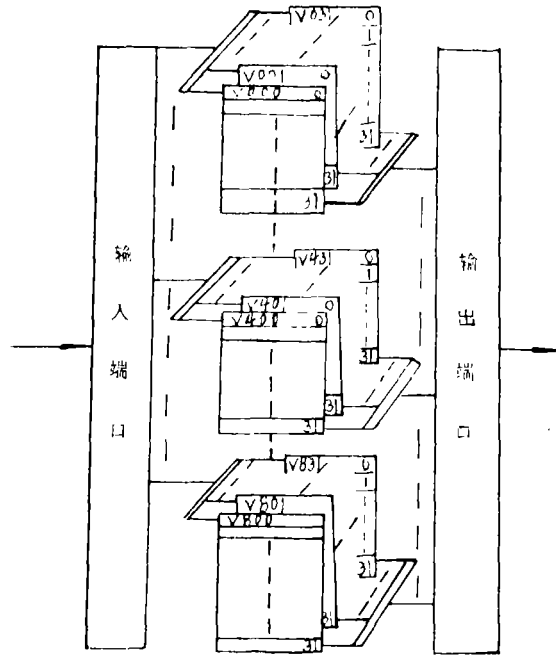
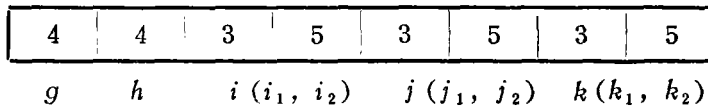


图 1 三维 (8, 32, 32) 结构向量寄存器框图

址(n), 一部分为组内寄存器号(m)。

若指令字长为32位(当机器字长为64位时, 该指令系统可以设计得与CRAY-1兼容), 其向量型算术/逻辑运算指令格式为:



其中 g 、 h (各4位) 为指令操作定义符;

i (8位) 为结果数地址。 i_1 (3位) 指明向量寄存器组号, i_2 (5位) 指明组内的寄存器号, j 、 k (各8位) 为两个操作数的地址码, j_1 、 k_1 (各3位) 分别指明两操作数的向量寄存器组号, j_2 、 k_2 (各5位) 分别指明两操作数的组内寄存器号。

约定: 若 $j_2 \neq k_2$, 则不允许 $j_1 = k_1$ 。由于同组内的各寄存器只能分时占用同一输入及输出端口, 所以这个约定是必要的。

由于这种设计允许一个向量寄存器在一拍之内完成读和写两种功能, 这使得“链接”可在任意时刻进行。则用文件寄存器构成V寄存器时即可解决, 或者用双体结构, 输入和输出端口分别设计也可解决。这样可大大提高功能部件操作的并行度。

所谓动态结构, 便是根据向量长度寄存器VL所指示的向量长度将向量寄存器构成相应的寄存器组, VL不同, 则使用的寄存器个数不一样。例如(VL)大于基本V寄存器片所包含的单元数, 则重新构成的寄存器组中一个V寄存器, 相当于基本V寄存器顺序使用下一个基本寄存器的诸单元。如此下去直到包含(VL)个单元为止, 此时, 所占的若干基本寄存器片看成一个V寄存器。而这样的几个寄存器就构成为寄存器组。所以用户只需给出所要加工的向量长度(VL), 机器便能自动组合成适当的组, 以达到最好的处理效果。如果一段程序中使用了不同的向量长度参数VL, 向量寄存器就自动的随其变

化而变化。

具有上述指令系统的计算机,设有8组,每组32个向量寄存器,若每个向量寄存器基本单元数为32,每个单元为64位,则允许一条指令可处理的向量长度最大为1024 (32×32),可用的向量寄存器为8组,每组是一个1024单元的寄存器,对应 $(VL) = 1024$,记为 $8 \times 32 \times 32 \times 64$ 。若每个向量寄存器基本单元数64,则记为 $8 \times 64 \times 16 \times 64$,适用于一般处理。以后者为准,当 $(VL) = 256$ 时,则可提供的寄存器组为16。概括之:

(VL)	64	128	256	512	1024
寄存器个数	128	64	32	16	8

总容量为64KB。上述组合可随 (VL) 而自动组合,非常方便、灵活,以适应用户不同程序中对不同单元数(即向量长度)和对不同向量寄存器个数的要求。所以比固定结构的向量寄存器可取得更高的效率。而工程实现上,控制起来与固定结构时同样简单。

如果每个向量寄存器 (V) 以双体结构设计时,则各有独立的输入、输出端口,这时当输入端口工作,输出端口也可以同时工作。所以虽然向量寄存器组数为8,实际有可16个寄存器处于工作状态:8个处于读、8个处于写状态,这样不但达到每拍可以“链接”,而且提高了并行度,例如:

$$V_{500} + V_{600} \longrightarrow V_{710}$$

$$V_{321} + V_{416} \longrightarrow V_{600}$$

二条指令可在两拍内相继流出,而不会因第二条指令的结果数地址是第一条指令的操作数地址而延缓流出。

在设计中,端口本身并不另加器材,只需 V 本体能够一拍内完成对一个 V 的读操作和写操作。这可通过利用RAM组件读/写时序关系,在时间上进行配合而降低对RAM组件访问速度的要求。

三、设计举例

例1 应用立体动态结构向量寄存器的思想,我们设计了一个大型高速向量寄存器,其基本指标为:

- 1) 寄存器总容量8192字(64KB),组数固定为8,每个组内寄存器个数可为:1, 2, 4, 8, 16, 32,每个寄存器的单元数相应为:1024, 512, 256, 128, 64, 32.
- 2) 每组的输入端口和输出端口是独立的,所以可同时进行读、写两种操作,在60MC主频下,每拍可分别写入和读出一个数,“链接”操作可在该指令执行中的任意CP进行。CRAY-1所具有的“循环特性”没有了,而代之以常规定义下的操作。于是用过去之后不再使用的操作数可被立即冲掉。例如相邻的指令是逻辑加和浮加指令时,

$$V_{000} + V_{100} \longrightarrow V_{000}$$

$$V_{200} + F V_{(300)} \longrightarrow V_{100}$$

两条指令可在连续两个CP内流出执行,指令执行完后 V_{000} , V_{100} 存放的是指令操作符到定义的操作结果,而非原来所存放的操作数。

该向量寄存器的总体结构见图 1。选用读写周期为 5ns 的 $1K \times 4$ 的 RAM 存储组件构成 V 寄存器本体，输入和输出端口、读和写控制电路以及读地址和写地址寄存器，都分别设计。每个 CP 前半拍读出、后半拍写入。其框图见图 2。

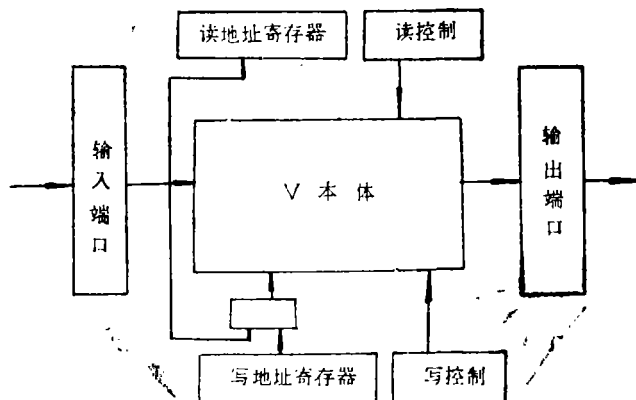


图 2 某组寄存器单体结构框图

每组寄存器为双体结构。对甲体读的同时，可对乙体写，对乙体写的时候，可对甲体读。输入和输出端口是独立的，以使其能同时进行读和写两种操作，其框图见图 3。

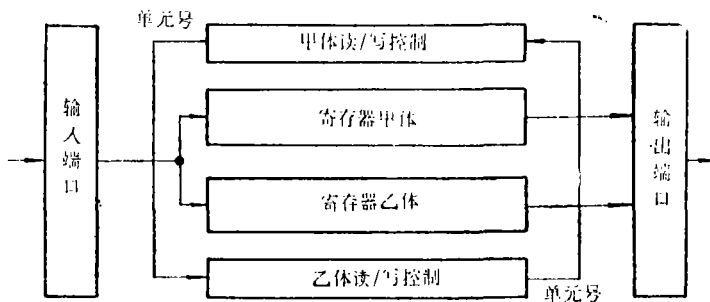


图 3 某组寄存器双体结构框图

当选用文件寄存器构成寄存器本体时，其框图与图 2 类似，只是读、写地址直接分别送入寄存器本体，不需要分时传送，因而控制要简单些。一般文件寄存器的容量不太大，因而用的组件数量要多些。

例 2 我们以 CRAY-1 机向量寄存器为基础，在硬件改动不大、而且指令兼容(和原机比较，只扩充了两条指令的功能)的条件下，将原机的 8 组乘 1 个 64 单元的平面结构向量寄存器，稍加改动而扩展为 8 组乘 4 个 64 单元的立体结构的向量寄存器，寄存器体积和成本均没有增加(原机使用 16×1 位的 RAM 组件，现改为用 64×1 位的 RAM 组件)，而且寄存器的容量大了四倍，程序可址编的寄存器个数，从 8 个增加到 32 个。如果使用更大容量的组件，其向量寄存器容量还可扩大，且体积不会增大。

上述二例说明，立体动态结构向量寄存器的设计思想用于实际，其向量寄存器具有体积小、成本低、效率高的优点。从而可以预料，在新的高性能巨型机的研制中，设置超大容量向量寄存器的方案变得现实而可行。

胡守仁教授、李勇副教授对本文的研究给予重视和支持,周兴铭副教授曾审阅了全文,并提出改进意见。作者一并致以谢意。

参 考 文 献

- [1] CRAY 研究公司, CRAY-1 计算机硬件参考手册, 1977。
- [2] 慈云桂、胡守仁, 亿次级巨型机技术发展近况, 计算机工程与科学, 1983。
- [3] 胡守仁、周兴铭、张德芳, 巨型机系统结构, 国防科大研究所, 1981。
- [4] 金兰等, 并行处理机系统结构, 国防工业出版社, 1981。
- [5] 张民选、刘毅之、刘进学, 一种扩充向量寄存器的新方法, 全国中大型计算机应用与开发学术交流会议论文, 1983。

Design Idea of the Large-capacity Vectorial Register and its Realization

Zhang Minxuan, Liu Yizhi, Liu Jinxue

Abstract

This paper advanced a design idea to design the vectorial register of the 3-dimensional structure, described its operational principles, moreover, gave two practical examples of the application of this idea to design and to expand CRAY-1 type huge computer's vectorial register.