

# 大规模并行处理器阵列中的 光纤互连网络

周革 何晓清 张以谟 刘文耀 华锋

(天津大学现代光学仪器研究所, 天津 300072)

**摘 要** 详细介绍了具有  $64 \times 64$  I/O 通道的光纤互连网络的设计原理、结构、制作技术和实验结果。此网络应用在两个并行处理器阵列之间, 具有可重构功能, 实验结果验证了光纤互连的优越性和可行性。

**关键词** 光纤互连网络, 并行计算, 处理器阵列.

## 1 引 言

在计算机系统中采用光互连的设想提出已有十年了<sup>[1]</sup>。十年来经过人们的不断努力, 在体系结构、互连模式、实现方法、光电子集成器件等方面已取得了许多进展。目前光互连技术正受到人们的关注。由于光互连技术在传输速率、带宽和抗干扰性等方面的许多优点, 使其能为大规模并行阵列之间的通信提供很多益处<sup>[2]</sup>。从光互连技术的发展来看可以分成两种主要的研究方面<sup>[4,5]</sup>, 一是用光学方法代替电子互连; 另一种则是光学和电子的混和体, 充分发挥光和电各自的优点。前一种方法由于器件缺乏或者系统庞大等原因, 达到实用还需要做更多的工作。后一种方法比较现实, 在目前的条件下更容易实现, 从而被人们普遍接受。本文将详细介绍基于光电混合思想实现的光纤互连系统。

## 2 并行处理器阵列与光纤互连系统的作用

并行处理技术是当前提高计算机计算能力的重要手段之一, 在并行系统中, 互连网络占有重要地位。当并行阵列规模较大、传输速率较高时, 互连网络将变得非常复杂, 这就给网络的设计和制作提出了更高的要求。光互连技术为复杂网络的实现提供了新的技术手段。本文研制了光电混和并行处理系统, 并采用光互连技术进行数据通讯。

光电混和并行处理系统(如图 1 所示)共有五个部分组成: 即处理器阵列、光纤互连网络、数据采集系统、共享存储器、管理主机。其中前两部分为主体, 后三部分为系统服务。下面重点说明阵列的组成和光纤网络的结构与功能。

整个系统有两个处理器阵列, 每个阵列包含  $8 \times 8$  个处理器单元(PE), 故整个计算系统

\* 国家科委 863 高科技项目。

收稿日期: 1994 年 11 月 22 日; 收到修改稿日期: 1995 年 1 月 5 日

共有 128 个处理器单元。每个阵列为一个局部整体，其内部互连用电子线路完成；两个阵列之间用光纤互连网络连接。

由于系统中互连网络的容量和复杂程度随着每个阵列处理器单元数目的增长成倍数增加，若 128 个处理器单元组成全互连结构，则需要  $128 \times 128$  (I/O) 交换网。为了减少互连网络的硬件开销，在系统中采用了空间复用和时间重叠相结合的并行计算模式<sup>[3]</sup>。这样在一个阵列内部可将任务分割成 64 块，由 64 个处理器单元进行相同的运算(即 SIMD 运算模式)或不同的运算(即 MIMD 运算模式)。在阵列之间采用流水线工作方式，将时序过程分成若干个子过程，前一个阵列的输出作为后一个阵列的输入；同样，后面的计算结果也可以反馈给前面的阵列进行再次运算，且在后面阵列计算的同时，前面阵列可执行另一个任务的运算。这样，整个系统内部互连只需 2 个  $64 \times 64$  交换网，从而使系统既具有很高的并行度，又可平衡硬件的开销。

系统的互连分为两个层次，一个是阵列内部的电互连，它由一个  $64 \times 64$  (I/O) 的 Crossbar 交换开关构成，如图 2 所示。另一个层次为阵列之间的光互连网络，它在管理主机的控制下完成阵列之间数据的双向交换。

光互连网络的结构采用双向折叠式<sup>[2]</sup>，由电子交换开关和光纤传输介质组成，如图 1 所示。光互连网络构成了系统中的第三维互连方向，它可以使两个阵列的相对位置不再受导线长度的限制(系统中电路连线长度不能超出 30 cm)，同时光纤具有良好的抗干扰特性，保证了系统运行的稳定性。

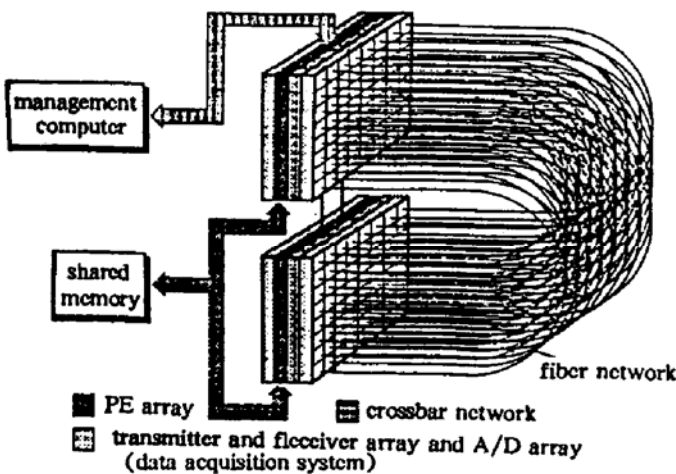


Fig. 1 The architecture of the optoelectrical hybrid parallel processing system

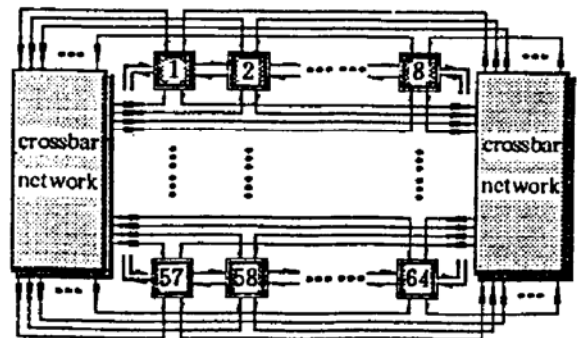


Fig. 2 The interconnection within an optoelectrical hybrid parallel processor array

### 3 光纤互连网络的设计与制作

光互连网络的任务是将前一个阵列上的第  $(i, j)$  个处理器单元的数据直接送到后一个阵列的第  $(l, m)$  个处理器单元中，即

$$PE_1(i, j) \Leftrightarrow PE_2(l, m), \quad j, i, l, m = 1, \dots, 8 \quad (1)$$

这种传送功能是双向的，且  $i, j, l, m$  可随意选择。根据这个要求和系统运算模式，综合考虑减小系统的硬件开销，光互连网络采用了与阵列内部互连共用 Crossbar 网络的形式(参见图 3)。在 Crossbar 的输出端口上采用一个复合开关做信道转换，当阵列内部运算时，复合开关的输

出接向阵列内部；当阵列之间进行数据交换时，复合开关接向光互连网络的输入端口。

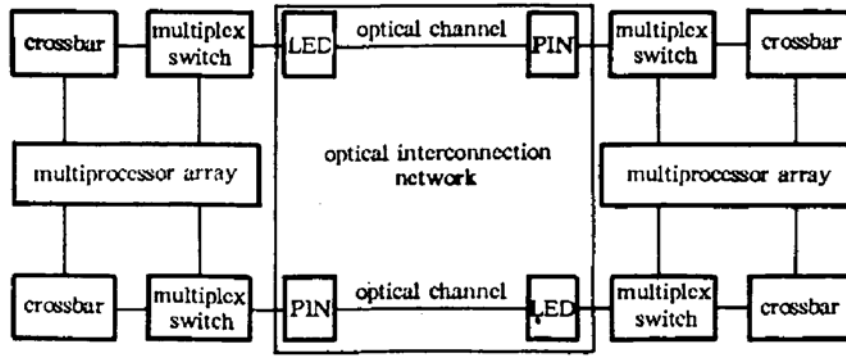


Fig. 3 The interconnection network of optoelectronic hybrid parallel processing system

Crossbar 开关网络的通道数为  $64 \times 64$  (I/O)。将交换开关与光学通道结合后，便可实现 (1) 式的功能。PE<sub>1</sub>(i, j) 将数据输出至 Crossbar 的 (i, j) 输入端口，在 Crossbar 中进行通道变换，Crossbar 的 (l, m) 输出端口与光学信道连接，经光学信道把数据送到 PE<sub>2</sub>(l, m) 中。由于光学信道是一对一连接的，保证了互连的准确性和可靠性，并具有实用价值。

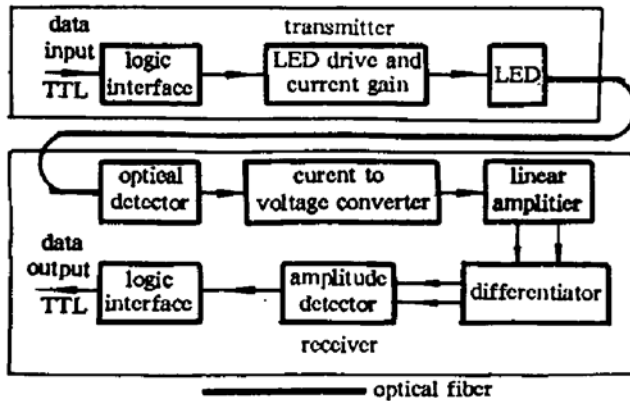


Fig. 4 The layout of optical interconnection network

光互连网络由电光转换(E-O)部分、发光器件、传输介质、通道交换部分、光电转换部分(O-E)及信号整形部分组成，如图 4 所示。发射端的输入信号和接收端的输出信号均为 TTL 电平，因此，光学信道的两端可以直接与数字芯片连接。

根据处理器列阵所采用的计算机芯片的要求，光学信道的传输速率为 20 Mbit/s，因此发射端采用光发射二极管发光器件，并直接用晶体管-晶体管逻辑(TTL)芯片驱动，芯片的响应速率、信号延迟和电流驱动能力均可满足要求，而且成本较低。在接收端，光电探测器件为 PIN 型光电二极管，其输出信号先由一个宽带差分放大器放大，再通过双线接收器将信号整形为 TTL 方波。

光互连网络中的另一个组成部分为 Crossbar 开关，它的交换原理如图 5 所示。系统中的交换网是由 6 个  $32 \times 32$  (I/O) 的电子 Crossbar 开关器件构成的  $64 \times 64$  (I/O) 开关网，如图 6 所示。该网络在管理主机控制下进行拓扑结构的编程重构，从而完成系统中互连网络的程控变换，使光纤互连网络具有灵活重构的特性。

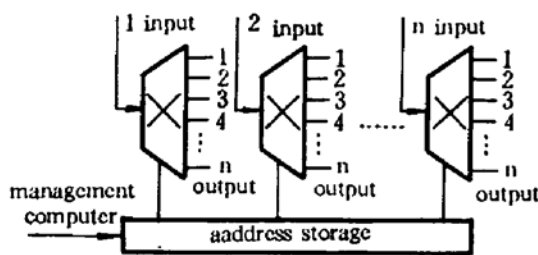


Fig. 5 The principle of programmable crossbar network

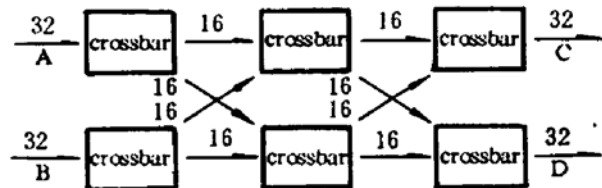


Fig. 6 The implement of programmable crossbar network with  $64 \times 64$  I/O

### 4 光纤互连网络的调试与实验

本文对所研制的光纤互连网络进行了性能测试、联机实验及网络重构实验。

#### 4.1 传输速率测试

本项实验的目的是检测光学信道所用的发光器件、光接收器件及电子线路的速率响应特性。用一个信号源产生一个含有 16 位二进制码的周期信号，码的频率为 20 Mbit/s，且码值可变。以此信号作为信号源对光互连信道进行测试。实验结果如图 7 所示，其中上面是信道输入码波形，下面是输出码波形。连续码的传输实验表明，互连网络的传输速率和码的波形均能达到阵列系统的要求。

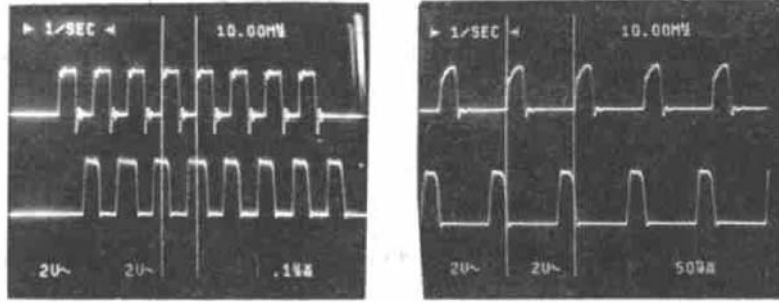


Fig. 7 The result of periodical singal transmission test

The above one is the waveshape of input singal and the below is output one

#### 4.2 误码率的测试

光互连网络误码率的测量是采用联机测试的方法。如图 8 所示，用光纤通道连接两个处理器单元，其中 PE<sub>1</sub> 产生 32 bits 的随机码，然后通过光纤信道送给 PE<sub>2</sub>，PE<sub>2</sub> 通过另一光纤信道将应答信号传给 PE<sub>1</sub>，PE<sub>1</sub> 接到应答信号后，再发出另一组 32 bits 的随机码，这样连续地传送。为了检验光纤通道传输结果正确与否，在 PE<sub>1</sub> 和 PE<sub>2</sub> 之间用一对很短的导线同时传输相应的随机码，光、电两条信道的传输结果在 PE<sub>2</sub> 中进行比较，PE<sub>2</sub> 将记录误码次数，以供最后显示。

实验结果表明，互连网络传输 320 兆每组 32 位的随机码没有出现误码，因此光互连网络的误码率小于 10<sup>-10</sup>。实验分别在 10 Mbit/s 和 20 Mbit/s 速率下进行，其中 20 Mbit/s 的传输波形如图 9 所示，图中由于测试波形是不固定的随机码，示波器屏幕上扫描线重叠出现。

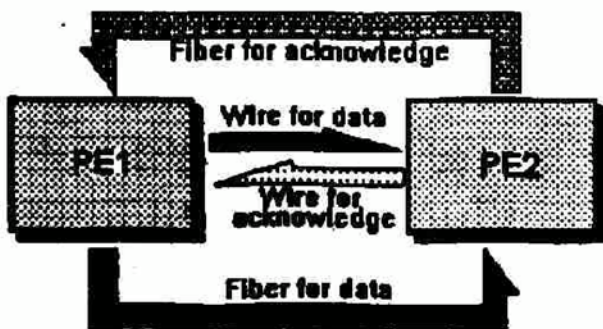


Fig. 8 The BER test of optical interconnection network

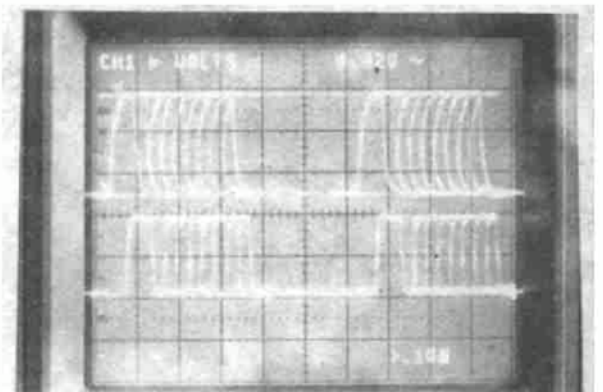


Fig. 9 The result of on-line transmission test. The above one is the waveshape of input signal and the below is output one

此外，本文还对光纤互连网络进行了拓扑重构实验，实验结果表明了光互连网络可进行程控拓扑结构变换，使并行处理器阵列系统可适用于多种计算模式的高度并行的数据处理和

信息处理。

**结 论** 本文作者已成功地研制出了具有  $64 \times 64$  I/O 交换功能的光纤互连网络, 并把它应用于并行计算机阵列系统中。互连网络采用了光子传输与电子交换相结合的方法, 拓扑结构具有灵活的编程重构特性, 从而满足并行阵列系统构成各种并行计算模式的需要。网络的 I/O 接口可直接与数字芯片相连接, 其单通道传输速率为 20 Mbit/s, 整个互连网络的并行传输速率为 1.28 Gbit/s, 网络的误码率小于  $10^{-10}$ 。

此项研究结果表明, 光互连网络的带宽不受传输长度的影响, 且具有很强的抗电磁干扰能力, 从而使光互连技术在并行处理器阵列系统中具有很大的应用潜力, 为并行处理器阵列中的高速数据通信和结构设计提供了方便。目前网络的传输速率主要取决于组成阵列的处理器单元的通信速率, 随着系统升级光互连网络的速率还可进一步提高。

### 参 考 文 献

- [1] J. W. Goodman *et al.*, Optical interconnections for VLSI system. *Proc. IEEE*, 1984, 72(7): 850~860
- [2] Edward E. E. Frietman *et al.*, Parallel optical interconnects; implementation of optoelectronics in multiprocessor architectures. *Appl. Opt.*, 1990, 29(8): 1161~1177
- [3] Harold S. Stone, John Cocke, Computer Architecture in the 1990s. *IEEE Trans. Computer*, 1991, (1): 30~38
- [4] Dean Z. Tsang, Thomas J. Goblick, Free-space optical interconnection technology in parallel processing systems. *Opt. Engng.*, 1994, 33(5): 1524~1531
- [5] Toshikazu Sakano *et al.*, Multiprocessor system using an automatically rearrangeable free-space multichannel optical switch; COSINE-2. *Appl. Opt.*, 1993, 32(20): 3690~3699

## Optical Fiber Interconnection Network in Massively Parallel Processor Arrays

Zhou Ge    He Xiaoqing    Zhang Yimo    Liu Wenyao    Hua Feng

(The Institute of Optoelectronics & Precision Eng., Tianjin University, Tianjin 300072)

(Received 22 November 1994; revised 5 January 1995)

**Abstract** The principle and architecture of a fiber optical interconnection network with  $64 \times 64$  I/O channels and its implementing means are presented. The network is applied to interconnect the two parallel processor arrays, and can be reconfigured programmably. The advantages and practicability of optical interconnection have been shown by the experimental results.

**Key words** fiber interconnection, reconfiguration, parallel computing, processor array.