

文章编号: 1001-3806(2003)03-0265-03

基于 VCSEL/MSM-CMOS 技术的并行光互连设计

陈涛 曹明翠 罗风光 袁菁 王江义

(华中科技大学激光技术国家重点实验室, 武汉, 430074)

摘要: 提出了一种基于 VCSEL/MSM-CMOS 技术的并行光互连系统, 详细介绍了系统整体设计, 着重对光网络通信接口卡中的 PCF-HIPPI 接口模块、HIPPI-FIBER 接口模块、光总线模块设计进行了阐述。

关键词: 并行光互连; 垂直腔表面发射激光器; 接口

中图分类号: TP393 **文献标识码:** A

Design of parallel optical interconnection based on VCSEL/MSM-CMOS technology

Chen Tao, Cao Mingcui, Luo Fengguang, Yuan Jing, Wang Jiangyi

(National Laboratory of Laser Technology, HUST, Wuhan, 430074)

Abstract: A novel parallel optical interconnection system based on VCSEL/MSM-CMOS technology is presented in this paper. The system design is introduced in detail. And the highlight is laid on PCF-HIPPI interface module, HIPPI-FIBER interface module and optical bus module in the interface card of optical communication network.

Key words: parallel optical interconnection; VCSEL; interface

引言

在并行计算系统中, 有两个因素对系统的计算性能有重要影响。一是微处理器的速度, 一是计算节点之间的互连网络性能。未来几年内, 微电子技术将推动计算机微处理器中互连密度和长度呈数量级的提高^[1]。不过物理极限(如电子衍射)的存在最终将使得 VLSI(very large scale integration)的规模和内部时钟频率达到上限。作为并行计算系统另一个关键技术的互连网络, 正在经历由电互连向光互连的转变。“光互连”方式具有极高的时间空间带宽积、高速、高密度、无干扰和传输功耗极低等优点, 发展宽带高速光互连网络系统已成为人们的共识。

提出的并行计算平台光互连系统采用并行光纤带构成光学广播网络, 用 VCSEL 列阵集成光发射模块, PIN 列阵集成光接收模块, 高精度光纤列阵 I/O 接口, 使得整个互连网络具有集成度高、体积小等优点; 在各个计算终端上设计了专用的通信接口

卡, 实现选路通信功能, 充分利用现有硬件资源快速互连构成并行计算环境。

1 系统设计

1.1 现有光互连并行处理系统方案

加拿大 McGill 大学的 Szymanski 等人提出了一种基于中央星型拓扑结构的多处理系统 Terabit 光局域网^[2], 每台工作站通过并行光纤带(parallel fiber ribbon)与交换核直接相连。每个连接可提供 6.4 Gbits/s 带宽, 运用 CMOS IC's 互连 32 个工作站的中央电交换核提供了 204.8 Gbits/s 的带宽, 能够在单个 PCB 上实现。系统采用了 Motorola 的 Optobus 平行光纤带, 当前的平行光纤带可以提供 10 根光纤, 光时钟频率为 800MHz。由于电总线速率典型只有 100MHz, 因此, E/O 模块的作用是提供不同时钟频率的光/电之间的接口。Optobus 收发模块中实现了光电转换。

天津大学井文才等人提出了一种光互连环网机群系统^[3], 数据在网络中以数据包的形式进行传输, 各结点通过网络接口卡连到光纤环中。数据包由数据传输的发送源地址、目的地址、数据长度标识、可变长度的数据段和循环冗余校验(CRC)段组成。结点的地址编码存储在各结点网络接口卡的

作者简介: 陈涛, 男, 1978 年 2 月出生。硕士研究生。主要从事宽带光通信、光互连网络及并行处理计算机技术的研究工作。

收稿日期: 2002-05-01; 收到修改稿日期: 2002-07-10

编程逻辑器件中,地址编码在整个网络中是唯一的。

1.2 互连链路设计

上边提及的并行计算互连系统由于结构不同有不同的特点。中心交换核结构,其所有的节点都通过光链路连接到中心交换核,系统的规模完全取决于交换核的端口数,一旦端口数确定,在最大规模数以下,可以方便地增加减少节点。采用这种结构,在具有 N 个独立节点的系统中,只需要 $2N$ 条单向光链路。但这种结构由于中心交换核的数据流量大,对硬件设计要求很高,需采用专用大规模集成电路芯片来实现。光纤环网简单易行,但是系统脆弱、可靠性差,当环出现断路时,整个系统通信瘫痪,且当系统节点数增加时,点到点通信延时成比例增加。

笔者提出的系统结构如图 1 所示,系统中互连的 4 台计算机,每台都有一块接口卡(card1 ~ card4),

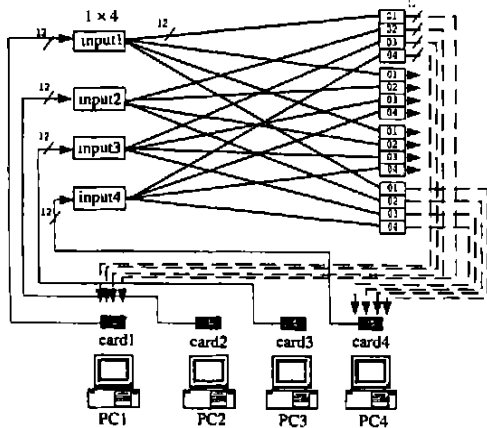


Fig. 1 (4 × 12) × (4 × 12) optical interconnect network

卡上集成有 1 × 12 的垂直腔表面发射激光器 (VCSEL) 列阵和 4 × 12 MSM 探测器列阵收发模块。该系统可实现 4 台计算机之间的双向光互连,每台计算机接口位宽为 12bit。例如,计算机 PC1 中的数据信号通过并/串转换进入 card1 上 VCSEL 的 12 条发送通道,由 1 × 12 VCSEL 发射窗口发出的数据光信号通过光纤列阵接口器件耦合到光纤列阵中进行并行传输,输入到光互连光学数据总线模块的 input1 中。光互连数据总线由 48 个 1 × 4 分束器构成,经分束后形成 4 × 48 的输出信号光,其中每个 input 包含 12 个 1 × 4 分束器。input1 输出 4 × (1 × 12) 的光信号,每个 (1 × 12) 的输出光信号与其余 3 个 input 的每个 (1 × 12) 输出光信号形成 1 × 48 的光信号,最终形成 4 × (1 × 48) 输出光信号。此时再利用光纤将输出信号光与 4 × 48 高密度规则排列的多层 2 维光纤列阵光电收发接口器件耦合后,分别输入到每台计算机的 1 × 48 MSM 探测器的窗口列

阵上,从而将任一计算机的信号送入其余计算机中。然后,经光电转换后形成的电信号通过与 MSM 光电探测器列阵芯片相连的 CMOS 逻辑选通电路,由 CMOS 逻辑选通电路选择其中一组 1 × 12 并行输入数据信号,输入到该计算机,从而可实现任意计算机之间信号的双向互连通信。该系统结构具有 4 台计算机完全无阻塞发送接收信息、传送速率快,结构简单可靠等优点,可同时 4 × 12 路数据通道的信号进行互连交换。

2 光互连网络通信接口卡设计

采用 HIPPI (high performance parallel interface) 接口作为计算机 PCI 总线与光纤之间的互连通道,每块接口卡由 PCI-HIPPI 转换模块、HIPPI-FIBER 转换模块以及光电转换模块构成 (见图 2)。

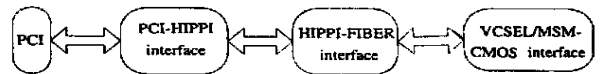


Fig. 2 Interface principle

2.1 PCF HIPPI 接口模块

PCF HIPPI 模块完成接口卡和计算机总线之间的接口逻辑,采用美国 Applied Micro Circuits Corporation (AMCC) 公司的专用芯片,外围控制逻辑采用美国 Altera 公司的 EPLD 逻辑器件。数据收发模块主要完成 PCI 接口到 HIPPI 接口的协议转换,DC/DC 模块产生系统需要的电压源。通过该模块,实现了数据转换功能。

2.2 HIPPI FIBER 接口模块

HIPPI 接口向光纤接口转化模块,主要是将 HIPPI 接口传送的信号通过一定的转换,转变成 12 路电信号,然后通过 E/O 转换进入光互连总线模块 (见图 3)。

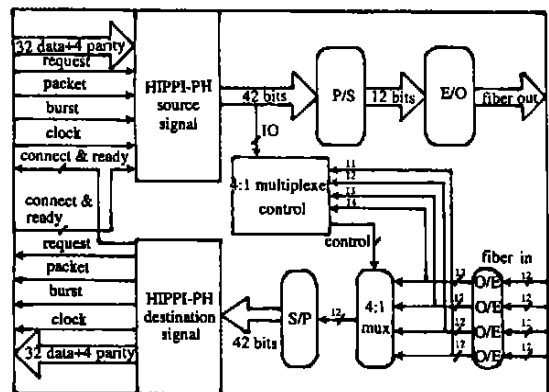


Fig. 3 HIPPI-FIBER interface

HIPPI 接口模块同时还要完成仲裁选通等一系

列的控制功能。仲裁机构用于控制 4 1 复用器,它提取出源模块输出的 I-field 信号和目标模块接收的 4 路 I-field 信号,并根据 I-field 信号的定义,产生相应的控制信号送入 4 1 复用器。4 1 复用器根据仲裁机构的控制信号,进行相应的通道选择,在通信结束前,仲裁机构对输入信号不响应。串/并转换部分将接收的 12 路信号转换为原始发送信号,送入 HIPPI 目标模块,从而完成计算机间的通信。该模块的串/并、并/串、4 1 选路控制功能均由 Altera 公司的可编程器件实现。

2.3 光总线模块

光总线互连模块主要完成光信号的广播、传输等功能,是一个 4 × 4 光学广播网络。采用分光的方法,由分布式的 1 × 12 VCSEL 光发射阵列和 4 块 1 × 48 MSM 探测阵列组成,把每台计算机出来的电信号由 VCSEL 器件转换为光信号,利用分束器分为 4 路,再分别送回 4 台计算机,实现计算机之间的互连,完成信息交换。

垂直腔表面发射激光器(VCSEL)与边发射半导体激光器相比,具有较小的远场发散角、发射光束既窄又圆、阈值电流低、功耗低、调制频率高、在很宽的温度和电流范围内都以单纵模工作,而且不必解理,就能完成工艺制作和检测,制造成本低廉。根据 Dennard 提出的器件尺寸按比例缩小原理,随着器件尺寸的缩小,集成密度呈平方关系提高,速度和功耗大幅度改善,易于实现大规模激光阵列及光电集成等优点。利用 VCSEL 阵列和光纤带耦合形成的并行数据传输链路,在并行光处理中得到了较为广泛的应用。

(上接第 264 页)

3 结 论

在详细讨论入射基频光的振幅调制和位相扰动对三倍频转换效率和动态范围影响的基础上,并采用 / 类偏振失配的 KDP 晶体三倍频方案,对倍频晶体的参数进行了优化设计。

研究结果表明,由于振幅调制和位相扰动的存在,致使三倍频转换效率明显下降和倍频系统的动态范围大大减小。在入射基频光功率密度较大的情况下,振幅调制和位相扰动对倍频器转换效率和动态范围的影响尤为明显。通过适当减小倍频晶体的

2 维光纤阵列收发接口器件(见图 4)用于实现 VCSEL/MSM-CMOS 光电集成器件的耦合输入输出,并提供光互连网络模块和计算机之间的光互连并行信号通道。笔者研制了 4 层、每层光纤数为 48 的 4 × 48 的高密度多层 2 维光纤阵列收发接口器件,用于计算机的光信号收发接口耦合输入输出。每层光纤阵列内相邻光纤的间距为 125 μm,相邻层光纤阵列间光纤的间距为 250 μm。

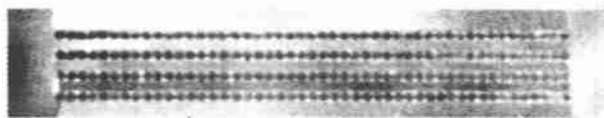


Fig. 4 4 × 48 high density multilayer 2D fiber array transceiver interface

3 结 论

比较了已有的光互连系统方案,并在此基础上利用 HIPPI 协议、VCSEL 集成器件提出了一种新的基于 VCSEL/MSM-CMOS 技术的光互连并行处理方案。同时设计了与之相对应的接口卡和接口器件,实现了 4 台计算机、每台接口 12bit 位宽的并行连接。在利用现有资源的基础上,实现了并行处理环境。

参 考 文 献

- [1] Cangellaris A C. MPP01 '98, Proceeding of the fifth international conference on (MPP01 '98). Las Vegas: Institute of Electrical and Electronics Engineers, 1998: 96.
- [2] Szymanski T H, Au A, Lafreniere-Roula M. Appl Opt, 1998, 37 (2): 264 ~ 275.
- [3] 井文才, 田劲东, 刘 卫 *et al.* 光电子·激光, 2000, 11(1): 7 ~ 10.

厚度使三倍频转换效率有一定程度的提高,且在振幅调制和位相扰动情况下倍频系统还获得了较大的动态范围。

参 考 文 献

- [1] Wegner P J, Hensian M A, Speck D R *et al.* Appl Opt, 1992, 3 (30): 6414 ~ 6426.
- [2] Craxton R S. IEEE J Q E, 1981, QE17(9): 1771 ~ 1782.
- [3] 秘国江, 魏晓峰, 袁晓东 *et al.* 强激光与粒子束, 1999, 11(3): 313 ~ 316.
- [4] Eimerl D, Auerbach J M. J Modern Opt, 1995, 42(5): 1037 ~ 1067.
- [5] Craxton R S, Jaccobs S D, Rizzo J E *et al.* IEEE J Q E, 1981, QE17(9): 1782 ~ 1786.