具有旋转连接功能的双层光互连网络设计*

贾大功 刘 琨 井文才 张以谟 周 革
(天津大学精密仪器与光电子工程学院,光电信息技术科学教育部重点实验室,天津 300072)

摘 要 在设计的双层光互连网络中,上层网络是星型连接,依靠数字路由结点进行通信. 单通道最大传输速率为 1.4 Gbps,数字路由结点吞吐率大于 10 Gbps,底层为网络接口卡和结点机连成的环形网,峰值传输速率为 1.056 Gbps. 在底层网络中个别结点机与传感器之间配置有光纤旋转连接器,可传输动态数据. 经计算,环网内最大通信延迟时间小于 5.292 μs,互连网络的平均通信延迟时间为 11.03 μs,环形网络的最大数据传输带宽为 50 Mbit/s.

关键词 光互连;动态连接;光纤旋转连接器;时间延迟;传输带宽
中图分类号 TN929 **文献标识码** A

0 引言

随着并行计算系统的发展,人们对系统性能的要求越来越高,原有的电互连技术已经成为系统性能提高的障碍[1~3]. 高效的光互连网络为大规模并行处理数据提供了有力支持,特别是在机群系统中采用光互连链路可有效提高网络的带宽、速率和吞吐量[4,5]. 目前,光互连技术研究热点主要集中在芯片间与芯片内的光互连[6~8]. 天津大学光互连实验室这几年分别研制出具有 64 个处理器的光电混合并行处理系统和应用于机群系统的 Gbit/s 光互连链路[9].

在海底探测系统中,大量传感器采集到的数据经过安装在转动绞盘上的光纤旋转连接器后[10],到达数据处理中心,由多个计算机并行处理数据. 这些数据流的传输是一种动态、实时的传输. 这就需要光互连网络不仅能高效处理数据而且能够动态传输数据. 为此,本文设计了一种具有旋转连接功能的光互连网络,并对网络的性能进行了分析.

1 光互连网络结构

图 1 是具有旋转连接功能的双层并行光互连网络. 它采用星型与环形结构相结合的网络拓扑结构,上层是星型结构,下层是环形结构. 结点机(Node Computer,NC)通过光网络接口卡(Optical Network Interface Card,ONIC)与网络相连. 底层环网中一台 NC 上有两块 ONIC,一块与数字路由结点(Digital Routing Node,DRN)相连实现环网转发,另一块与环网相连实现网内通信. 设计时,安排 1 个或 2 个网络进行数据采集,其余网络进行数据处理. 在数据采集环网中,传感器(Sensor)通过光纤旋转连接器(Fiber Optic Rotary Joint,FORJ)与 NC 相连. 这样,该网络能够完成多点、多方式同时测量和并行计算.

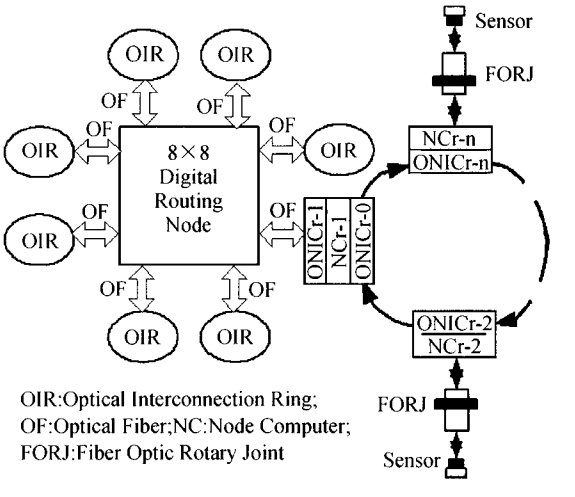

图 1 具有动态数据传输功能的双层光互连网络
Fig. 1 Structure of two-layer parallel optical interconnection network

DRN 采用光传输和硬件路由技术设计,其上集成有光电(Optical/Electrical, OE)和电光(Electrical/Optical, EO)转换模块,交叉开关(Cross-point switch),逻辑仲裁(Arbitration logic)模块,可实现路径判断和数据交换功能. 交叉开关单通道最大传输速率为 1.4 Gbps,且支持实时重构以实现 8 个端口的全双工传输,因此数字路由结点的吞吐率大于 10 Gbps.

ONIC 原理结构图见图 2. 按照接口卡上各模块的功能不同可分为:PCI 接口控制功能模块、核心控制模块(FPGA)、高速缓存(FIFO)、串/并转换和并/串转换模块、电/光和光/电转换模块五个部分. FPGA 的主要功能是完成数据的复用和解复用,对

* 国家自然科学基金(60377031,60577013)和 973 计划子课题(2003CB314900)资助
Tel:022-27403147 Email:dagongjia@tju.edu.cn
收稿日期:2005-07-18

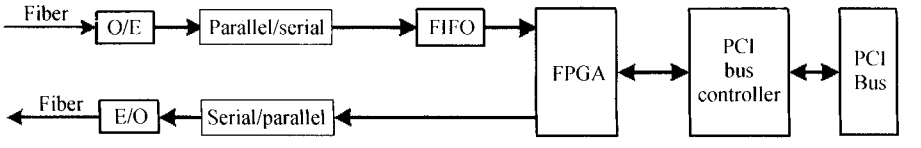

图 2　光链路接口卡原理

Fig. 2　Schematic of optical network interface card

接收的数据包进行地址判断后接收或转发数据包，协调控制相邻芯片的工作状态，发送结点机的数据包．该卡采用 32 位数据线，工作在 33 MHz 时钟频率下，支持 132 Mbit/s 峰值速率数据传输，接口卡的峰值带宽可达 1.056 Gbit/s.

FORJ 是一个具有旋转耦合面的光纤器件，即在信号传输过程中，有时需要信号从一端静止的平台，通过一个旋转平面不断地传输到另一端转动的平台[11].

FORJ 的结构如图 3. 图中旋转连接器主要由两大模块组成，分别是定子(Stator)和转子(Rotor)(图中虚线框部分)[12]，在定子和转子上分别装有光准直系统，每个系统由单模光纤(SMF)、C-lens 透镜和连接器(Connector)构成[13]，其作用是传输进来的光信号经过该系统扩束，在旋转连接器内部耦合通过旋转面后，再由转子中的准直系统接收，从而持续传输光信号．转子部分依靠球轴承(ball bearing)绕旋转轴转动，在转子的外部装有套筒(House)，防止外部光线对光信号产生干扰．该 FORJ 具有双向传输光信号的能力，解决了上行和下行信号传输的问题．整个系统中，外部转动的传感器不断地把探测到的信号通过旋转连接器传送到中央并行网络计算中心，同时也把控制信号传送到传感器的控制系统中．

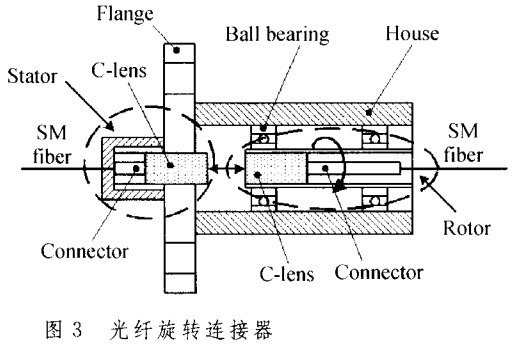

图 3　光纤旋转连接器

Fig. 3　Structure of fiber optic rotary joint

2　性能分析

通常影响光互连网络性能有以下几个因素：数据传输的通信延迟、网络传输带宽和传输链路长度等，其中前两个因素是网路性能的衡量参量．在分析时，所设计的光纤旋转连接器为无源光纤器件且光纤长度有限(约 0.4 m)对通信延迟和传输带宽的影响可忽略不计．本文中互连网络由星型和环型两级构成.

2.1　时间延迟

数据包在网络中的传输形式有两种：环内传输和环间传输．下面就这两种传输形式的通信延迟进行分析．环网内通信采用单向循环方式，同一时刻只有一个结点机发送数据，其他结点处于接收或转发状态．对于由 n 个结点机组成的环网，假设 ONIC 发送数据延迟为 Δt_1，数据包接收和转发的通信延迟为 Δt_2，数据包在光纤中的传播延迟为 Δt_3，则环网内最大通信延迟为

$$L_{\max}=(n-1)(\Delta t_1+\Delta t_2+\Delta t_3) \tag{1}$$

利用逻辑分析仪测试可以得到 ONIC 发送数据延迟时间 Δt_1 为 210 ns，结点接收和转发延迟 Δt_2 小于 328 ns[14]．设相邻结点机之间的光纤长度为 10 m，光在光纤内的传输速度为 C/n_0，其中 n_0 是芯部的折射率，计算后可得延迟时间 Δt_3 为 50 ns. 由于通信延迟与结点机的数量呈线性关系，为了保证延迟时间在理想范围内，结点机的数量不能太多，这里设 $n=10$，根据式(1)便可计算得到环网内最大通信延迟时间为 5.292 μs，远小于 LAN 的时间延迟(延迟时间为 ms 级)[15].

环内传输的平均通信延迟为

$$L_p=\frac{n(\Delta t_1+\Delta t_2+\Delta t_3)}{2} \tag{2}$$

由式(2)计算后，得到环网内平均通信延迟约为 2.6 μs.

环网之间的通信延迟包括三部分，第一部分是环网内通信延迟 Δt；第二部分是路由结点进行地址判断、数据交换和交叉开关配置的通信延迟 Δt_r；第三部分是数据包在光纤链路中的通信延迟 Δt_f.

若光互连网络中由 $m(\leqslant 8)$ 个环组成，在每个环中有 n 台结点机．在互连网络中，路由结点进行地址判断和数据交换的通信延迟约为 210 ns，交叉开光配置的时间为 40 ns，所以 $\Delta t_r=250$ ns；环网与路由器之间光纤链路的长度为 5 m，计算可得 $\Delta t_f=25$ ns. 网络的平均通信延迟时间为

$$L=\Delta t_r+\sum_{i=1}^{m}\left|\begin{array}{l}\dfrac{n}{2}(\Delta t_1+\Delta t_2+\Delta t_3)+2\cdot\\ \Delta t_f+\dfrac{n}{2}(\Delta t_1+\Delta t_2+\Delta t_3)\end{array}\right|=$$

$$\Delta t_r+\sum_{i=1}^{m}[n(\Delta t_1+\Delta t_2+\Delta t_3)+2\cdot\Delta t_f] \quad (3)$$

根据式(3)可以估计出双层网络的平均通信延迟时间为 11.03 μs(设 $m=4,n=10$). 所以所设计的光互连网络可用于计算机机群系统中结点机之间的通信.

2.2 数据传输带宽

在设计的光互连网络中,不同结点机之间靠 ONIC、光纤链路和路由器进行互连. 在底层网络中 ONIC 的数据链路层支持 DMA 传输和 I/O 传输两种方式[16],其中 DMA 传输速度快,传输不占 CPU 资源等优点使其适合大数量数据传输. I/O 传输带宽较低,传输时占用 CPU 资源,但可随时进行传输的特点使其适合小数据量传输.

使用逻辑分析仪对数据传输带宽进行测试,经测试可得:DMA传输模式下有效传输带宽可达 50 Mbit/s,在32 bits I/O传输模式下带宽约为 10 Mbit/s.

3 结论

本文设计了一种具有旋转连接功能的双层光互连网络,该网络由数字路由结点、光网络接口卡、结点机和光纤旋转连接器构成. 网络的上层是星型连接,网络吞吐量大于 10 Gbps;下层是环形连接,接口卡的峰值带宽可达 1.056 Gbps;文中从通信延迟和数据传输带宽两个因素分析了双层互连网络的性能. 通过计算可知,网络的通信延迟能够满足机群系统中网络通信的要求;在底层网络内部,最大数据传输速率约为 50 Mbit/s.

参考文献

1 Patel R R, Bond S W, Pocha M D, *et al*. Multi wavelength parallel optical interconnections for massively parallel processing. *IEEE Journal of Selected Topics In Quantum Electronics*, 2003, **9**(2): 657～666

2 李燕,许迈,马少杰,等. 用于光互连的一种非线性调制光栅. 光子学报,2000,**29**(Z1):434～436
Li Y, Wu M, Ma S J, *et al*. *Acta Photonica Sinica*, 2000, **29**(Z1): 434～436

3 刘中林,曹明翠,李洪谱,等. 自由空间光交换网络中 FET-SEED 灵巧像素(2,2,2)光开关结点. 光子学报,1996,**25**(4):289～294
Liu Z L, Cao M C, Li H P, *et al*. *Acta Photonica Sinica*, 1996, **25**(4): 289～294

4 井文才,张以谟,周革,等. 用光学分布时钟实现同步的光互连网络设计. 光电子·激光,2002,**13**(1):37～39
Jing W C, Zhang Y M, Zhou G, *et al*. *Journal of Optoelectronics · Laser*, 2002, **13**(1): 37～39

5 井文才,周革,张以谟,等. 用波分复用技术设计和实现的双波长光纤环网. 光电子·激光,2001,**12**(8):773～776
Jing W C, Zhou G, Zhang Y M, *et al*. *Journal of Optoelectronics · Laser*, 2001, **12**(8): 773～776

6 Christensen M P, Milojkovic P, Mcfadden M J, *et al*. Multi scale optical design for global chip-to-chip optical interconnections and misalignment tolerant packaging. *IEEE Journal of Selected Topics In Quantum Electronics*, 2003, **9**(2): 548～556

7 Kodi A K, Louri A. A scalable architecture for distributed shared memory multi processors using optical interconnects. *Proceedings of IEEE*, 2004: 11

8 Yoshimura T, Arai Y, Kurokawa H, *et al*. Predicted insertion loss reductions achieved by implementing three-dimensional microoptical network in chip-scale optical interconnects. *Photonics Technology Letters, IEEE*, 2004, **16**(2): 647～649

9 Zhang Y M, Jing W C, Zhou G. The application of optical interconnection for computer system. *Proceedings of IEEE*, 2003: 70～73

10 Yamaguchi M. Application of fiber optics for deep-sea exploration systems. *Proceedings of IEEE*, 1990, **15**(3): 238～243

11 贾大功,张以谟,井文才,等. 无源多通道光纤旋转连接器的研制. 天大学报,2004,**37**(5):382～385
Jia D G, Zhang Y M, Jing W C, *et al*. *Journal of Tianjin University*, 2004, **37**(5): 382～385

12 Dorsey G F. Fiber optic rotary joint-a review. *Proceedings of IEEE*, 1982, **5**(1): 37～41

13 Jing W C, Jia D G, Tang F, *et al*. Design and implementation of a broadband optical rotary joint using C-lenses. *Optics Express*, 2004, **12**(17): 4088～4093

14 唐峰,井文才,张以谟,等. 基于 MEMS 光开关的机群系统光互连网络. 光学技术,2003,**29**(5):537～540
Tang F, Jing W C, Zhang Y M, *et al*. *Optical Technique*, 2003, **29**(5): 537～540

15 李贵山,康继昌. 一种可扩展机群系统的研究. 微电子学与计算机,1999,**16**(1):44-47
Li G S, Kang J C. *Microelectronics & Computer*, 1999, **16**(1): 44～47

16 田劲东,周革,井文才,等. Linux 操作系统下并行光互连链路驱动软件的研究. 高技术通信,2000,**10**(12):98～102
Tian J D, Zhou G, Jing W C, *et al*. *High Technology Letters*, 2000, **10**(12): 98～102

Design and Implementation of Two-layer Optical Interconnection Network with Rotating Connection

Jia Dagong, Liu Kun, Jing Wencai, Zhang Yimo, Zhou Ge

Key Laboratory of Opto-electronics Information and Technical Science College of Precision Instrument & Opto-electronics Engineering, Tianjin University, Tianjin 300072

Received date: 2005-07-18

Abstract A two-layer parallel optical interconnection network has been designed. The top layer was a star network, which was made up of digital routing nod, node computer and optical network interface cards. The maximum transmission rate of single channel was 1.4 Gbit/s and the total throughput of network was more than 10 Gbit/s with eight channels. The bottom layer was a ring comprising optical network interface cards and node computer. The peak transmission rate of optical interface cards was 1.065 Gbit/s. Fiber optic rotary joints that can transmit dynamic data were assembled in sub-layer network between the node computer and sensors. The maximum communication latency inside a ring was less than 5.292 μs. The average communication latency of parallel optical interconnection network was 11.03 μs, and the bandwidth of the ring network was less than 50 Mbit/s.

Keywords Optical interconnection; Dynamic connection; Fiber optic rotary joint; Communication latency; Transmitting bandwidth

Jia Dagong was born in March 1972, in Shaanxi. He graduated form College of Precision Instrument & Opto-electronics Engineering, Tianjin University and completed the Ph. D. degree in Optical Engineering in 2004. His research interests include optical interconnection, optical communications and passive optical component. He is currently working for the Post-doctor degree at the Tianjin University.