Development of Low Cost Optical Interconnection Data Link

Ge Zhou, Wei Liu, Wencai Jing, Jindong Tian, Haifeng Li, Nan Zhang
College of Precision Instrument and Optoelectronics Engineering,
Tianjin University, Tianjin 300072, P.R.China
(email: gzhou@tju.edu.cn，Tel: 86-22-27403147, Fax: 86-22-27404547)

Abstract: A low cost fiber-optic data link based on PCI Local Bus is presented. The link is a 32bit-virtual-channel fiber-optic computer bus used only a pair of OE devices and fibers. The technology to match bus rate with optical link bandwidth, and the hardware/software co-design interface to make low overhead and latency between optics and electronics are described.
Key words: optical fiber link, optical interconnection, scalable parallel computing, computing cluster, virtual parallel, and PCI Local Bus.

I. INTRODUCTION

The development of optical fiber link will bring an efficient increase in the interconnection capacity of a scalable parallel computing system[1]-[4]. In fact, parallel fiber links have offered an excellent interconnect medium for the computer clusters[5][6]. This paper contributes to construct a low cost fiber-optic data link for the computing cluster whose node servers are based on PCI Local Bus [6][7]. We address how to match accessing bandwidth between computer bus and high-speed optical interconnection, and how to make low overhead and latency interfaces between optics and electronics. It appears that optical interconnection link is not as convenient as electronic one in practical application. One of reasons may be that computer architecture is usually designed based on electronic interconnection, and optical interconnection is only required to substitute for wire without any structure modification. To facilitate application, fiber link made in this paper is integrated with the popular PCI Bus Interface, it is fabricated as an available peripheral device that can been inserted into the bus slots of commercial computers directly, and can operate under managed by PCI Bus.

II. OPTICAL LINK OVER PCI LOCAL BUS

PCI Local Bus is an industry specification and is widely used in computer or workstation at present. PCI specifies 32bit-width and 64bit-width bus, 33MHz rates, and 66MHz in the future. Total bandwidth is 1.056Gbps based on 32bit×33MHz. Complying with the specification and to transmit the 32bits parallel data, a low cost virtual parallel optical fiber link (VPOFLink) by means of TDM technology has been developed in our laboratory.

The VPOFLink is a 32bit-virtual-channel fibre-optic bus used only a pair of OE devices and fibres. It is composed of the two parts, the virtual parallel OE module and the PCI Bus Interface, as shown in Fig. 1. The virtual parallel OE module is to perform data transformation and conversion between optical and electronic signals, it includes an optoelectronic transceiver, TDM transmitter (TX) and receiver (RX) processing unit, and functional unit exploited by a FPGA device. RX and TX carry out the channel multiplexing of 16bits parallel data to serial one and reversal demultiplexing processing, respectively, and provide synchronous clock extraction, data encode/decode, transfer status handshake, etc. An external extended FIFO（First In First Out）is

used as data receiving buffer. PCI Bus Interface is the bridge between the link´s parallel interface and computer bus. PCI Bus Interface offers accessing data for three physical address spaces such as configuration address space, memory address space and I/O address space, automatically configures bus address space resource, processes bus software/hardware interrupter, and performs data transfer.

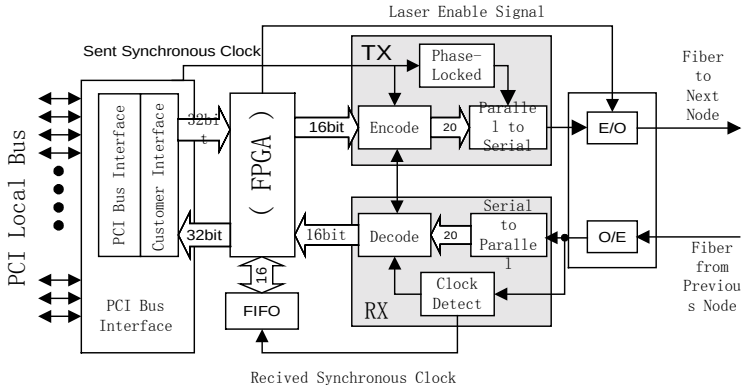

Fig. 1 Interface of virtual parallel optical fiber link

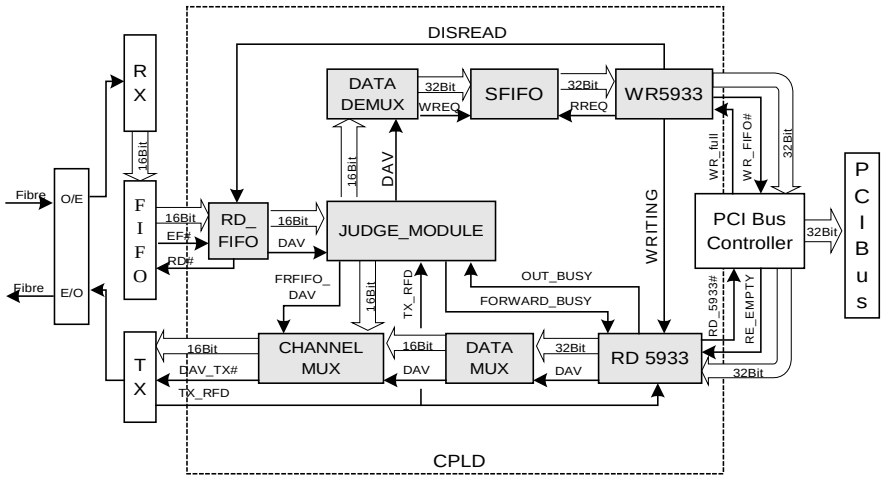

Fig.2 Performance implemented by FPGA

From 32bit parallel channels to one serial channel, the serialising goes through two-step transformations. Since TX and RX only provide 20bit parallel interface while PCI Local Bus provides 32bits I/O, a FPGA logic device is exploited to match the bandwidth between the bus and TX or RX, i.e. the first step of serialising transformation. The FPGA device connects the bus-bridge with TX and RX, and converts 32bit×33MHz from the bus-bridge into 16bit×66MHz, then sends that to TX. After encoding and inserting control signal, TX converts parallel data at the rate of 16bit×66MHz to 1.2Gbps serial data to drive LD to emit light. As receiving message, FPGA will convert 16bit×66MHz from RX to 32bit×33MHz, and send the data to the bus-bridge. The multiplexing and demultiplexing processing is an equal-flow transformation.

The link supports computer to take two kind of data transmission modes as I/O and DMA (Direct Memory Access). The tests of VPOFLink have been made, and the data transformation

latency from the parallel input of link to the parallel output at the other end of the link is 3 clock cycles. the available peak rate of the link is up to 131.1Mbyte/s under DMA (Direct Memory Access) transmission model, while the peak rate of PCI Bus is 132Mbyte/s, as shown in Fig 3.

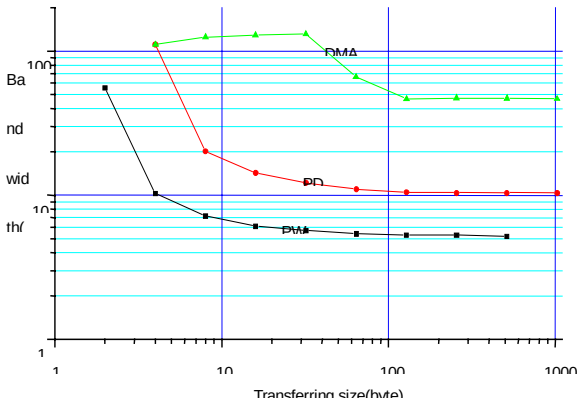

Fig. 3 The testing results of VPOFL

III. FIBER LINK FOR A RING NETWORK

Using the virtual parallel fiber link, an optical interconnection ring network is established to test the efficiency of the VPOFLink, as shown in Fig. 4. In the ring network, Data are transmitted in data packet, and the data packet is composed of destination address, source address, data size label, varied size data sector and cyclic redundancy check (CRC) sector, shown as Fig. 4(c). Every node has its unique address code，and 128 nodes will be permits to be hold in a ring. A packet buffer is used as a virtual straight-connect channel to avoid the collision inside the ring network, see Fig. 4(b). Data packet from up-stream (Server k-1 in Fig 10(b)) are first stored in the buffer and sent out until output channel is available, e.g. no input/output operating at Server k.

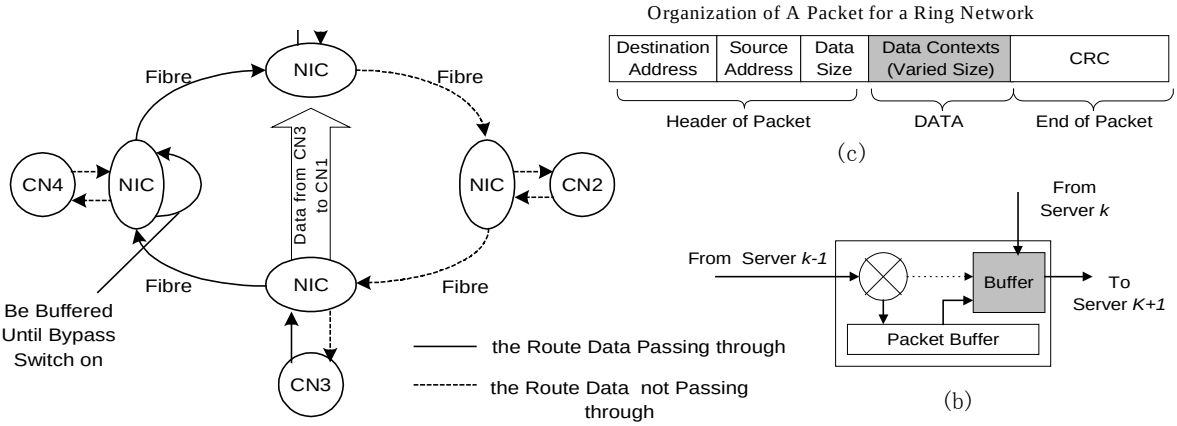

Fig. 4 The ring architecture, a packet buffer (b) and a packet content (c)

From point-to-point to ring interconnect, the link is required to process some protocol functions such as recognizing the destination address of data packet, buffer control of data packet, receiving and transmitting control of data, and so on. Four more main logic units are exploited in a FPGA device, as shown in Fig. 2, which are data transfer control logic unit, data output control logic unit, switching unit, data multiplexing unit. The first unit carries out reading/writing control on receiving package of RX, detects the destination address and size of receiving data packet, automatically distinguishes head or end of a packet, and controls its routing direction. The second one performs sending control of the packet coming from local server node, recognizes output data length, and distinguishes head or end of this packet. The third one carries out the routing of input packet, namely, receiving or transfer, the fourth one manages output of sending packet from local

server node and retransmission packet.

VPOFLink not only offers advantages over copper-based links, but also is easy to be fabricated, and it is lower cost than other parallel fiber links. Since integrated with the popular PCI Bus Interface, the VPOFLink is well suited to a scalable cluster system consisting of commercial computers. Even though the peak rate of single link is equal to that of Gigabit Ethernet and Fiber Channel, VPOFLink is much more flexible, and the good capability provides great potential for computer system researchers and scientists to construct a special scalable parallel computing system. Moreover, the cooperation of two VPOFLinks is able to construct 32bits interconnection channel at 2.5Gbp.

IV. CONCLUSIONS

In this paper, we present the design method of a fiber-optic data link. The fiber link is fabricated by commercial IC device, optoelectronic device and field programmable gate array (FPGA) logic device. The link integrates a virtual parallel technology with PCI Bus Interface technology, and has become a peripheral device of PCI Bus. Our work also exhibits that a well-designed interface between an optical interconnection and a computer bus is the key to fully exploiting the transmission bandwidth of an optical interconnection.

REFERENCES

1. L.J.Norton, “OPTOBUS: A Parallel Interconnection Solution”, in Proc. LEOS’97, Vol.2, p.189~190.
2. Mitsuo Usui, et al, “700-Mb/s×40-channel Parallel Optical Interconnection Module using VCSEL Array and Bare Fiber Connectors”, in Proc. LEOS’97, Vol.1, p.51~52.
3. S. Scott, “The giga-ring channel”, IEEE Micro, Feb. 96, 27~34.
4. Y. M. Wong, D. J. Muehlner, C. C. Faudskar, et al, “Technology development of a high-density 32-channel 16-Gb/s optical data link for optical interconnection application for the optoelectronics technology consortium (OETC)”, J. Lightwave Technolgy, Vol. 13, p.995~1013, 1995.
5. D. R. Engebrestsen, D. M. Kuchta, R. C. Booth, et al, “Parallel fiber-optic SCI links”, IEEE Micro, p.20~26, Feb. 1996.
6. T. Yoshikawa, “Skewless multimode fiber parallel optical interconnection for the parallel processor RWC-1”, Proc. SPIE, Vol. 3632, Optoelectronic Interconnects VI, p. 322~330, 1999.
7. G. Zhou, W. C. Jing, Y. M. Zhang, et al, “Design of Giga-bit/s optical interconnection network for computer cluster”, 5th International Conference on Massively Parallel Processing Using Optical Interconnections (MPPOI’98), p.194~199, 1998.
8. G. Zhou, W. Liu, W. C. Jing, et al, “The twin-wavelength ring optical network for computer cluster”, Proc. SPIE, Vol.3632, p. 96~102, 1999.