



Strathprints Institutional Repository

Glesk, Ivan and Runser, R.J. and Deng, K.-L. and Prucnal, P.R. (2000) *100 Gbit/s computer optical interconnect*. Acta Physica Slovaca, 50 (2). pp. 229-234. ISSN 0323-0465

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<http://strathprints.strath.ac.uk/>) and the content of this paper for research or study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to Strathprints administrator: <mailto:strathprints@strath.ac.uk>

100 Gbit/s COMPUTER OPTICAL INTERCONNECT**Ivan Glesk¹, Robert J. Runser, Kung-Li Deng, and Paul R. Prucnal***Department of Electrical Engineering, Princeton University, Princeton, NJ08544*

Received 23 September 1999, accepted 16 November 1999

An experimental demonstration of an error free 100Gbit/s optical time division multiplexing (OTDM) broadcast star computer interconnect is presented. A highly scalable novel node design provides rapid inter-channel switching capability on the order of the single channel bit period (1.6 ns).

PACS: 42.79.Sz, 42.79.Ta

1 Introduction

Although lightwave technology is meeting the demand for point-to-point and long-haul transport of digital information, routing packets at the nodes of the network has typically been carried out using electronically switched backplane routers. The growing capacity on the Internet is placing an ever greater demand on electronic routing technologies. While WDM can support large aggregate traffic bandwidths, it is difficult to perform routing functions which may involve challenging techniques such as dense wavelength conversion. Additionally, present WDM laser and filter tuning techniques rely upon slow technologies which increase the channel access latency and reduce the effective network bandwidth.

Recent advances in optical time division multiplexing (OTDM) have proven this technology's capability to handle the switching and routing needs for future. Channel access in OTDM networks is achieved by using time slot tuners and all-optical demultiplexers. Timing precision of less than 1ps is required to tune, multiplex, and demultiplex individual channels within the OTDM frame.

The computer interconnect we are constructing is based upon an OTDM broadcast star architecture. The high-level architecture and node design is shown in Fig. 1. Nodes transmit information at a slow data rate, B , by modulating picosecond optical pulses. By using a scalable time slot tuner, the pulse is appropriately delayed to correspond to the desired destination time slot. Data pulses from all nodes are multiplexed into a time frame with an aggregate bandwidth of NB , where N is the number of nodes in the network. The pulse spacing between adjacent channels is $(NB)^{-1}$ or typically less than 10ps to achieve 100+ Gbit/s. Ultrafast all-optical demultiplexers like the TOAD are used to extract the desired channel from the high capacity OTDM

¹E-mail address: glesk@ee.princeton.edu

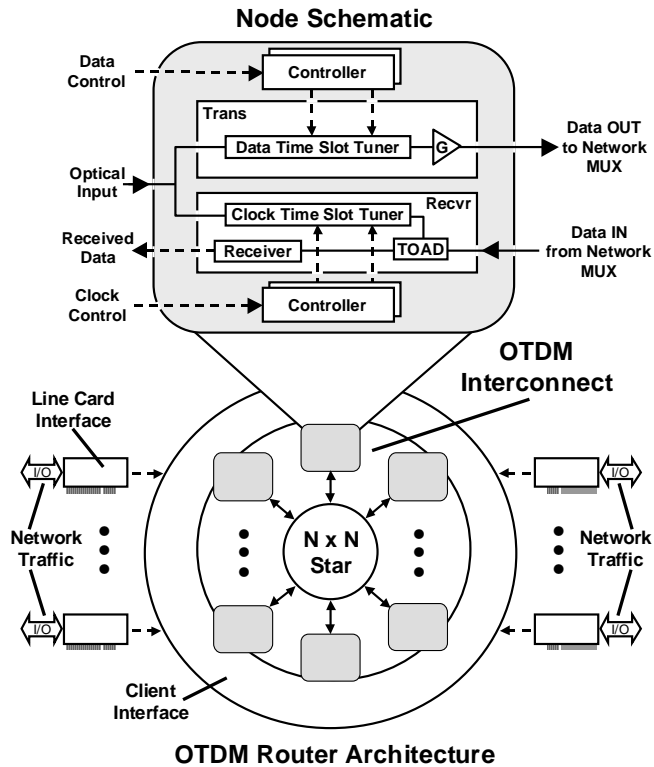


Fig. 1. OTDM network and node architecture

frame at the node receivers. Nodes can select the received time slot by using a time slot tuner to align the clock with an incoming time slot within the frame for all-optical demultiplexing.

To perform the functionality of a router, addresses are mapped to specific time slots within the network. Routing is achieved by sending each bit of the packet in a unique time slot corresponding to its destination node. All nodes in the network are synchronized by splitting and amplifying the optical output of a single modelocked fiber laser. Packet routing is performed by rapidly changing the state of the time slot tuner to transmit into time slots corresponding to destination addresses on the network. Recently, several experimental demonstrations [1-3] have shown that OTDM can meet many of the demanding needs of a router and a multiprocessor interconnect system which include full connectivity, low latency, and high aggregate throughput, reliability, and scalability. We report the demonstration of a testbed for a bit-interleaved 100-Gbit/s OTDM broadcast star architecture that was previously proposed [4]. Unique to our network is a highly scalable, novel node design that provides inter-channel switching within the single channel bit period (1.6 ns). By combining this hardware with a highly efficient arbitration protocol [4], near lossless channel allocation with low latency is achievable for high speed switching applications such as future all-optical routers.

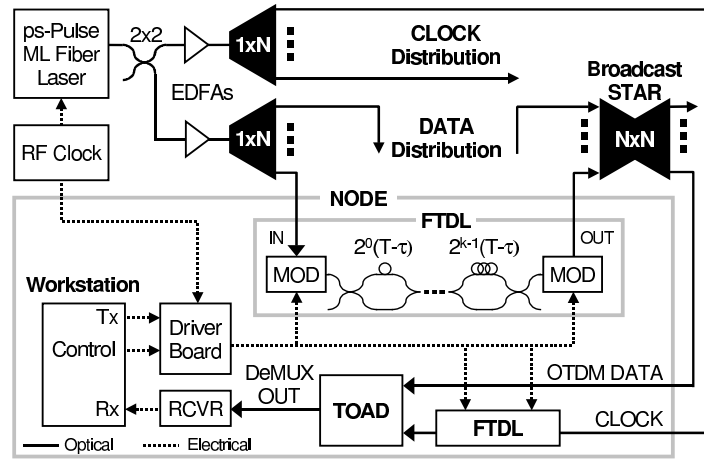


Fig. 2. Experimental OTDM computer interconnect and node architecture

2 Experimental Demonstration and Results

Fig. 2 shows the network and novel node architecture experimental setup. The two key optical components of the node are the recently developed fast tunable delay line (FTDL) [5] and the terahertz optical asymmetric demultiplexer (TOAD) [6]. A controller card residing in a workstation sends electronic NRZ data at the single channel bit rate, B , and control bits to the driver board specially designed to control the two FTDLs on the clock and data fibers. The FTDLs consist of cascaded feed-forward Mach-Zehnder fiber delay lattices designed to produce optical copies of the incoming pulse stream organized into 2^k -bit subcells spaced by T with inter-subcell bit spacing τ [5]. The two modulators controlled by the driver board select one of the $2^k \times 2^k (= N)$ time slots into which one of the copies is transmitted. The FTDLs in the node are used to transmit data into a selected time slot within the OTDM frame and align the clock with a given time slot for optical demultiplexing. Ultimately, the dimensionality of the network, N , is determined by k , the number of stages in the FTDL. The intermediate processing bandwidth, $B' (= 1/T)$, of the driver controller and the electro-optic modulators is designed to match the repetition rate of the picosecond pulsed fiber laser source and is related to the single channel bit rate as $B' = 2^k B$. Pulses are amplified by EDFAs and distributed to the individual nodes by $1 \times N$ splitters. After node data modulation and time slot selection, the data is multiplexed by precision fiber delays feeding an $N \times N$ star coupler. The high bandwidth OTDM frame is broadcast to all nodes in the network. Each node can demultiplex any single channel from the frame using an FTDL on the clock and a TOAD.

In our experimental testbed, we populated 16 ($= N$) time slots in the OTDM frame by constructing 2 ($= k$) stage FTDLs. The single channel data rate was chosen to match the OC-12 rate ($B = 622.08$ Mbit/s). The 2-ps pulsed 1550-nm fiber laser repetition rate and intermediate

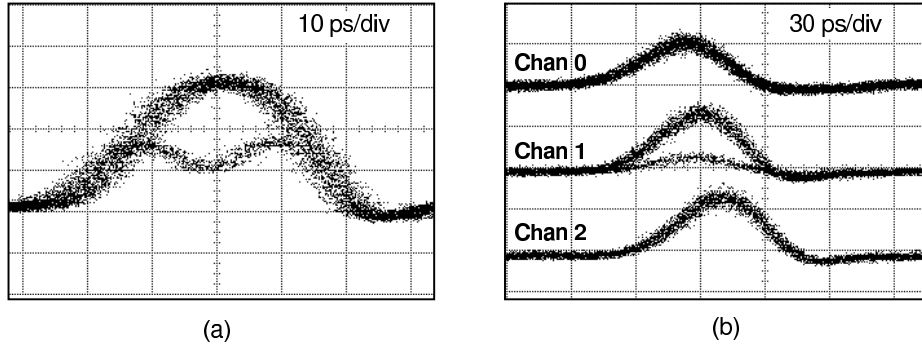


Fig. 3. 100 Gbit/s multiplexed data OTDM subcell eye diagram on bandwidth limited detector, and demultiplexed TOAD output eye diagrams for three channels in subcell. (a) 100 Gbit/s multiplexed data OTDM subcell eye diagram (b) Demultiplexed TOADF output eye diagrams.

electronic processing bandwidth were set to the OC-48 rate ($B' = 1/T = 2.48832$ GHz). The simple electronic design of the driver board permits the rapid control of the FTDL and provides low latency, arbitrary channel selection. The driver board was constructed using 4-bit electronic multiplexers (Vitesse) and simple logic operating at the OC-48 rate. To produce an OTDM frame with an aggregate bit rate of 100 Gbit/s, $\tau = 10$ ps was chosen. Each TOAD was designed with a demultiplexing window width of about 10 ps at FWHM and a polarization splitter was used to separate data from clock at the output.

The 100-Gbit/s multiplexing and demultiplexing experimental results are shown in Fig. 3. According to the design of the FTDL, the 16 time slots in our OTDM frame are arranged in 4 subcells each containing 4 time slots spaced by 10 ps. Our network demonstration focused on one of the subcells within the frame. Fig. 3a shows the aggregate eye diagram for a subcell with multiplexed data from 4 nodes with a fixed pattern, 1 - *pseudorandom* - 1 - 0, on a bandwidth limited detector (34-GHz photodetector, 50-GHz oscilloscope). Upon demultiplexing by TOADs tuned to the individual channels, each is resolved in Fig. 3b (the 4th time slot is omitted as it is 0).

We constructed two fully functional nodes to measure the bit error rate (BER) and demonstrate the rapid inter-channel switching capability of the network nodes using an arbitration protocol. These experiments were performed using adjacent channels in the same 100-Gbit/s subcell (Channels 0 and 1). Fig. 4a shows a plot of the BER versus the single channel average data input power at the TOAD when Chan 0 and Chan 1 were modulated with pseudorandom data. For average data and clock input powers greater than -21 dBm (13 fJ pulse energy) and -8 dBm (250 fJ pulse energy) respectively, several hours of error free operation have been achieved. Additionally, we have observed that the TOAD can provide gain to the demultiplexed signal. The inset to Fig. 4a shows the eye diagram of the data input (upper trace) and demultiplexed output (lower

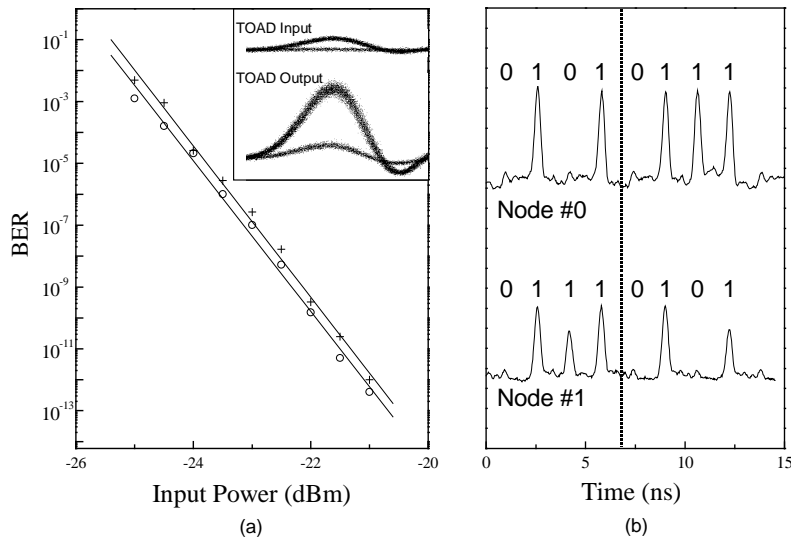


Fig. 4. BER of channels 0 and 1 against average single channel input power, and demonstration of rapid channel selection on bandwidth limited analogue detector. (a) BER of channels 0 and 1 against average single channel input power. Inset: TOAD input and output eye diagrams demonstrating gain. \circ - channel 0, $+$ - channel 1. (b) Demonstration of rapid channel selection.

trace) of a TOAD demultiplexing a single channel of pseudorandom data with identical oscilloscope settings. The demultiplexed output is larger in amplitude than the input by approximately 6 dB.

The fast inter-channel switching capability of the network was also demonstrated by using a previously proposed, low latency arbitration protocol [4] and two nodes of the network. The receivers of both nodes are fixed to listen to their own time slots. Each node transmits its binary address at the single OC-12 channel rate into its own time slot. If successfully received, each node then transmits its address into the time slot of the other node. Fig. 4b shows a demonstration of the protocol using two nodes in the network whose time slots are adjacent in the 100-Gbit/s subcell. The addresses assigned to Node 0 and Node 1 were 0101 and 0111 respectively. The traces shown are the demultiplexed TOAD outputs directly from the analog output of the receivers for the two nodes. After each node successfully receives its own address, the FTDLs rapidly reconfigure within a single bit period to transmit into the time slot of the other node. Note that each node now successfully receives the address of the other in its own time slot. The FTDLs and driver board electronics are capable of tuning to any one of the 16 time slots in the network within 1.6 ns, greatly reducing the hardware latency of the protocol.

3 Conclusion

We have demonstrated a fully connected 100-Gbit/s OTDM network architecture that offers fast switching among data channels with reliable, error free operation and low latency. Since the active components of the FTDLs do not scale with the number of nodes [5], simply adding another stage, $k = 3$, (3 dB additional loss per node), scales the interconnect up to 64 ($= N$) nodes without taxing the power budget significantly. If OC-24 ($B = 1.24416$ GHz) is chosen as the single channel data rate and 10-GHz ($= B'$) intermediate processing bandwidth electronics are used, an 80-Gbit/s interconnect with a rapid inter-channel switching speed of 800 ps is feasible. In such a 64-processor architecture, coherent crosstalk does not limit the BER performance significantly [7]. Since the demultiplexer [8] and other optical components in the node can be integrated, we believe this network is practical for future, high-speed multiprocessor interconnect systems.

Acknowledgement This work has been supported by DARPA Contract No. F30602-97-2-0316.

References

- [1] R.A. Barry, V.W.S. Chan, K.L. Hall, E.S. Kintzer, J.D. Moores, et al.: *IEEE J. Sel. Areas Commun.* **14** (1996) 999
- [2] J.K. Lucek, P. Gunning, D.G. Moodie, K. Smith, D. Pitcher: *Electron. Lett* **33** (1997) 887
- [3] M. Tsukada, W. De zhong, T. Matsunaga, A. Masaki, T. Oohara: *J. Lightwave Technol* **14** (1996) 1979
- [4] A.G. Nowatzky P.R. Prucnal: *Proc. 22nd Int. Symp. On Computer Arch.*, Santa Margherita, Italy (1995)
- [5] K.-L. Deng, K.I. Kang, I. Glesk, P.R. Prucnal: *IEEE Photon. Technol. Lett.* **9** (1997) 1496
- [6] I. Glesk, J. P. Sokoloff, P. R. Prucnal: *Electron. Lett.* **30** (1994) 339
- [7] K.-L. Deng, I. Glesk, K.I. Kang, P.R. Prucnal: *IEEE Photon. Technol. Lett.* **10** (1998) 1039
- [8] E. Jahn, N. Agrawal, M. Arbert, H.-J. Ehrke, D. Franke, et al.: *Electron. Lett.* **31** (1995) 1857