

High-Performance Optoelectronic Hierarchical Bus System

Jinghuai Fa, Chunhe Zhao, Jian Liu and Ray T. Chen

Microelectronics Research Center
Department of Electrical and Computer Engineering
University of Texas at Austin, Texas 78712-1084

ABSTRACT

In order to utilize the high speed of optical interconnects and overcome the latency problem of a large bus structure, we propose an architecture of optoelectronic hierarchical bus system. Waveguide hologram implementations, and its associated cache coherence problem are addressed. In our configuration, bus hierarchy is controlled with electronic programmable plates. Optical signals can be transmitted on all-optical paths without intermediate conversions.

Keywords: High-performance Bus, Backplane Bus, Optoelectronic Bus, Hierarchical Bus

1. INTRODUCTION

Due to the fast growing computer technology and the data rate, the GHz speed will be the norm for computer industry. The important problem emerging is the traffic of interconnects. We definitely need a better interconnection technique in the near future to fulfill the increasing demand in market.

Comparing with electronic interconnection, optical interconnection has the advantages of high speed--fast data rate and of longer communication distances. One concern, however, is the latency problem. In this paper, we propose a low latency optoelectronic hierarchical structure, companioned with cache and pipeline techniques.

There are several interconnection architectures exist for optical interconnections. We can classify them as static interconnections and dynamic interconnections. In static interconnections, the connected paths are fixed. For example, arrays, rings, meshes, cubes. On the other hand, dynamic interconnections are dynamic. They are busses, crossbars, multistage network, in which the right and path for two components to interconnect changes dynamically. This paper deals with a dynamic interconnection- based on an innovative bus structure.

With the high speed property of an optical bus structure, the latency becomes a major problem. The reduction of latency is achieved mainly in three ways. The first one is the pipeline method by which several processors split and pipeline their transactions with bus simultaneously. The second method to achieve high speed and low latency is by using cache. The idea behind this technique is based on the existing relationship between instructions and data--locality. Space locality means each instruction/data located near the currently used instruction/data is going to be used with high probability. While, the time locality means each currently used instruction/data is going to be used soon. Adding small cache locally results in the reducing traffic and latency. Cache and bus structure are mutually benefited. A bus needs cache to reduce the occupation probability of each attached processor. On the other hand, cache coherence protocol need bus structure which has broadcast property.

The third method is the hierarchical method. It is the most important method to build up a large network. This method shortens a long bus by hierarchically organizing short buses. In this paper, we present an optoelectronically controlled hierarchical bus system. With the presence of all-optical instruction/data paths and the electronic programmable control, conversions between optical signals and electronic signals are not needed, i.e., optical data can be transmitted on all-optical paths without intermediate conversions which is the primary sources of the increase of latency.

For the first generation bus interconnection, 10 to 20 processors can be attached to. A crossbar multistage network can accommodate about 100 processors. The hierarchical bus system with cache and pipeline, can have more than 1000 processors [8].

We emphasize here the importance of protocol design in optoelectrical bus system, see [1] for review. For other features of the optoelectronic bus, see [2][7][9].

Section 2 discuss optoelectronic bus including polymer based two dimensional waveguide hologram bus. Section 3 provide the hierarchical optoelectronic bus structure and its implementation. Section 4 analyses the projected system performance.

2. OPTOELECTRONIC BUS

For optics to be integrated into an electronic system, three basic components are necessary: (a) Transmitter blocks, including light sources, drivers and modulation circuits, which convert electrical signals to optical signals, and the optical signals can be internally or externally modulated. (b) Optical channel, consisting of free-space, guided-wave, or optical fibers, which transmit and distribute the optical signals. (c) Receiver blocks, including photodetectors, (pre-)amplifier circuits, and demodulators (analog optical receiver) or filters and decision circuits (digital optical receivers), which detects the optical signal and converts a data signal transmitted as amplitude modulation of an optical beam into an electrical data signal. A basic optoelectronic structure is shown in Fig. 1.

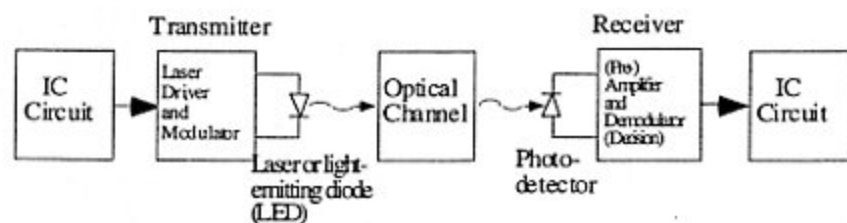


Fig. 1 Optical channel and opto-electrical conversion

Another important factor in integrating photonic devices with electronics is that the optoelectronic system must be compatible with the state-of-the-art VLSI technology. As the performance of current electronic circuits increases, more layers have been added on a semiconductor chip which contains the circuit. Especially, with the advancement of vertical cavity surface emitting lasers (VCSELs), integration of the lasers with other optoelectronic devices such as detectors and transistors on a single semiconductor chip is made possible.

One Dimensional Bi-directional Bus

One dimensional bus structure is the most popular bus configuration. Each processor can broadcast

signal to all the other processors which receive data via bus. Using polymer-based multiplexed waveguide holograms in conjunction with a thin waveguide substrate, a bi-directional optical backplane bus can be realized.

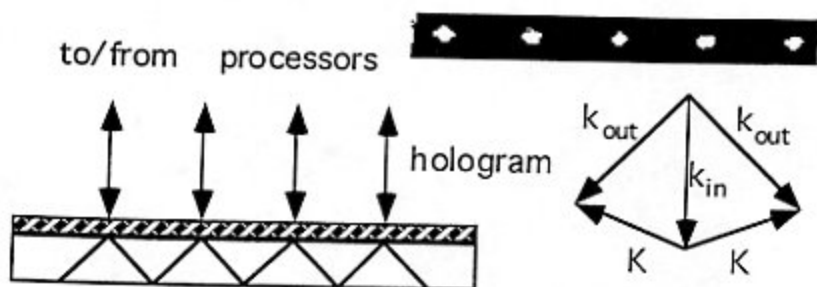


Fig. 2 Optical waveguide holograms implementation

Fig 2 shows the optical bus structure, vector relation (between grating vectors, input vectors and output vectors) and experimental result. In bi-directional optical bus, we can get a large number of fan-out easily.

Two Dimensional Bus

In massive elements interconnects, it is more favorable to attach processors onto backplane bus .

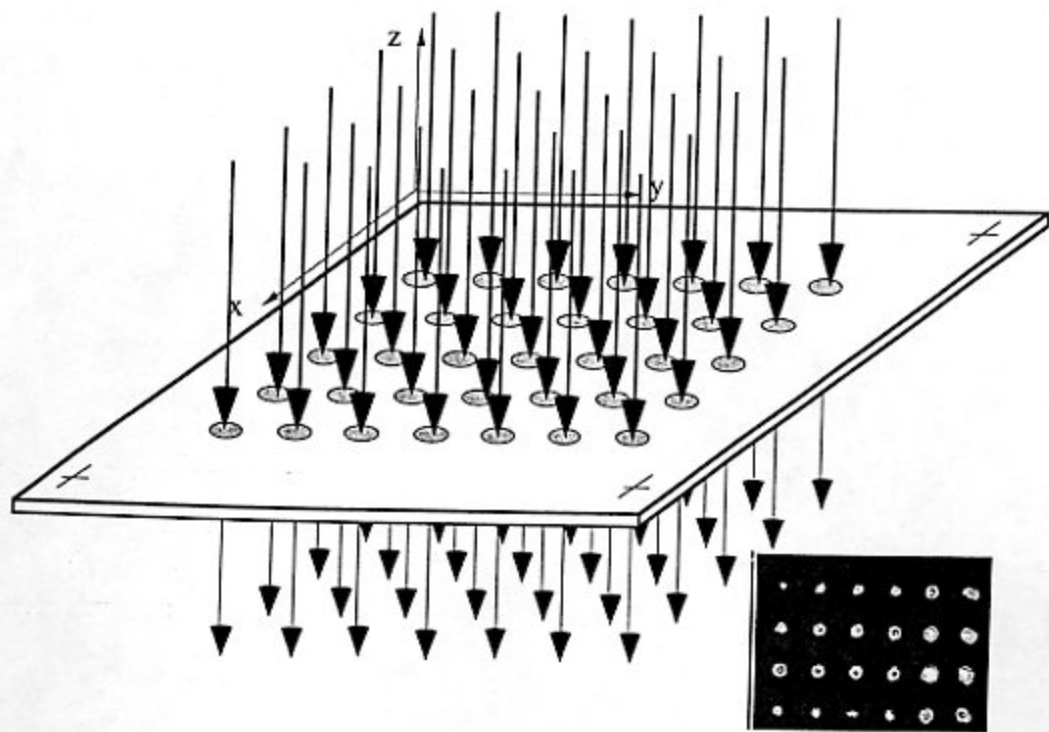


Fig. 3 Two dimensional optical bus implementation

We show how to extend the simple bus structure to two dimensional bus in a planar way. We consider here the polymer based waveguide holograms. Its relatively easier to understand the extension from one dimensional optical bus to two dimensional optical bus. But as to using polymer-based waveguide holograms, the multiplexing mechanism become more complicated[3]. We need four exposures on one polymer in four deferent diffraction directions.

In this paper, we use two overlapped polymers, each of them receives two exposures to get a multiplexed hologram grating to realize the two dimensional optical bus (fig. 3).

On one polymer, two exposure directions constitute a right angle. On the other polymer, two exposure directions also have a 90 degree, but in opposite to the first polymer in direction.

When the number of elements attached to bus becomes large, the latency problem will emerge since every element on the bus are trying to occupy the whole optical bandwidth. One can image a special situation: if token bus protocol is applied, then the token should go through every element.

In the following section, we give a flexible method to establish optical path between processing elements dynamically to reduce the latency problem.

3. HIERARCHICAL OPTOELECTRONIC BUS SYSTEM

In the previous section, processors contend with one another to occupy the optical bus. To reduce latency, we add a control between optical path. Although the control is by electrical mean, the physical paths of the buses are all optical. We keep the throughput while reducing the latency. The idea behind hierarchical bus structure is to shorten bus length and put them together in a dynamic way. We will explain such mechanism in this section.

Programmable Control

The electrical control device here has a polarization mechanism. It has a polarized plate and a programmable half-wave plate[5]. The polarization direction of the light going through the device can be controlled by the programmable half-wave plate. If the controlled polarization direction is coincident with the polarization direction of the polarization plate, the device is set to status "on"(light can go through), otherwise, the device is set to status "off"(light can't go through).

Fig. 4 shows fan-in from outside of the local bus to the processors on the local bus when the programmable plate P is set to "on", and fig 5 shows a fan-in from one processor on the local bus and fan-out to the other processors on the same bus as well as to the outside of the local bus when the programmable plate is set to "on".

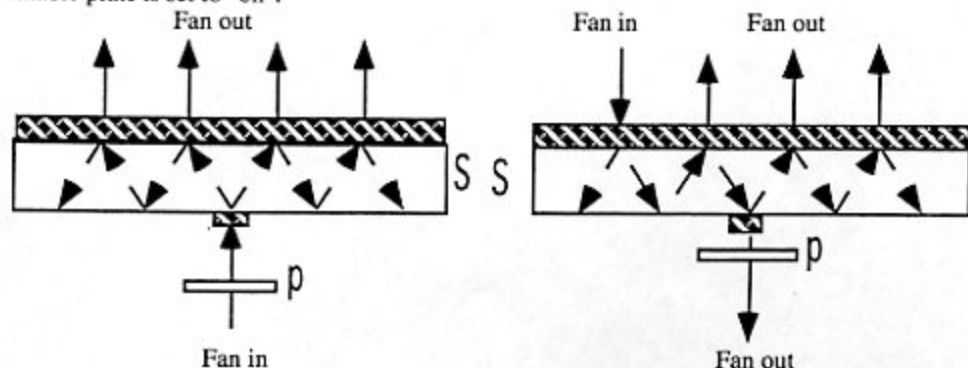


Fig. 4 Fan-in from outside to local bus.
P: Programmable plate
S: Substrate

Fig. 5 Fan-in from one local processor,
fan-out to other local processors
and outside local bus.

If the programmable plate is set to "off", the processors on the local bus can only communicate to each other, but not the outside of the local bus, and the contention domain is local. The programmable plate is set to "on" only when it is necessary in order to decompose the system. By dynamically changing the pattern of the programmable plates, speed will be high and latency will be low.

Two Level Hierarchical Bus System

Fig 6 shows a bus system integration--two level bus hierarchy. Four local buses are interconnected via the second level bus through their programmable plates. When the first local bus need to communicate with the last local bus, for instance, the first programmable plate and the last programmable plate are set to "on", they can communicate to each other. At the same time, each bus in the middle works only locally and simultaneously. Generally speaking, any subset of local buses may communicate with each other via the broadcast property of the second level bus, while each of the other buses works independently, locally, and concurrently. This mechanism greatly reduces the latency.

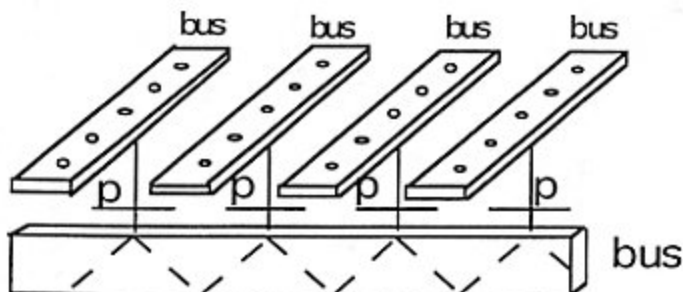


Fig. 6 Four buses are connected by the second level bus.
P: programmable plate.

Three Level Hierarchical Bus System

In order to interconnect the second level bus to the third level bus, we add a layer of polymer at the end of the second level bus (fig. 7). Since the grating is similar to other gratings, and the 45 degree holds at the endpoint, there results a surface normal fan-in fan-out at the ends of each second level buses(fig. 8).

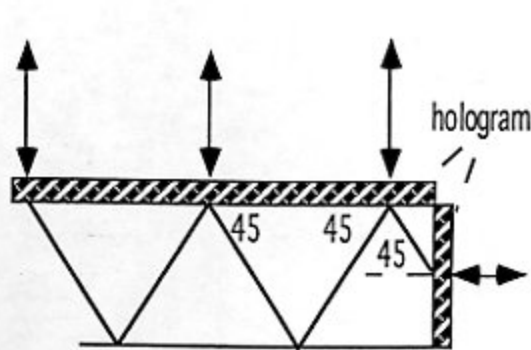


Fig. 7. Hologram grating principle

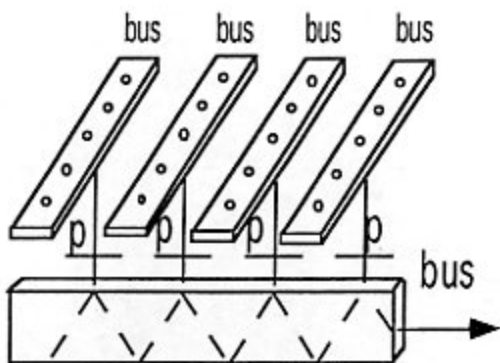


Fig. 8 Two level buses ready to connect with the third level bus

The interface between the second level buses and the third level bus also consists of a set of programmable plates, as shown in fig. 9 and fig. 10.

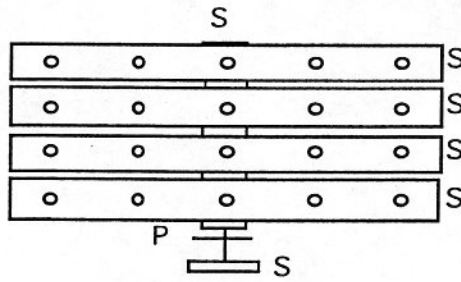


Fig. 9 Interconnection between two level bus system and the third bus.
P: programmable plate between the second and third bus.
S: Substrate.

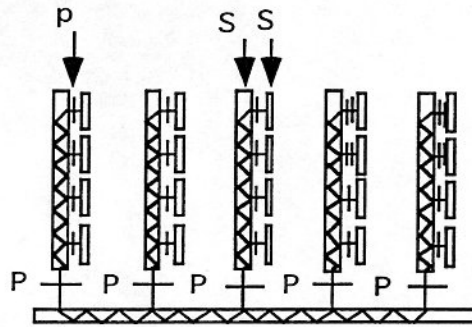


Fig. 10 Three level optical bus system integration (side view of fig 9).
P: programmable plate between first level bus, second level bus and the third level bus.

Multichannel Bus

The hierarchical bus system demonstrated in the previous part represents bus as with single lines. By incorporating arrays of transmitter and receiver multichip modules (MCMs) on each side of the waveguiding channel, a multi-bus line backplane architecture, say, with 32 bits or 64 bits parallel data transmission can be easily constructed. Fig. 11 shows the detailed diagram of the backplane with multi-bus lines using VCSEL and photodiode arrays, and also indicates the necessary components integrated into the transmitter and receiver multichip modules. Microlenses are used in the design to collimate and focus signal beams. Because the detector arrays just opposite the transmitter arrays are on the same board, they don't need to communicate with each other.

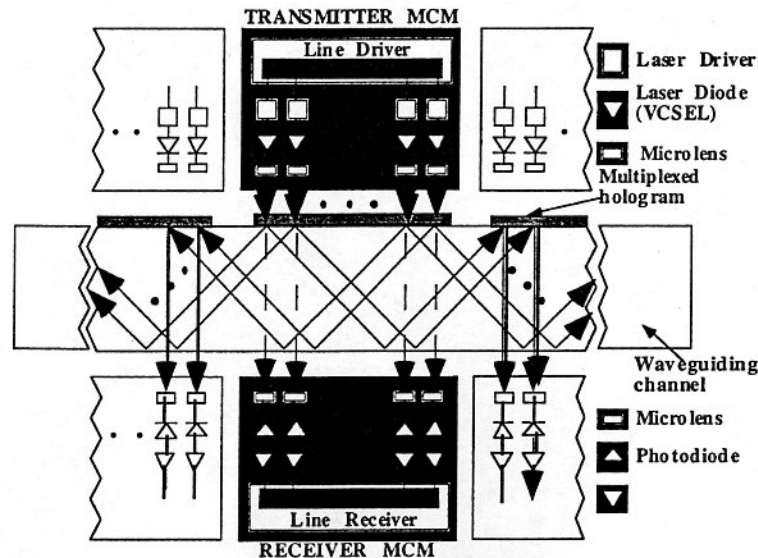


Fig. 11 Multichannel bus

Fig. 12 shows the picture of the VCSEL array we currently used in our lab. The device has a total of 32 VCSELs operating at a wavelength of $0.85\ \mu\text{m}$. Within the fabrication error, all the VCSELs in the array are identical. The array has $140\ \mu\text{m}$ pitch and typically has an output power about 1-2 mW at 10 mA operating current.

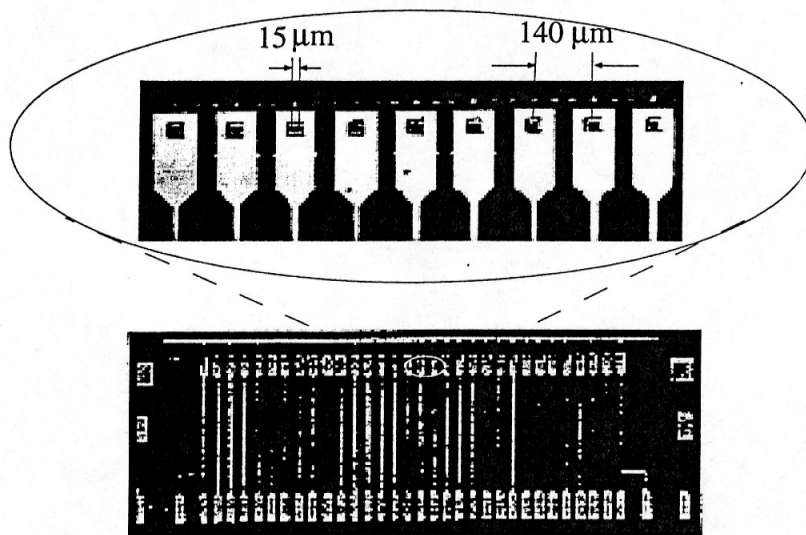


Fig 12 VCSEL diagram

A further improvement in the performance of the backplane with multi-bus lines can be achieved by using a 2-D VCSEL array and photodetector arrays instead of 1-D, so that the real estate of the backplane can be effectively utilized. This idea is demonstrated by the experimental results in Fig. 13, where the input/output configurations of a backplane with 4-bus lines are demonstrated.

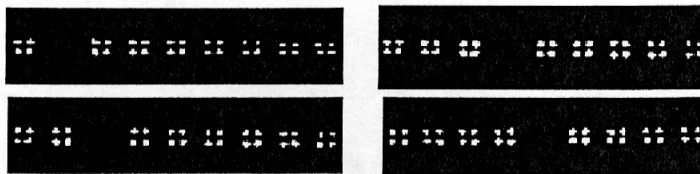


Fig. 13 Experimental result of parallel data bus

Eye Diagram

By overlaying sweeps of different segments of a long data stream driven by a master clock, we can get an eye diagram on the screen of a storage oscilloscope. In communication industry, the eye diagram is used to observe and analyze the performance of circuits that drive the transfer of digital data streams. Ideally, when many traces of randomly generated data series have been overlaid, positive- and negative-going pulses are superimposed on each other, a picture of rectangular box is resulted. In practice, especially when the displayed signal has traversed an imperfect communications channel, the traces do not lie perfectly atop one another, and a classic eye pattern results.

The experimental setup for measuring the eye diagram of our optical interconnection devices is shown in Fig. 14. The random bit pattern from a 3 GHz Pulse Generator (HP8133A) is used to current-

modulate an 822 nm semiconductor laser transmitter (FOT-FP-820-1M/3G-5/125-0 from Lawrence Labs, Ltd.).

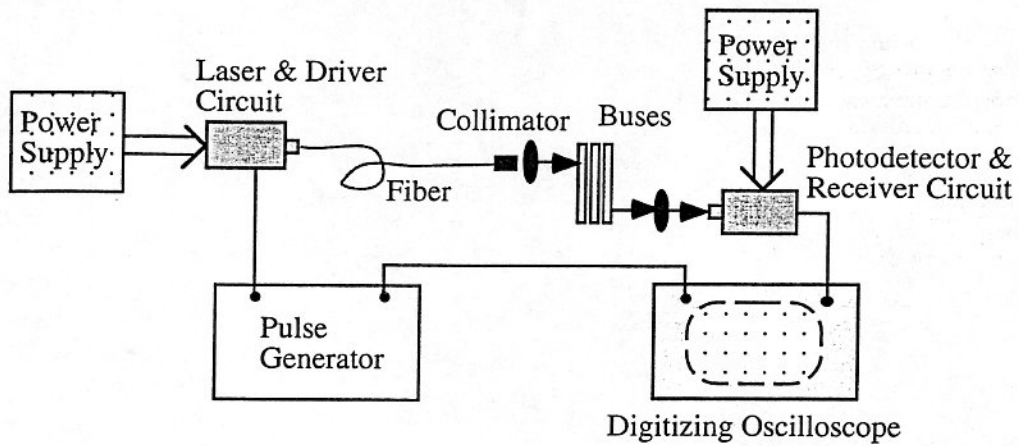


Fig 14 Setup for Eye Diagram measurement

After being collimated and passing through the device, the optical signal from the transmitter is focused onto a high-speed PIN silicon photodiode (Hamamatsu S4753). The output of the receiver is then analyzed by a Digitizing Oscilloscope (HP54120A).

To demonstrate the performance of our device, eye diagrams at the speeds of 500 MHz, 1 GHz,

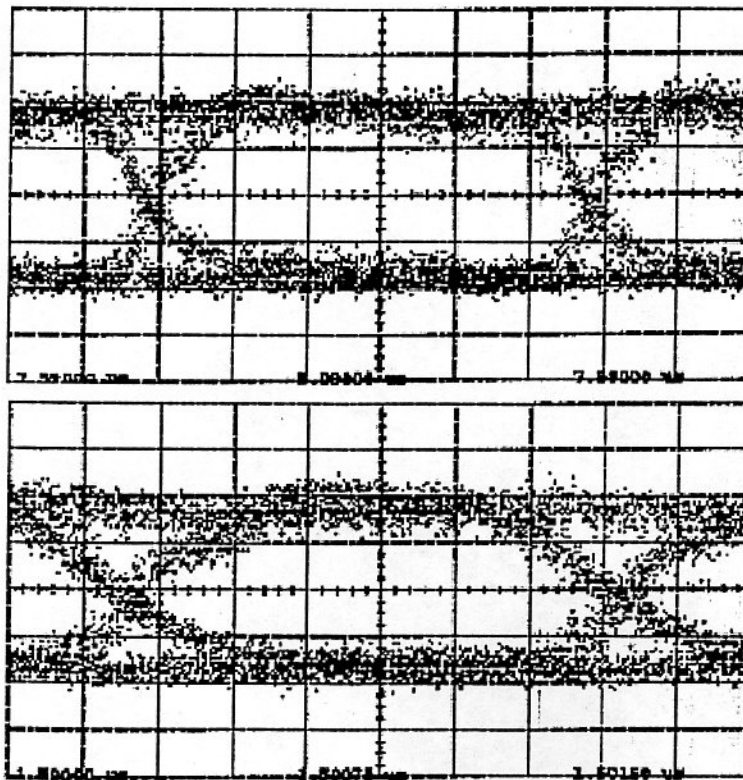


Fig. 15 Eye diagram under 1GHz(above) and 1.5GHz(below)

and 1.5 GHz were measured with and without the backplane (i.e. Buses between collimators in fig. 14) for comparison. The experimental results are shown in Fig. 15, with bus system. For speed higher than 1.5 GHz, the experiment is limited by the 1.5 GHz bandwidth of the photo receiver.

Much information can be gained from the measurement of the eye diagram. The horizontal dark areas of the display show the time-voltage combinations at which the signal spends most of its time, while the cross fainter areas correspond to less frequent (but possibly more troublesome) events. The clear, inside portion of the display is known as the eye. A very clean signal will have a large, clear eye, while a noisy, low-quality signal will have a smaller one. The eye can become completely closed if the data signal has a lot of timing jitter with respect to the master clock, if the pulse widths are incorrect, or if varying amounts of noise and attenuation cause the signal amplitude to vary excessively. Even up to a data speed as high as 1.5 GHz, our experiment shows very clear open eyes.

4. OPTO-ELECTRONICS INTERACTION

Optical transmission speed and electronic switching speed

As we mentioned at the beginning of this paper, instruction/data flows have a good property of locality both in time and in space. So, by adding cache to each bus(fig. 16), the traffic upon the switches will be substantially low.

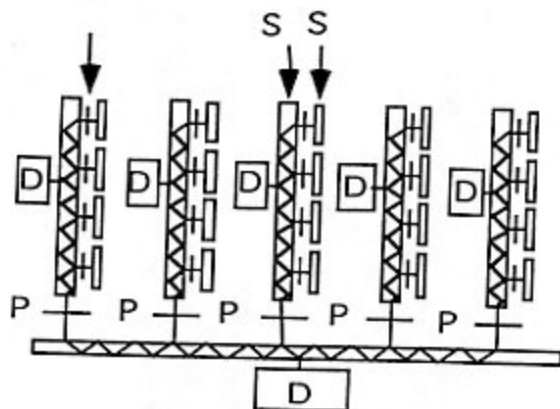


Fig. 16 Cache coherent protocol
D: cache.

Optical Transmission Speed and Electronic Transmission Speed

Suppose there are N processors in the system, and each bus has m processors. In two level bus case, there will be N/m switches to necessitate control. The electronic control channel must transport $\log_2(N/m)$ (addressing space) signaling bits for these switches. To simplify our analysis, we assume switches change in a synchronous way. That is, each cycle T needs $\log_2(N/m)$ bits. The electrical transmission bit rate B_c needed is $\log_2(N/m)/T$:

$$B_c = \log_2(N/m)/T.$$

On the other hand, $T=k/B$, where B is the optical bit rate, and k is a coefficient. Then, we have $B_c = \log_2(N/m)/T = B \log_2(N/m)/k$, and

$$B = \frac{k}{\log_2 N - \log_2 m} B_c.$$

For a fixed electrical bandwidth B_c , the optical bandwidth B slowly decreases as the number of processors N increase. The effect can be seen in fig. 17. Increase electrical bandwidth B_c can increase the optical bandwidth. We get a good optics--electronics relationship since $\log_2(N/m)$ grow slowly with respect to N .

5. CONCLUSION

High-performance optoelectrical hierarchical backplane bus has many outstanding features when compared with existing backplane schemes. It keeps the optical bandwidth while reducing the system latency by employing the techniques of: 1. Locality, 2. Concurrency, 3. all-optical transmission, and 4. Optimized electronic control.

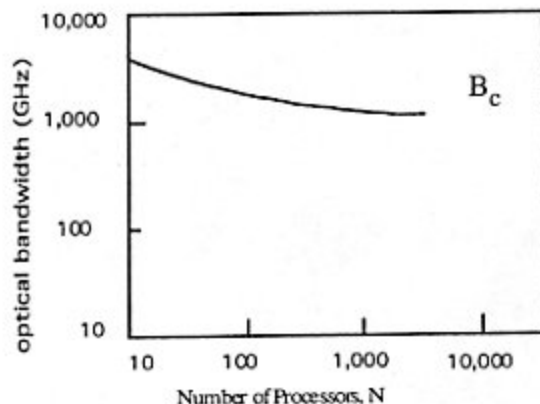


Fig. 17 For fixed electronic control signaling bit rate B_c , it shows optical bit rate decrease slowly with N .

This paper gives some illustration and demonstration to combine optics with electronics efficiently. Proper matching will arise an optimal benefit from both high speed of optics and high logic combinatorics of electronics.

REFERENCES

1. Ray Chen, "Architecture and building blocks for VME optical backplane bus", Optical Interconnects, SPIE Vol. 1894, 1993.
2. S. Natarajan, C. Zhao, and R. Chen, "Bi-directional optical backplane bus for general purpose multi-processor board-to-board optoelectronic interconnects", J. Lightwave Technology Vol. 13, No. 6, 1995.
3. M. Kao, Y. Huang, and R.K. Kostuk, "Multiplexed Substrate-mode Holograms", J. Opt. Soc. Am, A Vol.7, No.8, 1990.

4. H. Kogelnik, "Coupled wave theory for thick hologram gratings", *The Bell System Tech. J.*, Vol. 48, No. 9, 1969.
5. Y. Wu, L. Liu, and Z. Wang, "Optical crossbar elements used for switching networks", *Applied Optics*, Vol. 33, No.2, 1994.
6. J. Ha, T. M. Pinkston, "SPEED DMON: cache coherence on an optical multichannel interconnect architecture" *J. Parallel and Distributed Computing*, Vol. 41, pp. 78-91, 1997.
7. D. M. Chiarulli &, "Optoelectronic Buses for High-Performance Computing", *Proc. of the IEEE*, Vol. 82, No. 11, 1994.
8. D. Sima, T. Fountain, and P. Kacsuk, *Advanced Computer Architectures*, Addison-Wesley, 1997.
9. T. Yatagai, S. Kawai, and H. Huang, "Optical Computing and Interconnects", *Proc. of IEEE*, Vol. 84, No. 6, 1996.