LETTERS

A High Speed Optical Common Bus for a Multi-Processor System

Hiroaki TAJIMA, Yoshikuni OKADA and Koichiro TAMURA, Regular Members

UDC 681.322.012.078 : 631.391.63
621.375.856 : 681.7.062.44 : 621.383.52

SUMMARY One of the bottlenecks for the efficient parallel processing is the communication overheads among resources. To solve this problem, we have been developing a high speed optical communication bus, which consists of laser diodes for signal emission, APDs for signal reception and a cylindrical mirror for broadcasting. The experimental system reported here operates at the clock frequency of 50 MHz.

1. Introduction

We are developing a tightly coupled multiprocessor system which is aimed to be efficient for the logic programing [1]. In this system, the communication speed among the processor and memory units is one of the most critical factors for the high performance of the system. In this context, we have proposed a new type of the signal transmission method based on an optical signal transmission with a cylindrical mirror, which is effective to realize a high speed bus with clock frequency of several hundreds Mega Hertz and more than one thousand ports (the number of modules to be connected) [2].

In our optical bus, a bus master can transmit the signals to all the modules by emitting the light beam toward the cylindrical mirror. Hence the wiring for the bus is unnecessary, that will contribute to the considerable reduction in cost of a highly parallel processor system.

2. Basic Structure of Optical Bus

A common bus of the multiprocessor system shown in Fig. 1(a) can be realized with an optical bus shown in Fig. 1(b), where PUs (Processing Units) and MUs (Memory Units) receive and emit optical signals. We can determine the form of the arc CBD, as follows, such that an optical signal emitted from any unit on the arc and reflected by the mirror can be received by all the units located on the arc. Given the shape of the cylindrical mirror by the radius r, emitting angle of light θ and the region specified by the angle α where the units should locate, we can calculate the parameters of the arc CBD such as θ' and / in Fig. 1(b) [2].

Fig. 1 (a) Common bus of multiprocessor system.
(b) Basic structure of optical bus.

An example of the dimensions of such a system, which is used in our experimentation, is shown in Fig. 1.

Each PU or MU in Fig. 1 includes an optical transmitter/receiver and a communication control module [4]. The optical transmitters/receivers can be stacked to form a parallel bus which carries the signals such as basic clock, arbitration control [5] and data.

The advantages of this kind of optical bus are as follows:
(a) the structure is simple,
(b) the number of ports is large,
(c) the wiring among PUs and MUs are not necessary,
(d) the optical emitting angle of each transmitter can be made equal, and
(e) the transmission delay time from a transmitter to all the receivers can be made almost equal, that is advantageous for clock skewing problems in high speed signal exchange.

3. Experimental System

The feasibility of the optical bus has been studied by the one bit optical signal bus based on Fig. 1. The experimental system is shown in Fig. 2, where the LD (laser diode) is used for the optical signal emission, the APD (avalanche photo diode) for the optical signal reception. Each LD is driven by a pulse generator (PG) output which is triggered by a 32 bit word generator (WG).

The optical signal emitted through a lens is reflected by the cylindrical mirror. The signal is sensed by an APD through a lens and amplified at the reception buffer to the ECL level logic signal S1. The received signal is compared...
with S2 which is the delayed signal of the original signal (WG output). The amount of delay is adjusted at D to the transmission delay of the received signal. The result of the comparison S3 is sampled by S4 and counted to find the number of errors occurred.

4. Results of the Experimentation

The followings are the conditions and the results of the experimentation:

* Signal modulation — Return to Zero (duty 50%),
* Transmission clock frequency — 50 MHz,
* Optical output power of LD (=A) — 6 mW,
* Required minimum optical input power of APD (=B) — 4 micro watt,
* Maximum fanout (=A/B) — 1500,
* Input/output signal level to the module — standard 10k series ECL,
* Transmission delay between laser diode input and reception buffer output — 10 nano second,

* Transmission error — none in 4 hour continuous transmission (ten times).

5. Conclusion and Prospect

In this experimentation, we conclude that the reliability of the optical bus could be sufficient for use as the multiprocessor bus. The limit of the speed is mainly due to that of the amplifier in the buffer.

We are now fabricating a 5 bit parallel optical bus with the clock frequency of 100 MHz. If a laser diode array and an APD array in a single chip are available, a much more compact optical bus can be in reality. Moreover, it is desirable to integrate LDs, APDs and amplifiers in a single chip. We expect that the rapid progress of the opto-electronic technology will make this possible in the near future.

Acknowledgements

The authors wish to thank Dr. K. Sato, Deputy Director of ETL, and Mr. S. Wakamatsu, Chief of Automatic Control Division in ETL, for their support to this research. Our thanks also go to the members of Logical Systems Section for their cooperation in the experimentation.

References

Feasibility of Dialog.H for a Supercomputer

Yoshikuni OKADA, Koichiro TAMURA
Hiroaki TAJIMA and Youichi HAMAZAKI
Electrotechnical Laboratory

1. Introduction

The feasibility of Dialog.H is studied with foci on its throughput and physical realizability. Its experimental system is now under development in our laboratory. It was initially designed for the parallel processing of logic programming, but has generality in nature, which we believe should be very important to the future supercomputers.

2. System Concepts

As shown in Fig.1, the system consists of:
* PUs --- Processing units containing PE (Processing Elements), LM (local memory), WGS (Working Global Shared Memory) which has blocks of copies of GSM the contents of which are swapped in/out by hardware, and COMM (Communication controller) used to broadcast data to all or a set of PUs (block or word load through). PUs are arranged in matrix form (Npp, Nph).
* VB and HB --- Vertical and Horizontal buses connecting PUs and GSM. The numbers of VB and HB are Nph and Nh, respectively.
* GSM --- Global Shared Memory, which consists of MUs (Memory Units), arranged in matrix form (Nh, Nmh).
* CN --- Communication Network among PUs.
* GSL --- Global Status Line, which is the logical AND of a status bit in each PU and used to system synchronization.

The construction of Dialog.H is shown in Fig.2. The communication method among PUs or MUs through VB, HB and CN are all based on the unified approach: the bus arbitration by the multistage unary comparison method; the synchronous data transmission and the packet format. The VHBs are conventional wired buses, but the HBs are based on the high speed optical signal transmission. VHB is a buffer units between a VB and a HB.

The lines for CN are synchronous serial common buses utilizing basic clock of HB, which save the hardware costs and could fit for various topologies as shown in Fig.3.

3. Limits of the System Throughput

The limits of the system throughput are analyzed here from the view point of communication traffics. The three main limiting factors are the communication...
capacity of vertical and horizontal bus and the data input/output rates of MUs, denoted by $V_b$, $H_b$ and $W_b$, respectively. The proportion of the required data transfer capacity for a PU is $V_b*N_p*h : H_b*N_h : W_b*N_m*h*N_h$. Assumed that we have 1,000 PUs ($N_p*h=200$, $N_p*v=5$), $N_h=5$ and $N_m=20$, the typical examples of this proportion are:

- (1/100ns)*200:(1/10ns)*5:(1/100ns)*20*5 = 2 : 0.5 : 1, and
- (2/80ns)*200:(1/1ns)*5:(2/40ns)*20*5 = 5 : 5 : 5.

The first example shows that the main bottleneck is the capacity of HB and the second shows that the load distribution is well balanced with these parameter values.

The utilization factor of the HB is:

$$U(HB) = \frac{Rvh*X(WGSM)*S(WGSM)}{(1+Rhs)}$$

where $U(i)$, $X(i)$ and $S(i)$ denote the utilization factor, the output rate of requests and the mean service time of device $i$, respectively, and $Rvh=N_p*h/N_p*h$, $Rhs=U(COMM)/U(WGSM)$. The values $X(WGSM)*S(WGSM)$ vs Probability of page fault (Ppf) are plotted in Fig. 4 where it is assumed that $T_h$ (the clock period of HB) is 10ns or 1ns, the total number of PUs is 1,000, the arbitration clock period is 40 or 10ns, the page size is 512 or 1K bytes the probability of swapout is 0.7; and the number of data channel of HB is 32. From the figure, it seems that the acceptable processing speed of a PE in each PU is about 5 (Th=10ns) to 20 (Th=1ns) MIPS.

4. Feasibility of Physical Realization

The physical alignment of HB, VHB, VH, MU and PU are roughly sketched in Fig. 5. To implement a system including five or ten HB each of which contains 37 optical channel (32 for data and 5 for control), $N_p*h=1000$ and $N_m=100$, some examples of the required length of $(L_h, L_v)$ are (320mm, 740mm), (1100mm, 1480mm). To realize these, the lenses area for optical receiver and transmitter per channel must be about 6*4 and 10*4 mm^2, respectively. These are very small but are enough to guarantee that the amount of light power emitted by a laser diode and received at each APD is acceptable even if we use the devices of present technology. But the boards of the analog/digital circuit must be so small as to fit for these sizes and the cooling technology for this system must be developed. The easy method of tuning the lenses must also be developed.

5. Conclusion

The system architecture proposed here could fit to a multi-processor system containing 500 to 1,000 PUs, with the processing speed of each PU being up to 20 MIPS (maximum system throughput: 20 BIPS), if we can expect the ceaseless progress in VLSI and opto-electronic technologies in the coming five to ten years.

REFERENCES

4. Y.Hamazaki, Y.Okada, H.Tajima, & K.Tamura, "Implementation of 5 Channel and 100 Mega bit per Second Optical Bus", in this IPSJ.

![Fig.4 Horizontal Bus Traffic](image1)

![Fig.5 Physical construction](image2)