High Performance Multiprocessor Computing System using Optical Centralized Shared Bus

Hai Bi, Xuliang Han, and Ray T. Chen, *Fellow, OSA, IEEE*
Microelectronics Research Center, Department of Electrical and Computer Engineering
The University of Texas at Austin, PRC/MER 1.606G, 10100 Burnet Road,
Austin, TX, USA 78758

Multiprocessing designers have to face a challenge to fulfill effective interconnects employing parallel processing on multiple processors. Centralized shared-memory communications are usually used to simplify programming, to reduce communication overhead because multiple processors share the same physical memory address space. The performance of such multiprocessing systems critically depends on cache coherence maintenance. Shared bus topology was the preferred interconnect scheme because its broadcast nature can be effectively utilized to reduce communication latency to maintain cache coherence. When a processor writes to a shared memory in a broadcast manner, cache controller of each processor can determine whether or not they have to update the new value. However, the physical limitations of electrical bus result in a fact that the trend of computing speed outpacing electrical interconnect capacity is becoming more prominent. In order to operate at higher bit rate electrically, complicated switched medium has been used in backplane topology. In Ref [1], the statistics of the memory read latency in a medium and a large size switch-based multiprocessing system shows that the transaction to maintain cache coherence is a significant fraction. Furthermore, the switching delay for maintaining cache coherence increases with the system scale more rapidly than the wire delay.

The success of optical interconnects already emerged at the machine-to-machine level giving an example for replacing electrical interconnects. There are three basic board-level optical implementation techniques: waveguide, free-space, and substrate-guided optical interconnects. Essentially, the substrate-guided optical interconnect has a bus structure and can keep the data links away from environmental disturbance noise using TIR [2]. With appropriate design and integration of optical components, signal broadcast can be effectively implemented in an innovative architecture called optical centralized shared bus [3]. Fig. 1 illustrates a simplified architectural design of such a concept. A memory board is inserted into the central slot while the other slots are for processor boards. Electrical backplane provides interconnections for non-critical paths. VCSEL and photodiodes are integrated at the bottom of the electrical backplane and aligned with the underlying optical interconnect layer. Positions of the VCSEL and photodiode in the central slot are swapped as indicated in Fig. 1. The optical interconnect layer consists of a waveguiding plate with the properly designed volume holographic gratings integrated on its top surface as optical fan-in/fan-out devices. The diffraction properties in the Bragg regime can be analyzed using Kogelnik's Coupled Wave Theory [4]. Underlying the central slot is an equal-efficiency double-grating hologram, while the others are single-grating holograms. So if one board tries to deliver data to the main memory, the information carried by lightwave will be delivered to the photodiode in the central slot; then central VCSEL will broadcast the updated data on the shared bus. By snooping on the shared bus, all processor boards can immediately obtain the data update information so that cache coherence is consistently maintained across the whole system. By balancing the diffraction efficiency of the waveguide holograms in use, the bus fan-outs across the entire optical interconnect layer can be equalized [5] as demonstrated by the CCD photos in Fig. 2. This merit is highly desirable for system integration because the constraints on the dynamic ranges of the EO transceiver modules in use are reduced. The bandwidth capacity per substrate-guided optical line was experimentally characterized to be approximately 2.5THz [6]. Based on this architecture, we proposed in this paper to apply the optical centralized shared bus in the multiprocessing systems for performance improvement.

With the focus on demonstrating the technical feasibility in a general scenario, a conceptual emulation of the centralized shared-memory scheme was carried out on a generic PCI subsystem that incorporated an optical centralized shared bus. In this emulation, a PCI memory functions as the central memory, while a single 1.2GHz microprocessor and a Gigabit Ethernet Network Interface Card can communicate with each other through the shared PCI memory. The optical centralized shared bus was integrated underneath the PCI backplane. As a preliminary attempt, only PCI bus line AD05 was replaced by the optical interconnection link. Special extension interface to interpret PCI protocols was developed in order to determine data flow direction and appropriately coordinate the operations of the EO transceiver modules during the PCI data transfers. The signal waveforms captured from the extension modules were displayed on an oscilloscope as shown in Fig. 3 for the direct visualization of optical interconnection. Channel 2 displays the modulation current of the laser diode for NIC card. Channel 1 displays the signal generated by the edge detector for central memory board. The results verified the correct connectivity of the implemented optical interconnection link.

0-7803-8306-0/04/$20.00©2004 IEEE

The optical centralized shared bus utilizes the enormous bandwidth capacity of substrate-guided optical interconnects, while at the same time, retaining the desirable architectural features of the shared bus. Its unique topological configuration enables the fulfillment of equalized optical bus fan-outs and simplifies the system integration. As a preliminary attempt, a conceptual emulation of centralized shared-memory multiprocessing scheme carried out on a generic PCI subsystem proved the technical feasibility to eliminate interconnect bottleneck by employing the optical centralized shared bus. The photos of the system and further experimental results will be shown in the conference. This research is sponsored by DARPA, MDA and AFOSR.

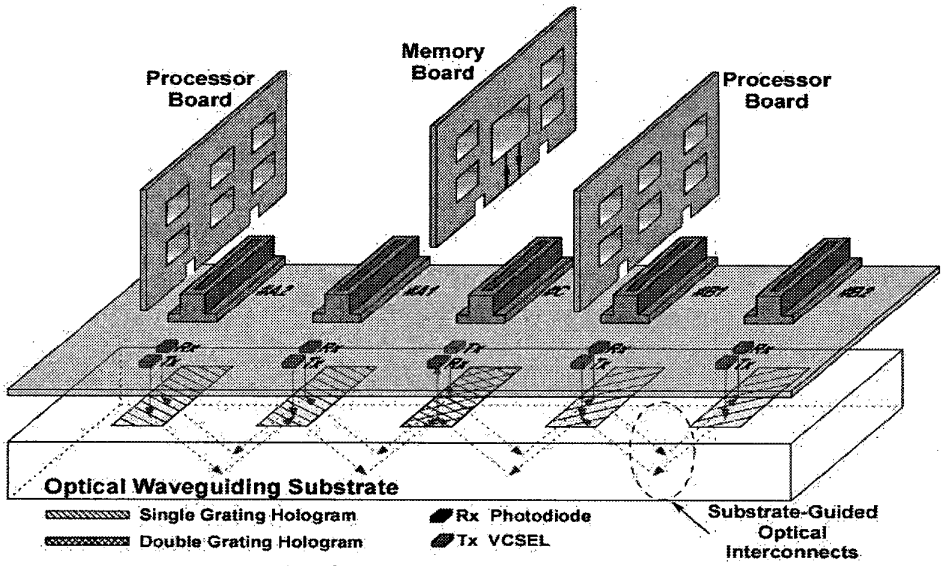

Fig.1. Optical centralized shared bus architecture

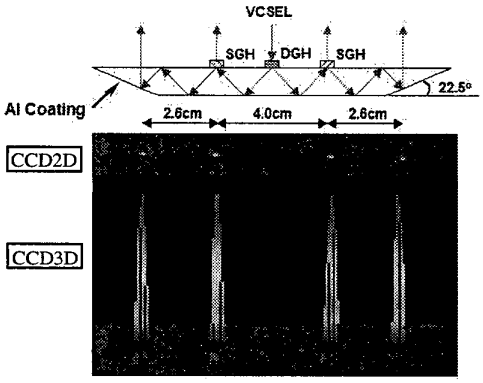

Fig. 2. Equalized bus fan-outs on an optical centralized shared bus

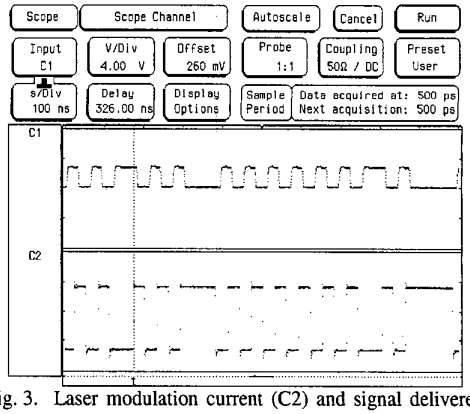

Fig. 3. Laser modulation current (C2) and signal delivered to memory board (C1) at pin AD05

References

1. D. Huang, T. Sze, A. Landin, R. Lytel, and H. Davidson, "Optical interconnects: out of the box forever?" *IEEE Journal on Selected Topics in Quantum Electronics*, vol. 9, pp. 614-623, March/April 2003.
2. K. Brenner, F. Sauer, "Diffractive-reflective optical interconnects," *Applied Optics*, vol. 27, pp. 4251-4254, October 1988.
3. X. Han, G. Kim, G .J. Lipovski, and R.T. Chen, "An optical centralized shared-bus architecture demonstrator for microprocessor-to-memory interconnects," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 9, pp. 512-517, March/April 2003
4. H. Kogelnik, "Coupled wave theory for thick hologram gratings," *The Bell System Technical Journal*, vol. 48, pp. 2909-2947, November 1969.
5. X. Han, G. Kim, and R. T. Chen, "Accurate diffraction efficiency control for multiplexed volume holographic gratings," *Optical Engineering*, vol. 41, pp. 2799-2802, November 2002
6. G. Kim, R. T. Chen, "Three-dimensionally interconnected bi-directional optical Backplane," *IEEE Photonics Technology Letters*, vol. 11, pp. 880-882, July 1999.