1. Technical Summary

This summary describes the technical accomplishments and milestones achieved by the University of Southern California under the POLO program from 7/94 to 10/97.

Early system demonstration:
During the POLO program USC has developed hardware and software to demonstrate the use of HP-POLO modules in system applications. An important early milestone demonstrated use of HP-POLO modules in a Gb/s ring network at USC. The network is used to connect a small cluster of HP workstations which use the UNIX operating system. This early system-level insertion of parallel optical modules provided important practical experience and information which was leveraged in subsequent hardware and software effort at USC.

High-performance CMOS design and implementation:
USC has developed and implemented GB/s ports in submicron CMOS technology. The high-speed analog design is needed to interface to the parallel optical modules developed as part of the POLO program. The integration of digital logic and high-speed analog on the same chip is also an important technical achievement.

Hardware and software system integration:
The Link Adapter chip developed as part of the POLO program have been used to bridge between Intel-based host systems and a network interconnected using HP-POLO modules. The Link Adapter serves as the link interface between an industry standard 32 bit host data bus and the HP-POLO-2 optical module. The HP-POLO-2 module contains the opto-electronic interface ICs as well as the VCSEL optical transmitters and PIN receivers. Each port consists of 10 channels and optical coupling is provided by a multimode ribbon fiber. Each channel supports a bandwidth of at least 1 Gb/s, giving an optical bisection bandwidth of 20 Gb/s.

The HP-POLO-2 modules and LA chip are integrated on the SurfBoard. System-level testing includes use of custom system debug software developed at USC and demonstration of the POLO workstation cluster environment for multimedia and medical visualization applications.

USC has also developed applications under the Microsoft Windows NT 4.0 operating system to test the user-perceived performance of the complete system. This work is based on remote medical imaging applications where a volume rendered image is passed from a server to a client via the experimental interconnect.

2. Link Hardware

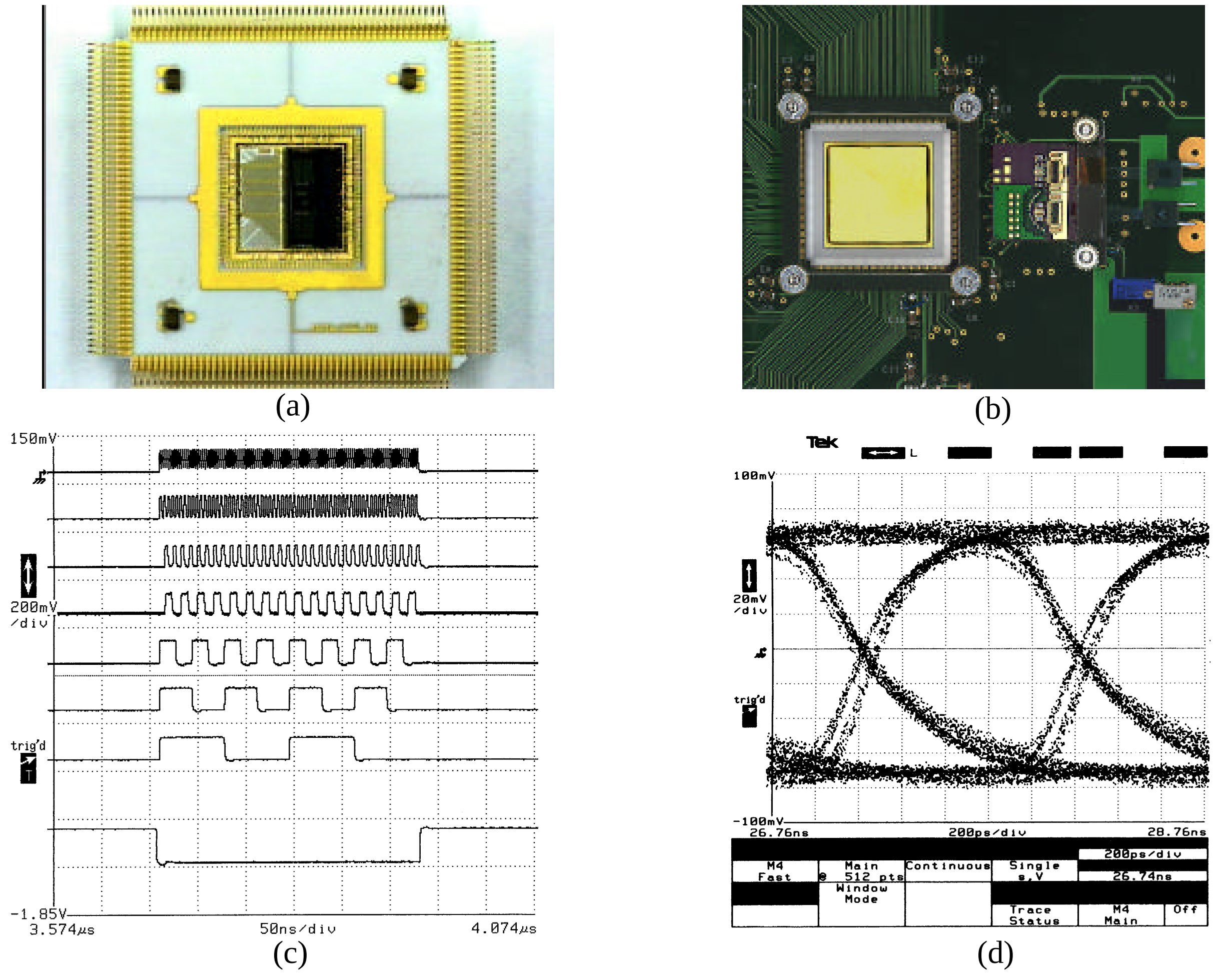

Figure 1: (a) 196/152 QFP with P2P IC bonded in 1x1 cm^2 cavity. (b) QFP and HP-POLO-2 module mounted on SurfBoard. (c) data channels (D0-6) and the frame control channel (FC) during the transmission of a data packet while the frame control line is enabled. D0 is upper most trace and FC is lowest trace. Data is clocked at 1Gb/s. (d) Eye diagram for one differential data channel (D0) operating at 1 Gb/s. The measured bit error ratio (BER) for a non-return to zero (NRZ) pseudo-random bit stream (PRBS) of 2^{32}-1 is less than 10^{-13}. Signals are attenuated by -20 dB so that the vertical scale for (c) is 2 Volts/div and 200 mVolts/div for (d).

Key components of this system are the P2P interface IC developed at USC and the Hewlett Packard HP-POLO-2 parallel opto-electronic transceiver module. The P2P chip is a synchronous serializer/deserializer bridge capable of a total signalling rate of 8 Gb/s, plus one clock and one control channel, giving a total electrical transceiver bisection bandwidth of 20 Gb/s. This IC serves as the link interface between a standard 32 bit host data bus and the HP-POLO-2 optical module. The HP-POLO-2 module contains the opto-electronic interface ICs as well as the optical transmitters (VCSELs) and receivers (PIN photodetectors). Each port consists of 10 channels and optical coupling is provided by a multimode ribbon fiber. Each channel supports a bandwidth of at least 1Gb/s, giving an optical bisection bandwidth of 20 Gb/s.

3. System architecture overview

3.1 Schematic diagram of architecture

Each node of the P2P system consists of an Intel Pentium host, an AMCC PCI Controller, the PCI to LA adapter (glue logic), and the LA chip and HP-POLO-2 module mounted on a SurfBoard.

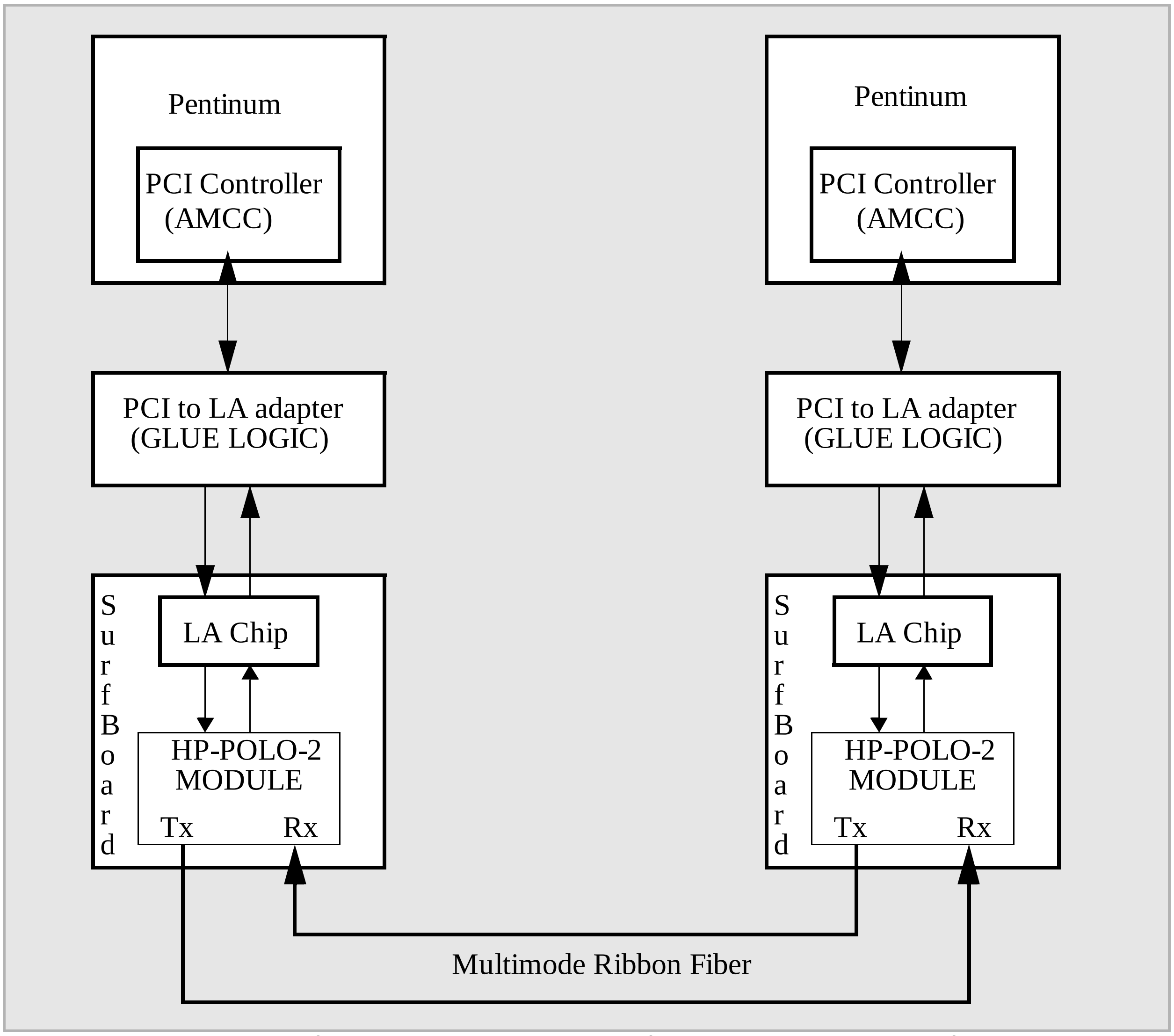

Figure 2. Schematic diagram for complete system test. Basic functional blocks consisting of host, glue logic, and SurfBoard are shown. Multimode ribbon fiber is the physical medium used to transport optical signals between the two hosts.

The core of POLO network test system is the POLO link adapter chip (LA chip) and HP-POLO module, however to implement and demonstrate this POLO networking system, a ‘PCI to LA adapter’ (glue logic) was built to provide the interface to Intel Pentium motherboards (see reference 9, ‘PCI to LA Chip Interface Specification’).

3.2 PCI to LA adapter board description

The PCI glue logic adaptor board is designed to enable basic synchronous data flow functionality between the AMCC PCI Controller chip and the LA chip. The AMCC S5933 PCI Controller chip performs all the critical timing and handshaking with the PCI bus. On the other hand, the PCI glue logic board performs the handshaking of control signals and the coordination of data transfer between the AMCC S5933 MatchMaker and LA chip. The PCI Board's FIFOs allow up to 1Kx32 bits (4kB) of elastic storage in each direction (both Tx and Rx). Thus, the two primary data transfer functions, namely: DMA Read, DMA Write, provide a fast and efficient method of rapidly transferring data between the host's local memory and the LA chip. Synchronous operation allows up to 1.06 Gb/s burst throughput on the 32-bit PCI interface clocked at 33 MHz and 2.1 Gb/s on the LA chip interface clocked at 33 MHz. The AMCC interface contains a 32-bit data bus and the needed additional control and flag bits. The LA chip interface provides separate Tx and Rx ports for simultaneous transceiver operation supporting a bisection bandwidth of up to 2.1 Gb/s for 33 MHz clock speeds. In the future, this interface can be reconfigured for 66 MHz operation allowing up to 4.2 Gb/s at the LA chip interface and 2 Gb/s over a 66 MHz 32-bit wide host interface.

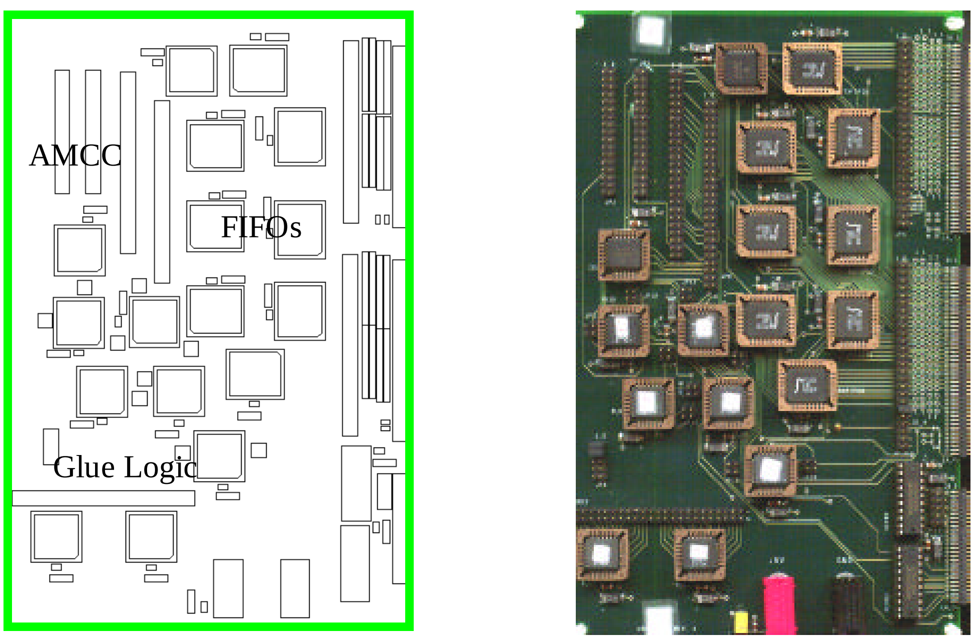

Figure 3. PCI to LA glue logic adapter board showing functional layout (left) and implementation (right). The AMCC interface contains a 32 bit data bus and additional control and flag bits. The LA chip interface provides separate Tx and Rx ports for simultaneous transceiver operation supporting bisection transfer rates of up to 2.1 Gb/s for 33 MHz clock speeds. Byte-wide FIFOs and fast, low complexity, PLAs were chosen to meet the timing requirements for synchronous operation. The programmable glue logic contains the state machines, counters, and decoding logic to handle host initiated bus-master synchronous DMA operations.

3.3 Communication between two PCI to LA interface boards

To demonstrate the functionality of the PCI to LA interface boards, two boards were connected back-to-back, and DMA transfers were performed on data packets from one Intel Pentium host to another. Figure 4 shows the configuration that was used to exhaustively test the glue logic for data transfer between the two hosts. Successful error-free data transfer was demonstrated.

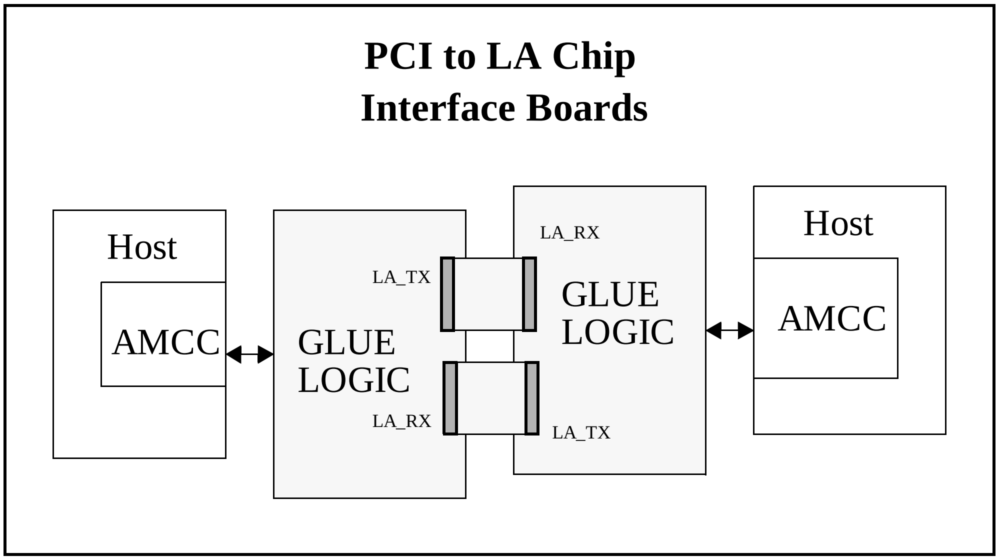

Figure 4. Schematic block diagram of the PCI to LA chip adapter board wired back-to-back.

We developed an assembly program to demonstrate high-sustained transfer rates. The total sustained throughput of approximately 380 Mb/s achieved is limited by host architecture implementation and not the performance of the LA chip interface boards.

We also developed an application under the Microsoft Windows NT 4.0 operating system to test the user-perceived performance of the complete system. This work is based on remote medical imaging applications where a volume rendered image is passed from a server to a client via the experimental interconnect. Figure 5 shows the Graphical User Interface (GUI) developed for this application. The application initially loads large image files into the server's main memory and transfers these files as 128 word packets over the PCI to LA interface boards to the client Intel Pentium computer. The received packets are then assembled and displayed on the remote computer in real time.

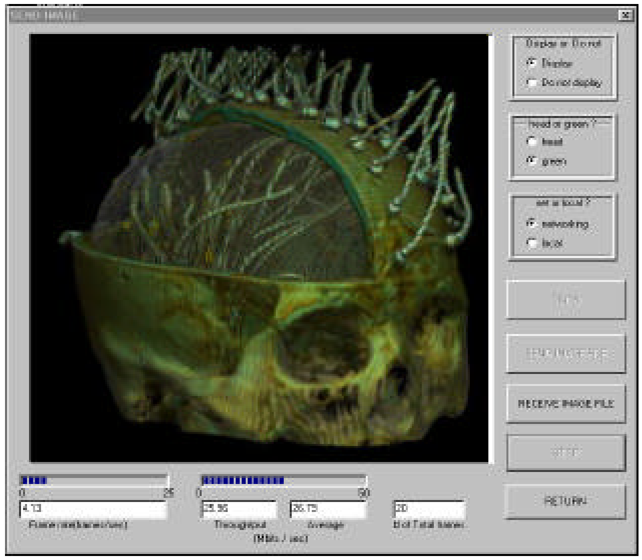

Figure 5. GUI for interconnect demonstration using the PCI to LA chip adapter board. The medical imaging application is developed using Microsoft VC++ 5 and the Microsoft Windows NT4.0 operating system.

3.4 SurfBoard communication

Two versions of the SurfBoard were implemented. Both versions provide support for the PCI to LA chip adapter board. For pure electrical testing of the LA chip, the electrical implementation of the SurfBoard is used. On the electrical SurfBoard, the LA chip's high-speed I/O pins are routed to SMA connectors instead of the HP-POLO-2 module BGA I/O pins. Electrical interconnection between SMA connectors is via 1m-long coaxial cables. This allows either full physical loop-back or back-to-back testing of the LA chip. For optoelectronic testing, the optical SurfBoard is used and the electrical interconnection between two LA chips is replaced by the HP-POLO-2 module. Multimode ribbon fiber is the physical medium used to transport optical signals between the two HP-POLO-2 modules.

Communication between the SurfBoards is initiated by the host. The initiator host sets up a DMA read transfer while the receiver host enables a DMA write transfer. Synchronous DMA transfers begin after the host's operation registers are programmed with the local host's memory address and data transfer lengths. Packet length counters on the PCI to LA interface board delimit transfer lengths to a preset value of up to 256 32-bit wide words. The process of terminating packets with End Of Packet (EOP) bits are transparent to the host, however, they are necessary for the LA chip. Therefore, a predetermined packet length of 128 words has been encoded in order to maximize the throughput. Upon receipt of 128 words by the LA chip, the data packet is serialized from 32 bits down to 8 bits (4-to-1) and sent out over the high-speed ports of the LA chip at the high speed clock rate. On the RX side of the LA chip, the high-speed serialized data is de-multiplexed and clocked into the RX FIFOs on the PCI to LA interface board. Simultaneous send and receive is possible on this interface due to the presence of both an RX and TX clock and the existence of two independent data ports. As soon as the receipt of data into the RX FIFOs occurs, a bit is written into the operation register of the receiver host indicating that a DMA write transfer should begin. This completes one phase of a packet transfer.

To test communication between the host and a SurfBoard, a program was developed in assembly code to handle all the packet data transfers and received data verification. This program can repeatedly send and receive an arbitrary number of data packets in loop-back mode as well as between hosts. The program was used to verify error-free data transmission between hosts using the LA chip mounted on both the electrical and optical SurfBoards.

3.5 Electrical SurfBoard physical loop-back test

The electrical version of the SurfBoard was first tested and verified in physical loop-back mode. An HP external pulse generator was used as a stable high-frequency clock source for high-speed electrical testing. Correct logic operation at 660 MHz was confirmed, and up to 1 GHz is possible with a 6.2 V power supply (Ref. 2) for the IC implemented in 0.8 µm CMOS. The electrical SurfBoard tests were performed both with connection to a host and with a logic analyzer to simulate the signaling interface of the host.

3.6 Communication between two electrical SurfBoards

This procedure is similar to physical loop-back. Coaxial cables connect the two electrical SurfBoards (as shown in Figure 6) and an external pulse generator is also used to provide the clock. The result were identical to the loop-back test. As described above, communication between the two SurfBoards is established by the host. The initiator host fills the TX FIFOs on the PCI to LA chip adapter board, while the LA chip's high-speed TX port waits for the completion of one packet. Data packets are sent along with a control frame signal over the coaxial cables, deserialized and latched into the receiver's RX FIFOs. The RX FIFOs retain the packet until the receiver host is ready to perform a DMA write transaction.

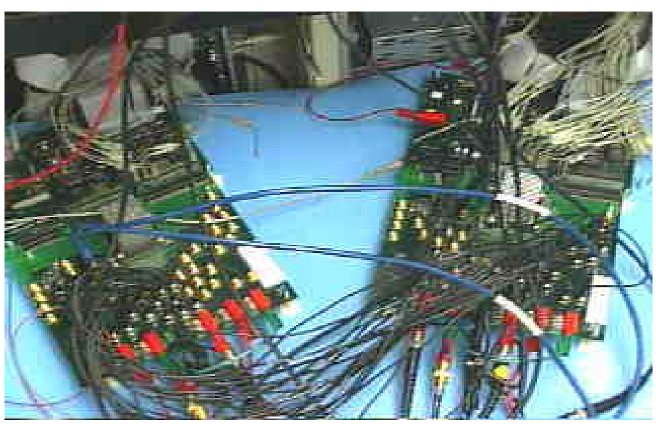

Figure 6. Experimental test configuration for communication between two electrical SurfBoards. Coaxial cable is used to interconnect the high-speed I/O of the electrical SurfBoards.

3.7 HP-POLO-2 on optical SurfBoard tests

The HP-POLO-2 module on the optical SurfBoard was tested and characterized for high-speed operation at data rates up to 1.25 Gb/s. An HP BER tester was used to generate 2^{31}-1 NRZ PRBS data. A pair of mini-coaxial cables were stripped and soldered onto vias in the path of the high-speed TX signal lines between the LA chip and the HP-POLO-2 mounted on the SurfBoard. The RX output of the HP-POLO-2 module was then coupled into a Bias-T and observed on a high-speed sampling oscilloscope.

Figure 8 shows the results of measuring eye diagrams for an HP-POLO-2 module mounted on an optical SurfBoard. Tests were performed in physical loop-back mode at a data rate of 1.25 Gb/s (800 ps per bit) using a NRZ 2^{31}-1 PRBS. In the Figure, channels CN, D0-7, and CLK have horizontal scale 200 ps/div and vertical scale 200 mV/div.

The physical loop-back tests on the HP-POLO-2 module verified that each channel in the combined Tx and Rx paths functioned correctly with measured minimum eye widths of 399 ps at 1.25 Gb/s. The center channels (D0 - D6) exhibited superior > 500 ps eye widths. During the course of our testing it became apparent that the optical outputs are very sensitive to the pre-bias voltage, LASERP. We found that the optimum setting for this value is LASERP = 0.6 V. The average optical power output of each channel with the inputs floating is 1.1 mW.

Bit error ratios of less than 10^{-12} were measured for data rates of up to 1.7 Gb/s. In addition, we observed no signs of an error floor caused by modal noise, making this optical link very suitable for our system application.

Following the verification of the physical loop-back test on the HP-POLO-2 module, an optical link was established using two separate SurfBoards and a pair of the multimode ribbon fiber interconnect. Again, using the HP BER as a data and clock source, and a high speed sampling oscilloscope as the receiver, each channel was individually tested to confirm functionality before integrating the SurfBoards into the experimental POLO network.

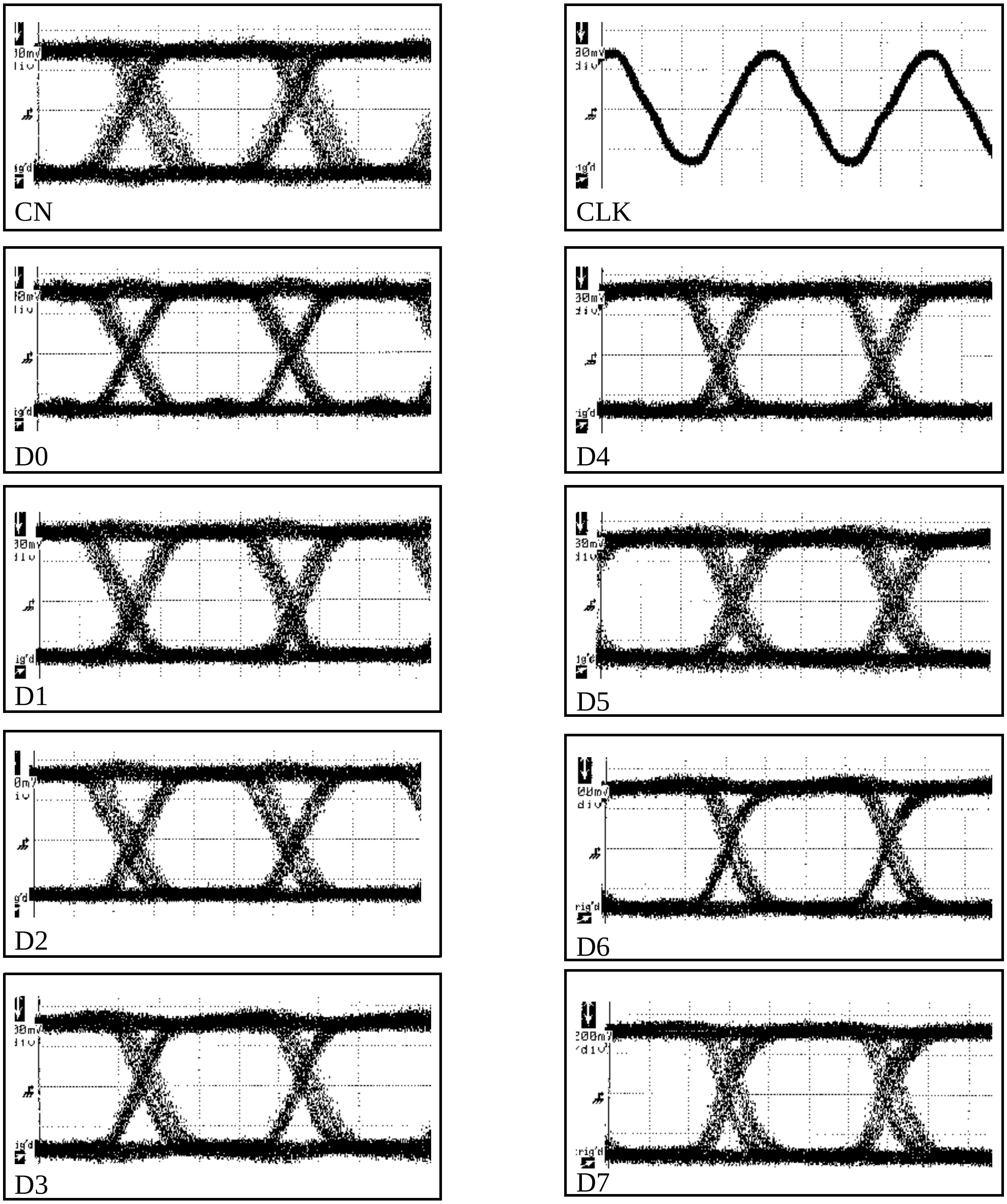

Figure 7. Eye diagrams for HP-POLO-2 module mounted on an optical SurfBoard. Tests were performed in physical loop-back mode at a data rate of 1.25 Gb/s (800 ps per bit) using a NRZ 2^{31}-1 PRBS. In the Figure, channels CN, D0-7, and CLK have horizontal scale 200 ps/div and vertical scale 200 mV/div. Our measurements confirmed that bit error ratios of less then 10^{-12} are achievable over all module channels for data rates up to 1.7 Gb/s.

3.8 Communication between two optical SurfBoards

After the verification of each component in the link, the entire parallel opto-electronic interconnect system is tested. Figure 8 shows the experimental arrangement when two optical SurfBoards are connected via multimode ribbon fiber. We have successfully demonstrated the full integration of a functional parallel fiber-optic link. The parallel fiber-optic link is a communication path between two Intel Pentium hosts. Each node consists of the Intel Pentium host, the PCI to LA adapter, the P2P IC, and the HP-POLO-2 module.

The experiments described in this report use a P2P IC implemented in 0.8 µm CMOS. As a result of our experiences developing and measuring the performance of our CMOS interface to parallel opto-electronics, we have identified a number of potential improvements in the IC design. In the near future, we expect to test an improved version of the LA chip. We anticipate that this improved P2P IC, which has been implemented in 0.5 µm CMOS technology, will be capable of operating at greater data transfer rates and much reduced power levels.

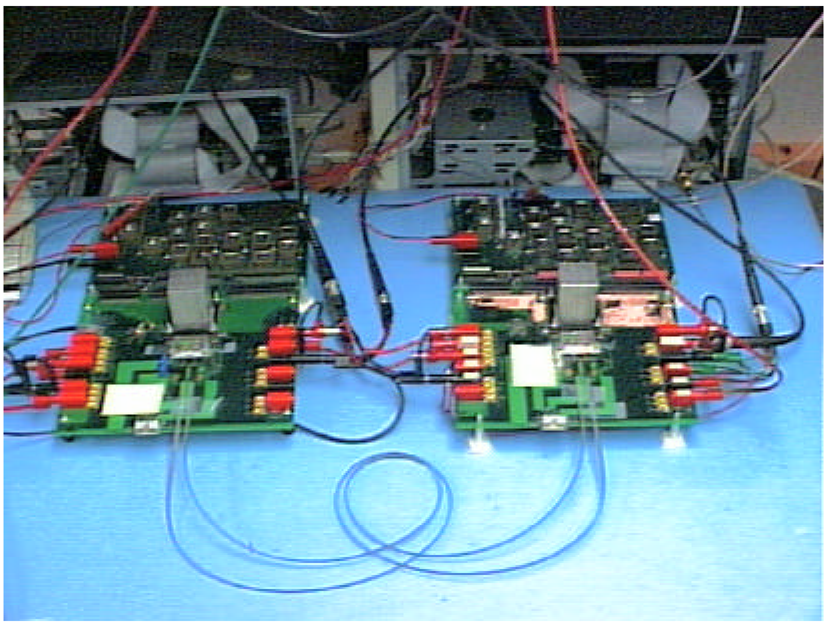

Figure 8. Two optical Surfboards connected via multimode optical ribbon fibers. Pictured behind the surfboards are the PCI to LA chip interface boards and the Intel Pentium host computers. Burst rates of up to 2.1 Gb/s are possible through the optical link.

References:

1. “Gigabyte/s Data Communications with the POLO Parallel Optical Link”, Hahn et. al, 46th Electronics Components and Technology Conference, 1996.
2. “8 Gbit/s CMOS interface for parallel fiber-optic interconnects”, B. Sano, B. Madhavan and A.F.J. Levi, Electronic Letters, Nov 21, 1996. v32 No.24.
3. AMCC PCI Matchmaker Developer’s Kit Technical Reference Manual (http://www.amcc.com)
4. AMCC Data Book: S5933 PCI Controller (http://www.amcc.com)
5. HP-POLO-2 SurfBoard Specification. (/home/polo/usc/POLO-DOC/dir.POLO_BGA/)
6. “ Hewlett Packard Afterburner Specification Rev 1.0”, Networks and Communications Lab, Hewlett Packard Labs, Filton Road, Stoke Gifford, Bristol, BS12 6QZ, England.
7. Intel Chipset Specifications (http://developer.intel.com/design/pcisets/)
8. BlueWater Systems WinDK Developer Kit. (http://www.bluewatersystems.com/)
9. PCI to LA Interface Specification. (/home/polo/usc/POLO-DOC/dir.POLO_PCI_adapter/Doc)