Networking with Free Space Optical Data Packets Using Carrier-Sense Multiple-Access with Collision Detection (CSMA/CD) Protocol

Jen-Ming Wu, C.B. Kuznia, Chih-Hao Chen, Bogdan Hoanca and A.A. Sawchuk
Signal and Image Processing Institute; University of Southern California
3740 McClintock Avenue; Los Angeles, CA 90089-2564
Tel: (213) 740-4143 Fax: (213) 740-4651 email: jwu@biron.usc.edu

1. TRANSPAR Network Introduction
We have developed a smart pixel networking scheme for distributing three-dimensional optical data packets among nodes. The system has a large array of parallel channels operating at on-chip clock rates, allowing for a potential throughput of $> 1$ T/s between VLSI chips [1]. The use of parallel data packets results in lower latency because each parallel packet spends less time on the network as compared to serial methods. Each network node implements a modified Carrier-Sense Multiple-Access with Collision Detection (CSMA/CD) media access protocol derived from the Ethernet standard and is extended to operate over ring networks that pass spatially parallel packets.

To demonstrate this novel networking concept, we created an optoelectronic VLSI chip TRANslucent Smart Pixel Array (TRANSPAR). The TRANSPAR chip is being fabricated through the DARPA/OMU/CO-OP foundry program using the Lucent Technologies OE/VLSI process [2]. This foundry offers flip-chip bonding of 2-D arrays of MQW modulators and detectors (that operate as free space optical data ports) onto 0.5 micron CMOS circuitry. The TRANSPAR chip contains a $5 \times 8$ array of identically replicated smart pixels in a $2 \times 2$ mm$^2$ chip. Each smart pixel contains digital circuitry (550 transistors/pixel), a one-bit optical transmitter and a one-bit optical receiver. A TRANSPAR network contains multiple copies of the TRANSPAR chip, with each chip acting a network node. TRANSPAR nodes are inter-linked by a two-dimensional array of free space optical channels as shown in Fig. 1.

The TRANSPAR transfers data between nodes using 3-D packets of bits called optical parallel data packets (OPDPs). An OPDP consists of multiple 2-D arrays of bits transmitted from optical I/O ports by smart pixels over several clock cycles. The OPDP contains a data payload, source and destination address information and an optical clock channel. When transferring an OPDP, the transmitters of a source node are synchronously linked to the receivers of a destination node using the optical clock channel. This allows TRANSPAR nodes to perform high-speed data transfers between nodes having independent (and widely varying) clock rates. The source and destination address fields are spatially encoded in the OPDP for parallel packet recognition. Because TRANSPAR processes packet header information in parallel, no packet buffering is required at any node.

2. CSMA/CD Networking with OPDPs
The TRANSPAR nodes are physically arranged as a ring network shown in Fig. 2. The TRANSPAR nodes have distributed control of the network, with attached host processors performing uploading/downloading of packets to/from the nodes.

Once transmitted from a source TRANSPAR node, an OPDP travels through the entire ring, optically propagating in free space between nodes and electrically propagating on the VLSI plane within nodes. All TRANSPAR nodes detect the OPDP and compare its destination address with their own

Figure 1: TRANSPAR nodes communicate using 3D Optical Parallel Data Packets (OPDPs). An OPDP contains payload data, packet source/destination address, and a data clock channel for synchronous transfer.
address, downloading on a match. The latency per node (including both the propagation time through the node and the optical propagation between nodes) in a typical physical system is less than 2 ns.

TRANSPAR has been designed and simulated to have a target on-chip clock rate of 200 MHz. Since the optical data channels are formed directly between the internal (or core) circuitry on communicating nodes, the data transfer rate is comparable to the on-chip clock rate. This contrasts with traditional computing architectures that use a much lower off-chip clock rate for data transfer between chips or boards.

Figure 3 shows the optical network interface circuitry within a smart pixel PE. This circuit contains an optical receiver (Rx), optical transmitter (Tx), and an 8-bit FIFO buffer for downloading OPDPs. The 5 x 8 array of smart pixels form an OPDP that is 5 x 8 bits in spatial dimension by 8 bits deep in the temporal dimension. Thus, TRANSPAR forms an 5 x 8 x 8 OPDP by having all smart pixels transmit data over eight clock cycles. Scaling this prototype to standard CMOS chip sizes creates OPDPs containing 4 Kbytes (64 x 64 array size) of data with a network throughput of over 800 Gb/s.

When a node is not adding an OPDP onto the network, all of its smart pixels are in transparent mode. In transparent mode, any signal entering the Rx passes directly on to the Tx, incurring only a few gate delays (~1 ns). When adding an OPDP onto the network, the PE is in opaque mode. Signals entering the Rx are blocked while the Tx transmits the OPDP onto the network. The transmitted packet travels through all other TRANSPARs almost instantaneously, since they are transparent. The smart pixels dedicated to destination address recognition determine if there is an address match and, if so, signal all smart pixels on the chip to download the OPDP. When downloading, a dedicated optical clock smart pixel detects the optical clock, supplied in the OPDP, and broadcasts the signal electronically to all PEs for synchronous transfer into their 8-bit FIFO buffers.

When a source node wants to add an OPDP onto the network, it ‘senses’ the network for inactivity. There is a chance that two or more TRANSPAR nodes will detect an idle network at nearly the same moment and upload OPDPs that will cause contention (or a collision, in Ethernet terms). Each node contains circuitry to detect such a collision and reset the network.

3. Conclusions and Acknowledgments
The TRANSPAR chips are being assembled into a packaged optoelectronic system. We will present additional details of their protocol and operational parameters. This work was supported by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, with additional support from the Annenberg Center for Communication at the University of Southern California and the California Trade and Commerce Agency. It was also supported by the Joint Services Electronics Program through the Air Force Office of Scientific Research.

4. References