A HIGH-SPEED OPTICAL MULTIDROP BUS FOR COMPUTER INTERCONNECTIONS

Signal integrity constraints of high-speed electronics have made multidrop electrical buses infeasible. This high-speed alternative uses hollow metal waveguides and pellicle beam splitters that interconnect modules attached to the bus. With 1 mW of laser power, the bus can interconnect eight modules at 10 Gbps per channel and achieves an aggregate bandwidth of more than 25 Gbytes per second with 10-bit-wide signaling paths.

Michael R.T. Tan
Paul Rosenberg
Jong-Souk Yeo
Moray McLaren
Sagi Mathai
Terry Morris
Huei Pei Kuo
Joseph Straznicky
Norman P. Jouppi
Shih-Yuan Wang
Hewlett-Packard

Multidrop buses find many applications in computer systems, particularly in the I/O and memory systems. They let architects create readily expandable systems with uniform access times for all devices. A good example is expanding a computer’s main memory by adding more memory modules. For applications such as front-side buses, the ability to broadcast simplifies the design of coherency protocols. Increasingly, however, networks of point-to-point links are displacing buses—as seen, for example, in the migration of PCI to PCI Express—driven primarily by signal integrity considerations. The impedance discontinuities inherent in multidrop electrical buses have made it impossible to scale their data rates to track improvements in processor performance. To keep system design and provisioning flexible, we’d like a high-speed alternative to traditional electrical buses.

We can easily avoid the difficulty of creating a multidrop bus electrically by using optics. Optics provides numerous benefits over copper:

- a high carrier frequency of $10^{15}$ Hz, resulting in no signal degradation with increased modulation frequency;
- low propagation loss ($< 0.1$ dB/km for glass fiber);
- broadband impedance matching (anti-reflection coatings, beam splitters) with low loss;
- no loading effects;
- electromagnetic interference immunity; and
- low power consumption.

Moreover, optical interconnects have higher density than copper interconnects at high data rates.

**Multidrop optical bus**

Figure 1 shows a master-slave parallel optical bus. It consists of two unidirectional
signal buses with four or more modules attached. The master module broadcasts signals on the bus, where any module can receive them and send data back to the master. Two sets of parallel optical waveguides, one a 10-bits-wide fan-out and the other a 10-bits-wide fan-in, interconnect the modules attached to the bus. The master module transmits optical signals received by all the slave modules. Optical beam splitters (BS₁...BS₄) at each T junction interconnection node tap power from the bus. To distribute transmitted optical power equally among interconnected modules, each beam splitter has a different reflectivity $R_i$ and transmissivity $T_i$.

The relationship between the beam splitters’ reflectivities for the case of no loss in the beam splitters (that is, $T_i = 1 - R_i$) is

$$R_n = \frac{R_1}{\prod_{m=2}^{n} (1 - R_{m-1})^k}$$

$$= \frac{R_1}{1 + R_1 - nR_1}$$

for $n > 1$ and $k = 1$ where $R_1$ is the reflectivity of the first beam splitter, $n$ is the tap's position along the bus, and $k = e^{-\alpha L}$ takes into account the propagation loss of distance $L$ between taps. The $k$ factor can include excess loss from the splitters. The reflectivity of the first beam splitter $R_1$ is determined by the power the receiver requires to maintain a reliable link (receiver sensitivity) with a reasonable link margin. Each slave module in turn sends data back on the fan-in bus to the master module using identical beam splitter junctions. As Figure 2 shows, on the return path, the selected beam splitter ratios ensure that the master receives the same optical power from each slave module no matter how many stages away the slave module is. The reflectivity values of the beam splitters on the return path are the same as for the output path.

The master module performs an electrical-to-optical conversion using a transmitter...
(Tx) IC that drives a $1 \times 10$ 850-nm vertical-cavity surface-emitting laser array capable of 10-Gbps modulation. A 90-degree turning mirror and a microlens array couple light from the vertical-cavity surface-emitting laser (VCSEL) array to the waveguide array. The beam splitters are used to broadcast the optical signal to the different receiver modules. At each slave module, an identical microlens array focuses the tapped optical signal onto a high-speed photodetector array, and a receiver (Rx) IC converts the signal back to an electrical signal. Each slave module sends data back to the master module via an identical set of optoelectronic components. Only one slave can send data back to the master at a time. We can ensure this either by having the master exactly schedule all slave transfers, or by adding an arbitration mechanism between slaves.

Transmit power, receiver sensitivity, and system losses determine the maximum optical fan-out. At 10 Gbps, the receiver sensitivity of commercially available transimpedance amplifier (TIA) receivers is around $-14$ dBm (40 μW). The major system losses are waveguide propagation loss, coupling loss, and beam splitter losses. Table 1 shows a detailed power link budget for the optical bus. The budget assumes a waveguide loss of $-0.1$ dB/cm with a propagation length of 3 cm between taps and an excess loss of $-0.3$ dB at the splitters. We used a receiver sensitivity of $-14$ dBm. We assumed an additional 0.5 dB of coupling loss into the receiver. With 0.6 dBm of transmit power, the maximum fan-out is 7. Increasing the input power to $+1$ dBm increases the link margin to $\sim 3.45$ dB per tap.

Decreasing the input power to $-0.5$ dBm decreases the link margin to $\sim 1.95$ dB. If the propagation loss and excess loss per tap decrease to $-0.07$ dB/cm and $-0.25$ dB, respectively, the optical fan-out increases to 8 with a link margin of 3 dB. The beam splitter power reflectivities for this case would be 0.082, 0.100, 0.123, 0.155, 0.204, 0.284, 0.439, and 0.867. Thus, with about 1 mW of laser power, we can interconnect seven to eight modules at 10 Gbps per channel. Increasing the laser power allows larger link margins and/or longer bus lengths. An aggregate bandwidth of more than 25 Gbytes per second (GBps) is achievable with a 10-bit wide fan-in and fan-out bus.

For the fan-in bus, each module transmits the same amount of optical power. At each tap, the tap reflectance reflects power into the waveguide, while the rest (tap transmission) is discarded. The power coupled into the waveguide passes through the remaining intermediate taps before reaching the receiver at the master module. The overall transmission loss is the product of the tap reflectance and the transmittance of the subsequent taps traversed. The total loss consists of the transmission loss plus the total waveguide loss and the excess loss from each tap junction traversed. Table 2 details the $1 \times 7$ fan-in link budget.

### Competing technology

For bus applications at high data rates, one electronic option is to use daisy chains of point-to-point links. The fully buffered dual inline memory module (FBDIMM) standard is an example of this technique applied to memory systems.

<table>
<thead>
<tr>
<th>Tap number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam splitter reflectivity</td>
<td>0.09</td>
<td>0.11</td>
<td>0.15</td>
<td>0.20</td>
<td>0.28</td>
<td>0.45</td>
<td>0.96</td>
</tr>
<tr>
<td>Tap input power (dBm)</td>
<td>$+0.6$</td>
<td>$-0.4$</td>
<td>$-1.5$</td>
<td>$-2.8$</td>
<td>$-4.4$</td>
<td>$-6.4$</td>
<td>$-9.7$</td>
</tr>
<tr>
<td>Loss (dB)</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
</tr>
<tr>
<td>Excess loss (dB)</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
<td>$-0.3$</td>
</tr>
<tr>
<td>Reflected power (dBm)</td>
<td>$-10.5$</td>
<td>$-10.5$</td>
<td>$-10.5$</td>
<td>$-10.5$</td>
<td>$-10.5$</td>
<td>$-10.5$</td>
<td>$-10.5$</td>
</tr>
<tr>
<td>Transmit power (dBm)</td>
<td>$-0.4$</td>
<td>$-1.5$</td>
<td>$-2.8$</td>
<td>$-4.4$</td>
<td>$-6.4$</td>
<td>$-9.7$</td>
<td>$-25$</td>
</tr>
<tr>
<td>Receiver coupling loss</td>
<td>$-0.5$</td>
<td>$-0.5$</td>
<td>$-0.5$</td>
<td>$-0.5$</td>
<td>$-0.5$</td>
<td>$-0.5$</td>
<td>$-0.5$</td>
</tr>
<tr>
<td>Receiver power (dBm)</td>
<td>$-11$</td>
<td>$-11$</td>
<td>$-11$</td>
<td>$-11$</td>
<td>$-11$</td>
<td>$-11$</td>
<td>$-11$</td>
</tr>
<tr>
<td>Link margin (dB)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
disadvantages. First, the FBDIMM access latency increases proportionally to the number of hops. Although the memory modules have a very low pass-through latency, the round-trip delay starts to approach the base memory access time for systems with four or more FBDIMMs.\(^3\) Second, having to retransmit the data at each intermediate FBDIMM module incurs a significant power penalty. The advanced memory buffer in the FBDIMM adds 6 W to each DIMM’s power irrespective of whether DRAMS on that DIMM are accessed. It would be possible to implement the FBDIMM protocol with point-to-point optical links at each stage of the outbound and return daisy chains, allowing for longer lengths. However, this would increase overall delay because the delay of the optical-to-electrical and electrical-to-optical conversions would occur on every hop, rather than once in each direction for the outbound and inbound broadcast structure. This arrangement also requires twice the number of optical-electrical converters and twice the power for an equivalent bandwidth.

Modifying the bus protocol to use a single ring would yield a solution that uses the same number of optical-electronic converters as the optical bus. But this solution would halve the maximum throughput because there no longer would be independent outbound and return data paths.

Optical buses for computer interconnections have been studied previously.\(^4\) Most results have been limited to low data rates or a low number of fan-out and fan-in nodes resulting from high system loss.\(^5-9\) The IBM/Agilent Terabus program\(^10-12\) is a complete optical system-interconnect solution consisting of optical transceivers, package-to-board connectors, and waveguide technology that can be integrated with standard FR4 (flame-retardant) board-manufacturing processes. However, Terabus links are point-to-point, and the polymer-waveguide technology’s relatively high losses and numerical aperture make it extremely difficult to add a broadcast capability with a reasonably high fan-out.

### Novel low-cost technologies

We used two novel technologies to realize the optical bus: hollow metal waveguides (HMWGs)\(^13\) and nonpolarizing pellicle beam splitters (NPPBSs). HMWGs have several interesting properties that make them ideal candidates for use in intraboard interconnections:

- low propagation loss (< 0.05 dB/cm),
- fabrication ease,
- low numerical aperture (NA < 0.01), and
- an effective index (\(n_{ef}\)) of ~1, which yields zero skew between waveguide channels and the lowest latency (0.033 ns/cm) per unit length.

Figure 3 shows the properties of HMWGs for the lowest order, low loss, \(EH_{11}\) mode. Many other higher order, low loss modes exist inside the HMWGs. The HMWGs are air-core light pipes with a square cross section and a high reflectivity coating in the interior. The metallic walls behave as a high index layer, guiding the light in the hollow pipe’s interior cross section. The low loss modes can be described by

<table>
<thead>
<tr>
<th>Tap no.</th>
<th>Beam splitter reflectance</th>
<th>Transmissivity of cascaded taps</th>
<th>Total transmittance (dB)</th>
<th>Loss (dB)</th>
<th>Total loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.96</td>
<td>0.55 0.72 0.80 0.85 0.89 0.91</td>
<td>-6.9</td>
<td>-4.2</td>
<td>-11.1</td>
</tr>
<tr>
<td>6</td>
<td>0.45</td>
<td>0.72 0.80 0.85 0.89 0.91</td>
<td>-7.5</td>
<td>-3.6</td>
<td>-11.1</td>
</tr>
<tr>
<td>5</td>
<td>0.28</td>
<td>0.80 0.85 0.89 0.91</td>
<td>-8.1</td>
<td>-3.0</td>
<td>-11.1</td>
</tr>
<tr>
<td>4</td>
<td>0.20</td>
<td>0.85 0.89 0.91</td>
<td>-8.7</td>
<td>-2.4</td>
<td>-11.1</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>0.89 0.91</td>
<td>-9.3</td>
<td>-1.8</td>
<td>-11.1</td>
</tr>
<tr>
<td>2</td>
<td>0.11</td>
<td>0.91</td>
<td>-9.9</td>
<td>-1.2</td>
<td>-11.1</td>
</tr>
<tr>
<td>1</td>
<td>0.09</td>
<td></td>
<td>-10.5</td>
<td>-0.6</td>
<td>-11.1</td>
</tr>
</tbody>
</table>
rays that bounce back and forth at near grazing incidence to the metal walls as they propagate down the waveguide. To excite the low-loss modes in the HMWGs, we use a lens to collimate the input beam (NA < 0.02) and control the launch angle to within ±0.5 degrees of the waveguide axis. We accommodate a lateral xy misalignment tolerance of ±15 microns by collimating the input beam to ~100 microns.

The HMWG's low NA allows the insertion of taps or beam splitters into the HMWG with little excess loss. We do this by inserting the beam splitters at 45 degrees into slots or gaps cut into the waveguides. Figure 4 shows a Monte Carlo ray trace simulation of the excess loss dependence with gap spacing between two 150-micron square HMWGs.

Figure 3. Properties of hollow metal waveguides (HMWGs) for the EH11 mode: waveguide cross section and propagation characteristics (a) and electric field intensity profile (b).

Figure 4. Calculated excess loss dependence with gap spacing between two 150-micron square HMWGs.
loss as a function of gap separation between two 150-micron core HMWGs. A 1-mm gap introduces an excess loss of only $-0.34$ dB.

Marcatili and Schmeltzer estimated the HMWG’s theoretical loss for a circular cross section. For square waveguides whose core dimension $a$ is much larger than the wavelength $\lambda$, we approximate attenuation by

$$\alpha = \left(\frac{\lambda^2}{2a^2}\right) \frac{\operatorname{Im}(1 + n_{\text{clad}}^2)}{2\sqrt{1 - n_{\text{clad}}^2}}$$

where $n_{\text{clad}}$ is the refractive index of the metal cladding layer. This expression shows that we can make the loss small by choosing the waveguide dimension $a \gg \lambda$. The cubic dependence of the loss limits the HMWG’s useful core size. For a 150-micron square waveguide with silver cladding ($n_{\text{clad}} = 0.15 + 5.68$), the theoretical loss is approximately $-0.0015$ dB/cm at 850 nm. A 50-micron cross section HMWG would have 27 times more loss. We can estimate the modal dispersion of these waveguides simply with the ray model. The difference in propagation delay between the on-axis ray and the grazing ray (indicative of higher-order modes) is

$$\Delta T = \frac{L}{c} \left(\frac{1}{\cos(\theta_{\text{max}})} - 1\right)$$

where $L$ is the propagation length, $c$ is the speed of light, and $\theta_{\text{max}}$ is the maximum angle between the ray and the waveguide axis. For a path length of 30 cm and an NA of 0.01 ($\theta_{\text{max}} \approx 1.1^\circ$), the differential delay is $\approx 20$ femtoseconds. Also, because the temperature dependence of the refractive index of air is about 1 ppm/°C, the propagation delay of these waveguides will change approximately 1 part in $10^4$ for a 100°C change in temperature. These properties make HMWGs a good choice for intra-board chip-to-chip interconnections that can scale with bandwidth.

The optical bus’s other essential component is the optical tap (the direct analog to an electrical stub in a transmission line). We’ve described the reflectivity required for each tap. Moreover, the taps should introduce little or no perturbation to the optical mode in the waveguide. Pellicle beam splitters less than 10 microns thick are best because they minimize etalon or ghosting effects (from back reflection) and, more importantly, beam walk-off. Beam walk-off is due to the optical beam’s refraction as it exits the beam splitter’s finite thickness, resulting in the beam’s displacement from its original path. A portion of the displaced beam can miss the waveguide, introducing loss. The beam walk-off at 45 degrees for a 100-micron-thick BK7 glass beam splitter is about 30 microns.

An NPPBS is necessary because of the multimode VCSEL’s random polarization state. This is due to the VCSEL’s circular symmetry and isotropic gain profile. Figure 5 shows the construction of an NPPBS and its insertion in the HMWG. NPPBSs consist of a free-standing membrane on which is deposited a multilayer optical coating. We choose a pellicle size much larger than the waveguide core so that the pellicle requires only precise placement at a 45-degree angle along the waveguide axis. We place the complexity on the optical coating to yield the correct reflectivity and transmissivity. The multilayer coating covers a thin (250 nm) supporting membrane of Si$_3$N$_4$ on silicon. The layer indexes and thicknesses yield a
desired reflectivity and transmissivity for both s and p polarizations at a 45-degree angle over a spectral bandwidth of 40 nm. We release the pellicle by patterning a hole on the backside of the Si wafer and etching the wafer in potassium hydroxide (KOH), resulting in a free-standing ~2-micron-thick pellicle beam splitter. An antireflection coating on the backside is unnecessary because the Si3N4 film is part of the coating stack.

Figure 6 shows reflectance as a function of wavelength for both the s and p polarizations of an 11-percent beam splitter designed for a 45-degree angle. We produced the layer thicknesses with a commercial optical thin-film design program, Software Spectra’s TFCalc (www.sspectra.com). The reflectances for both polarizations are matched to within less than ±0.5 percent over a 40-nm wavelength range centered on 850 nm. To relax the VCSEL requirements, we need a 40-nm wavelength spectral bandwidth. The designs require accurate control of the indexes and thicknesses of optical layers and the base SiN film. The optical coating’s residual stress must also be controlled so that the released pellicle remains flat. We can relax the accuracy needed to achieve the correct beam splitter ratio by providing a larger power budget per tap.

Construction of the optical bus
The optical bus consists of a 12.5-cm-long 1 × 4 array of HMWGs with a set of optical taps spaced at 1.4 cm. A 1 × 4 850-nm VCSEL array plus driver IC on a Zarlink evaluation board serves as the laser source. A corresponding 1 × 4 GaAs PIN optical receiver (~20 μW sensitivity) on a separate Zarlink evaluation board is the receiver. An Agilent Parallel BERT (bit error ratio tester) drives the VCSELs at a data rate of 5 Gbps. We use an Agilent Infinium DCA (digital communications analyzer) to display the receiver’s output. We fabricated the 1 × 4 HMWG array with a 1-mm pitch on a suitable substrate material (silicon) using a dicing saw to create 150-micron square channels 12.5 cm long. We accommodate insertion of the optical taps by cutting 0.8-mm-wide by 0.59-mm-deep triangular slots into the waveguides with a specially designed dicing saw blade. The blade is designed to create a reference 45-degree (±0.2 degree) surface on which the pellicle
A beam splitter is mounted. Figure 7 shows a drawing and a photo of the HMWG.

The pellicle’s clear aperture measures 0.5 mm x 4 mm, wide enough to encompass all four waveguides. For ease of handling, the finished pellicle remains in a 14-mm-long, 1-mm-wide, and 0.25-mm-thick Si frame. After defining the channels and slots, we deposit a high-reflectivity silver coating and a dielectric protective layer over the whole substrate, coating the channel walls. A thin, gold-coated metal sheet with 0.95-mm-wide and 4.6-mm long openings at the tap spacing serves as the cover layer. The openings allow light to be coupled out from the beam splitters into the receiver. In the first slot of the bus, a 100-percent mirror rotated 90 degrees with respect to the optical taps serves as the input coupler to the bus. The last slot in the bus also contains a mirror, which couples all the remaining light out of the waveguides. After inserting the mirrors and taps, we clamp the two pieces of the HMWG together in a metal frame. We then mount the finished waveguide assembly on the demo board together with the optic and Tx and Rx evaluation boards. Figure 8 shows the demo board assembly with the HMWG and the Tx and Rx evaluation boards.

A three-lens bulk optic reduces the VCSEL beam’s NA from 0.26 to 0.065, a 4x magnification. At this magnification, the 1 x 4 VCSEL array with a 250-micron pitch is imaged to four spots spaced 1 mm apart, setting the waveguide pitch. The nonoptimal collimation (NA ~ 0.065) of the input beam results in a higher propagation loss, limiting the number of taps on the bus. A higher magnification would make the bulk optic large and the waveguide spacing unwieldy.

We measured the propagation loss through the 12.5-cm-long waveguide with nine 0.8-mm-wide slots (but with no mirrors or taps inserted) at about -0.1 dB/cm. Straight waveguides without slots exhibit propagation losses of approximately -0.05 dB/cm. These numbers include the coupling loss from the 500-micron ball lens into the HMWGs. This corresponds to an excess loss per gap of approximately -0.15 dB, assuming that each gap introduces the same loss. This value is lower than the -0.27 dB predicted by ray tracing. The dicing process yields a relatively rough surface on the channel walls, suggesting that lower losses (and lower cost) are possible with smoother finishes from injection-molded plastic parts.

We inserted a turning mirror consisting of gold-coated silicon at the Tx side and four pellicle beam splitters, each consisting of a single 250-nm-thick layer of SiN, into the optical waveguide starting from the last tap. The reflectivity of the single-layer film is polarization sensitive, with \( R_p \) approximately 23 percent and \( R_s \) almost 5 percent. This large difference in reflectance is due to the 45-degree proximity to the Brewster, or polarizing, angle (63.4 degrees) of SiN. At the Brewster angle, the \( p \) polarized light passes without attenuation, whereas the \( s \)
polarized light is totally reflected. This discrepancy and the bulk optic’s poor coupling efficiency caused an uneven power distribution among the taps. As a result, only two taps had enough power to create an open eye.

To increase the coupled power into the waveguide, we replaced the transmitter board with a single VCSEL collimated by a 500-micron ball lens, which produced a collimated beam with an NA < 0.02. The overall coupling efficiency into the waveguide increased from 50 to 67 percent. Figure 9 shows the results. We coupled an input power of 1.3 mW into the waveguide. The outcoupled power from the four nonoptimal SiN beam splitters measured 164, 110, 84, and 60 μW, respectively. At the waveguide output, 450 μW of optical power was left over—enough to yield three to four more taps, assuming we have the correct beam splitter ratios.

Translating the receiver evaluation board under each of the taps resulted in clean open-eye diagrams at 5 Gbps. The speed of the evaluation board electronics limited the data rate. These results demonstrate for the first time a viable optical multidrop bus.

We constructed a 30-cm-long optical fan-out bus in the same manner. It consists of a 30-cm-long 1 × 4 array of HMWGs with a set of optical taps spaced at 3 cm. The HMWGs have a 150-micron square cross section and a 250-micron pitch. They were fabricated by dicing saw on an 8-inch Si substrate to create the 150-micron square, 15-cm-long channels. Two butt-joined 15-cm-long sections form the 30-cm-long optical bus. A microlens array collimates the VCSEL output into the optical bus (NA ~ 0.02). Eight different pellicle beam splitters were inserted in the optical bus. The overall propagation loss was less than 0.05 dB/cm, and the excess loss at each tap was approximately ~0.15 dB. The tap reflectance variation resulted in a tap-to-tap power variation on the order of 3 dB.

To test the optical bus’s efficiency, a VCSEL was directly modulated from an Agilent ParBert and coupled to one of the bus waveguide channels using a ball lens. The tap outputs were coupled into a multimode fiber and detected with a high-speed Agilent optical receiver. Figure 10 shows eye diagrams at 10 Gbps from all eight taps. The taps had 2.1 mW of total laser power. The tap powers, from the first to the eighth tap, were 110, 182, 135, 100, 112, 120, 208, and 109 μW, corresponding to a total system efficiency of 51 percent.
Further applications

A possible application of the optical bus is to enable large disaggregated memory using a $1 \times 8$ fan-out by interconnecting several memory bridge chips together via the optical bus. Figure 11 illustrates this application. The northbound bus sends command, address, and write signals, and the southbound bus sends back read data. The memory bridge chips serve as the interface to a standard DDR3 DRAM interface.

The basic optical bus structure provides a master/slave bus in which one node is responsible for scheduling all transfers. By adding appropriate bus sequencing and arbitration protocols, we can use the optical bus in multimaster configurations. In the current configuration, only the master node can broadcast; slave nodes have only point-to-point communication back to the master.

The characteristics of the optical bus are well suited to parallel links with source-synchronous clocking. The HMWG transmission medium has essentially no crosstalk, and the electronic components are the only source of crosstalk in the system. The propagation delays down the parallel paths are closely matched and stable. Source-synchronous clocking is highly desirable from the perspective of the fan-in bus because it permits rapid bus turnaround between transmitting modules. A source-synchronous optical communication link could have significant advantages in terms of both latency and power. Other possible applications for the optical bus include expansion buses in modular switch architectures, midplanes for blade systems, and I/O systems.

The main impediments to bringing the multidrop bus into wide use are the optoelectronic engine’s cost and reliability and the need to reintroduce support for broadcast protocols in high-speed interfaces.

Future demonstrators will include both northbound and southbound buses running at 10 Gbps. With further optimization, the optical bus will provide improved bandwidth and connectivity with reduced power and lower latency over its electrical counterpart.

Acknowledgments

We acknowledge Scott Lerner, Pavel Kornilovich, Neal Meyer, Robert Bicknell, Charles Otis, Srim Ramamoorthi, and Len Seals at AMS/TDO for their valuable technical contributions; and Lennie Kiyama, Eric Montgomery, Plary Mendoza, and Gil Perusa at HP Labs for their technical support. We also thank Laura King and Stan Williams for their continued support in this project.
References

Michael R.T. Tan is a distinguished technologist at Hewlett-Packard Laboratories. His research interests include surface acoustic wave (SAW), RF, and mm-wave devices and high-speed vertical-cavity surface-emitting lasers (VCSELS) for data communications and laser printing applications. Tan has a PhD in electrical engineering from Stanford University. He is a member of the IEEE.

Figure 11. Application of an optical bus to a large disaggregated memory system. Memory requests are broadcast from the processor to several memory bridge ASICs on the optical bus.
Paul Rosenberg is an optoelectronic engineer at Hewlett-Packard Laboratories. His research interests include advanced optomechanical packaging and thermal design for low cost optical transceivers. Rosenberg has an MS in materials science and engineering from Stanford University.

Jong-Souk Yeo is a research scientist at Hewlett-Packard. His research interests include laser micromachining, nanoscale devices, integrated photonics, flexible electronics, and displays. Yeo has a PhD in materials science and engineering from Stanford University.

Moray McLaren is a distinguished technologist in the Exascale Computing Laboratory of Hewlett-Packard Laboratories. His research interests include the impact of nanophotonics on future computer architectures with a focus on high-speed networking and memory architectures. McLaren has an honors degree in microelectronics from the University of Edinburgh.

Sagi Mathai is a researcher in the Information and Quantum Systems Lab at Hewlett-Packard Laboratories. His research interests include parallel optical interconnects and nanowire optoelectronics. Mathai has a PhD in electrical engineering from UCLA. He is a member of the IEEE Photonics Society and the Optical Society of America (OSA).

Terry Morris is an HP Fellow at Enterprise Storage and Servers. His research interests include high-speed interconnections for computer systems ranging from GaAs supercomputers to blades.

Huei Pei Kuo is a master scientist in the Information and Quantum Systems Lab at Hewlett-Packard Laboratories. His research interests include off-chip optical interconnect technology. Kuo has a PhD in applied and engineering physics from Cornell University.

Joseph Straznicky is a researcher in the Information and Quantum Systems Lab at Hewlett-Packard Laboratories. His research interests include molecular electronics and optical communication. Straznicky has a Dipl.-Ing. in electrical engineering from Technical University Brno, Czechoslovakia. He is a member of the IEEE.

Norman P. Jouppi is an HP Fellow and director of HP’s Exascale Computing Laboratory. His research interests include computer memory systems, including stream prefetch buffers, victim caching, and development of the CACTI tool. Jouppi has a PhD in electrical engineering from Stanford University. He is a Fellow of the ACM and the IEEE.

Shih-Yuan Wang is a distinguished technologist at Hewlett-Packard Laboratories. His research interests include VCSELs, photodetectors, modulators, GaN LED, nanophotonics/electronics, and nanowire devices. Wang has a PhD in electrical engineering and computer sciences from the University of California, Berkeley. He is a Fellow of the IEEE and the Optical Society of America.

Direct questions and comments to Michael Tan, Information and Quantum Systems Lab, Hewlett-Packard Laboratories, 1501 Page Mill Rd., MS 1123, Palo Alto, CA 94304; mike.tan@hp.com.

For more information on this or any other computing topic, please visit our Digital Library at http://computer.org/csdl.