Progress in Low-Power Switched Optical Interconnects

Ashok V. Krishnamoorthy, *Senior Member, IEEE*, Keith W. Goossen, *Senior Member, IEEE*, William Jan, Xuezhe Zheng, *Senior Member, IEEE*, Ron Ho, *Senior Member, IEEE*, Guoliang Li, Richard Rozier, Frankie Liu, *Member, IEEE*, Dinesh Patil, *Member, IEEE*, Jon Lexau, Herb Schwetman, Dazeng Feng, Mehdi Asghari, Thierry Pinguet, and John E. Cunningham

(Invited Paper)

***Abstract*—Optical links have successfully displaced electrical links when their aggregated bandwidth–distance product exceeds ∼100 Gb/s-m because their link energy per bit per unit distance is lower. Optical links will continue to be adopted at distances of 1 m and below if link power falls below 1 pJ/bit/m. Providing optical links directly to a switching/routing chip can significantly improve the switched energy/bit. We present an early experimental switched CMOS-vertical-cavity surface-emitting laser (VCSEL) system operating at Gigabit Ethernet line rates that achieves a switched interconnect energy of less than 19 pJ/bit for a fully nonblocking network with 16 ports and an aggregate capacity of 20 Gb/s/port. The CMOS-VCSEL switch achieves an optical bandwidth density of 37 Gb/s/mm^2 even when operating at a modest line rate of 1.25 Gb/s and is capable of scaling to much higher peak bandwidth densities (∼350 Gb/s/mm^2) with 5–10 pJ/switched bit. We also review a silicon photonic system design that will lower link energies to 300 fJ/bit, while providing multiterabits per second per square millimeter bandwidth densities. This system will ultimately provide switched optical interconnect at less than a picojoule per switched bit and computer/router system energies of tens of picojoule per bit. We review progress made to date on the silicon photonic components and analyze an energy and bandwidth–density roadmap for future advances toward these goals.**

***Index Terms*—CMOS, optical interconnects, routers, silicon photonics, switching, vertical-cavity surface-emitting lasers (VCSELs) very large scale integration (VLSI).**

I. Introduction

COMPUTING and communications performance must continue to scale to meet the growing proliferation of capability and reductions in price expected from systems manufacturers. Indeed, the past three decades have witnessed over five orders of magnitude improvement in data processing performance-per-dollar, in large part due to the investments and improvements in very large scale integration (VLSI) technologies. However, for next-generation systems, designer focus has turned to also reducing the power of computation and communication.

Today, the cost of computing, switching, and routing equipment is no longer dominated by its acquisition, but rather by its infrastructure: the costs of its powering, operation, maintenance, and administration. As a simplistic example, a $10 K high-performance switch that depreciates over five years costs $2 K per year. If it consumes 1 kW, at an average urban cost of 12/kWh, it would cost over $1 K per year to power, assuming a 100% efficient data center. Because data centers are often 50% (or less) power efficient, this switch's operation cost is already comparable to its acquisition expense. Furthermore, at 1-kW consumption, and thus heat dissipation, and a cooling coefficient of performance of 2, it requires another 0.5 kW to operate the air-conditioning. Because the performance–cost ratio of data switching and routing systems must increase to satisfy the exponential growth in networking high-bandwidth applications, reducing power per unit of switched or routed information will become the paramount design challenge we face. This financial pressure, added to environmental concerns related to power generation and usage, as well as the difficulty and cost of building power plants will necessitate a significant reduction in the energy per switched, processed, and/or routed bit.

Unfortunately, in switches, trends in energy per processed bit have not followed the exponential improvements in performance/cost discussed earlier. Today, individual switching chips consume about 20–40 pJ/switched bit.[1] The dominant energy consumers in these switches are not the switching circuits themselves, which can be implemented with <1 pJ/switched bit (for a simple crosspoint switch), but rather by the electrical I/Os of the chips. These I/Os transfer data to and from the VLSI switching chips through a multilevel interconnect package hierarchy that can include several electrical interconnect interfaces. For instance, two chips on separate boards might need to communicate through a chip, package, board, connector, backplane, connector, board, package, and chip. In addition, switching systems

Manuscript received April 18, 2010; revised August 4, 2010; accepted August 12, 2010. Date of publication December 5, 2010; date of current version April 6, 2011. This work was supported in part by the Defense Advanced Research Project Agency Microsystem Technology Office under Agreement HR0011–08-09–0001.

A. V. Krishnamoorthy, X. Zheng, G. Li, and J. E. Cunningham are with Sun Laboratories, Oracle, San Diego, CA 92121 USA (e-mail: ashok.krishnamoorthy@oracle.com).

K. W. Goossen is with University of Delaware, Newark, DE 19716 USA.

W. Jan is with Bridge Semiconductor, Pittsburg, PA 15235 USA.

R. Ho, F. Liu, D. Patil, and J. Lexau are with Sun Laboratories, Oracle, Menlo Park, CA 94025 USA.

R. Rozier is with Hovey Williams, LLP, Kansas City, MO 64106 USA.

H. Schwetman is with Sun Laboratories, Oracle, Austin, TX 78727 USA.

D. Feng and M. Asghari is with Kotura Inc., Los Angeles, CA 91754 USA.

T. Pinguet is with Luxtera Carlsbad, CA 92011 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTQE.2010.2081350

[1] As an example of a commercially available switch chip; see, for instance, Vitesse VSC3144 Crosspoint Switch. http://www.vitesse.com/products/download.php?fid=4349&number=VSC3144

1077-260X/$26.00 © 2010 IEEE

typically concatenate multiple stages (typically, 3–5) of chips to build a large switch, multiplying the required energy/switched bit commensurately. One simple way of reducing total system power is to maximize the size of the switch in each chip and reduce switch chip count, thereby cutting the number of power-hungry chip-to-chip electrical links needed in the system. However, the size of switch that can be implemented on a single chip is already limited by both the area required by I/O drivers and the chip I/O pin count. Hence, system designers cannot easily scale down system power in a manner commensurate with improvements in chip processing capability.

Photonics can help. Optical links directly integrated into switch chips would remove the need for electrical chip I/O. This, of course, is beneficial only if dense photonic integration—beyond what is expected with scaled VLSI technology—is possible and if the integrated photonic links consume less power than electrical ones. We believe both requirements can be met. In Section V, we discuss examples of such technologies and present results from a prototype vertical-cavity surface-emitting laser (VCSEL)-on-CMOS switch that accomplishes these objectives. Such a photonics-based approach also allows the size of the switch on the VLSI chip to be maximized, improving power, system complexity, and chip area savings.

The aforementioned discussion relates to simple data switching applications. The energy/bit efficiency for packet-based data routers is far worse. The energy per routed bit is not only much higher today, but also will likely increase. The largest, most efficient commercial internet routers consume several nanojoules per bit.[2] This is because routing functions are far more complex than a simple crosspoint switch and involve much more involved processing and data movement to and from memory. This is exacerbated when the percentage of small packets that the router must individually handle increases. Given this complexity and the tremendous progress in fiber transport, a single fiber can now carry more data than the largest router can effectively process. A detailed discussion of routers is beyond the scope of this paper but can be found elsewhere [2], [3]. In [3], the authors show that router capacity increases 2× every 18 months, and that multistage, multichip fabrics are needed to optimize router scheduling. This doubling in router capacity comes with a 1.4× increase in router power each generation [2]. While both papers recognize the potential benefit of optical technologies, particularly passive [microelectromechanical system (MEMS)-based] switches, it is evident that the energy per routed bit will be bounded from below (even with zero-static power switches) by transceiver energy efficiencies.

Continuing the earlier discussion, suitably low-power silicon-photonic links can reduce the communication energy between chips comprising the router fabric and simultaneously minimize the I/O costs to and from the router.

Even greater power savings from silicon photonic links can be had if processing, interconnect, and memory functions can be optically enabled and brought in close proximity to each other to create a multireticle-sized "chip" [4] that can house a significant amount of switching, processing, routing, and stacked memory functions. Such a compressed system would not only reduce the interconnect latency between chips, but also lower the power of the interconnected chips. In Section VI, we review the architectural and packaging concepts of a "macrochip," which provides dense, low-power links between chips placed close to one another. This arrangement can increase the density and lower the power dissipation of any existing optoelectronic (OE) system by combining all of the aforementioned benefits: direct areal photonic I/O to chips, low-area, high-density wavelength multiplexed silicon photonic links, and a dense collection of functional chips closely packed on an optically interconnect silicon substrate. Such a platform may ultimately provide the lowest power dissipation for a complex system such as an electronic internet router.

For this vision to eventually become a reality, a number of silicon photonic and packaging advances are anticipated. We motivate and review progress made to date on the silicon photonic link component technologies in Section VII. In Section VIII, we analyze a roadmap for future advances in low-power optical interconnects. Specifically, we review progress toward ultralow power CMOS-compatible silicon photonic transmitters and receivers, with the potential to provide <1 pJ/switched bit when directly integrated with switching chips. This is significant because it allows an order of magnitude reduction in energy per switched bit without compromising any other aspect of the electronic switching fabric including its data rate, reconfigurability, or switching time. Again, this also permits optimization of the number and complexity of switch ports per chip based not only on I/O limits but also on compute and memory capacity paving the way for complex computing and/or routing systems to be incorporated into this macrochip.

II. Penetration of Optics Into Communication: 100 Gb/s-m Distance–Bandwidth Product

Over the past 30 years, optical data links have been increasingly adopted into communications systems. Fig. 1 shows the commercial penetration of optics versus the link distance and the bandwidth. The approximate trend, across five orders of magnitude of bit rates and distance, shows that over the past several decades, optical links have achieved bandwidth-reach product performance exceeding 10^6 Gb/s-m with single-mode solutions and 10^3 Gb/s-m with multimode fiber solutions. On the electrical side, copper-link solutions, including digital subscriber line (DSL), asymmetric DSL, and other twisted pair solutions, have generally followed the 10 Mb/s-m performance trend. A general conclusion supported by this data is that the minimum successful technology crossover from electrical to optical links, from a commercial perspective, can be represented by a bandwidth–distance product of 100 Gb/s-m. There are only a few exceptions to this trend, such as early long-range electrical coaxial solutions from 100 Kb/s to 1 Mb/s deployed in the 1970s. On the optical side, exceptions include avionic and automobile applications, where other properties of optical cables, such as weight, electromagnetic interference emissions or susceptibility, can drive adoption at a

[2] An example of an internet router is the Cisco CRS-3 carrier routing systems. http://newsroom.cisco.com/dlls/2010/prod_030910.html

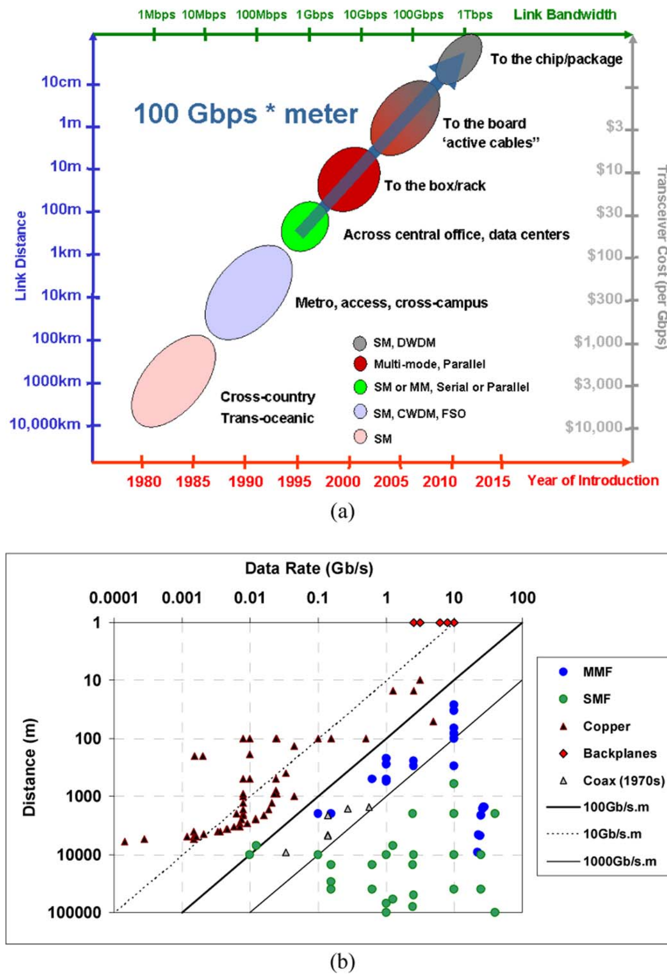


Fig. 1. Penetration of optical links into communications. (a) Historical roadmap for introduction of optical interconnections into digital systems showing link distance versus link bandwidth, approximate cost per unit gigabit per second of the link and approximate date of commercial introduction [1]. (b) Data rate versus distance of commercial electrical versus optical links (data after [46]). Electrical links are bounded by the 100 Gb/s-m bandwidth-distance product.

reduced performance level. Projecting this trend line to higher bandwidths and shorter distances, we begin to approach the traditional sweet spot for commoditized serial intrasystem electrical links.

Active optical cables were recently introduced as a new form of transceiver that includes OE components in the end points of the cable; therefore, the user could plug the active optical cable into an existing socket intended for an electrical transceiver. Power for the optical transceiver is provided from the connection to the socket. This eliminates the need to redesign the system for an optical part, and also enables new optical technologies to compete with existing optical and electrical technologies in a manner transparent to the user. As we approach the sub-10-m distance range, fiber solutions have become embedded into active cables, which make system adoption easier because the electrical and optical cables share a common signaling and connector interface. Widely used active optical cables today include infiniband cable solutions at 4× double data rate (4 channels at 5 Gb/s/channel), 4× quad-data rate (4 channels at 10 Gb/s/channel), and 12× quad-data rate (12 channels at 10 Gb/s/channel). Active cables will scale by adding more channels and increasing the data rate, potentially pushing down the minimum adoption of the optical links as determined by the 100 Gb/s-m trend line shown in Fig. 1.

Continuing this trend suggests that during the next five years, optical links can be expected to reach right into the chip-scale package on a printed circuit board (PCB), representing link distances below 1 m. Although some may view this prediction as a radical departure from current design norms, we will support the view that the required technologies are well under development. Central to this case are optical links that show a power dissipation benefit over electrical solutions. Although specific technology choices will depend on bit rates, loss metrics, and application performance requirements; in the following section, we examine a bit-rate-independent metric to establish a general guideline for optical interconnect adoption.

III. Opportunities for Photonics in Extremely Short-Reach Interconnect: Bit-Energy per Unit Distance Metric

Ultrashort reach interconnects are dominated by electrical links, primarily due to the relative ease of designing and building many electrical interconnect drivers and receivers on a silicon CMOS VLSI chip. In addition, serial-link circuit techniques have been developed to overcome signal attenuation and other effects of electrical channels (including bandwidth limits due to the skin effect, dielectric losses, and impedance discontinuities from vias and connectors) that have typically limited signal bandwidth on PCBs. Hence, the continued scaling of VLSI technologies to finer lithography and larger silicon wafers has driven the density and bit rates of such high-speed electrical signaling up and simultaneously driven down the cost per gigabits per second.

Today, we confront power delivery as another design driver. In the past, electrical power delivery and thermal and cooling aspects of system design were relegated to the end of the system design cycle. Electrical power and thermal engineers were thus forced to find acceptable power delivery and cooling solutions in a highly constrained environment with little ability to alter the logical system design or to influence the design of the microchips or of the system packaging.

This is changing. The advent of 100 W+ superprocessors means that over 100 A of current must be delivered with low-resistive losses and low inductance to a single-chip socket in a PCB. Hence, including power delivery and thermal considerations early in the design of the chips is critical. Similarly at the system level, power densities of equipment racks approaching 10 kW/ft^2 make thermal considerations central to the system designer. Finally, as power utilization becomes ever more expensive and total compute capacities become power constrained, customers and end users are forced to become aware of total system power dissipation (see Section I).

The power costs for using electrical links at short-range distances (1 m and shorter) center around two aspects of their design. First, electrical circuits employ a number of techniques

to overcome channel losses and transmission nonuniformities, both of which degrade with link distance and bitrate. These techniques include transmit preemphasis, receiver equalization, and potentially sending multiple bits per hertz, and represent a significant cost in power dissipation, design complexity, and silicon area. Material improvements to aid electrical-link performance, such as better dielectrics, blind vias, or improved cabling, also mean unwelcome cost increases. Second, the circuits need to align the transmitted data to a synchronizing signal, such as a clock, at both transmit and receive ends of the link. Clock-data recovery and alignment at the receiving chip, involving acquisition of the data's frequency and its phase relative to a clock, can represent a large fraction of the total power in a modern link.

However, many of these design complexities can be tied directly to the high data bitrate. Data signals have high-frequency content, leading to more channel loss and nonuniformity. Also, clock rates on the link are much faster than system clocks, leading to complex clock generators and very tight timing margins.

The reason electrical links push so heavily on data bitrate is due to the size of I/O solder connectors, at a 0.15-mm pitch on today's chips. Because a single link requires two such solder balls for differential data, plus a fraction of other solder balls for power and ground, a single chip, with most of its solder balls devoted to power supplies, can support only a few hundred such data links. Thus, to get sufficient bandwidth onto and off of chips, electrical links must run at data rates much higher than clock frequencies. Such overclocking of these serial links leads directly to their high power and complexity costs.

Compared to electrical links, optical channels represent two benefits that can be readily exploited. First, an optical channel has substantially greater bandwidth and response uniformity than a comparable electrical channel, allowing designers to avoid the complexity and costs of channel equalization. Second, optical channels using parallel fibers and wavelength division multiplexing (WDM) and small connectors to VLSI chips can dramatically increase the number of discrete channels to a chip. This allows each channel to be run at the most efficient bitrate (not constrained by number of pins or timing complexities), significantly simplifying the clock-data recovery problem and lowering its energy cost.

Because electrical-to-optical conversion technologies are now quite well understood and within the production ability of standard CMOS foundries, the use of optical interconnects in ultrashort reach applications is increasingly plausible. In fact, with suitably integrated optical modulators, lasers, and detectors in standard silicon CMOS circuits, an optical interconnect can be viewed as simply a standard electrical interconnect (when viewing the endpoints) with efficient optical-to-electrical converters (and *vice versa*) designed to minimize the overall cost, weight, and size of each link. Opportunities for such system-level efficiency can span applications from high-performance computing and telecommunications systems to high-volume consumer applications.

As detailed in the following section, commercial optical links today provide lower bit energy per unit distance than their electrical counterparts. While this is true today only for the longer lengths (>10 m) with VCSEL links, it is interesting to note that this advantage of optical links can be exploited in the future even for much shorter links because the communication bit-energy per unit distance of electrical links does not scale down as one approaches chip-to-chip and even short on-chip length scales. Hence, the total interconnect energy can be reduced with photonics for a given system configuration. As shown in Section IV, this power advantage is already being widely exploited in systems links over 10 m, where the bandwidth–distance of the links exceeds 100 Gb/s-m.

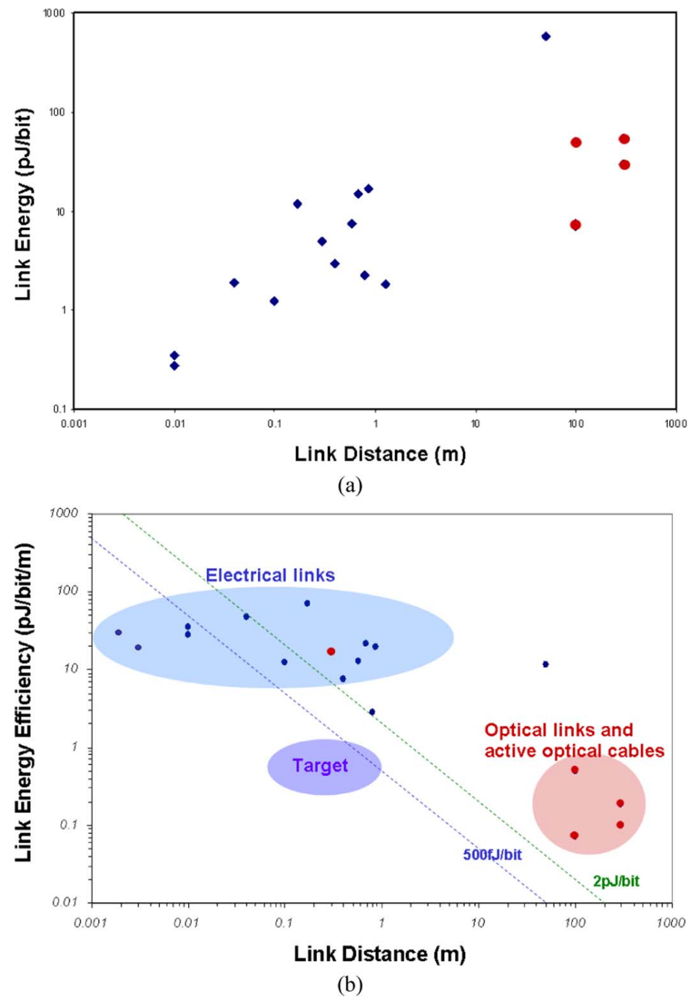


Fig. 2. Performance metrics for commercial and research electrical links (blue diamonds) and commercial optical links (red circles) based primarily on VCSELs and CMOS photonic modulators. (a) Energy per bit versus maximum reach of the link. (b) Bit energy per meter versus the maximum reach of the link. Optical links do not include retiming circuitry. Unlike electrical links, the efficiency of optical links is not highly sensitive to distance. However, the potential exists to lower the energy of optical links to below a picojoule per bit to service link lengths ~1 m.

IV. Scaling to High Bandwidth With Ultralow Bit-Energy Per Meter

The performance metrics for commercial and research electrical links published or deployed in the past five years versus commercial active optical cables based on VCSELs (multimode fiber) and CMOS photonic modulators (single-mode fiber) are depicted in Fig. 2. Fig. 2(a) shows the energy per bit versus the

maximum reach of the link in the 10 cm to 1 km range. Fig. 2(b) scales the energy per bit by the maximum reach of the link. A general trend is that even efficient electrical links typically consume in excess of 10 pJ/bit/m of reach in the commercial space—with some research and upcoming commercial results about 3× better than this in the one-meter range.

Another interesting point is that for on-chip interconnect lengths, the energy per bit is generally in the range of 10–100 pJ/bit/m. For an on-chip electrical wire, energy per bit depends simply on the capacitance of the wire, the voltage swing of the wire, and the power supply. On today's CMOS chips, wire capacitance can range from 0.2 to 0.4 pF/mm, depending on the layout. Assuming 0.3 pF/mm, with a voltage supply and voltage swing of 1.1 V each, and accounting for the energy cost to drive the wire drivers, we get 360 fJ/bit/mm of energy cost to charge up an on-chip wire. The energy cost to drive a data stream on this wire is about one-fourth of this total: half of the time, the wire voltage is constant and does not change, and half of the rest of time the wire is being discharged, which costs no energy. Note that this 90 fJ/bit/mm (90 pJ/bit/m) result is more accurately expressed as "energy per bit time," although it gives a result consistent with traditional link energy metrics. As technologies scale and voltage supplies and swings reduce to 0.7 V, then the switching energy on average reduces to below 50 fJ/bit/mm (50 pJ/bit/m). These numbers can of course be further reduced by lowering the voltage swing on the wire. Such techniques add design and verification complexity, but can reduce the energy per bit significantly. In the absence of receiver offset compensation, transistor mismatch will set a lower limit of voltage swing to around 100 mV. In practice, differential wiring is needed to mitigate noise effects; therefore, doubling wire capacitance, and more complex receivers may slightly increase the circuit overhead. Hence, average switching energy may reduce to 20 fJ/bit/mm in a 1 V technology or closer to 10 fJ/bit/mm (10 pJ/bit/m) in a 0.7 V technology.

Individual and parallel optical interconnects based on VCSELs are widely deployed today in switching and routing systems, local area networks, central offices, data centers, and high-end computing systems in the 1–600 m range. This penetration resulted from many well-known performance and cost advantages of the VCSEL device: its low-drive current, favorable high-speed modulation characteristics, high wall–plug efficiency, wafer-level testability, and its ability to tailor light output to improve coupling to optical fiber. Its low cost and high yield is also evident in its choice as the preferred laser in optical mice.

Active optical cable links provide 5–10 pJ/bit and can be deployed at reaches up to 10–50 m, leading to 0.1 to 1 pJ/bit/m energy metrics even though these optical links do not typically include retiming circuitry, which will increase the power dissipation. By contrast, serial electrical links usually include this function. Thus, simply scaling a VCSEL or CMOS photonic active cable-link architecture to shorter distances would not necessarily produce an optical link competitive (on an energy per bit per meter metric) with an electrical link, unless the electrical driver and receiver to synchronize to clocked digital logic are also considered.

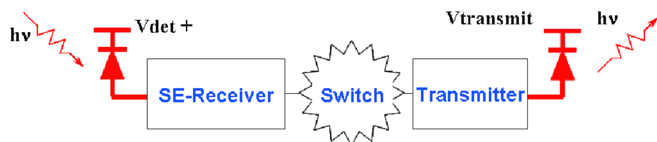


Fig. 3. Single chip with receivers, switches, and transmitters (either VCSELs or low-power silicon photonic devices) can significantly improve the energy efficiency of switch chips to under 1 pJ/switched bit.

Also important are techniques that increase the number of effective optical channels. Today, these techniques make use of dense VCSELs and photodetectors (PDs) integrated with silicon, and tomorrow will use silicon devices and waveguides with WDM to increase the number of data channels per optical waveguide. Just as the need to share amplifiers was one of the drivers for WDM in long-distance transmission, the need to share fiber terminations, connectors, and to increase the interconnect density will be the drivers for WDM in the data center. Silicon photonic interconnect, by virtue of its abilities to achieve dense integration, promote electrical–optical symbiosis, utilize wafer-scale silicon manufacturing, and achieve WDM or other efficient encoding techniques, appears to be well positioned to ultimately drive to sub-picojoule per bit per meter links for scalable, extremely short distance switched interconnects (see Fig. 3).

V. Energy-Efficient VCSEL-Based OE Switch

Today's VCSEL device has broad industry acceptance, in part due to its ability to be directly connected to electrical circuits at the chip and wafer levels and directly modulated at speeds in excess of 15 Gb/s [5]–[9] with 25 Gb/s devices being actively pursued. Arrays of VCSELs have been bonded directly to CMOS VLSI chips, with each VCSEL capable of up to 15 Gb/s modulation by the CMOS circuits with wide-open eyes.

Here, we describe a process of bonding VCSELs and detectors to a chip that involved two flip-chip attachments and substrate removals; the result was two interleaved 16 × 16 arrays of high-speed detectors and VCSELs, with a total of 512 devices attached in an area of less than 2.2 mm × 3.8 mm (see Fig. 4). This effectively integrated a transmitter array, a receiver array, and switching circuits capable of 320 Gb/s (256 × 1.25 Gb/s) with an areal bandwidth density exceeding 37 Gb/s/mm^2. The substrate removal of the devices was important to reduce their thermal crosstalk and to allow the high-speed lasers to be attached to the circuits with a linear pitch of 144 microns. Interleaved devices allowed a compact crosspoint switch design as described in the following, and completed a hexagonal close-packed device structure that represented the most compact spacing manufacturable for the multimode-fiber bundle that interfaced the surface–normal optical I/O to the chip. The detectors used light-blocking hemispherical contacts [see Fig. 4(b) and (c)] to avoid optical crosstalk from neighboring VCSELs and attached to CMOS transimpedance receiver circuits for optical-to-electrical conversion.

The arrays employed 850 nm bottom-emitting VCSELs developed by Avalon Photonics (now Bookham, San Jose, CA)

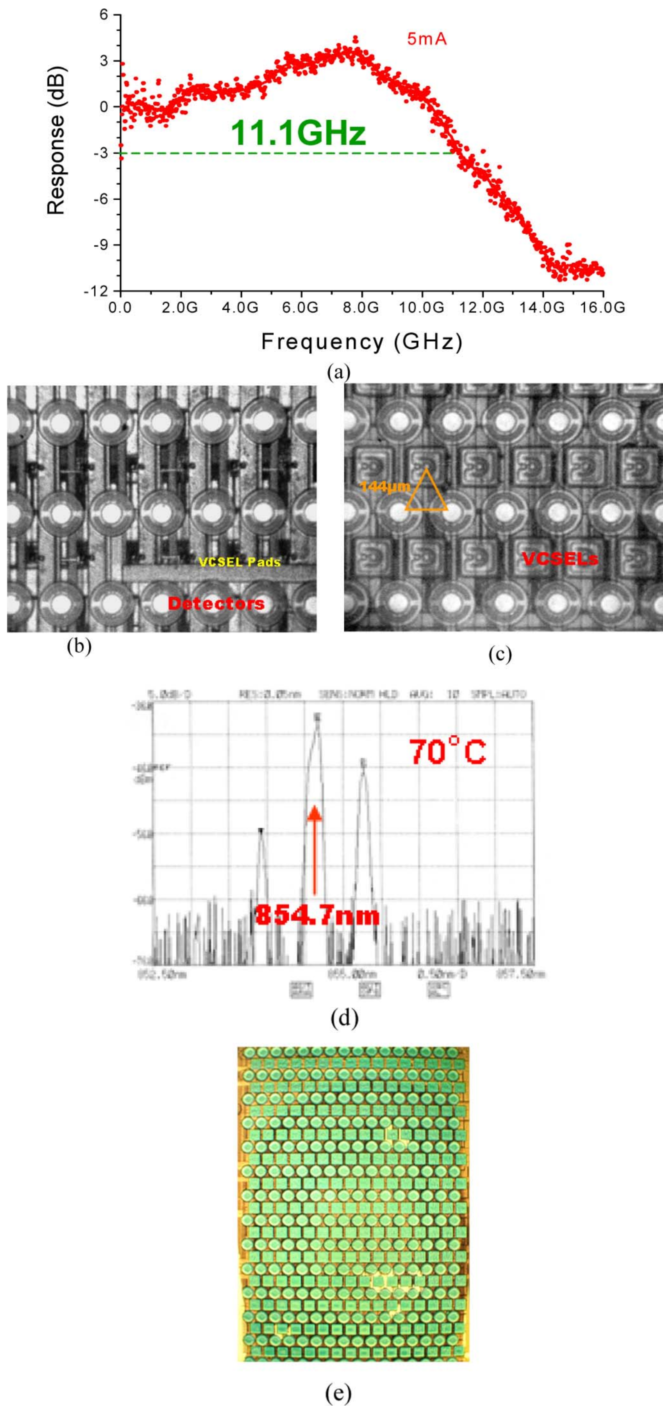


Fig. 4. (a) S12 measurement result of small signal bandwidth of an early flip-chip bonded, substrate-removed, bottom-emitting, 850 nm VCSEL when bonded to a silicon tester die. (b) Micrograph of a CMOS switch chip with an array of 256 850 nm detectors (circular mesas) bonded to the chip followed by substrate removal of the detector array substrate to expose secondary bonding pads on the CMOS chip. (c) Array of 256 VCSELs (square mesas) subsequently bonded to the CMOS switch chip followed by substrate removal of the VCSEL array substrate to create dual-bonded, interleaved 850 nm VCSELs and detectors hexagonally close packed on the CMOS switch chip with a center to center spacing of 144 microns. (d) Flip-chip bonded VCSELs were multimoded and operated with a center wavelength of 854.7 nm at 70 °C. (e) Microphotograph of the array of 512 optical devices after dual-bonding, dual-substrate removal, and antireflection coating on the CMOS chip occupying an area under 2.2 mm × 3.8 mm.

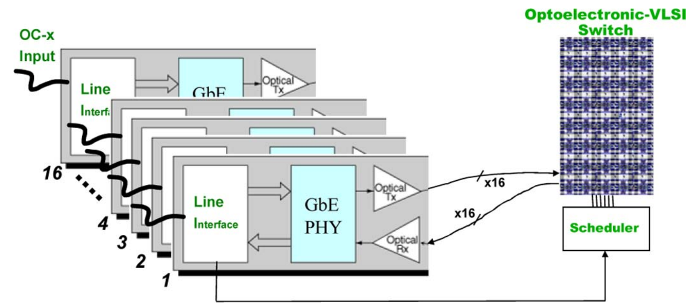


Fig. 5. System concept for the OE switch. Ethernet line interfaces with VCSEL-based parallel optics interfaced directly to the integrated CMOS-VCSEL switch.

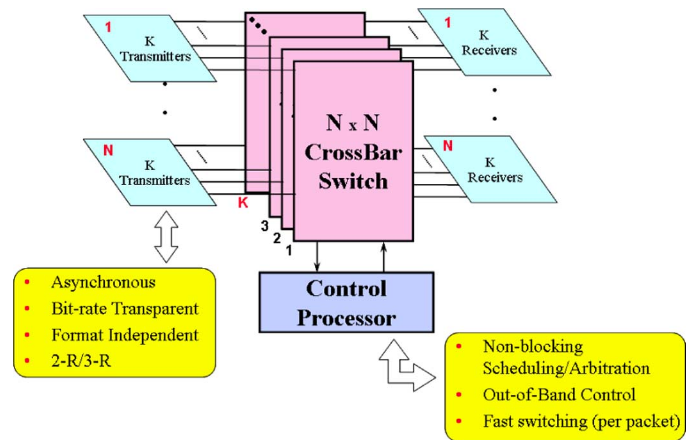


Fig. 6. Logical view of the OE switch; $K = 16$ copies of an $N \times N$ crossbar ($N = 16$) switch comprised the switch fabric. The switch could operate in an asynchronous mode or in a digitally clocked mode. All switch control was performed out-of-band. High-speed electrical I/O to the chip was used to control the switch settings. The 4 × 6 mm switch chip was built in a 250 nm CMOS technology (circa 1999) and contained approximately 100 K transistors.

with a threshold current less than 1 mA. They produced over 1 mW of output power at 5 mA of current, with just over 2 V of forward-bias voltage across the diode at room temperature. The effective slope efficiency of the device (at room temperature) was approximately 10%. The VCSEL also produced approximately 0.5 mW of optical power under these same bias conditions at a substrate temperature of 80 °C. In the following, we describe an early switching system that utilized this photonics-on-VLSI technology to create a very low power switching chip.

This OE-VLSI chip consisted of 16 independent channels, each with 16 optical input and 16 optical output lines. Electrical nonblocking crossbars, one per channel, provided the switching capabilities for the system (see Fig. 5). These 16 crossbar switches together created a fully nonblocking 16-port switch, with each port providing 16 × 1.25 Gb/s or 20 Gb/s per port, as shown in Fig. 6. The entire switch thus had a total of 256 inputs and 256 outputs. An array of registers (not shown) placed at the output channels synchronized the switch's operation. The switch, designed and implemented in Lucent's 250 nm CMOS process, contained 100 K transistors and consumed 4 × 6 mm including all digital circuits and electrical I/O pads.

The chip's power dissipation had three major contributors: 6 mW × 256 = 1.54 W optical receivers; 12 mW × 256 = 3.1 W VCSELs plus drivers; and a 1.4 W switching fabric. A total maximum throughput of 320 Gb/s was possible with this chip. Several chips and system prototypes were assembled. Although the detectors routinely had 100% yield, the best OE-VLSI chip had one nonfunctional VCSEL so that actual total throughput was slightly less than 320 Gb/s. Gigabit Ethernet data traffic at a relatively low bit rate of 1.25Gb/s was transmitted through the switch. Even so, the system provided a switched energy of 19 pJ/bit including the optical input and VCSEL outputs. As shown in Fig. 4(a), each VCSEL was capable of supporting 10x faster baud rates.

As mentioned earlier, the optical I/O and the switching circuits were contained in an area 2.2 mm × 3.8 mm. This area was bounded by the compactness of the fiber-to-fiber spacing in the bundle, and not by the following circuits. While the transceiver itself offered 37.5 Gb/s/mm^2 of bidirectional I/O density, when we include the area for the switching fabric, clocking, control logic, and peripheral power pads of the 4 × 6 mm chip, we achieved a switching bandwidth density of over 13 Gb/s/mm^2. This transceiver bandwidth density could be greatly enhanced by increasing the bit rate of the VCSELs, which were capable of over 12.5 Gb/s [see Fig. 4(a)], even though the switching circuits in this chip were not. On this basis, we project that a transceiver-limited I/O bandwidth density of over 350 Gb/s/mm^2 is possible with bit rates above 10 Gb/s per VCSEL.

The switching fabric was built using digital multiplexers (muxes). Each of the 16 output channels used a 16b mux to select among the 16 input channels, using four binary-coded select lines. The 64 total select lines were controlled externally and driven onto the chip electrically.

In order to equalize the switch latency for all inputs, each 16-to-1 mux was constructed from four stages of identical 2-to-1 muxes. The first stage used eight muxes, the second stage four muxes, the third stage two muxes, and the fourth stage a single mux. A compact 2-D switch layout was enabled by the interleaved detectors and lasers, by inserting the mux stages in between input optical receivers and output VCSELs. The optical channel spacing was sufficient to fit the muxes for each output channel in close proximity. This also enabled us to minimize skew between channels and between ports.

The task of packaging the system involved multiple optical alignments. The fiber bundle and optical packaging is shown in Fig. 7. A 512-fiber bundle with hexagonal close packing at a 144-micron pitch was the densest that could be reasonably manufactured at this bundle size. On the other end, the bundle was terminated into 64 ribbons of eight fibers each. A 50-micron graded-index core fiber represented a good optimization of the tolerances between optical output to the detectors and optical input from the VCSEL array at the fiber bundle face. The detectors presented a more critical alignment concern, but we used 60 micron active area detectors so that the array alignment was relatively straightforward: over 10 microns of x–y displacement could be tolerated, while collecting most of the light.

The switch chip was mounted directly to a PCB that provided power, fiber alignment and attach, and I/O interfaces. Fig. 8 shows a picture of the board after being populated with the OE-VLSI chip. This board was then placed on a copper block, with a heat sink attached to the copper plate. Because of the relatively low power of the entire switch, active cooling was not needed to operate the system (see Fig. 9). The face of the fiber bundle was actively aligned to the chip using four corner VCSELs, and then, attached to the PCB. The packaged switch chip, PCB, and fiber bundle were then placed into a standard 19″ rack-mountable enclosure with a face plate containing dual 64 ribbon connectors of fiber connectors. Power was routed through cables in the back of the box, as were control signals for the switch (see Fig. 9). The scheduler was implemented externally and was not placed in the box for the prototype.

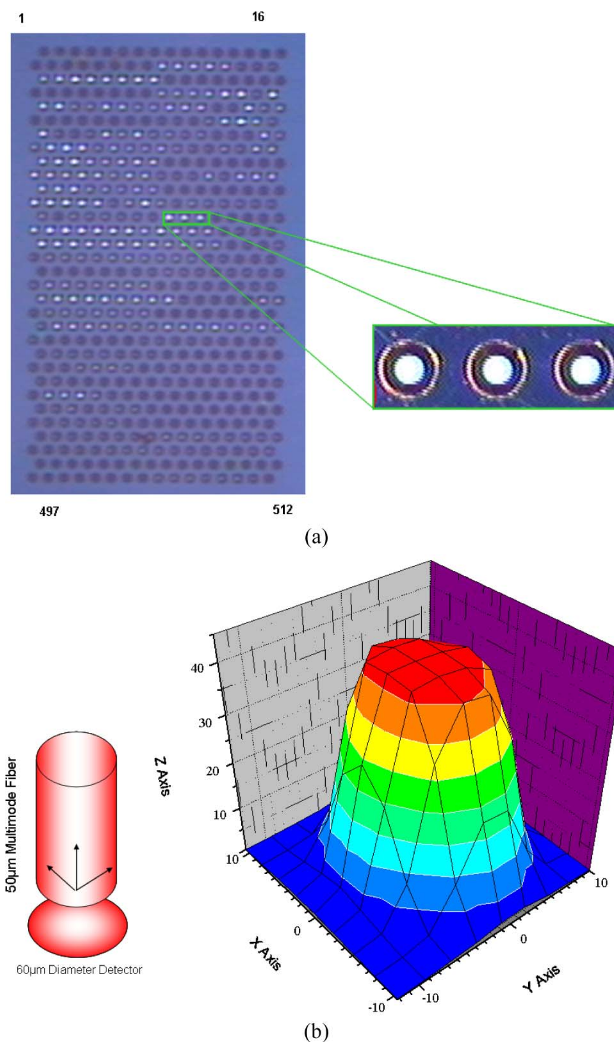


Fig. 7. (a) Fiber bundle containing 512 multimode fibers. Each fiber was a 50-μm core graded index fiber. Fibers were hexagonal-close packed with 144 μm pitch, matching the optical device array. The fiber bundle was terminated into 64 8-fiber ribbon multimode connectors. (b) Critical alignment to the chip was determined by the optical input to the flip-chip bonded detectors on the chip. To reduce alignment sensitivity, each detector had a 60-μm diameter active area. This enabled the fiber bundle to be placed up to 40 μm away with a lateral tolerance of 10 μm with less than 0.5 dB loss. This relaxed alignment tolerance, even at very dense pitches, is one of the key benefits of VCSEL-based multimode photonics.

Fig. 10 shows optical data propagated from a line card, through the switch, and out the VCSEL array. Optical input to the switch was provided by electrically modulating a gigabit-ethernet module, which was coupled through the fiber bundle

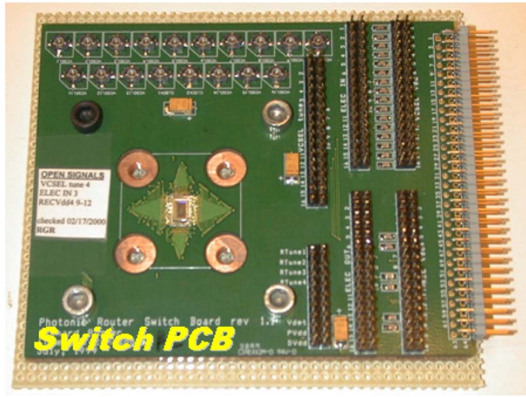


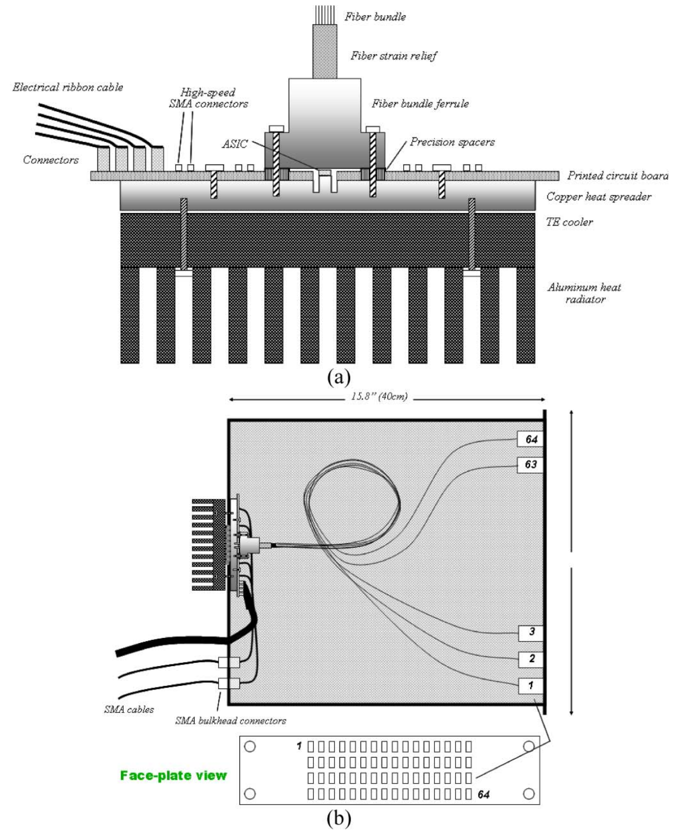
Fig. 8. Photograph of the switch PCB and connector—all data I/O was optical, and all crossbar scheduling and control was electrical.

to the appropriate on-chip detector and receiver. The data was passed through the selected switch fabric path, and then, sent to the corresponding output port on-chip driver and VCSEL. The output waveform of the modulated VCSELs were directly measured using an external multimode analog receiver: 1) output at 1.25 Gb/s through the switch at approximately 18.5 pJ/bit, and 2) clocked operation with retimed data through the switch at half rate (625 Mb/s). In both cases, the bit error rate (BER) was measured below 10^{-13} with error-free operation during a full day of testing. The switching speed was also demonstrated at less than 5 ns, as depicted in Fig. 10(c), more than adequate for the system to be operated as a high-speed packet switch.

VI. Silicon Photonics and the "Macrochip"

So far we have focused on integrating optical interconnect to the chip, and assumed that conventional waveguides or fibers would connect chips to their distant data origins and destinations. The ability of optics to improve bandwidth and energy at long distances compared to electrical links was particularly appealing for certain system applications, such as circuit switching, or other applications that did not require significant functionality beyond a single chip or package. But another opportunity exists: to bring close together large collections of chips and stitch them with very high bisection bandwidth to provide high levels of system performance. If done right, this can provide very high system performance and reduced latency across a large-scale multiple-chip system with fundamentally low-energy consumption [4]. This will be appealing not only for computing systems but also to a large class of packet switching and packet routing applications that require the complexity of a microsystem containing many chips.

The "macrochip" is a logically contiguous piece of photonically interconnected silicon that integrates CPUs, memory, and a system-wide interconnect to provide significant advantages in computational density, energy efficiency, bisection bandwidth, and reduced message latency over traditional integrated multi-chip systems. Optical proximity communication (OPxC) connects the CPUs distributed on the macrochip to optical routing layers; this support low-latency, WDM optical links between sites using silicon-on-insulator (SOI) photonic waveguides (see



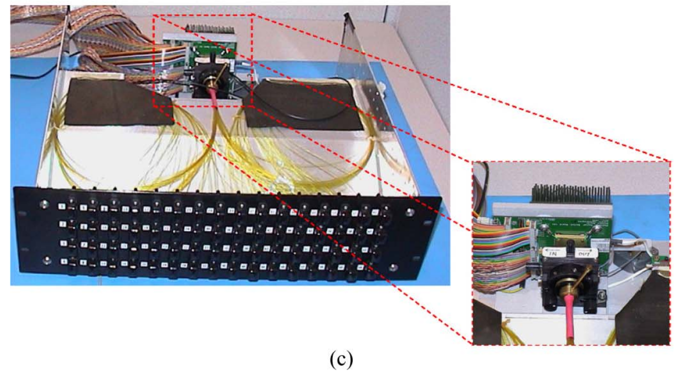


Fig. 9. (a) Schematic of packaged OE switch. (b) Schematic of prototype switch in a rack-mount unit. (c) System photogram showing fiber connectors, bundle, and packaged switch. Scheduler was implemented in software and was not placed in the box for the prototype. Inset shows *in situ* alignment system for the fiber-bundle attached to board.

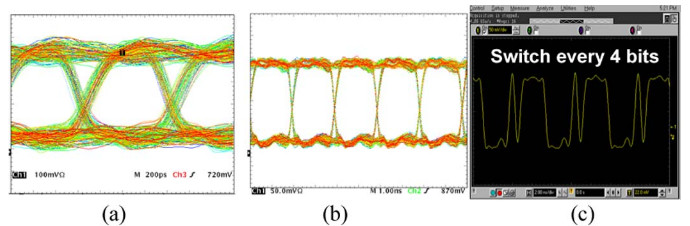


Fig. 10. Operation of the OE-VLSI switching system: (a) Optical-in, optical-out data transmitted asynchronously through the switch at 1.25 Gb/s. (b) Optical-in, optical out data clocked through the switch at the maximum clock speed of 625 Mb/s limited by the on-chip clock distribution circuitry. (c) Channel 1 output when switching between two separate inputs: channel 0 and channel 1. Input channel 0 data pattern: 0001; input channel 1 data pattern: 1110. The switching of 5 ns was achieved for every 4 bits at the data rate of 1.25 Gb/s.

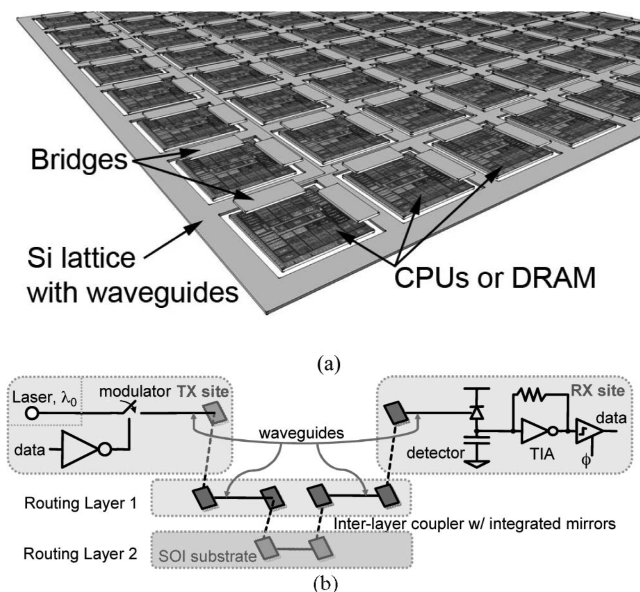


Fig. 11. (a) Macrochip concept and (b) canonical representation of a macrochip WDM link.

Fig. 11). This interconnection network provides point-to-point, nonblocking links for every pair of CPU or memory sites on the macrochip and allows system designers to mix, match, and replace processor or memory die in a modular fashion. A macrochip enables dense single-node integration of processor and system interconnects with highly connected topologies (such as all-to-all connections), offering breakthrough performance per unit of power and floor space. A system can be as small as a single macrochip, or can combine many macrochips in a multinode system with a densely connected fiber network.

One of the macrochip's key advantages over conventional multichip modules is its use of optical proximity coupling with WDM waveguides. This provides high-density optical communication and permits the chips to be replaceable. This means that the chips can be properly tested in the system and replaced if faulty—solving the uncertainty of knowing if the constituent chips in an array are good before packaging them together; aka the "known-good die" issue that limits the yield of large arrays of chips. Additionally, the macrochip is an all-silicon package, which has many advantages related to powering and cooling of the chips. It uses planar packaging, which allows delivery of power to the chip faces and removal of heat from their backs. The multilayer substrate holding the chips, unlike conventional packaging materials like ceramic, are outfitted with passive SOI waveguides, optical proximity mirrors, and v-groove couplers to interface to fibers. In addition, silicon can be chemically and mechanically treated to create precise alignment structures to locally and globally align chips. Chip-size holes may also be etched into the macrochip substrate to enable multichip stacking at each site. Finally, the planarity of the silicon substrate can be controlled to a much finer level than possible with other multichip packaging materials or methods. These are some of the qualitative benefits; a complete discussion of macrochip packaging is beyond the scope of this paper, but will be visited in the future.

The photonic network on a macrochip provides low power, high bandwidth, and high-density communication between sites. Every site in a macrochip is interconnected to every other site via WDM links that run in orthogonal directions on two routing layers. The optical signals from the sites are coupled into, and between, the routing layers through OPxC face-to-face optical couplers. We anticipate a bisection bandwidth on the order of 10 TB/s traversing the macrochip.

Although a number of different network topologies may be possible for site-to-site routing on the macrochip, a preferred on-macrochip photonic network uses a point-to-point topology with static WDM routing. Each waveguide from a site connects to a different column of sites. A source site connects to a destination site by first selecting the waveguide connected to the destination column, and then, selecting the wavelength within that waveguide to address the site (row) within that column. This point-to-point network offers low-optical loss between any two points by avoiding splitters and reducing the number of optical components. It also incurs no setup delays and provides the highest bisection bandwidth for a fixed number of transmitters and receivers. The network is transparent to data rate and communication protocols, and exploits the low latency, high density, and long reach benefits of optical technology. While several network topologies have been considered for the macrochip, a point-to-point interconnect topology appears to be well suited to this system architecture.

A key requirement for the macrochip is that the energy cost of optical communication must be dramatically reduced from that in present day systems. Optical links will not replace electrical links in such dense systems unless their per-bit energy costs are much lower. Indeed, the simple all-to-all optical network, which provides an overprovisioned total network bandwidth in order to eliminate network switching, is largely predicated on a very low per-link power cost.

For a macrochip in the 2015–2018 timeframe, we expect optical macrochip links with a 20 Gb/s channel data rate and an on-chip power dissipation under 300 fJ/bit. This number includes the on-chip photon loss, but excludes the electrical-to-optical laser conversion efficiency and is discussed further in Section VIII. We also expect the worst case (or longest) link loss of an 8×8 macro chip to be about 17 dB. No amplifier is needed (or can be afforded given current amplifier power efficiencies), and we assume a 0 dB·m laser source and an optical receiver with –21 dB·m sensitivity. A minimum of eight-channel WDM components are assumed at a 200 GHz channel spacing—scalable to a larger number of wavelengths per channel as needed. In the following, we review components that will allow such WDM optical links to be intimately integrated and comanufactured with the switching electronics and concurrently lower link energies to a few hundred femtojoule per bit. Such a system could achieve switched optical interconnect at $\leq$1 pJ/bit, provide waveguide transport bandwidth densities in excess of 10 Tb/s/mm, and achieve surface normal optical transmission as needed from these waveguides at a density up to 100 Tb/s/mm^2,

with a communication I/O density ultimately limited by the area requirements of these WDM-link components.

VII. Recent Progress in Low-Power Photonic Interconnects

As shown in Fig. 11, a macrochip WDM link consists of an off-chip laser source, a transmitter, a receiver, and WDM optics components including mux/demux, waveguides, and OPxCs. The total on-chip link power consumption can be expressed as follows.

$$\mathrm{Power}^{\mathrm{on\text{-}chip}}_{\mathrm{dissipation}} = P_{\mathrm{receiver}} + P_{\mathrm{transmitter}} + P_{\mathrm{WDM}} + P_{\mathrm{optical\,loss}} \quad (1)$$

Here, the total power represents the effects of the receiver, transmitter, WDM mux/demux, and the optical (photon) loss on the macrochip. To achieve energy-efficient links, high-performance WDM passive components with low-insertion loss and low-tuning power, low-power transmitters with low-insertion loss, as well as low-power receivers with high sensitivity are needed.

A. Ultralow Power Transmitter Using Reverse-Biased Ring Modulator

A low-power transmitter requires three critical elements: an energy-efficient modulator, a low-power driver circuit, and intimate integration of these two components with low parasitics. We optimized all three components to develop ultralow power silicon photonic transmitter.

Although monolithic integration is ideal for low-parasitic integration, photonic devices, and circuits presently have contradictory requirements on substrates. We chose hybrid integration to allow the optimization of the photonics and circuits on separate substrates without compromising either one. We employed a flip-chip integration technique modified for silicon-to-SOI chip integration, using microsolder bump technology [4]. In order to retain the performance advantages gained from separate optimization of the CMOS circuits and the silicon nanophotonic devices, we reduced the parasitics to a total pad + bump capacitance estimated between 20–25 fF.

A depletion ring modulator was chosen for its potential to achieve high bandwidth, low-power modulation with a CMOS compatible drive voltage. To tradeoff device stability versus modulation quality, we developed a device with a "medium" Q measured to be about 8300 using the Luxtera-Freescale 130-nm SOI CMOS process. The device exhibited superior stability and had a total capacitance of about 50 fF. It achieved a 3 dB electrooptic (E/O) bandwidth of 15 GHz, a wavelength shift of 10 pm/V, and an extinction ratio (ER) of 3 dB with 6 dB insertion loss with a voltage swing of 2 V [10].

We built a reliable low-power driver using a 90-nm CMOS process. A cascode voltage driver was designed to issue an effective voltage swing up to 2 V across the device while maintaining a voltage no more than 1 V across each transistor [11], at a data rate up to 5 Gb/s.

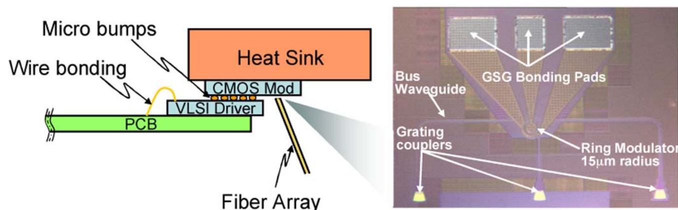


Fig. 12. All-CMOS transmitter consisting of a CMOS-photonic ring modulator built in the Luxtera-Freescale 130 nm SOI CMOS process flip-chip integrated with a 90-nm CMOS driver circuit [10].

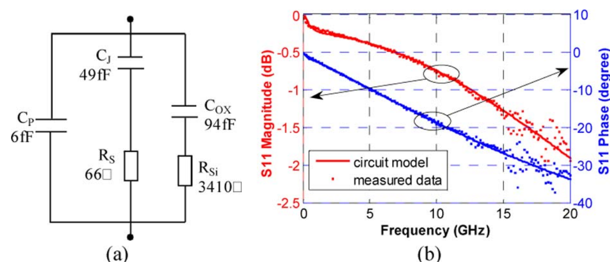


Fig. 13. (a) Small-signal circuit model for reverse-biased ring modulators built on a SOI substrate. The marked parameter values are for 0 V bias. (b) Curve-fitting of the measured S11 at 0 V using the circuit model in (a).

The optimized ring modulator and low-power driver circuits were then integrated in a diving-board configuration using the low-parasitic microsolder bumps by flip-chip bonding. The assembled chips were then die-attached and wire-bonded to a test board for high-speed testing, as shown in Fig. 12. A simple copper heat sink was used in direct contact with the modulator chip to maintain the thermal stability together with controlled air flow to the test setup. Using an off-the-shelf receiver to receive the modulated optical signal, error-free transmission was achieved at 5 Gb/s in over 3.5 days during which more than 10^{15} bits of data were transmitted. The total transmitter electrical power consumption was only 2 mW, yielding an energy efficiency of 400 fJ/bit [10].

It is possible to further optimize the ring modulator for better modulation quality. This was accomplished by Kotura using its in-house photonic foundry. The device maximized the E/O modulation by using an optimized waveguide design and an asymmetric p–n junction design, with a 5×10^{17} cm^{-3} p-doping concentration, a 1×10^{18} cm^{-3} n-doping concentration, and a junction offset of 50 nm to maximize the mode overlap with the depletion region for maximized index change [12].

A complete device model includes parasitic capacitance and impedance beyond the junction capacitance. We studied the high-speed behavior of the ring modulator using a circuit model extracted by curve fitting the measured S11 data. In this circuit model shown in Fig. 13(a), C_P represents the capacitance between the electrodes (mostly due to the contact pads) through the top dielectrics and the air, C_J denotes the capacitance in the reverse-biased diode junction, R_S denotes the diode-series resistance, C_{OX} denotes the capacitance through the dielectric and Si layers, and R_{Si} is the resistance in Si layer. The parameter values in Fig. 13(a) were extracted at 0 V, indicating a junction capacitance of 49 fF. The curve-fitting result is shown in Fig. 13(b).

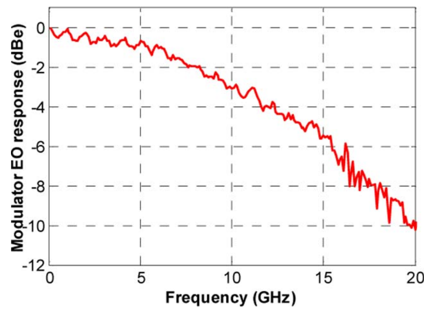


Fig. 14. Measured frequency response for the 15-μm radius ring modulator biased at 0 V.

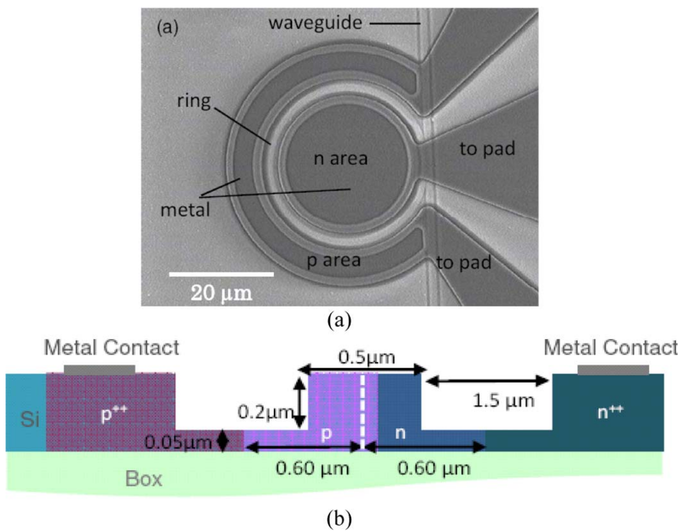


Fig. 15. (a) Structure and (b) micro photograph of a custom silicon photonic ring modulator [12].

The same data analysis at 1 and 2 V bias indicated a slightly smaller junction capacitance, consistent with our expectation. Using the extracted circuit model, we calculated the modulation energy of the ring modulator at ∼100 fJ/bit (twice the junction modulation energy) when driven by a pseudorandom data with a 2 V swing at 5 Gb/s (note that $C_{\rm OX}$ can be partially charged and discharged during the modulation). This switching energy can be readily lowered by reducing the ring modulator size.

The small-signal modulation bandwidth of the ring modulator was measured using a microwave network analyzer and a reference detector with known frequency response [13]. The measured result in Fig. 14 indicates a 3 dB bandwidth of 10 GHz at 0 V, with the wavelength set at ∼1550.6 nm. The modulation bandwidth of the device is subject to both the *RC* limit, which is ∼24 GHz estimated from the circuit model (with a 50 Ω source), and the photon lifetime limit, which is ∼11.4 GHz based on the measured quality factor. Hence, the photon lifetime represents the primary limit to the modulation bandwidth for this device.

The resulting device was a compact ring modulator with high-modulation bandwidth sufficient for 15 Gb/s modulation, low-voltage swing of 2 V, high ER (∼7 dB), and low-optical loss (∼2 dB at on-state). This device (shown in Fig. 15) was then flip-chip integrated to the modulator driver circuit described earlier.

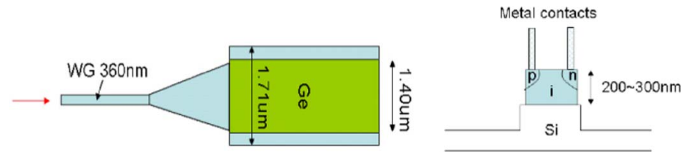


Fig. 16. Structure of high-speed Ge detector compatible with CMOS.

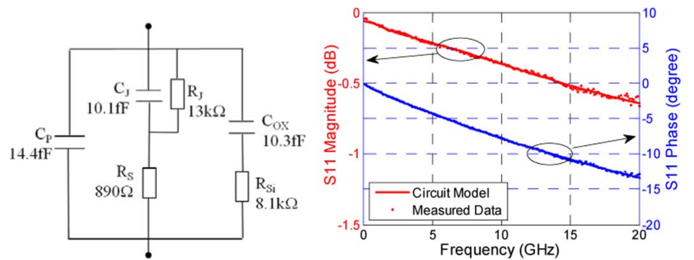


Fig. 17. (a) Small-signal circuit model for a Ge PD at a reverse bias of 0.5 V. (b) Curve fitting of the measured S11 using the circuit model in (a).

Integrated with the 5 Gb/s CMOS driver, a record low-power silicon photonic modulator-based transmitter was obtained with an on-chip power consumption of only 1.6 mW, or 320 fJ/bit in terms of energy efficiency [13]. Further energy reductions will be possible by scaling the circuit technology node and increasing the operating bit rate to 15 Gb/s and higher.

B. Ultralow Power Receiver Using Ge Waveguide PD

A high-sensitivity, low-power receiver also requires several key elements: a PD with high bandwidth, low parasitics, high responsivity, and low dark current, a current amplifier circuits with low power and low excess noise, and the intimate integration of these two subcomponents with low parasitics. As mentioned earlier, we used different CMOS platforms to make the waveguide PD devices and the associated CMOS receiver circuit.

Ge-p-i-n waveguide diodes were developed for high-speed signal detection. They were fabricated using Luxtera's Ge-enabled OE process integrated in Freescale's HIP7_SOI 130 nm CMOS. Fig. 16 shows the schematics of the Ge waveguide PD design. A 200–300-nm-thick Ge film is grown on top of the Si waveguide. The edges of the Ge strip are doped to form a p-i-n diode and to form the contacts. Light coming in from the Si waveguide is evanescently coupled into the Ge region and converted to photocurrent. To make the Ge fabrication compatible with CMOS process, the Ge film was epitaxially grown using a reduced pressure CVD (RPCVD) technique at 350 °C. The single-step low-temperature Ge growth minimizes the thermal impact on the transistors, but it may result in more defects in the Ge film, which can lead to larger dark current. The fabricated PD achieved a high responsivity of ∼0.7 A/W at 1550 nm that is insensitive to the applied reverse-bias voltage. To lower dark current, we chose a low 0.5 V reverse bias for the Ge PD used in our integrated receiver. At 0.5 V bias, the Ge PDs have a dark current of 1–3 μA, and a 3-dB electrical bandwidth of >10 GHz measured in a 50 Ω system.

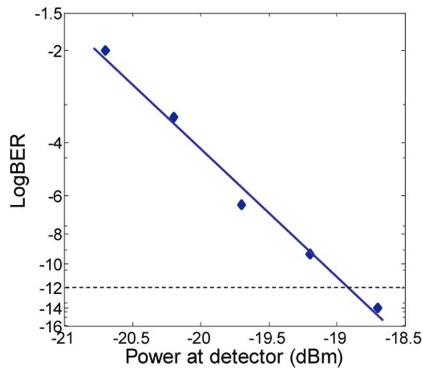


Fig. 18. BER versus received power at 5 Gb/s for the subpicojoule per bit receiver reported in [14].

To understand the details of the high-speed behavior of the Ge PD, and to help the receiver circuit design, we measured the S11 response of the device and extracted a circuit model using curve fitting. The Ge PD circuit model shown in Fig. 17(a) was extracted for 0.5 V reverse-bias voltage, in which resistance R_J was used to model the junction leakage path. The curve-fitting result is shown in Fig. 17(b), from which a very small junction capacitance of 10 fF and a reasonable parasitic capacitance of 14.4 fF were deduced. A small PD capacitance is critical for the receiver circuit to achieve low power and high speed. It is worth noting that this total capacitance is smaller than that of any other monolithic or hybrid receiver technology to date, and has the potential to be reduced even further as detector and waveguide geometries are shrunk. The extracted circuit model also indicates that the Ge PD has a large series resistance, although this did not limit the device speed in our case. Using the circuit model, we can estimate that the *RC*-limited PD bandwidth is $\sim$16 GHz in a 50 Ω system.

The receiver circuit amplifies the small photocurrent from a PD to a correct (error-free) data symbol sequence. It generally consists of a transimpedance amplifier (TIA) front end and multistage of post amplifiers to provide enough transimpedance gain, followed by limiting amplifiers and threshold/decision circuits to convert the data signal back to the digital domain. For low-power considerations, we designed the receiver with a three-stage TIA followed by a slicer, where a simple and power-efficient sense amplifier is used [11]. The total TIA current is about 1 mA and has a bandwidth exceeding 4 GHz. It provides 10 kΩ transimpedance gain, which would converts a 20 μA photocurrent signal to a 200 mV voltage swing. The clocked sense amplifier then subsequently amplifies the TIA output to a full-swing CMOS signal. The transistor mismatch ($1\sigma \sim$10 mV for 90 nm CMOS), and the residual noise in the detector bias path limit the sensitivity of the receiver. The input referred noise current was calculated to be 1.1 μA.

The receiver circuits chip and the PD chip were post processed, and flip-chip bonded together using the microsolder bumps. The hybrid chip assembly was die-attached and wire bonded to a test PCB for power, control, and high-speed digital I/O connections. With a modulated 5 Gb/s optical signal at 18 μA average photocurrent (equivalent to an optical power of –16 dB·m for detector responsivity of 0.7 A/W), we obtained error-free operation for more than 6 h, indicating a BER better than 10^{-14}. Measurements of the receiver BER for different average input power levels, as plotted in Fig. 18, indicates a receiver sensitivity of about –18.9 dB·m for a BER of 10^{-12} at data rate of 5 Gb/s.

A receiver power consumption of 3.45 mW (excluding the power consumed by the digital data buffers, and the on-chip clock distribution) was measured during extended operation by measuring the supply voltages and currents, which corresponds to a receiver energy efficiency of 690 fJ/bit and represents the entire power of the PD and its CMOS receiver, as well as any excess power required to drive circuit parasitics including internal wiring and flip-chip pads [14]. This was the first all-CMOS subpicojoule per bit optical receiver reported and further reductions in energy-per-bit are expected at higher bitrates using lower capacitance detectors and finer linewidth CMOS technologies.

C. Energy Efficiency of Integrated CMOS Photonic Devices Versus Hybrid III–V Devices on CMOS

It is instructive to compare this energy with previous work in low-parasitic III–V modulators on CMOS. One of the simplest methods for measuring the switching energy of a quantum well modulator is to optically address the device. In this sense, the symmetric self-electro-optic effect device [15], [16] was such a modulator. There were several references to ''optical'' switching energy, but few to electrical switching energy. From the optical switching energy measurements, it is possible to go back and estimate the electrical energy.

The expression for the optical switching energy is *CV*/*S*, where *C* is the total capacitance being switched by the device, *V* is the voltage being switched, and *S* is the responsivity in amperes per watt. If the responsivity is not constant as a function of voltage, as in the case, where a multiple quantum well (MQW) modulator is also used as the detector, it is best to use the integral form of the equation to solve for the switching time $v(t) = 1/c$ int ($S^*P(t)$), the 10–90% rise time can be found by inspection from the voltage waveform, and the optical switching energy is merely the integral of the optical power times over time. From this, we can derive the optical switching energy as a function of capacitance, or solve for the capacitance from the measured optical switching energy.

The quantum well modulator in [15] had logic 0 and 1 transmission states of about 12% and 50% at 15 V. The absorbed light is then 88% and 50%, and the average is about 65%, because it takes longer to switch with less current. The maximum possible responsivity is 0.69 A/W for 850 nm optical input (l/1.24), and therefore, the "average" responsivity in the experiments was about 0.69×0.65 or 0.45 A/W.

The measured optical switching energy was about 3.6 pJ, so the capacitance can be estimated to be $C = ES/V$ or 95.3 fF for a 17 V swing, including 1 V for the built in field. Because there were two diodes, the capacitance per diode was 47 fF.

Now, the average electrical switching energy per bit is $\frac{1}{4}\,CV^2$, which can then be calculated as 688 fJ. The reason for the energy to be so large is that the voltage was increased to an unusually high value to meet the minimum bistability requirements.

Two subsequent devices were made with much lower modulation energies. One of these, the diode-clamped device [16], essentially operated at only plus and minus the built-in field for a total voltage swing of 2 V. As a result, the optical switching energies were reduced to 340–580 fJ/bit when measured over almost six orders of magnitude in optical power. We can use the same arguments as earlier to estimate the capacitance at 65 fF per diode, using the most pessimistic number earlier. The diode capacitances were higher, because there was an extra set of electrical diodes and the window size was larger on this device. This device does have an offset field of 3.5 V, so in this case, the electrical energy per bit is estimated to be 347 fJ/bit.

The other device used asymmetric MQWs [17], but the switching energy was never reported. This said the static field of the device was not as large as the static field of the diode-clamped device, so it was plausible that the energy was $\leq$100 fJ/bit assuming the capacitance was similar to the two devices earlier.

Although these devices were quite efficient, it was only after connecting them to CMOS circuits [18] that the integrated transmitter switching energy could be measured. To this end, the size of the quantum well modulators and the flip-chip pads were minimized, and subsequently, connected to CMOS circuits. The integrated CMOS/MQW modulator had a total capacitance of approximately 50 fF [19], competitive with a monolithic silicon photonic modulator. But the required voltage swing was higher (in this case over 2 V). Even so, an early CMOS-MQW transmitter was reported with approximately 1.6 pJ/bit in 800 nm CMOS [20].

In terms of energy-efficient receivers, MOSFET-based TIAs with active feedback have been investigated, since the early 1980s for low-cost and low-area receivers [21]–[23]. Most early work was based on discrete detectors, but efforts to flip-chip integrate III–V compound semiconductor detectors to CMOS TIAs with low parasitics were launched in the early 1990s [24], [25], leading to receiver efficiencies of approximately 3 pJ/bit [26].

Since silicon can also be used as a detector at short wavelengths, the use of Si MOS-compatible detectors has also been widely investigated and is worth reviewing in this context. Early work in this area traded sensitivity for speed to compensate for the long absorption length in silicon [27], [28], but even so, was eventually able to provide gigabit speeds from standard silicon CMOS or SOI foundry detectors monolithically integrated with TIA-based receivers [29], [30]. In [30], a monolithic CMOS receiver and Si detector demonstrated an energy efficiency as good as 1.5 pJ/bit at 1 Gb/s, but required external biasing of the TIA.

Meanwhile, commercial-grade receivers with on-chip biasing and offset correction were also built at gigabit speeds, but resulted in an order of magnitude larger energy per bit [31]. The use of dc-balanced input data, however, allowed novel, lower power circuits to be explored. For instance, TIA-less CMOS receivers with energies less than 2 pJ/bit were recently demonstrated at 1.6 Gb/s [32]. This result used a III–V flip-chip detector and an on-chip bias correction circuit and eliminated the need for external TIA biasing. In [33], Guckenberger *et al.* capacitively coupled a discrete PIN detector to a TIA to significantly reduce parasitic loading and simplify TIA biasing, albeit at a somewhat reduced sensitivity, to achieve a record energy efficiency of 1 pJ/bit at 10 Gb/s.

More recently, evanescently coupled waveguide-based Ge detectors have been incorporated into SOI CMOS [34], [35] to achieve the combined benefits of high-speed operation, good responsivity, and full-CMOS integration. This approach has gained momentum and several groups have demonstrated Ge waveguide PDs with capacitances as low as a few femtofarad [36], bandwidths above 40 GHz [37], and excellent responsivity. Initial receiver work with such Ge waveguide detectors has focused on achieving high-speed operation (10 Gb/s and higher) [38], rather than energy efficiency. The integrated CMOS receiver reviewed in the previous section [14], directly benefits from the small size and low-parasitic capacitance of the Ge detector and achieved the lowest reported energy to date.

In terms of switched interconnects, a notable early example of an OE-VLSI switching chip with over 1 Tb/s total throughput was described in [39]. This single chip contained 1024 optical receivers on the 0.8-μm CMOS IC, operating at a bit rate of 625 Mb/s. The static dissipation through the entire chip was $\sim$ 5 W, which was dominated by the transimpedance receivers. This provided a switched link with an on-chip power of less than 9 pJ/bit not including the off-chip laser power. Although this link required an external laser and hence cannot be directly compared with the 19 pJ/switched-bit VCSEL switching system described in Section V, it is worth noting that in both cases, the dissipation was dominated by the optical transmitters and receivers and not by the switching circuits, which we expect can be reduced to the order of 100 fJ at 15–20 Gb/s. This early example also supports the view that a switch chip with integrated photonics and low-power optical I/O can significantly reduce the total energy costs of communication and switching to below 1pJ/bit—perhaps even lower that a passive optical switch because the direct energy costs of bit switching can be quite low if the optical transceiver power can be suitably reduced. We discuss exactly how much the photonic transceiver energy per bit can be expected to drop in Section VIII.

D. Low-Loss Broadband OPxC

The optical proximity coupler is one of the most important passive optical components in the macrochip. This coupler enables broadband vertical coupling of optical signal between the bridge and routing wafer, and between routing layers. Not only does it enable multiple photonic layers to be seamlessly interconnected, but it also allows the implementation of the complex photonic WDM network on macrochip with no waveguide crossings. Optical proximity couplers can be mirror based or grating-coupler based.

To create a mirror-based coupler, a microreflector was created in SOI using silicon micromachining with an anisotropic wet etch to create (1 1 1) facets that form interplanar angle of 54.7° with the (1 0 0) surface of SOI, with lithographically defined alignment to terminated silicon waveguides with mode tapers at the ends. When two chips with this OPxC structure are placed face to face in alignment, optical signals in the waveguide of one

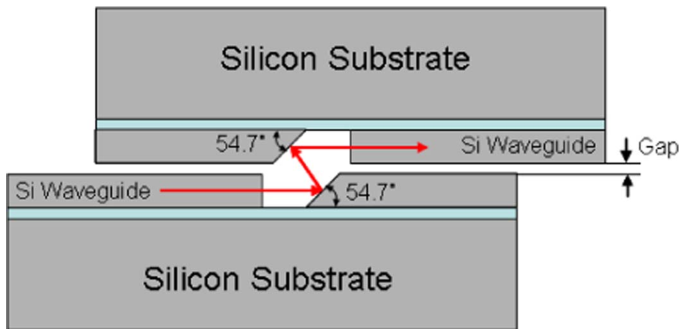


Fig. 19. Concept for OPxC on a macrochip using etched mirrors.

chip are coupled to the waveguide on the facing chip through the pair of the parallel reflectors, as shown in Fig. 19.

Pyramidal pits were cofabricated simultaneously with the couplers. Using the kinetic mating of spherical balls with the pyramidal pits, the optical couplers can be registered together with very precise alignment. With off-the-shelf 400 μm precision sapphire balls and pyramidal pits of matching size, we demonstrated packaged OPxC with a chip-to-chip coupling loss of less than 4 dB over 100 nm wavelength range. Error-free transmission with negligible link penalty was achieved for high-speed data transmission at 10 Gb/s through the interface, indicating the optical coupler essentially appears as a passive loss component to the WDM macrochip links [40]. To enable lower power macrochip WDM links, the coupling loss of the optical proximity coupler must be improved, and the density of the optical transceiver circuits must be improved. Nevertheless, it appears that such couplers can support chip-to-chip optical communication densities in excess of 100 Tb/s per sq. mm.

E. 1 × 4 Tuneable Ring Mux/Demux

Si photonic interconnects require silicon real estate for its signal routing using silicon waveguides. WDM is a natural solution to effectively reduce the number of interconnect waveguides, and consequently, improve the integration density. Ring-resonator-based add/drop filters using high-index contrast Si waveguides have the potential to make very compact mux/demux with desirable optical performance.

One critical hurdle, however, for ring-resonator-based WDM filters is its center wavelength accuracy, as well as the accurate channel spacing for multichannel muxs or demuxs. Due to manufacturing tolerances, the effective index of the Si waveguide varies due to silicon layer thickness variation on the SOI substrate, waveguide width variation, etch depth variation for ridge waveguide structures, as well as the residual stress on the substrate. Each factor increases the required tuning range. In addition, ambient temperature changes also affect the waveguide effective index. All these variations cause significant wavelength shift for ring-resonator-based WDM filters. Although special fabrication techniques have been reported as effective in achieving accurate channel spacing for multichannel devices, dynamic tuning would still be required to align the filter center wavelengths with preselected wavelength channels.

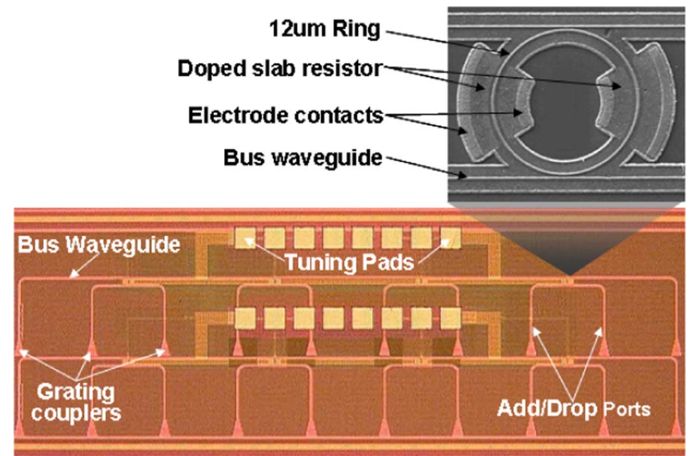


Fig. 20. Fabricated 1×4 mux/demux by cascading four ring add/drop filters with integrated doped resistor thermal tuner (shown in the inset SEM picture) using the Luxtera-FreeScale 130 nm SOI CMOS process. Grating couplers are used for optical I/Os [41].

We designed a four-channel mux/demux using 12 μm radius rings with integrated thermal tuners. As shown in Fig. 20, the doped resistor heaters were integrated directly on the slab of the ring waveguide for high-tuning efficiency. With optimized waveguide and coupling design, the fabricated 1 × 4 ring mux/demux device achieved a uniform low insertion loss of less than 1 dB, 3 dB optical bandwidth of 0.4 nm, and better than 16 dB channel isolation at 1.6 nm channel spacing. A uniform and linear tuning efficiency of 90 pm/mW was also achieved with the integrated resistor tuners. High-speed data transmission at 10 Gb/s showed a power penalty as small as 0.6 dB [41].

Although encouraging, a significant improvement in tuning efficiency was still needed because existing fabrication tolerances result in a maximum tuning range up to a full free-spectral range (FSR). The use of substrate etching [42] and top-side silicon undercut-etching techniques [43] was successfully used to substantially reduce the power to tune an entire FSR of the ring-based mux/demux to only a few milliwatts. This represents over a 10× improvement in ring mux/demux tuning energy, which is an important step toward enabling practical subpicojoule per bit silicon photonic WDM links in the macrochip.

VIII. Scaling Down the Power of OE Transceivers: Roadmap for Low-Power Silicon Photonics

As discussed earlier, low-power WDM optical links require extremely efficient transmitters, receivers, WDM components, and waveguides. In the previous section, we reviewed recent demonstrations of such efficient devices, in particular, a transmitter with 0.32 pJ/bit and a receiver with 0.69 pJ/bit efficiencies, both running at 5 Gb/s.

The low power of these integrated electrical–optical transceiver components results from the intimate integration of optical devices to an electronic logic substrate. The transmitter expends energy to change the charge on a modulator's capacitance, and this energy is minimized when small-dimension resonant rings or electroabsorptive structures are bonded directly to a CMOS driver's output. For the receiver, this hybrid attachment

is even more critical, because it is where the PD's small-output current gets converted into a CMOS-compatible voltage swing, and any parasitic capacitances there directly degrade the performance of the circuit.

Aggressive targets for transceiver energy can be used to drive improvements in devices and circuits, and force designers to rethink system architectures. In this section, we will explore designs for a scaled goal in 28 nm CMOS of 0.3 pJ/bit at 15 Gb/s for the combined transmitter and receiver. The following example assumes approximately 0.1 pJ/bit for the transmitter and approximately 0.2 pJ/bit for the receiver, although other systems, with different optical devices, may employ a different power division.

A. Efficient Transmitters

Highly efficient optical transmitters employ a compact modulator—either a resonant ring or an electroabsorptive device—in reverse bias, so the modulator appears to the driving circuit as a simple capacitance with no shunt current. Its energy cost is proportional to $C^* V^* DV$, where C is the total capacitance, V is the power supply of the transmitter, and DV is the voltage swing across the modulator. To get from 0.4 pJ/bit in a contemporary 90 nm CMOS circuit down to 0.1 pJ/bit in a scaled silicon technology requires scaling the voltage supply and/or the capacitance. Complicating matters is the requirement of certain modulators for a large voltage swing: the 0.4 pJ/bit transmitter described earlier generated a 2 V swing. It therefore employed 1 V transistors in a stacked topology to prevent device breakdown when subjected to the large 2 V supply voltage [11].

One simple solution for scaling energy would be to lower the modulation voltage from 2 to 1 V, resulting in a quadratic reduction of the energy per bit from 0.4 pJ/bit down to 0.1 pJ/bit. However, a closer look at the energy breakdown is instructive. The 0.4 pJ/bit demonstration discussed earlier used a modulator with approximately 120 fF of total load, including bonding pads and parasitic routing. That load was driven by a 2 V supply through a 2 V swing. For nonreturn-to-zero (NRZ) data with a random distribution of zeros and ones, the average energy per bit is $C^* V^* DV/4$, so the modulator itself accounted for 120 fJ of energy. The remaining 280 fJ of energy was consumed in the predriver circuit: about 400 fF of capacitance in a circuit path with a 2 V supply, but a 1 V swing, generating 200 fJ of energy, and another 320 fF of capacitance in a circuit path with a 1 V supply and a 1 V swing, generating the last 80 fJ of energy. Both of these paths had high-capacitive loads partly due to a pulsed circuit topology, and eliminating that pulse driver reduces the predriver's capacitance and energy by a factor of two. The other reason for high-driver capacitance was to enable fast edge rates on a large (120 fF) load, reducing the load would reduce the rest of the transmitter energy linearly.

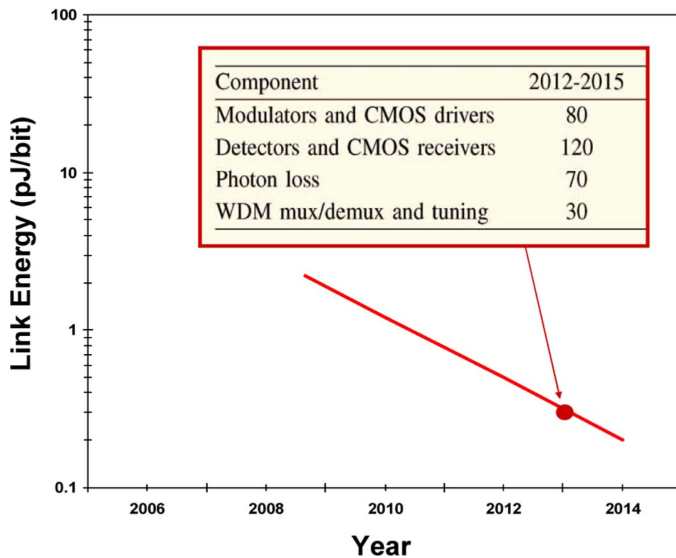
Thus, even without a reduction in modulation voltage in the transmitter, a roadmap for energy reduction from 0.4 to 0.1 pJ/bit could be managed through discarding a pulse circuit topology in favor of a static CMOS design and reducing the total-driven modulator load by about a factor of two. The latter might be accomplished by reducing the bonding parasitics through smaller pads and cutting routing capacitance on the optical device. Any further reduction in modulation voltage would be a direct improvement on the net consumed energy of the transmitter. Because the scaled 0.1 pJ/bit design targets 15 Gb/s operation or higher, it would also require scaling the silicon technology from a 90 to 28 nm CMOS process, and using the improvements in saturation drain current to maintain signal edge rates at the higher bandwidth.

B. Efficient Receivers

Creating efficient receivers is harder than creating efficient transmitters, because receivers convert small photocurrents into CMOS-level voltages and must tradeoff energy consumption, performance, and SNR against each other. Because SNRs correspond to certain BERs, the SNR becomes a primary design target for optical receivers.

With a sufficiently small photodiode capacitance, direct integration of the photocurrent onto the diode capacitor can lead to sufficiently large voltage swings, and thus, obviate a standard receiver [44]. Such systems could reduce or eliminate the electrical power in the receivers using additional optical power. Exploiting this tradeoff would require fully integrating optical devices into a CMOS process technology and minimizing all device and contact capacitance and parasitics to a few femtofarad, which could prove difficult. Additionally, any hybrid integrated optical link with a directly bonded PD, even the low-capacitance devices considered here, is likely to have too much capacitance to make such receiverless systems beneficial. Furthermore, this full integration is impeded by many technology and economic constraints. Hence, in this section, we consider a more traditional TIA receiver and consider its power and performance scaling in the context of required SNR along the lines of [45] with updated assumptions for the silicon photonic components discussed in Section VII. We note that while this paper has thus far assumed a CMOS photonics platform built in an SOI process [35], the discussion on circuit power scaling presented below is generally true for a variety of proposed silicon photonics integration strategies including photonics in bulk silicon [47] and polymer modulators on CMOS [48], each of which can be parameterized by a characteristic capacitance associated with path from the circuit to the optical device.

The circuits considered here generally consist of a TIA with a resistive (shunt–shunt) feedback. The feedback resistor, which provides transimpedance gain, is formed through a transistor. The output of the TIA drives a clocked-sense amplifier, which is strobed by the output of a delay-locked loop (DLL), which periodically acquires phase from the input data stream [49]. A digital engine employs a digital-to-analog converter (DAC) to set the DLL, to properly bias the analog data path, and to cancel the offsets of the sense amplifier. In the following discussion, we present simulation results on a 45-nm CMOS technology running at 10 Gb/s. We do this because the 0.69 pJ/bit, 90 nm CMOS, 5 Gb/s receiver discussed earlier did not include every receiver component. Notably, it omitted the DLL, because the timing margins at 5 Gb/s were wide enough for a shared global clock; and it omitted the sense-amplifier offset compensation, because it assumed a large enough transimpedance gain to



Component	2012-2015
Modulators and CMOS drivers	80
Detectors and CMOS receivers	120
Photon loss	70
WDM mux/demux and tuning	30

Fig. 21. Energy-per-bit targets for silicon photonic links. Inset shows the budget of the individual components of the link. The expectation for a macrochip-wide link (excluding laser efficiency) is to reach the few hundred femtojoule per bit range in the 2015 timeframe [4].

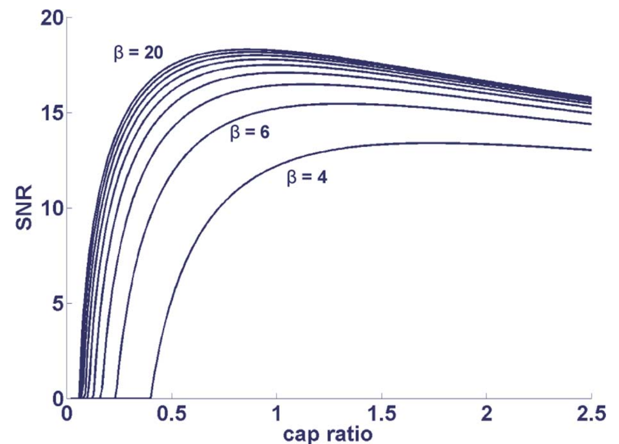


Fig. 22. SNR versus receiver TIA size (input capacitance normalized to the photodiode) for different values of β (the ratio of intrinsic transistor speed w_t to the bandwidth of the TIA). Larger amplifiers increase the power and, beyond a point do not increase SNR. To optimize performance and power, the receiver is designed with an amplifier input capacitance is between 0.5–1.25 the photodiode capacitance, depending on the amplifier bandwidth desired.

overcome six sigma of input-referred voltage offset. Both of these blocks are in the 45 nm simulations.

In a 45-nm design, the energy consumption of the receiver topology can be broken into the TIA portion, discussed in the following, and the rest of the receiver: DLL, sense-amplifier, bias/offset cancellation, and a digital finite-state machine to control them. Simulations show that in a 45-nm 1 V CMOS technology running at 10 Gb/s, the "non-TIA" components consume around 0.18 pJ/bit. Migrating to a 28-nm technology, scaling the voltage down to 0.75 V, and scaling up the DAC resistor sizes allows us to reasonably project that the non-TIA energy could scale down to approximately 0.1 pJ/bit.

For the TIA itself, the total energy is directly proportional to its size. As a result, its energy scaling will be constrained by its SNR: too-small amplifiers will generate too-small output signal levels, suffer from low SNR, and hence, suffer an unacceptable BER. Increasing the amplifier size would improve SNR, but only to a point. Beyond this size (when the amplifier's input capacitance becomes approximately equal to the photodiode capacitance), further increases in amplifier size reduce SNR. This is because increasing amplifier size means that total input capacitance is increasing, and to maintain bandwidth, the transimpedance feedback would need to be appropriately reduced. However, this leads to increased thermal noise from the transimpedance, and hence, a falling SNR. Therefore, a medium amplifier size (whose input capacitance is roughly half the photodiode capacitance) leads to a maximum SNR.

Assuming two thermal noise sources (one from the transimpedance device, and the other from the amplifier stage's output) that integrate over the amplifier's output bandwidth, one can generate a following plot like Fig. 22. The *x*-axis shows amplifier size (normalized to the size of the photodiode), and the different parameters β represent the ratio of intrinsic transistor speed w_t to the bandwidth of the TIA: the lower the *b*, the harder we are pushing the technology.

One can also derive the maximum SNR as proportional to $(\mathrm{BW}^2 \times C_{\mathrm{diode}})^{-1}$, where BW is the bandwidth of the TIA, and C_{diode} is the device parasitic. This relationship shows that if our receiver were optimized for maximum SNR, maintaining it would require C_{diode} to scale by the square of the bandwidth scaling: if speeds double, then C_{diode} would need to be cut by $4\times$. Even if receivers do not operate at a maximum SNR condition (for example, if they are power constrained), C_{diode} is still critical: as process technologies scale and data rates increase, a constant or slow-scaling C_{diode} will require an increasingly larger TIA to hit data rates and SNR, increasing TIA power, and eventually, this strategy will run out of SNR headroom. By contrast, a C_{diode} that scales faster than the data rate bandwidth will allow clear power reduction in the TIA, not only from maintaining an SNR but also from allowing designers to remove stages in the TIA. Simulations show that depending on the scaling trends of C_{diode}, the power of the TIA will scale down to 0.06 to 0.13 pJ/bit, leading to a reasonable projection for 0.1 pJ/bit. Combined with the non-TIA circuits and the transmitter, this gives a total of 0.3 pJ/bit. For the 300 fJ/bit link energy budget proposed in [4], this requires further scaling of transmitter plus receiver energy to 0.2 pJ/bit and leaves approximately 0.1 pJ/bit for the laser power and tuning the WDM components, which represent significant challenges and opportunities for power reduction in these components (see Fig. 21).

Beyond the energy reductions possible with silicon photonics, it is also useful to characterize the density improvements associated with monolithic integration of photonic devices with silicon waveguides and their associated reduction in size to subwavelength geometries because this directly defines the utility of providing optical I/O to silicon VLSI chips. The integration with silicon-based waveguides not only helps the energy and area efficiency of the optical devices but it also significantly improves the density of electrical driver and receiver circuits because of the reduction of device capacitance and other

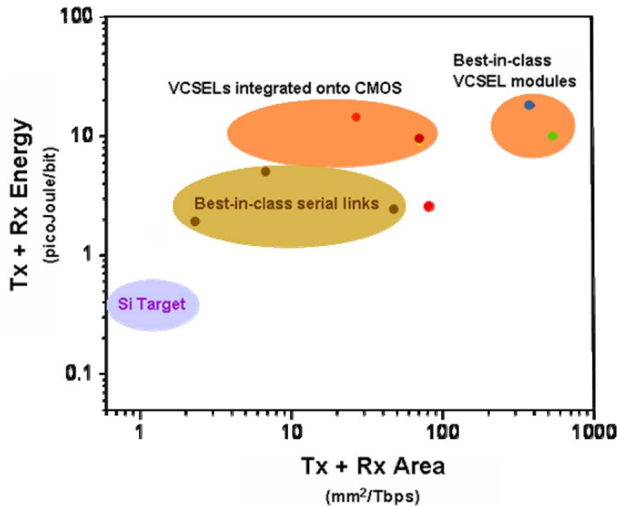


Fig. 23. Transceiver power efficiency (energy per bit) versus area efficiency (area per 1 Tb/s) of commercial and research optical and electrical technologies.

electrical parasitics related to the integration (and these effects in turn held reduce power as described earlier). Fig. 23 depicts this improvement by plotting the energy-per-bit targets as well the area required to achieve 1 Tb/s of aggregated throughput of known best-in-class optical products relying on VCSELs (active optical cables and modules), VCSELs integrated to CMOS (Tx plus Rx), serial electrical links, and silicon targets. Recall that the VCSEL-based products do not include retiming.

IX. Summary and Conclusion

Optical links have successfully displaced electrical links for application domains when the aggregated bandwidth–distance product of the domain has exceeded ∼100 Gb/s-m. It is also evident that commercial optical links provide lower bit energy per unit distance than their electrical counterparts. Today optical links based on VCSELs provide a bit energy per unit distance of less than 0.1 pJ/bit/m, whereas electrical links at competing distances provide 1–10 pJ/bit/m and closer to 10 pJ/bit/m even at very short on-chip links at the millimeter scale and below. In Sections II and III, we reviewed these general trends that have held true for length scales varying across four orders of magnitude and described how this power advantage is already being widely exploited in systems links over 10 m, where the bandwidth–distance product of the links exceeds 100 Gb/s-m.

While the aforementioned is true today only for the longer lengths (>10 m) with VCSEL links, it is interesting to note that this advantage of optical links can be exploited in the future even for much shorter links because the communication bit energy per unit distance of electrical links does not scale down as one approaches chip-to-chip and even on-chip length scales. Hence, there is an opportunity to reduce total interconnect energy with photonics for a given system configuration. In Section IV, we suggested that a maximum optical system link power including all static, dynamic, and laser power, as well as bit-level retiming power below ∼1 pJ/bit/m represents an appropriate value of an energy-per-bit per unit distance which, when taken together with the 100 Gb/s-m bandwidth–distance product, will allow optical links to continue to successfully penetrate even further into the digital communications hierarchy.

For switched or routed links, the integration of the photonics and electronics is vital to power savings. Simply providing optical links directly to a chip can not only alleviate constraints on electrical chip I/O but can significantly improve the switched energy per bit. Several experiments in integrating VCSEL-based links at the system level have been undertaken in an effort to increase bandwidth density and reduce system complexity and power. In Section V, we presented results from an early experimental CMOS-VCSEL system operating at Gigabit Ethernet line rates that achieves a switched interconnect energy of less than 19 pJ/bit for a fully nonblocking switched network with 16 ports and an aggregate capacity of 20 Gb/s/port. This was based on multimode link with a maximum link distance on the order of 300 m, corresponding to switched bit energy per unit distance of <0.1 pJ/bit/m. The switch made use of ultradense integration of VCSEL-based photonics on CMOS demonstrating an optical bandwidth density of over 37 Gb/s/mm^2 even when operating at modest line rates of 1.25 Gb/s. This represents a first experiment in the aggressive integration of VCSELs, detectors, and a VLSI switch, but, nonetheless, provides a record system power for a fully switched VCSEL-based link and a record density for direct integration of lasers and detectors onto VLSI circuits. Extrapolating these results, we expect this technology to be capable of peak bandwidth densities in the range of 350 to 500 Gb/s/mm^2 with energy efficiencies on the order of 5–10 pJ/switched bit.

Even greater power savings from photonic links could be had, if processing, interconnect, and memory functions could be optically enabled and brought in close proximity to each other to create a multireticle-sized "macrochip" that supports switching, processing, routing, and stacked memory functions. In Section VI, we reviewed the architectural and packaging concepts of a "macrochip" that could potentially provide <1 pJ/switched bit when directly integrated with switching chips. This is significant because it allows an order of magnitude reduction in energy per switched bit and an order of magnitude increase in bandwidth–density without compromising any other aspect of the electronic switching fabric including its data rate, reconfigurability, or switching time. The macrochip reduces latency between chips, while still retaining a relatively conventional packaging methodology by combining all of the aforementioned benefits: direct areal photonic I/O to chips; low area, high-density wavelength multiplexed silicon photonic links, and a dense collection of functional chips closely packed on an optically interconnect silicon substrate.

In Section VII, we reviewed progress toward ultralow power and ultradense CMOS-compatible silicon photonic WDM transmitters and receivers needed for the macrochip. We discussed appropriate energy per bit target values for short-distance silicon photonic links in Section VIII and reviewed an energy and bandwidth–density roadmap for efficient photonic transceivers for chip-to-chip and intrachip communication, stating an ultimate link efficiency goal of 300 fJ/bit and a link I/O density of in excess of 1 Tb/s/mm^2.

ACKNOWLEDGMENT

The authors would like to thank many previous colleagues including J. E. Ford, A. L. Lentine, J. A. Walker, D. T. Neilson, L. M. F. Chirovsky, T. K. Woodward, F. E. Kiamilev, and B. Tseng at Bell Labs, A. Persuad, R. Krause, V. Sinyansky, and H. H. Kim at AraLight, and D. Huang, T. Sze, H. Davidson, and R. Lytel at Sun Microsystems for their input to this work, their comments on the manuscript, and for many stimulating and informative discussions on optical and electrical link technologies. They would also like to thank Dr. P. Dong, Dr. D. Lee, and the rest of the Kotura team for their work on the silicon-ring modulator and the optical proximity coupler reviewed in Section VII-A and D, and the Luxtera team for their work on the CMOS photonic modulator, detector, and mux discussed in Section VII-A, B, and E, respectively. They gratefully acknowledge Dr. K. H. Gulden and his team formerly of Avalon Photonics (currently Bookham), for fabricating the flip-chip bondable 850 nm VCSEL arrays used in the VCSEL-switch experiments. The views, opinions, and/or findings contained in this article/presentation are those of the author/presenter and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

REFERENCES

[1] A. V. Krishnamoorthy, "The intimate integration of photonics and electronics," in *Advances in Information Optics and Photonics*, A. T. Friberg and R. Dändliker, Eds. Bellingham, WA: SPIE Press, 2008, pp. 581–598.

[2] D. T. Neilson, "Photonics for switching and routing," *IEEE J. Sel. Topics Quantum Electron.*, vol. 12, no. 14, pp. 669–678, Jul./Aug. 2006.

[3] I. Keslassy, S.-T. Chuang, K. Yu, D. Miller, M. Horowitz, O. Solgaard, and N. McKeown, "Scaling Internet routers using optics," in *Proc. ACM SIGCOMM 2003*, Karlsruhe, Germany, Aug.

[4] A. V. Krishnamoorthy, R. Ho, X. Zheng, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, and J. E. Cunningham, "Computer systems based on silicon photonic interconnect," *Proc. IEEE*, vol. 97, no. 7, pp. 1337–1361, Jul. 2010.

[5] A. V. Krishnamoorthy, K. W. Goossen, L. M. F. Chirovsky, R. G. Rozier, P. Chandramani, W. S. Hobson, S. P. Hui, J. Lopata, J. A. Walker, and L. A. D'Asaro, "16 × 16 VCSEL array flip-chip bonded to CMOS VLSI circuit," *IEEE Photon. Technol. Lett.*, vol. 12, no. 8, pp. 1073–1075, Aug. 2000.

[6] C. Cook, J. E. Cunningham, A. Hargrove, G. G. Ger, K. W. Goossen, W. Y. Jan, H. H. Kim, R. Krause, M. Manges, M. Morrissey, M. Perinpanayagam, A. Persuad, G. J. Shevchuk, V. Sinyanski, and A. V. Krishnamoorthy, "A 36-channel parallel optical module based on optoelectronics-on-VLSI technology," *IEEE J. Sel. Topics Quantum Electron.*, vol. 9, no. 2, pp. 387–399, Mar./Apr. 2003.

[7] C. Kuznia, J. Ahadian, M. Englekirk, M. Wong, J. Richaud, M. Pendleton, D. Pommer, and R. Reedy, "Ultra-thin silicon-on-sapphire component technology for short reach optical parallel optical interconnects," in *Proc. IEEE 2003 Custom Integrated Circuits Conf.*, Sept., pp. 335–338.

[8] L. A. B. Windover, J. N. Simon, S. A. Rosenau, K. S. Giboney, G. M. Flower, L. W. Mirkarimi, A. Grot, B. Law, C.-K. Lin, A. Tandon, R. W. Gruhlke, H. Xia, G. Rankin, M. R. T. Tan, and D. W. Dolfi, "Parallel-optical interconnects >100 Gb/s," *J. Lightw. Technol.*, vol. 22, no. 9, pp. 2055–2056, Sep. 2004.

[9] C. L. Schow, F. E. Doany, C. W. Baks, Y. H. Kwar, D. M. Kuchta, and J. A. Kash, "A single-chip CMOS-based parallel optical transceiver capable of 240 Gb/s bidirectional data rates," *J. Lightw. Technol.*, vol. 27, no. 7, pp. 915–929, Apr. 2009.

[10] X. Zheng, J. Lexau, Y. Luo, H. Thacker, T. Pinguet, A. Mekis, G. Li, J. Shi, P. Amberg, N. Pinckney, K. Raj, R. Ho, J. E. Cunningham, and A. V. Krishnamoorthy, "An ultra-low all CMOS Si photonic transmitter," *Opt. Express*, vol. 18, no. 1, pp. 204–211, Jan. 2010.

[11] R. Ho, J. Lexau, F. Liu, D. Patil, R. Hopkins, E. Alon, N. Pinckney, P. Amberg, X. Zheng, J. E. Cunningham, and A. V. Krishnamoorthy, "Circuits for silicon photonics on a macrochip," in *Proc. Asian Solid-State Circuits Conf.*, Taipei, Taiwan, Nov. 1–4, 2009.

[12] P. Dong, S. Liao, D. Feng, H. Liang, D. Zheng, R. Shafiiha, C.-C. Kung, W. Qian, G. Li, X. Zheng, A. V. Krishnamoorthy, and M. Asghari, "Low Vpp, ultralow-energy, compact, high-speed silicon electro-optic modulator," *Opt. Express*, vol. 17, no. 25, pp. 22484–22490, Dec. 2009.

[13] G. Li, X. Zheng, J. Lexau, Y. Luo, H. Thacker, P. Dong, S. Liao, D. Feng, M. Asghari, J. Yao, J. Shi, P. Amberg, N. Pinckney, K. Raj, R. Ho, J. E. Cunningham, and A. V. Krishnamoorthy, "Ultralow-power, high-performance Si photonic transmitter," in *Proc. OFC 2010*, paper OMI2, San Diego, Mar., pp. 1–3.

[14] X. Zheng, F. Liu, D. Patil, H. Thacker, Y. Luo, T. Pinguet, A. Mekis, J. Yao, G. Li, J. Shi, K. Raj, J. Lexau, E. Alon, R. Ho, J. E. Cunningham, and A. V. Krishnamoorthy, "A sub-picojoule-per-bit CMOS photonic receiver for densely integrated systems," *Opt. Express*, vol. 18, no. 5, pp. 5151–5160, Mar. 2010.

[15] A. L. Lentine, L. M. F. Chirovsky, L. A. D'Asaro, C. W. Tu, and D. A. B. Miller, "Energy scaling and subnanosecond switching of symmetric self-electrooptic effect devices," *IEEE Photon. Technol. Lett.*, vol. 1, no. 6, pp. 129–131, Jun. 1989.

[16] A. L. Lentine, L. M. F. Chirovsku, M. W Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, G. J. Pryzylek, and L. E. Smith, "Diode clamped symmetric self electro-optic effect devices with sub-picojoule switching energies," *Appl. Phys. Lett.*, vol. 60, no. 15, pp. 1809–1811, 1992.

[17] K. W. Goossen, D. A. B. Miller, J. E. Cunningham, W. Y. Jan, A. L. Lentine, A. M. Fox, and N. K. Ailawadi, "Self-Biased S-SEED using highly electroabsorptive slightly asymmetric coupled quantum wells," in *Proc. IEEE Lasers Electro-Optics Society Annu, Meet.*, pp. 310–311, 1990.

[18] A. V. Krishnamoorthy, A. Lentine, K. Goossen, J. Walker, T. Woodward, J. Ford, G. Aplin, L. D'Asaro, S. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, L. Chirovsky, and D. Miller, "3-D integration of MQW modulators over active submicron CMOS circuits: 375 Mb/s transimpedance receiver-transmitter circuit," *IEEE Photon. Technol. Lett.*, vol. 7, no. 11, pp. 1288–1290, Nov. 1995.

[19] A. V. Krishnamoorthy, T. K. Woodward, R. Novotny, K. Goossen, J. Walker, A. Lentine, L. D'Asaro, S. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, L. Chirovsky, G. Aplin, R. Rozier, F. Kiamilev, and D. Miller, "Ring oscillators with optical and electrical readout based on hybrid GaAs MQW modulators bonded to 0.8 micron silicon VLSI circuits," *Electron. Lett.*, vol. 31, no. 22, pp. 1917–1918, Oct. 1995.

[20] A. V. Krishnamoorthy, T. K. Woodward, K. Goossen, J. Walker, A. Lentine, L. Chirovsky, S. Hui, B. Tseng, R. Leibenguth, J. Cunningham, and W. Jan, "Operation of a single-ended 550 Mb/s, 41 fJ, hybrid CMOS/MQW receiver-transmitter," *Electron. Lett.*, vol. 32, no. 8, pp. 764–766, Apr. 1996.

[21] G. Williams, "Wide dynamic range fiber optics receivers," in *Proc. IEEE Solid State Circuits Conf.*, Feb. 1982, pp. 80–81.

[22] D. Fraser Jr., G. Williams, R. Jindal, R. Kushner, and B. Owen, "A single-chip NMOS preamplifier for optical fiber receivers," in *Proc. IEEE Solid State Circuits Conf.*, Feb. 1983, pp. 80–81.

[23] A. A. Abidi, "Gigahertz transresistance amplifiers in fine line NMOS," *IEEE J. Solid-State Circuits*, vol. SC-19, no. 6, pp. 986–994, Dec. 1984.

[24] A. V. Krishnamoorthy, A. L. Lentine, K. W. Goossen, J. A. Walker, T. K. Woodward, J. E. Ford, G. F. Aplin, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossivcs, D. Dahringer, L. M. F. Chirovsky, and D. A. B. Miller, "3-D integration of MQW modulators over active submicron CMOS circuits: 375 Mb/s transimpedance receiver transmitter circuit," *IEEE Photon. Technol. Lett.*, vol. 7, no. 11, pp. 1288–1290, Nov. 1995.

[25] O. Vendier, S. W. Bond, L. Myunghee, J. Sungyung, M. Brooke, N. M. Jokerst, and R. P. Leavitt, "Stacked silicon CMOS circuits with a 40 Mb/s through-silicon optical interconnect," *IEEE Photon. Technol. Lett.*, vol. 10, no. 4, pp. 606–608, Apr. 1998.

[26] T. K. Woodward and A. V. Krishnamoorthy, "Optical receivers for optoelectronic VLSI," *IEEE J. Sel. Topics Quantum Electron.*, vol. 2, no. 1, pp. 106–116, Apr. 1996.

[27] D. H. Hartman, M. K. Grace, and C. R. Ryan, "A monolithic silicon photodetector/amplifier IC for fiber and integrated optics application," *J. Lightw. Technol.*, vol. LT-3, no. 4, pp. 729–738, Aug. 1985.

[28] P. J.-W. Lim, A. Y. C. Tzeng, H. L. Chuang, and S. A. St. Onge, "A 3.3 V monolithic photodetector/CMOS preamplifier for 531 Mb/s optical data

link applications," in *Proc. Tech. Dig. 1993 Int. Solid State Circuits Conf.*, 1993, pp. 96–97.

[29] T. K. Woodward and A. V. Krishnamoorthy, "1 Gb/s CMOS photoreceiver with integrated detector operating at 850 nm," *Electron. Lett.*, vol. 34, pp. 1252–1253, 1998.

[30] C. L. Schow, J. D. Schaub, R. Li, J. Qi, and J. C. Campbell, "A 1-Gb/s monolithically integrated silicon NMOS optical receiver," *IEEE J. Sel. Topics Quantum Electron.*, vol. 4, no. 6, pp. 1035–1039, Nov./Dec. 1998.

[31] M. Ingels and S. J. Steyaert, "A 1-Gb/s, 0.7-m CMOS optical receiver with full rail-to-rail output swing," *IEEE J. Solid-State Circuits*, vol. 34, no. 7, pp. 971–977, Jul. 1999.

[32] A. Emami-Neyestanak, D. Liu, G. Keeler, N. Helman, and M. Horowitz, "A l.6 Gb/s, 3 mW CMOS receiver for optical communication," in *Proc. 2002 Symp. VLSl Circuits Digest Tech. Papers*, 2002, pp. 84–87.

[33] D. Guckenberger, J. D. Schaub, D. Kucharski, and K. T. Kornegay, "A 1 V, 10 mW, 10 Gb/s CMOS optical receiver front-end," in *Proc. IEEE Radio Frequency Int. Circuits Conf.*, Jun. 2005, pp. 309–312.

[34] G. Masini, G. Capellini, J. Witzens, and C. Gunn, "High-speed, monolithic CMOS receivers at 1550 nm with Ge on Si waveguide photodetectors," in *Proc. IEEE LEOS*, New York, 2007, pp. 848–849.

[35] C. Gunn, "CMOS photonics for high-speed interconnects," *IEEE Micro*, vol. 26, no. 2, pp. 58–66, Mar. 2006.

[36] L. Chen and M. Lipson, "Ultra-low capacitance and high speed germanium photodetectors on silicon," *Opt. Express*, vol. 17, no. 10, pp. 7901–7906, 2009.

[37] L. Vivien, J. Osmond, J.-M. Fedeli, D. Marris-Morini, P. Crozat, J.-F. Damlencourt, E. Cassan, Y. Lecunff, and S. Laval, "42 GHz p.i.n Germanium photodetector integrated in a silicon-on-inculator waveguide," *Opt. Express*, vol. 17, no. 8, pp. 6252–6257, Apr. 2009.

[38] T. Pinguet, B. Analui, E. Balmater, D. Guckenberger, M. Harrison, R. Koumans, D. Kucharski, Y. Liang, G. Masini, A. Mekis, S. Mirsaidi, A. Narasimha, M. Peterson, D. Rines, V. Sadagopan, S. Sahni, T. J. Sleboda, D. Song, Y. Wang, B. Welch, J. Witzens, J. Yao, S. Abdalla, S. Gloeckner, and P. De Dobbelaere, "Monolithically integrated high-speed CMOS photonic iransceivers," in *Proc. IEEE Conf. Group IV Photonics*, New York, 2008, pp. 362–364.

[39] A. L. Lentine, K. W. Goossen, J. A. Walker, J. E. Cunningham, W. Y. Jan, T. K. Woodward, A. V. Krishnamoorthy, B. J. Tseng, S. P. Hui, R. E. Leibenguth, L. M. F. Chirovsky, R. A. Novotny, D. B. Buchholz, and R. L. Morrison, "Optoelectronic VLSI switching chip with greater than 1 Tbit/s potential optical I/O bandwidth," *Electron. Lett.*, vol. 33, no. 10, pp. 894–895, 1997.

[40] A. V. Krishnamoorthy, J. E. Cunningham, X. Zheng, I. Shubin, J. Simons, D. Feng, H. Liang, C.-C. Kung, and M. Asghari, "Optical proximity communication with passively aligned silicon photonic chips," *IEEE J. Quantum Electron.*, vol. 45, no. 4, pp. 409–414, Apr. 2009.

[41] X. Zheng, I. Shubin, G. Li, T. Pinguet, A. Mekis, J. Yao, H. Thacker, Y. Luo, J. Costa, K. Raj, J. E. Cunningham, and A. V. Krishnamoorthy, "A tunable 1×4 CMOS photonic wavelength multiplexer/demultiplexer for dense optical interconnects," *Opt. Express*, vol. 18, no. 5, pp. 5151–5160, Mar. 2010.

[42] J. E. Cunningham, I. Shubin, X. Zheng, T. Pinguet, A. Mekis, Y. Luo, H. Thacker, G. Li, J. Yao, K. Raj, J. E. Cunningham, and A. V. Krishnamoorthy, "Highly-efficient thermally-tuned resonant optical filters," *Opt. Express*, vol. 18, no. 18, pp. 19055–19063, Aug. 2010.

[43] P. Dong, W. Qian, H. Liang, R. Shafiiha, D. Feng, G. Li, J. E. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "Thermally tunable silicon racetrack resonators with ultralow tuning power," *Opt. Express*, vol. 18, no. 19, pp. 20298–20304, Sep. 2010.

[44] C. Debaes, A. Bhatnagar, D. Agarwal, R. Chen, G. A. Keeler, N. C. Helman, H. Thienpont, and D. A. B. Miller, "Receiver-less optical clock injection for clock distribution networks," *IEEE J. Sel. Topics Quantum Electron.*, vol. 9, no. 2, pp. 400–409, Mar. 2003.

[45] A. V. Krishnamoorthy and D. A. B. Miller, "Scaling optoelectronic-VLSI circuits into the 21st century," *IEEE J. Sel. Topics Quantum Electron.*, vol. 2, no. 1, pp. 55–76, Apr. 1996.

[46] D. T. Neilson, D. Stiliadis, and P. Bernasconi, "Ultra-high-capacity optical IP routers for the networks of tomorrow: IRIS project," in *Proc. Europ. Conf. Opt. Comm.*, Glasgow, 2005, vol. 5, pp. 45–48.

[47] J. S. Orcutt, A. Khilo, M. A. Popovic, C. W. Holzwarth, H. Li, J. Sun, B. Moss, M. S. Dahlem, E. P. Ippen, J. L. Hoyt, V. Stojanovic, F. X. Kartner, H. I. Smith, and R. J. Ram, "Photonic integration in a commercial scaled bulk-CMOS process," in *Proc. IEEE Int. Conf. Photonics Switching*, Pisa, Italy, Sep. 2009, pp. 1–2.

[48] I. Young, E. Mohammed, J. Liao, A. Kern, S. Palermo, B. Block, M. Reshotko, and P. Chang, "Optical I/O for tera-scale computing," *IEEE J. Solid-State Circuits*, vol. 45, no. 1, pp. 235–248, Jan. 2010.

[49] R. Ho, F. Liu, D. Patil, X. Zheng, G. Li, I. Shubin, E. Alon, J. Lexau, H. Schwetman, J. E. Cunningham, and A. V. Krishnamoorthy, "Optical interconnect for high-end computer systems," *IEEE Design Test Computers*, vol. 27, no. 4, pp. 10–19, Jul./Aug. 2010.

Ashok V. Krishnamoorthy (M'93–SM'10) received the B.S. (Hons.) degree in engineering from the California Institute of Technology, Pasadena, CA, the M.S. degree in electrical engineering from the University of Southern California, Los Angeles, CA, and the Ph.D. degree in applied physics from the University of California, San Diego, CA.

He is currently a Hardware Architect and a Principal Investigator for the Oracle-DARPA initiative on ultraperformance nanophotonic intrachip communication and related photonics R&D at Oracle, San Diego. Previously, he was a Distinguished Engineer and a Director at the Sun Microsystems Microelectronics Physical Sciences Center, San Diego. He was also the President and Chief Technical Officer of AraLight, a Lucent spinout. He was also an Entrepreneur-in-Residence at Lucent's New Venture group, and a member of technical staff in the Advanced Photonics Research Department, Bell Labs, where he investigated methods of integrating optical devices to silicon very large scale integration circuits. He has authored or coauthored more than 185 technical publications, eight book chapters, and 75 conference invited talks. He holds 48 U.S. patents.

Dr. Krishnamoorthy is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He was the recipient of several individual and team honors including an Eta Kappa Nu Outstanding Electrical Engineer Award, an IEEE Distinguished Lecturer Award, the 2004 International Commission of Optics prize in optics, the Sun Microsystems Chairman's Award for Innovation, and the Best Paper Award at the International Microelectronics and Packaging Society (IMAPS) meeting 2009. He has also been engaged as Guest Editor for several journals and Program Committee Member or Chair of more than 20 internationals conferences, most recently as a Co-Chair for the 2010 IEEE Photonics Society summer topical meeting on Optical networks and Devices for Data Centers. He also serves on the advisory board of several early-stage technology companies.

Keith W. Goossen (S'83–M'88–SM'10) received the B.S. degree from University of California at Santa Barbara, in 1983, and the Ph.D. degree from Princeton University, Princeton, NJ, in 1988, both in electrical engineering.

In 2002, he joined the University of Delaware, Newark, DE, where he is currently an Associate Professor in the Department of Electrical and Computer Engineering. He has international stature in the fields of optoelectronics and optical fiber communication, including more than 130 refereed journal papers with an h-index of 31. He holds 82 patents. He was a member of Technical Staff at Bell Laboratories, first as part of AT&T, then Lucent Technologies. He was an Inventor and a Demonstrator of the world fastest microelectromechanical system (MEMS) optical modulator, which was used in system demonstrations for ultra broadband fiber-to-the-home networks and dynamic wavelength management in wavelength division multiplexing networks. He has also invented and demonstrated flip-chip bonding techniques for integrating electronic and photonic chips, particularly short-wavelength vertical cavity surface emitting laser chips, which resulted in the first demonstration of optical transceiver chips capable of transmitting, receiving, and processing up to a terabit/second of data. In 2000, he was a Co-Founder of Aralight, Inc. as a Chief Scientist to commercialize this technology for optical backplane applications, resulting in a full-product demonstration two years later of 36 channel, 90 Gb/s transmit/receive parallel optical modules. His research interests include optical MEMS and optical interconnects, as well as fiber optic sensors, and modulators for free-space optical communications.

Prof. Goossen was a Director of the Mid-Atlantic Industrial Assessment Center in 2006, funded by the Department of Energy, where he was the Leader of teams of students to perform energy efficiency audits of mid-sized industries.

William Jan received the M.S. degree in material science from Steven's Institute of Technology, Hoboken, NJ.

He was at AT&T and Lucent Technologies Bell Laboratories and a number of startups. He has 25 years of semiconductor processing experience from crystal growth to device processing to packaging. He is currently at Bridge Semiconductor, Pittsburg, PA, where he is engaged in a hybrid infrared focal plane array camera.

Xuezhe Zheng (M'03–SM'03) received the B.S., M.S., and Ph.D. degrees in optical instruments from Tsinghua University, Beijing, China in 1993 and 1997, respectively.

He was with Sun Microsystems as a Senior Staff Engineer, where he was engaged on advanced optical interconnects. He is currently a Principal Hardware Engineer with Sun Laboratories, Oracle, San Diego, CA. From 2000 to 2005, he was with Calient Networks Incorporated, San Jose, CA, where he was a Manager of optical engineering engaged on 3-D microelectromechanical systems-based photonic switching and its application in wavelength division multiplexing (WDM) networks. From 1997 to 1999, he was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, University of California, San Diego, where he was involved in high-speed, high-density free-space optical interconnects. He has extensive experiences in photonic switching and optical cross connect, fiber-optic components, dense WDM (DWDM) optical networks, and optical interconnections. He has authored or coauthored more than 60 papers in technical journals. He holds 10 U.S. patents. His current research interests include WDM Si photonics for advanced inter/intrachip interconnects.

Dr. Zheng is a recipient of the Science and Technology Development Award from the National Education Committee of China.

Ron Ho (S'92–M'93–SM'08) received the Ph.D. degree in electrical engineering from Stanford University, Palo Alto, CA.

He is currently a Hardware Architect at Sun Laboratories, Oracle, Menlo Park, CA, where he is engaged in the Very Large Scale Integration Research Group. He was with Intel Corporation for ten years. His research interests include chip-to-chip and on-chip communication technologies, and memory designs.

Guoliang Li received the Ph.D. degree in electrical and computer engineering from the University of California, San Diego, CA, in 2002.

In March 2008, he joined Sun Microsystems Laboratories, Oracle, San Diego, CA, where he is currently a Principal Hardware Engineer and is engaged in the ultraperformance nanophotonic intrachip communication project. In January 2001, he joined the Semiconductor Research Department, Bell Labs, which was subsequently spun off with Agere Systems, where he was the Lead Designer for 40 Gb/s InP Mach–Zehnder modulators and 40 Gb/s electroabsorption modulated lasers (EML). In 2003, he joined Luxtera, where he was involved in the development of electronic and photonic integrated circuits, and successfully developed world's first 10 Gb/s Si optical modulator on CMOS fabrication platform. In 2006 and 2007, he was with the Optical Platform Division, Intel Corporation, where he was engaged in the SFP+ and X2 transceiver development for 10GBASE-LRM and 10GBASE-LR applications. In March 2008, he joined Sun Microsystems to work on the UNIC project. He has coauthored two book chapters and more than 40 peer-reviewed technical papers. He holds six U.S. patents with eight more pending. His research interests include high-speed optical devices, Si photonics, and optical interconnects.

Dr. Li was engaged with the technical program committee of the IEEE Microwave Photonics Conference. He is also an External Reviewer for the Research Grants Council of Hong Kong.

Richard Rozier received the B.S.E.E., M.S.E.E., and Ph.D. degrees from the University of North Carolina, Charlotte, in 1992, 1996, and 1998, respectively.

He was at Bell Labs Innovations, Holmdel, NJ as a Member of Technical Staff, where he was in the Optical LAN Research Department from 1998 to 2000, and performed research and development of optoelectronic very large scale integration multichannel crossbar switching circuits to be utilized with optical switching and routing systems. From 2000 to 2004, he was also at Xanoptix, Inc., Merrimack, NH as an IC Development and Test Engineer. He is currently with Hovey Williams, LLP, Kansas City, MO as a Registered Patent Agent.

Frankie Liu (M'06) received the Ph.D. degree in electrical engineering from Stanford University, Palo Alto, CA.

He is currently a Designer of analog and digital circuits in Very Large Scale Integration Research Group, Sun Laboratories, Oracle, Menlo Park, CA. His research interests include applied physics and mathematics.

Dinesh Patil (S'03–M'07) received the Ph.D. degree from Stanford University, Palo Alto, CA, in 2008.

He is currently in Very Large Scale Integration Research Group, Sun Laboratories, Oracle, Menlo Park, CA. His research interests include chip-to-chip and on-chip communication technologies and circuit optimization tools for energy-efficient designs.

Jon Lexau received the B.S. degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, NY, in 1989, and the M.S. degree in electrical engineering from Stanford University, Palo Alto, CA, in 1994.

From 1989 to 1993, he was with the Amdahl Corporation, where he was designing high-performance mainframe CPUs. Since 1994, he has been with Sun Laboratories, Oracle Menlo Park, CA, where he is engaged in the Very Large Scale Integration Research Group. His current research interests include capacitively coupled proximity communication and electrical transceivers for optical interconnect.

Herb Schwetman received the Ph.D. degree in computer sciences from the University of Texas at Austin.

He is currently a Senior Staff Engineer in Sun Laboratories, Oracle, Austin, TX. Since 2007, he has been a Member of architecture/performance team in the Oracle-DARPA ultraperformance nanophotonic intrachip communication project. In 1994, he was the Founder and CEO of Mesquite Software, Inc. In 1984, he was also an Adjunct Professor in the Department of Computer Sciences, The University of Texas. He was also with MCC, Austin, an R and D consortium. He was a Member of the faculty in the Department of Computer Sciences, Purdue University, West Lafayette, IN.

Dazeng Feng received the B.S. and Ph.D. degrees in optics from Fudan University, Shanghai, China, in 1986 and 1992, respectively.

From 1992 to 1995, he was a Faculty Member at Fudan University. From 1995 to 2000, he was a Research Scientist at Optiwave Corp. In 2000, he jointed Kotura Inc., Los Angeles, CA, where he is currently engaged in silicon photonics as a Senior Principle Engineer. His research interests include silicon waveguides, photonics, and optics.

Mehdi Asghari received the B.A. degree in engineering from Cambridge University, Cambridge, U.K., the M.Sc. degree in optoelectronics and laser devices from the Heriot-Watt University, Edinburgh, Scotland, and St. Andrews University, Scotland, U.K., and the Ph.D. degree in optoelectronics from the University of Bath, Bath, U.K.

He is currently the Chief Technical Officer at Kotura Inc., Los Angeles, CA. He has more than 15 years of progressive engineering and management experience within the silicon photonics and optoelectronics industries. He was a Vice-President Research and Development of Bookham, Inc. He has authored or coauthored more than 70 publications. He holds more than 15 patents in silicon photonics and optoelectronics.

Thierry Pinguet, photograph and biography not available at the time of publication.

John E. Cunningham received the B.S. degree in physics at the University of Tennessee, Knoxville, and the M.S. and Ph.D. degrees in physics from University of Illinois at Champaign-Urbana.

He is currently a Consulting Member of Technical Staff at Sun Laboratories, Oracle, San Diego, CA, where he is engaged in packaging initiatives to develop interchip proximity communication and Si nanophotonics solutions for data communications within computers. He was a Veteran Research Scientist for more than 25 years with University, Bell Labs. He was involved in the area of optoelectronic and semiconductor materials and devices used within optical networks. He was the Chief Scientist at Aralight, where he developed products based on the integration of vertical-cavity surface-emitting lasers and photodetectors with CMOS, a technology spun out of Bell Laboratories. While at Bell Laboratories, he pioneered eight world records on various types of quantum mechanically engineered nanodevices and materials. He was a Member of the Research Faculty in the Physics Department, University of Illinois, where he pioneered metals molecular beam epitaxy. He has coauthored more than 360 journal papers. He holds 30 U.S. patents.

Dr. Cunningham was the recipient of the Chairman's Award at Sun Microsystems.