Optical interconnects to electronic chips

David A. B. Miller

Ginzton Laboratory, Stanford University, Stanford California 94305, USA (dabm@ee.stanford.edu)

Received 24 February 2010; revised 9 June 2010; accepted 21 June 2010;
posted 24 June 2010 (Doc. ID 124561); published 14 July 2010

Optical interconnects are progressively replacing wires at shorter and shorter distances in information processing machines. This paper summarizes the progress toward and prospects for the penetration of optics all the way to the silicon chip.

OCIS codes: 200.4650, 060.4510.

1. Introduction

In the 50 years since the introduction of the laser, there can be no doubt that it has transformed our ability to send information. When coupled with optical fiber, lasers handle essentially all of our long distance telecommunications and have led to dramatic increases in capacity. In recent years, too, the use of optics for shorter distances has steadily increased, with optical data links now extensively employed between cabinets in large systems. But, should we expect to use optics at ever shorter distances, all the way down to the electronic chips themselves or even possibly for connections on the chips? And, in the coming decades, will such use transform our ability to handle information in some substantial way?

In this paper, we will summarize the physical reasons, the potential, the challenges, and the prospects for making the transition to optical interconnects to chips in the coming decades. There is not space here to give an exhaustive review or to explore the technical points in any detail; instead we will give an overview of some arguments and conclusions. Most of these opinions are largely contained in greater detail in other work by this author [1–8]. In the interests of readability, we will make only limited references in the main body of the paper, concentrating on the arguments and conclusions themselves rather than how they evolved. In an Appendix, we will give some historical context and a representative roadmap through the large volume of literature in this field.

0003-6935/10/250F59-12$15.00/0

2. Physics of Electrical Interconnects

Electrical wiring, especially when integrated in electronic chips and circuit boards, has been extraordinarily successful in enabling complex information processing at remarkably low cost. Displacing wiring at any level of the system therefore requires compelling reasons. The problems of electrical wires are, however, limiting information processing systems [3,9–11]. One example of a specific problem from interconnects is that it is increasingly difficult to keep the balance between the logic operations (e.g., floating point operations or "FLOPs") on chips and the corresponding read and write operations of bytes of information into and out of the memories (see the discussion in Ref. [3]). One rule of thumb is to try to keep a ratio of 1 byte/FLOP [12]. Modern machines increasingly cannot achieve anything like this ratio, and the prediction for electrical interconnects on and off chips is that this ratio is going to become progressively smaller because of the limitations of wiring.

To understand why optical connections may be important at short distances, we need first to understand the limitations of electrical wiring and where they come from.

A. Interconnect Density

The most obvious problem limiting interconnects over long wires is signal attenuation. Such attenuation is unavoidable in wires because of the resistance of the conductors and other factors such as dielectric loss. Since these losses generally become

progressively worse for higher frequencies, they also lead to distortion of signals—the high frequency components of a signal will have greater attenuation than the lower frequency parts and will also have different phase shifts. The distortion can be compensated to some extent by equalization, which, at its simplest, makes the loss equal at all the frequencies of interest. Attenuation can be compensated up to a point by amplification, though eventually such approaches reach limits imposed by noise. Advanced signal formats and signal processing can be used to maximize the information capacity of the lines in the presence of these imperfections, up to the Shannon limit that is determined by the physical noise limitations. All of these compensations increase the complexity of the system, however, and that complexity will increase the cost and the power dissipation.

The simplest solution to increase the information capacity for a given wire is to increase its cross-sectional size, which reduces the resistance. For long lines, such larger wires increase cost. Once we get inside large complex systems, the cross-sectional size of wires becomes a problem in itself, limiting the density of wiring.

Indeed, there is a qualitative change that takes place in the problems of wired connections once we move inside information processing systems themselves [3]. There, space is limited, and all the available space tends to become filled with wiring. Competing with the space for the interconnect wires is the space required to deliver power and remove heat from the system. Hence, information processing systems tend to become simultaneously limited by wiring density and power.

Once we get to the point of the wiring filling all available space, then a relatively simple scaling starts to take over with electrical wiring. Somewhat surprisingly, once we reach this point, we cannot substantially increase the capacity of the links either by miniaturizing the system (e.g., pushing the wires onto the chips) or making it bigger (pushing the information processing elements further apart to get the wires in) [3]. So, relatively suddenly with increasing information capacity, there is a wiring problem at all levels of the information processing system, from the cabinets down to the surface of the chips. We have arguably reached this point with many or even most high performance information switching and processing systems.

We can understand this scaling limitation of wires by looking at the physics [3]. In broad terms, there are two kinds of electrical wiring—resistive-capacitive (RC) lines and inductive-capacitive (LC) lines. Whether a line is RC or LC at a given frequency is primarily determined by its cross-sectional size because that determines whether resistive or inductive impedance dominates at a given frequency (resistance per unit length increases for wires with smaller cross-sections, though inductance and capacitance per unit length remain the same since they depend on the geometry of the line's cross-section, not its cross-sectional size). At the gigahertz frequencies of modern information processing systems, interconnect lines on chips are primarily RC, limited by the bulk resistance of the metals, and lines off chips are primarily LC transmission lines, limited by the resistance through the skin effect (the crowding of the current towards the surface of the wire at high frequencies). (The crossover from bulk resistance to skin effect resistance occurs at roughly the same frequency and with the same scaling as the crossover from RC to LC behavior—hence there are essentially only bulk-resistance-limited RC lines and skin-effect-resistance limited LC lines [3].)

The physics of this scaling problem is quite easy to understand for the RC-limited wiring on chips themselves. In particular, neither the capacitance nor the resistance of wires scales well at smaller sizes. Obviously, if we scale down the cross-sectional size of a wire, its resistance will go up. Less obviously, the capacitance per unit length of a wire does not change at all as we scale down the cross-section because it depends only on the shape of the cross-section. Nearly all useful wires will in practice have a capacitance of roughly a few picofarads per centimeter; there is no practical prospect of reducing this substantially because the capacitance scales only logarithmically in the ratio of the wire size to wire separation (a simple consequence of the logarithmic fall-off in electrostatic potential as we move away from a charged wire). For simple on-off signaling, taking the resistance-capacitance (RC) product as the characteristic minimum allowable bit time, we can easily conclude that scaling down a wire in all three dimensions leaves the RC product the same, and hence does not change the number of bits per second we can get down the wire [3,9].

Figure 1 illustrates the scaling for the case of a simple electrical wire. As the wire is scaled down in all three dimensions, the resistive-capacitive time constant of the wire does not change at all (assuming simple scaling of resistance with size—actually, because of surface scattering, small wires may have even larger resistance than such scaling suggests).

One might imagine that we could change from the RC-limited wires to inductive-capacitive (LC) lines to get higher information capacities; perhaps

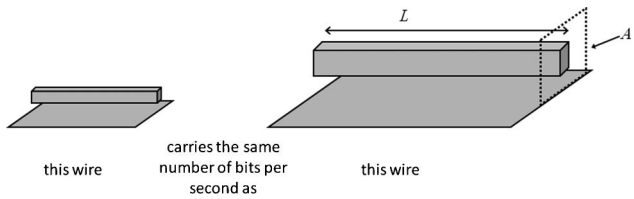

Fig. 1. Simple scaling of electrical wiring shows that the resistance × capacitance product would not change at all as the wire is scaled down in all three dimensions. Hence, for simple signaling where the resistance × capacitance time constant of the wire would set the shortest pulse that could readily be sent on the line, the bit rate capacity of the line is independent of the size scale of the line [3].

surprisingly, though, the scaling of the information capacity of such LC lines for simple on-off signaling obeys a similar scaling law, with actually slightly smaller information capacity for a given cross-sectional area and length. In fact, at least with a simple model of on-off signaling, the capacity of electrical lines from such resistive limits can be written approximately as

$$B \le B_o \frac{A}{L^2}, \tag{1}$$

where A is the cross-sectional area of the wiring, L is the length of the wires, and B_o is a constant [3]. $B_o \sim 10^{16}$ bits/s for RC-limited lines that are typical on a chip, a slightly smaller number holds for inductive-capacitive lines with resistive loss (RLC lines), and $B_o \sim 10^{17}$ bits/s for off-chip equalized RLC lines.

Equation (1) also holds if A is the total cross-sectional area of the wires, not just of one wire. The dimensionless nature of the ratio A/L^2 is what makes the capacity of a wiring system independent of the actual size of the whole system, explaining why we cannot increase the capacity of the wiring either by making the system smaller or bigger once we have filled all available space. This simple argument does not take account of additional losses from surface scattering in small wires nor does it include dielectric loss, all of which make the information carrying capacity of electrical wires smaller than this simple scaling would suggest.

Present electronic systems illustrate this difficulty —they are experiencing the limitations of wired interconnects at all levels at least from the global wires that must go across chips all the way to the electrical backplanes between boards. (At longer distances beyond these, we already change over to optics.) Of course, as mentioned above, we can go past simple signaling approaches, using equalization of the lines, multilevel signaling, and/or sophisticated signal processing to expand the information capacity. On chips themselves, we also have the option of breaking the line into multiple shorter sections with repeater amplifiers, hence circumventing the density limitations of long wires to some degree (see, e.g., Refs. [5,10] for discussions). Recently, Kim and Stojanovich [13] have considered the use of equalized lines also on chips. (Interestingly, these apparently also obey [7] the scaling of Eq. (1).) It is possible as well to stack chips in a three-dimensional (3D) fashion, which can help with local connections to memory, for example [14]. There is, however, no question that the technical cost of using wired interconnections rises steeply as we try to push past this scaling limitation to the information density in electrical lines.

The physics of electrical lines leads also to several other problems in signal integrity for dense, high-speed wiring. For lines on chips, even with repeater amplifiers, the effective signal propagation velocities are slow (e.g., ~5% of the velocity of light), and adjacent lines can readily "cross-talk" to one another through capacitive or inductive coupling. Backplanes can have substantial difficulties with wave reflections from connectors and other discontinuities in the transmission lines. The limited bandwidth of electrical connections also limits the precision of timing edges for clock distribution.

B. Interconnect Energy

The performance of modern information processing systems is increasingly limited by energy. Chips cannot readily dissipate any more power, for example, with power per chip saturating at ~100–200 W from this point on, according to current projections [11]. The overall energy consumption of information and communication technologies is also becoming ecologically significant, already being percents of total electric power consumption and carbon footprint (see Ref. [7] for a discussion). Energy may well be the ultimate physical bound on our ability to handle information.

There are many sources of energy dissipation in information processing. We expect to be performing logic operations, we must read and write to memory, and we must send and receive information through interconnects, both within the system and to the outside world; all of these operations must consume energy. In addition to these core operations, there are other dissipations, such as those associated with leakage and subthreshold currents in modern logic devices. For each unit of energy spent on the core operations, we will have an additional multiplier (e.g., a factor of 2 or so) associated with supplying and conditioning the power and sinking the thermal dissipation.

One might imagine that the dominant energy in the core operations would be in performing the actual logic switching, or in the energy stored in the memory cells that hold the information. In fact, however, of these three core operations of logic switching, memory, and interconnect, it is the interconnect that accounts for most of the energy dissipation, and that energy is almost all associated with charging and discharging the capacitance of signal lines.

To understand this dominance of interconnect over logic energies, we can examine the capacitances in a complementary metal-oxide semiconductor (CMOS) gate [8]. In a gate made with minimum-size transistors, both now and in expected future technologies, the capacitance of the transistors is approximately equal to the capacitance of a wire that is the same length as the distance from one side of the gate circuit to the other. Hence, even if all a gate did was to pass the result of the logic operation to an adjacent gate, the energy associated with charging or discharging the transistor capacitances would be on the same scale as that involved in charging or discharging that short interconnect line. Since information will often or even usually have to be sent farther than that, the energy for charging or discharging the interconnect line can easily exceed that required just for the switching of a logic gate.

The situation is likely even worse in memory banks, such as dynamic random access memory (DRAM) and flash memories. Such memories are typically addressed in a matrix fashion, in which whole sets of lines are charged or discharged even if all one wants to do is read or write a single cell. The capacitances of those lines greatly exceed the capacitances of the memory cells themselves, and hence the reading and writing energy is much greater than the energy required in the memory cell itself just to retain a bit of information reliably.

Hence, other than possible static dissipations like leakage currents, the energy in electronic information processing is spent very substantially to charge the lines that carry information from one place to another. Interconnect power can dominate the power dissipation of information processing, and any way of reducing that power could be very important.

Emerging nanotechnologies, such as the use of carbon nanotubes, could have a role to play in future interconnects, but it does not appear that they solve these core problems of density and energy, at the very least not for longer lines [15,16].

3. Physics of Optical Interconnects

In long distance telecommunications, the very low loss and dispersion of optical fiber allows individual sections of fiber to be of ~100 km in length between repeater stations. Minimizing the number of such stations is important financially. The technical optimization is therefore largely one of designing the fiber system to operate over the longest possible distance with the highest possible bandwidth. The size, power dissipation, and even cost of the optical transmitters and receivers are of secondary importance.

As we move to shorter distances, however, the reasons for the use of optics and the constraints on the interconnect systems both change significantly [4–7]. Optics can help in interconnect density, in interconnect power, and in other ways, such as improved signal integrity and timing [4,17–20], and possibly also through optical switching and routing [21–23]. The comparison between the requirements on and benefits of optical technology at long and short distances is summarized in Table 1. We see some characteristics in common—for example, the use of wavelength-division multiplexing for increasing the information capacity of a spatial channel, such as an optical fiber, and for permitting physical routing of entire channels based on their wavelength. But, density and total power dissipation are particularly important at short distances, neither of which was critical at long distances.

We can also see additional potential benefits in signal integrity [4]. Optics can deliver particularly precise timing of signals because optical channels have low dispersion, allowing the propagation of short pulses over relatively long distances without the pulses being substantially broadened, and because there is actually less temperature dependence of propagation velocity in optical signals than with electrical ones. In optics, too, though there can be cross-talk between channels, such cross-talk is not essentially dependent on bit rate because the bit rate is very much lower than the carrier frequency. Phenomena such as signal reflections can also be minimized through the use of antireflection coatings, an option not practically available in electrical lines. Any optical connection also automatically provides voltage isolation. Though such signal integrity benefits may become important and lead to considerable simplification in the design of optical interconnect systems compared to electrical busses, interconnect density and energy may be the dominant reasons for the introduction of optical interconnects.

Table 1. Comparison of the Reasons and Requirements for the Use of Optics to Communicate over Long and Short Distances

Long Distance	Short Distance
Maximum bit/s over longest possible span. Longest length of connection.	Maximum connection density. Smallest cross-section for a given bandwidth and length.
Total power not critical. Design for minimum received energy.	Minimize power dissipation. Design for minimum total energy per bit.
Wavelength division multiplexing. Reuse one fiber for many channels. Allow wavelength switching and routing.	Wavelength division multiplexing. Give higher density of connections. Allow wavelength switching and routing.
	Signal integrity. Improved timing precision. Reduced reflections, cross-talk. Voltage isolation.

A. Interconnect Density

A major reason for the introduction of optics for connections between the cabinets of large machines is to improve the density of interconnections—for example, the number of bits per second that can flow off the edges of cards and backplanes. Electrical cables can carry substantial amounts of information over the distances between cabinets, but optical cables can do so with much smaller cable diameters and, consequently, with higher densities of connections also. The underlying reason for this density advantage is that optical fibers simply do not possess the resistive loss physics that limits the capacity of electrical cables [3].

Up to this point, and possibly for some time in the future, this density benefit is what has driven and is driving the introduction of optics for connections inside machines. This benefit remains even as we move to shorter distances, such as in connections to chips themselves. Because of the underlying scale invariance of the capacity of simple electrical interconnects discussed above, this density benefit drives us towards trying to use optics at ever shorter distances.

Optical fibers themselves can in principle carry extremely high densities of information. Single-mode telecommunications fibers, with a diameter of only

125 μm, can carry tens of terabits per second of information [24]. Of course, preparing the information in the right form to exploit that bandwidth is not a simple task, and would involve many high-speed transmitters and receivers as well as sophisticated wavelength-division multiplexing. But, the fiber itself is not practically any limit to the information capacity for the foreseeable future. Even fibers and waveguides with less exacting specifications, such as multimode systems, can carry very large bandwidths over substantial distances.

Another more radical option is to consider "free-space" connections where optical outputs on the surface of one chip are imaged using lenses to optical inputs on another chip. This approach can offer very large numbers of channels (e.g., tens of thousands) with a single optical system [25,26], and could avoid the use of wavelength-division multiplexing [27] for short links. This optical approach could be an interesting option further in the future if we want to continue to scale the performance while retaining high bytes per FLOP ratios.

B. Interconnect Energy

Since power is now such a substantial constraint on the performance of information processing machines, it is critical that any replacement technology, such as optical interconnects, must take less power than its electrical counterpart. At first glance, this does not appear to be an advantage for optics. The long distance telecommunications transmitters and receivers typically consume significant amounts of power. Trying to use that same technology at shorter distances appears to offer no advantage in power. Long distance optical technology was developed, however, to work with the minimum received optical power, not the minimum total energy per bit communicated. When looking at interconnects at short distances, it is that total energy per bit, including the power of both the transmitter and receiver, that becomes a dominant criterion.

In electrical systems, as mentioned above, the energy required is at least that needed to charge the line (or at least that section of the line that corresponds to the length of the electrical pulse) to the signal voltage. In optical systems, a different piece of physics becomes available that can avoid this energy. This physics is a quantum-mechanical phenomenon [1,4] that derives ultimately from the photoelectric effect.

The photoelectric effect is that remarkable phenomenon where, even as the power in a light beam is turned down, the voltage that can be generated in a photocell can remain essentially constant. That voltage is essentially numerically equal to the photon energy in electron volts. The classical voltage in the light beam is irrelevant to the voltage that can be generated, so even very small classical voltages can generate large signal voltages. Hence there is no need to "charge" the optical line to the signal voltage. One equivalent way of viewing this is to say that the quantum detection in the optical case has effectively matched the high impedance of small devices to the low impedance of electromagnetic propagation (~50 Ω in cables, ~377 Ω in free-space propagation), performing a "quantum impedance conversion" [1]. Figure 2 illustrates this process in two versions. In version (a), a hypothetical 1 nW light beam shines on a photodiode that is connected to a 1 GΩ resistor. With ~1 eV photons, ~1 V may be generated in the resistor. For comparison, the classical voltage in a 1 nW electromagnetic beam is ~600 μV (r.m.s.) from one side of the beam to the other. Hence the photodetector has converted ~600 μV propagating in 377 Ω into 1 V in 1 GΩ. In Fig. 2(b), we see the related process in which a 1 fJ light pulse generates a ~1 V swing in a 1fF capacitor because ~1 fC of charge (positive and negative) is generated in the photodetector.

It might seem obvious that, to minimize energy per bit, one should design for the minimum received energy (or number of photons), but that neglects the energy to run the receiver amplifier. The most sensitive amplifier must operate at the limit set by noise. A typical transimpedance receiver circuit may operate with quite a large channel current in the input transistor to reduce the effects of Johnson (thermal) noise. Such an operating mode likely does not minimize the total energy dissipation in systems like short interconnects, where the loss in the optical channel may not be very large. Minimizing total dissipated energy in a short link (including transmitter and receiver energy) may well result in a system where we deliberately use a large number of received

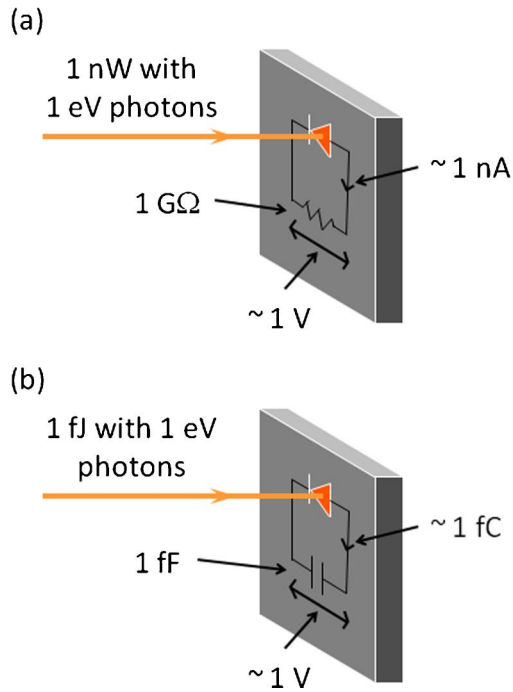

Fig. 2. (Color online) Illustration of quantum impedance conversion (a) for continuous beams and (b) for pulses.

photons so that the receiver is not operating in a noise-limited mode [28].

It is even possible that such a link might operate with no voltage amplification at all at the receiver ("receiver-less" operation [19]), with the input optical pulse having sufficient energy to swing the photodetector over a full logic voltage range. From the point of view of conventional long-distance optical link design, this would be very unusual; systems with only a small amount or even zero amplification can, by contrast, make sense for short interconnections because of the need to minimize total energy per bit [28]. The key additional point here is that interconnect receivers, to minimize total energy per bit, should have the lowest photodetector capacitance possible. With total detector and input transistor capacitance of 1 fF, for example, 1 fJ of 1 eV photons (i.e., ~6000 photons) would generate ~1 V swing in the input stage, as illustrated in Fig. 2. 1 fF of detector capacitance would be very small compared to typical picofarad capacitances of discrete photodetectors, but such low capacitances are conceivable in micrometer-sized detectors integrated beside or within [29] the receiver transistors. This potential energy benefit of optics from quantum impedance conversion can also, at least in principle, be realized at the transmitter side as well, though we would need optical output devices (lasers or modulators) that are efficient at very low energies.

Up to this time, this energy benefit of optics for interconnects remains largely unexploited, in part because it requires a different approach to the receiver and transmitter technologies, specifically one that emphasizes energy efficiency and integration with electronics rather than maximum transmitter power and minimum received signal.

4. Technological Requirements for Optical Interconnects

Given the many potential advantages of optics, we could ask why we are not already using it more at short distances. Obviously, the technology of electrical wiring is very advanced and inexpensive. Arguably, too, at short distances, wiring has been adequate for the interconnect needs of many systems until recently. A major reason, however, that optics is not used more at short distances is that the technology for dense, short-distance optical interconnects is immature.

We have argued above that the technology developed for long distances is not appropriate for short distances, especially for connections all the way to chips themselves. The power dissipation, the physical size, and the cost of longer distance optical technology are all too large for use at short distances. Because information processing systems are increasingly constrained by power, optical interconnects will not be implemented if they take more power than their electrical rivals.

On electrical backplanes, the total dissipations of present high-performance electrical interconnects are in the scale of a few tens of picojoules per bit. For connections on and off chips themselves, energies of several of picojoules per bit would characterize current low-energy interconnects. For the global interconnect lines on chips, ~1 pJ/bit might be a typical number. (See Ref. [7] for a discussion of these energies.) We should also presume that future electrical links will operate at somewhat lower energies per bit. Hence, in looking at solutions in optics, especially in research, we should be looking at target energies of perhaps ~1 pJ per bit for backplane connections, and ~100 fJ/bit for shorter connections to chips and for global on-chip wiring. Such targets, if met, would give sufficient headroom to justify seriously considering changing to optical interconnects at these shorter distances.

100 fJ/bit would be the total system energy, and so the energies of the devices—especially the optical output devices like lasers and/or modulators—would have to be in the scale of only 10 fJ/bit or at most a few tens of fJ/bit for such systems so that there is enough energy left to perform other functions such as driver, receiver, and, possibly also, timing circuits. With 10 fJ/bit transmitted optical energies, the received optical energy would be ~1 fJ/bit in a reasonable optical system allowing for various losses. As illustrated above, such a received energy could be adequate to avoid substantial (or possibly even any) voltage amplification in the receiver circuit, though very low capacitance in the detectors and the subsequent transistors would be essential. A 1 fF total photodetector-plus-transistor capacitance is only possible in a very well integrated optoelectronic technology. A connecting wire of only 10 μm length will already have a capacitance of a few femtofarads. Hence, integration is essential to take full advantage of small photodetector capacitance to minimize system energy per bit. The optical output device (such as a modulator or laser) would also have to be very well integrated to similarly minimize its capacitance—we could not afford to drive more than 10 fF at a 1 V drive if we are to retain 10 fJ/bit energy for that device. Hence again integration is essential.

The output device that offers such low energies per bit and that can be integrated efficiently onto electronic circuits has been and continues to be a substantial challenge. We have discussed such devices elsewhere [7], and will not repeat the details here. Light emitting diodes (LEDs) that are not lasers are likely not viable approaches unless they are single-mode LEDs because otherwise too much power is wasted in trying to couple into waveguides or small detectors. As far as laser sources are concerned, conventional edge-emitting lasers cannot meet such low energy targets. Vertical cavity surface emitting lasers could likely be adequate for 1 pJ/bit systems. 100 fJ/bit systems would require more radical laser sources such as nanocavity lasers. All such lasers have challenges in integrating them directly with silicon.

The other viable option for optical output devices is modulators. Modulators would require an external light source, of course, but that can actually have advantages. Wavelength stabilization can be done once away from the relatively hostile temperature environment of the surface of a chip, and multiwavelength sources could be used [27]. The excess power involved with generating the light is not dissipated on the chip itself. The light source to read out the modulators could be clocked, implicitly reducing timing problems in interconnects [18]. Modulators are not threshold devices like lasers, so they are easier to scale to very low operating powers. Modulators are more tolerant of crystal defects (such defects can lead to short lifetimes for lasers), and so it is easier in practice to integrate disparate materials (e.g., with different lattice constants) to make modulators [30].

Major advances have been made in recent years in the field of silicon photonics (see, e.g., Refs. [31–40]. Such an approach, where the optical, optoelectronic, and electronic components can be made in a way that is compatible with mainstream electronic manufacture (e.g., CMOS electronics), shows a path to low cost and dense integration. When combined with germanium detectors (see, e.g., Refs. [29,34,35]), which can be grown in CMOS-compatible processes, complete optoelectronic systems can be fabricated with integrated modulators. Silicon-based modulators still present challenges here, though, because silicon itself only has a relatively weak mechanism (a change in refractive index with changes in carrier density) for such modulators [31]. As a consequence, such silicon modulators either have to be relatively long (e.g., hundreds of micrometers or even millimeters [33]) or have to operate with high-Q resonators (e.g., 5000–10000) [41]. Such resonators need precise tuning, which may in turn require very precise temperature control, though these resonators have the advantage that they can simultaneously be used for wavelength channel selection. The power needed for the temperature control also has to be included in the overall power budget, however, in calculating the overall energy per bit.

Another modulator approach is to use electroabsorption associated with germanium, either the bulk Franz–Keldysh effect [38] or, with germanium quantum wells, the related [42] quantum-confined Stark effect (QCSE) [36,37,42]. The QCSE is a particularly strong mechanism, though it requires the relatively sophisticated growth of quantum well structures. Either with high-Q devices or with future optimized Franz–Keldysh or QCSE devices, there are possible paths to 10 fJ/bit devices, however.

In the optics itself, the silicon photonics approach largely solves the problem of dense waveguides on chips or possibly on somewhat larger optical units. For the densest future interconnects, or for direct "chip-to-network" connections for longer distances, wavelength division multiplexing may be required for waveguide systems [7]. Such wavelength multiplexing components still present significant challenges. Either high-Q resonators with precise control and tuning will be required, or new approaches, possibly based on nanophotonics, may have to be developed, though there are some promising approaches there [43–50]. Again, though, integration will be essential if optics is to achieve sufficient channels to take over the major task of interconnecting information. Hundreds to thousands of channels would be required for board-to-board or backplane interconnections and hundreds to thousands of channels would be needed for future connections to chips. To take full advantage of optics for interconnects, it will also be necessary to move to single-mode grade optics rather than the multimode optics of current data links. Only single mode optics easily supports the advanced methods for dense wavelength multiplexing, and only single mode optics allows efficient coupling in and out of sources and detectors that are sized for minimum power dissipation.

The technical challenges of moving to dense optics and optoelectronics, integrated with electronics for low energy performance and made in overall integrated systems for low cost and sufficient density, are substantial. Arguably the physics and the basic device concepts and approaches do exist and are advancing rapidly, however.

5. Prospects

The need for denser, faster, and lower energy interconnects inside electronic information processing systems is growing rapidly, at length scales all the way to the chips themselves or even on the chips. Electrical interconnect cannot scale to keep up with this demand. Without a change in interconnect technology, the performance of information processing systems will become bounded by these interconnects. The ratio of the number of bytes of memory access to the number of operations on the chip (e.g., the byte/FLOP ratio) will continue to decline.

When we look for other approaches that could solve this problem, arguably the only one that offers any physical solution, especially once we leave the silicon chip itself, is optics. It avoids the scaling problems of wires and offers various other advantages in signal integrity as well as the ability to connect to the optical networks already used for longer distances. Arguably, too, no physical breakthrough is required for us to have the devices that could meet the energy and size targets needed for optics to take over interconnects down to the level of the chips. The arguments for optics are quite difficult at short distances on chips where wires are small and very cheap and still consume low absolute energies. The arguments become stronger as we think about possible long interconnects on chips, though wires are still a formidable competitor. For interconnects to the environment outside the chip, the arguments for optics become progressively stronger.

Significant progress has been made in recent years, especially in silicon photonics and in optoelectronic devices that could be integrated with such an

approach. There are also many further possibilities for highly functional nanophotonic components that could be integrated with silicon, including nanometallics for deeply subwavelength devices [51,52] and advanced ideas like quantum dot structures combined with photonic crystals [53]. The technological challenges are, however, substantial. They include not only the development of devices and integration on the chips, but also the packaging at the next level up (e.g., on boards and backplanes), where single mode optics is not currently extensively developed. Novel optical components may still be required, such as compact wavelength division multiplexers [43–50]. Significant issues remain in areas such as temperature stabilization since many optical and optoelectronic approaches require temperature to be controlled to a greater or lesser degree.

The ultimate benefit to information processing from the incorporation of optics at shorter distances could, however, be radically important. The vision is to eliminate the problems of interconnect scaling so that future information processing can continue to expand into the future instead of strangling on its own wiring. Increasingly optics looks to be the key technology to enable this transformation.

This work has been supported by the Focus Center Research Program (FCRP) Interconnect Focus Research Center and Connectivity Center. I am pleased to acknowledge stimulating conversations with Ross Audet, Bill Dally, Ron Ho, and Azad Naeemi.

Appendix A: Historical Background

The history of the possible use of optics within digital computing systems dates from the 1960s. Here we will attempt to give a brief summary of the developments leading up to the current work in optical interconnects. A full literature review is, unfortunately, beyond the scope of this paper; at best, we will only be able to cite a few key papers together with a representative sampling of other work.

Early work first considered the possibility of optical or optoelectronic logic gates. The first analysis [54] was pessimistic, not least because of the energy required. Smith's later analysis on the possibilities for optical logic was more optimistic [55]. Interest in optical logic grew in the late 1970s and continued through the early 1990s, stimulated by observations of optically bistable switching in various systems [56–58], and inspired by the notions that optical devices might be faster and/or be capable of working in very highly parallel systems with large arrays of light beams. Work on optical logic has continued at a lower level since that time, though most proposals either do not have the required functional characteristics [54] or take too much energy to compete with CMOS for logic. In the late 1980s and early 1990s, free-space optics for handling large arrays of parallel beams became quite sophisticated [25,59] and several systems were demonstrated [26,60,61]. Some quite large (e.g., 65,000 light beam [26]) optical logic systems were successfully demonstrated using optoelectronic bistable device arrays [62] that did meet the functional requirements for logic systems (even if they were not competitive with the logic energies of electrical devices).

Alongside this logic work, the idea was growing that the benefits of optics lay in the ability to communicate or interconnect rather than in logic itself. MOS (metal-oxide-semiconductor) electronic technology was advancing continuously and rapidly. By the early 1980s [9], potential limitations from the wiring were being foreseen, such as the wiring delay dominating over the electronic logic gate switching time. The seminal paper by Goodman *et al.* [63] marked the first serious consideration of optics for interconnections to chips inside computers, and included discussions of optics as a solution to the delay problem, the idea of optical clock injection, and many different options in optics and optoelectronics for implementing interconnects. Arguably, though, the optoelectronic output devices that could enable any substantial density of optical interconnects did not exist at that time and wires were still good enough.

In the mid to late 1980s, however, two new devices emerged from the rapidly developing technology of layered semiconductor structures—the vertical cavity surface emitting laser (VCSEL) [64,65] and modulators based on the discovery of the quantum-confined Stark effect (QCSE) electroabsorption mechanism [66] in quantum wells. Both of these devices went on to have substantial and continuing use in communications—VCSELs in optical data links and quantum well modulators (integrated with lasers) in fiber telecommunications. Both offered the opportunity to have two-dimensional arrays of optical devices working with beams perpendicular to the surface of a chip. The quantum well electroabsorption effects were also used to make the optoelectronic logic devices in the larger optical logic system demonstrations (the self-electro-optic-effect devices—SEEDs) [62,67,68].

Various analyses were performed for the use of devices such as the VCSEL and quantum well modulators for optical interconnects from chips [1,69], with the conclusion that beyond a certain "break-even" distance, optics would be more attractive for power dissipation than electrical approaches. Such analyses have continued [1,3–5,7,15,17,28,41,60,69–71].

In the late 1980s, the underlying physical reason for this potential power benefit, the quantum impedance conversion discussed above [1], was also clarified.

The work in parallel optoelectronic logic progressively changed into work on parallel optical interconnects. With the successful hybridization of large two-dimensional arrays of quantum well modulators [72] and VCSELs [73] to silicon CMOS chips, the intellectual jump to using optics to overcome the limitations of electrical wiring, rather than as a competitive logic technology, was complete.

In the early to mid 1990s, a physical argument on the scaling of the density of information in electrical

wiring was completed [3], and the set of physical arguments for optical interconnects was understood in some detail [4]. The concepts for how to design interconnect links to minimize power dissipation overall in a technology integrated with CMOS were clarified [28]; a principal conclusion of that work was that integration of photodetectors with very low capacitance (e.g., in the range of femtofarads) offered substantial opportunities for very low power systems.

In the late 1990s into the 2000s, work on free-space parallel interconnects continued in several groups [74–78]. Guided wave optics using waveguides inside printed circuit boards was also investigated, primarily by industrial researchers [79]. Understanding of the specific benefits of optical interconnects to chips deepened [5]. The analysis of the limitations of wires on chips was expanded and clarified [10]. The idea of using wavelength division multiplexing for interconnects between chips was examined and demonstrated [27]. The use of short optical pulses in interconnects was investigated both for some direct benefits in interconnects [18], such as reducing timing skew, and for optical clock distribution to chips [19]. Optics was demonstrated to be able to deliver extremely precise clocking, with subpicosecond accuracy [20], and it could potentially be used for critical applications (such as triggering analog to digital conversion or clocking high speed links) or for removing some layers from the clock distribution tree [19]. The overall conclusion, however, was that there would not be enough available optical power to clock every point on the entire chip.

Another major technical development was the emergence of silicon photonics in the late 1990s and especially in the 2000s. Silicon itself is not generally good for optoelectronic emitters or modulators, in part because it has an indirect optical bandgap, with correspondingly weak emission. It does have a relatively weak but usable change in refractive index with carrier density, whose physics was characterized in the late 1980s [31]. With the emergence of silicon nanophotonic technologies, such as silicon waveguides with well-controlled, small cross-sections, it was possible to exploit this effect to make viable optical modulators [32,33], either as single-pass Mach–Zehnder interferometers or as silicon ring resonators. Germanium detectors, which give good absorption in the infrared regions in which silicon is transparent, could also be fabricated on silicon using variations of the existing complementary MOS (CMOS) integrated circuit fabrication techniques [34,35]. Hence, at least using external laser sources, it became possible to make complete optoelectronic systems with modulators, waveguides, and detectors in silicon in CMOS foundries and even integrated with silicon electronics. The discovery of strong quantum-confined Stark effect (QCSE) electroabsorption in germanium quantum wells grown on silicon [36,37] offered another, much stronger mechanism, and viable modulators were also demonstrated in bulk germanium using Franz–Keldysh electroabsorption [38] (which is the non-quantum-confined cousin of the QCSE [42]). Some possibilities also exist for putting lasers on chips, as in hybrid-bonded approaches [80], or in prospects for germanium lasers on silicon [81]. Hence there is a strong prospect of being able to integrate dense optical transmitters, receivers, and optical and electronic circuits using the CMOS platform.

In parallel with the evolution of ideas and technology for optical interconnects, there was growing analysis of the trends and projections for silicon integrated circuits, trends that are tracked by the International Technology Roadmap for Semiconductors (ITRS) [11], a document that is openly available and has been updated annually. It becomes increasingly clear from the 1990s onwards that interconnections to chips and even on chips are an increasing problem. These roadmap documents show no known solutions in many cases for future interconnect demands to chips. As discussed above, one metric at the system level is the number of bytes of memory access for each instruction, such as the byte per FLOP ratio. Large computing systems in general are not able to sustain this ratio as they advance, and future silicon chips are not predicted to be able to retain the current value of this ratio either. This progressive reduction in the ability to access memory has substantial consequences for future computer architectures and systems performance.

As also mentioned above, another key issue in information processing systems in general that has emerged especially in the 2000s is power consumption. The ITRS roadmap predicts that chip power will not increase beyond about 200 W. It was already the case in the early 2000s that more than half the power on a chip was actually used for interconnection rather than logic, and this number is expected to rise. In addition, the overall power dissipation of information processing technologies is becoming environmentally significant, accounting for percents of electric power dissipation and carbon footprint. This makes the potential power saving from optical interconnects yet more important.

Considerations of power dissipation, bytes per FLOP, free-space optics, waveguide optics, and the resulting requirements on future optical and optoelectronic devices for interconnects are discussed in [7]. Other recent analyses have argued for the incorporation of optical interconnections for the overall benefits it would offer to architectures [21–23]. Overall, there are now substantial technical reasons why a change to optical interconnects at least for distances beyond the size of the chip would be beneficial provided suitable technology can be generated in practice.

References

1. D. A. B. Miller, “Optics for low-energy communication inside digital processors: quantum detectors, sources, and

modulators as efficient impedance converters," Opt. Lett. **14**, 146–148 (1989).
2. D. A. B. Miller, "Device requirements for digital optical processing," in *Digital Optical Computing*, R. A. Athale, ed., SPIE Critical Reviews of Optical Science and Technology (SPIE, 1990), Vol. CR35, pp. 68–76.
3. D. A. B. Miller and H. M. Ozaktas, "Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of the system architecture," J. Parallel Distrib. Comput. **41**, 42–52 (1997).
4. D. A. B. Miller, "Physical reasons for optical interconnection," Int. J. Optoelectron. **11**(3), 155–168 (1997).
5. D. A. B. Miller, "Rationale and challenges for optical interconnects to electronic chips," Proc. IEEE **88**, 728–749 (2000).
6. D. A. B. Miller, "Optical interconnects to silicon," IEEE J. Sel. Top. Quantum Electron. **6**, 1312–1317 (2000).
7. D. A. B. Miller, "Device requirements for optical interconnects to silicon chips," Proc. IEEE **97**, 1166–1185 (2009).
8. D. A. B. Miller, "Are optical transistors the next logical step?," Nat. Photon. **4**, 3–5 (2010).
9. K. C. Saraswat and F. Mohammadi, "Effect of scaling of interconnections on the time delay of VLSI circuits," IEEE Trans. Electron Devices **29**, 645–650 (1982).
10. R. Ho, K. W. Mai, and M. A. Horowitz, "The future of wires," Proc. IEEE **89**, 490–504 (2001).
11. "International technology roadmap for semiconductors," http://www.itrs.net/.
12. D. E. Atkins, K. K. Droegemeier, S. I. Feldman, H. Garcia-Molina, M. L. Klein, D. G. Messerschmitt, P. Messina, J. P. Ostriker, and M. H. Wright, *Final Report of the NSF Blue Ribbon Advisory Panel on Cyberinfrastructure: Revolutionizing Science and Engineering Through Cyberinfrastructure* (NSF, January 2003), http://www.nsf.gov/cise/sci/reports/atkins.pdf.
13. B. Kim and V. Stojanovic, "Characterization of equalized and repeated interconnects for NoC applications," IEEE Design Test Comput. **25**, 430–439 (2008).
14. W. R. Davis, J. Wilson, S. Mick, J. Xu, H. Hua, C. Mineo, A. M. Sule, M. Steer, and P. D. Franzon, "Demystifying 3D ICs: the pros and cons of going vertical," IEEE Design Test Comput. **22**, 498–510 (2005).
15. K.-H. Koo, H. Cho, P. Kapur, and K. C. Saraswat, "Performance comparisons between carbon nanotubes, optical, and Cu for future high-performance on-chip interconnect applications," IEEE Trans. Electron Devices **54**, 3206–3215 (2007).
16. H. Cho, K.-H. Koo, P. Kapur, and D. C. Saraswat, "Performance comparisons between Cu/low-κ, carbon-nanotube, and optics for future on-chip interconnects," IEEE Electron Device Lett. **29**, 122–124 (2008).
17. K.-N. Chen, M. J. Kobrinsky, B. C. Barnett, and R. Reif, "Comparisons of conventional, 3-D, optical, and RF interconnects for on-chip clock distribution," IEEE Trans. Electron Devices **51**, 233–239 (2004).
18. G. A. Keeler, B. E. Nelson, D. Agarwal, C. Debaes, N. C. Helman, A. Bhatnagar, and D. A. B. Miller, "The benefits of ultrashort optical pulses in optically-interconnected systems," IEEE J. Sel. Top. Quantum Electron. **9**, 477–485 (2003).
19. 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. Top. Quantum Electron. **9**, 400–409 (2003).
20. D. A. B. Miller, A. Bhatnagar, S. Palermo, A. Emami-Neyestanak, and M. A. Horowitz, "Opportunities for optics in integrated circuits applications," *International Solid State Circuits Conference, 2005* (IEEE, 2005), paper 4.6, pp. 86–87.
21. A. Shacham, K. Bergman, and L. P. Carloni, "Photonic networks-on-chip for future generations of chip multiprocessors," IEEE Trans. Comput. **57**, 1246–1260 (2008).
22. R. G. Beausoleil, P. J. Kuekes, G. S. Snider, S.-Y. Wang, and R. S. Williams, "Nanoelectronic and nanophotonic interconnect," Proc. IEEE **96**, 230–247 (2008).
23. A. V. Krishnamoorthy, R. Ho, X. Zheng, H. Schwetman, J. Lexau, P. Koka, G. L. Li, I. Shubin, and J. E. Cunningham, "Computer systems based on silicon photonic interconnects," Proc. IEEE **97**, 1337–1361 (2009).
24. A. H. Gnauck, G. Charlet, P. Tran, P. J. Winzer, C. R. Doerr, J. C. Centanni, E. C. Burrows, T. Kawanishi, T. Sakamoto, and K. Higuma, "25.6 Tb/s WDM transmission of polarization-multiplexed RZ-DQPSK signals," J. Lightwave Technol. **26**, 79–84 (2008).
25. N. Streibl, K.-H. Brenner, A. Huang, J. Jahns, J. Jewell, A. W. Lohmann, D. A. B. Miller, M. Murdocca, M. E. Prise, and T. Sizer, "Digital optics," Proc. IEEE **77**, 1954–1969 (1989).
26. F. B. McCormick, T. J. Cloonan, F. A. P. Tooley, A. L. Lentine, J. M. Sasian, J. L. Brubaker, R. L. Morrison, S. L. Walker, R. J. Crisci, R. A. Novotny, S. J. Hinterlong, H. S. Hinton, and E. Kerbis, "Six-stage digital free-space optical switching network using symmetric self-electro-optic-effect devices," Appl. Opt. **32**, 5153–5171 (1993).
27. B. E. Nelson, G. A. Keeler, D. Agarwal, N. C. Helman, and D. A. B. Miller, "Wavelength division multiplexed optical interconnect using short pulses," IEEE J. Sel. Top. Quantum Electron. **9**, 486–491 (2003).
28. A V. Krishnamoorthy and D. A. B. Miller, "Scaling optoelectronic-VLSI circuits into the 21st century: a technology roadmap," IEEE J. Sel. Top. Quantum Electron. **2**, 55–76 (1996).
29. A. K. Okyay, D. Kuzum, S. Latif, D. A. B. Miller, and K. C. Saraswat, "Silicon germanium CMOS optoelectronic switching device: bringing light to latch," IEEE Trans. Electron Devices **54**, 3252–3259 (2007).
30. K. W. Goossen, G. D. Boyd, J. E. Cunningham, W. Y. Jan, D. A. B. Miller, D. S. Chemla, and R. M. Lum, "GaAs-AlGaAs multiquantum well reflection modulators grown on GaAs and silicon substrates," IEEE Photonics Technol. Lett. **1**, 304–306 (1989).
31. R. A. Soref and B. R. Bennett, "Electrooptical effects in silicon," IEEE J. Quantum Electron. **23**, 123–129 (1987).
32. M. Lipson, "Compact electro-optic modulators on a silicon chip," IEEE J. Sel. Top. Quantum Electron. **12**, 1520–1526 (2006).
33. A. Liu, L. Liao, D. Rubin, J. Basak, Y. Chetrit, H. Nguyen, R. Cohen, N. Izhaky, and M. Paniccia, "Recent development in a high-speed silicon optical modulator based on reverse-biased pn diode in a silicon waveguide," Semicond. Sci. Technol. **23**, 064001 (2008).
34. Z. Huang, N. Kong, X. Guo, M. Liu, N. Duan, A. L. Beck, S. K. Banerjee, and J. C. Campbell, "21 GHz-Bandwidth germanium-on-silicon photodiode using thin SiGe buffer layers," IEEE J. Sel. Top. Quantum Electron. **12**, 1450–1454 (2006).
35. D. Ahn, C. Hong, J. Liu, W. Giziewicz, M. Beals, L. C. Kimerling, J. Michel, J. Chen, and F. X. Kärtner, "High performance, waveguide integrated Ge photodetectors," Opt. Express **15**, 3916–3921 (2007).
36. Y.-H. Kuo, Y.-K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. B. Miller, and J. S. Harris, "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon," Nature **437**, 1334–1336 (2005).
37. J. E. Roth, O. Fidaner, R. K. Schaevitz, Y.-H. Kuo, T. I. Kamins, J. S. Harris, and D. A. B. Miller, "Optical modulator on silicon employing germanium quantum wells," Opt. Express **15**, 5851–5859 (2007).
38. J. Liu, M. Beals, A. Pomerene, S. Bernardis, R. Sun, J. Cheng, L. C. Kimerling, and J. Michel, "Waveguide-integrated, ultra-low-energy GeSi electro-absorption modulators," Nat. Photon. **2**, 433–437 (2008).

39. G. T. Reed and A. P. Knights, *Silicon Photonics* (Wiley, 2004).
40. L. Pavesi and D. J. Lockwood, eds., *Silicon Photonics* (Springer-Verlag, 2004).
41. G. I. Yayla, P. J. Marchand, and S. C. Esener, "Speed and energy analysis of digital interconnections: comparison of on-chip, off-chip, and free-space technologies," Appl. Opt. **37**, 205–227 (1998).
42. D. A. B. Miller, D. S. Chemla, and S. Schmitt-Rink, "Relation between electroabsorption in bulk semiconductors and in quantum wells: the quantum-confined Franz–Keldysh effect," Phys. Rev. B **33**, 6976–6982 (1986).
43. B. E. Nelson, M. Gerken, D. A. B. Miller, R. Piestun, C.-C. Lin, and J. S. Harris, Jr., "Use of a dielectric stack as a one-dimensional photonic crystal for wavelength demultiplexing by beam shifting," Opt. Lett. **25**, 1502–1504 (2000).
44. T. Baba and M. Nakamura, "Photonic crystal light deflection devices using the superprism effect," IEEE J. Quantum Electron. **38**, 909–914 (2002).
45. K. Jia, J. Yang, Y. Hao, X. Jiang, M. Wang, W. Wang, Y. Wu, and Y. Wang, "Turning-mirror-integrated arrayed-waveguide gratings on silicon-on-insulator," IEEE J. Sel. Top. Quantum Electron. **12**, 1329–1334 (2006).
46. S. Zheng, H. Chen, and A. W. Poon, "Microring-resonator cross-connect filters in silicon nitride: rib waveguide dimensions dependence," IEEE J. Sel. Top. Quantum Electron. **12**, 1380–1387 (2006).
47. F. Horst, W. M. J. Green, B. J. Offrein, and Y. Vlasov, "Echelle grating WDM (de-)multiplexers in SOI technology based on a design with two stigmatic points," Proc. SPIE **6996**, 69960R (2008).
48. J. Brouckaert, W. Bogaerts, S. Sevaraja, P. Dumon, R. Baets, and D. Van Thourhout, "Planar concave grating demultiplexer with high reflective Bragg reflector facets," IEEE Photonics Technol. Lett. **20**, 309–311 (2008).
49. M. Gerken and D. A. B. Miller, "Multilayer thin-film stacks With steplike spatial beam shifting," J. Lightwave Technol. **22**, 612–618 (2004).
50. M. Gerken and D. A. B. Miller, "Limits to the performance of dispersive thin-film stacks," Appl. Opt. **44**, 3349–3357 (2005).
51. L. Tang, S. E. Kocabas, S. Latif, A. K. Okyay, D.-S. Ly-Gagnon, K. C. Saraswat, and D. A. B. Miller, "Nanometre-scale germanium photodetector enhanced by a near-infrared dipole antenna," Nat. Photon. **2**, 226–229 (2008).
52. L. Tang, S. Latif, and D. A. B. Miller, "Plasmonic device in silicon CMOS," Electron. Lett. **45**, 706–708 (2009).
53. I. Fushman, D. Englund, A. Faraon, N. Stolz, P. Petroff, and J. Vuckovic, "Controlled phase shifts with a single quantum dot," Science **320**, 769–772 (2008).
54. R. W. Keyes and J. A. Armstrong, "Thermal limitations in optical logic," Appl. Opt. **8**, 2549–2552 (1969).
55. P. W. Smith, "On the physical limits of digital optical switching and logic elements," Bell Syst. Tech. J. **61**, 1975–1993 (1982).
56. H. M. Gibbs, S. L. McCall, T. N. C. Venkatesan, A. C. Gossard, A. Passner, and W. Wiegmann, "Optical bistability in semiconductors," Appl. Phys. Lett. **35**, 451–453 (1979).
57. D. A. B. Miller, S. D. Smith, and A. Johnston, "Optical bistability and signal amplification in a semiconductor crystal. Application of new low-power nonlinear effects in InSb," Appl. Phys. Lett. **35**, 658–660 (1979).
58. H. M. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic, 1985).
59. J. Jahns and A. Huang, "Planar integration of free—space optical components," Appl. Opt. **28**, 1602–1605 (1989).
60. M. E. Prise, N. C. Craft, M. M. Downs, R. E. LaMarche, L. A. D'Asaro, L. M. F. Chirovsky, and M. J. Murdocca, "Optical digital processor using arrays of symmetric self-electrooptic effect devices," Appl. Opt. **30**, 2287–2296 (1991).
61. R. G. A. Craig, B. S. Wherrett, A. C. Walker, F. A. P. Tooley, and S. D. Smith, "Optical cellular logic image processor: implementation and programming of a single channel digital optical circuit," Appl. Opt. **30**, 2297–2308 (1991).
62. A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham, and L. M. F. Chirovsky, "Symmetric self-electro-optic effect device: optical set-reset latch," Appl. Phys. Lett. **52**, 1419–1421 (1988).
63. J. W. Goodman, F. J. Leonberger, S. Y. Kung, and R. A. Athale, "Optical interconnections for VLSI systems," Proc. IEEE **72**, 850–866 (1984).
64. K. Iga, F. Koyama, and S. Konoshita, "Surface emitting semiconductor lasers," IEEE J. Quantum Electron. **24**, 1845–1855 (1988).
65. J. L. Jewell, A. Scherer, S. L. McCall, Y. H. Lee, S. Walker, J. P. Harbison, and L. T. Florez, "Low-threshold electrically pumped vertical-cavity surface-emitting microlasers," Electron. Lett. **25**, 1123–1124 (1989).
66. D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus, "Electric field dependence of optical absorption near the bandgap of quantum well structures," Phys. Rev. B **32**, 1043–1060 (1985).
67. D. A. B. Miller, "Quantum-well self-electro-optic effect devices," Opt. Quantum Electron. **22**, S61–S98 (1990).
68. A. L. Lentine and D. A. B. Miller, "Evolution of the SEED technology: bistable logic gates to optoelectronic smart pixels," IEEE J. Quantum Electron. **29**, 655–669 (1993).
69. M. R. Feldman, S. C. Esener, C. C. Guest, and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations," Appl. Opt. **27**, 1742–1751 (1988).
70. H. Cho, P. Kapur, and K. C. Saraswat, "Power comparison between high-speed electrical and optical interconnects for interchip communication," J. Lightwave Technol. **22**, 2021–2033 (2004).
71. A. Naeemi, J. Xu, A. V. Mule, T. K. Gaylord, and J. D. Meindl, "Optical and electrical interconnect partition length based on chip-to-chip bandwidth maximization," IEEE Photonics Technol. Lett. **16**, 1221–1223 (2004).
72. A. V. Krishnamoorthy and K. W. Goossen, "Optoelectronic-VLSI: photonics integrated with VLSI circuits," IEEE J. Sel. Top. Quantum Electron. **4**, 899–912 (1998).
73. A. V. Krishnamoorthy, L. M. F. Chirovsky, W. S. Hobson, R. E. Leibenguth, S. P. Hui, G. J. Zydzik, K. W. Goossen, J. D. Wynn, B. J. Tseng, J. Lopata, J. A. Walker, J. E. Cunningham, and L. A. D'Asaro, "Vertical-cavity surface-emitting lasers flip-chip bonded to gigabit-per-second CMOS circuits," IEEE Photonics Technol. Lett. **11**, 128–130 (1999).
74. M. W. Haney, M. P. Christensen, P. Milojkovic, J. Ekman, P. Chandramani, R. Rozier, F. Kiamilev, Y. Liu, and M. Hibbs-Brenner, "Multichip free-space global optical interconnection demonstration with integrated arrays of vertical-cavity surface-emitting lasers and photodetectors," Appl. Opt. **38**, 6190–6200 (1999).
75. A. C. Walker, M. P. Y. Desmulliez, M. G. Forbes, S. J. Fancey, G. S. Buller, M. R. Taghizadeh, J. A. B. Dines, C. R. Stanley, G. Pennelli, A. R. Boyd, P. Horan, D. Byrne, J. Hegarty, S. Eitel, H. P. Gauggel, K. H. Gulden, A. Gauthier, P. Benabes, J. L. Gutzwiller, M. Goetz, and M. P. Y. Desmulliez, "Design and construction of an optoelectronic crossbar switch containing a terabit per second free-space optical interconnect," IEEE J. Sel. Top. Quantum Electron. **5**, 236–249 (1999).
76. M. B. Venditti, E. Laprise, J. Faucher, P.-O Laprise, J. Eduardo, A. Lugo, and D. V. Plant, "Design and test of an optoelectronic-vlsi chip with 540-element receiver-transmitter arrays using differential optical signaling," IEEE J. Sel. Top. Quantum Electron. **9**, 361–379 (2003).

77. P. Lukowicz, J. Jahns, R. Barbieri, P. Benabes, T. Bierhoff, A. Gauthier, M. Jarczynski, G. A. Russell, J. Schrage, W. Sullau, J. F. Snowdon, M. Wirz, and G. Troster, "Optoelectronic interconnection technology in the Holms system," IEEE J. Sel. Top. Quantum Electron. **9**, 624–635 (2003).
78. R. Barbieri, P. Benabes, T. Bierhoff, J. J. Caswell, A. Gauthier, J. Jahns, M. Jarczynski, P. Lukowicz, J. Oksman, G. A. Russell, J. Schrage, J. F. Snowdon, O. Stübbe, G. Troster, and M. Wirz, "Design and construction of the high-speed optoelectronic memory system demonstrator," Appl. Opt. **47**, 3500–3512 (2008).
79. F. E. Doany, C. L. Schow, C. W. Baks, D. M. Kuchta, P. Pepeljugoski, L. Schares, R. Budd, F. Libsch, R. Dangel, F. Horst, B. J. Offrein, and J. A. Kash, "160 Gb/s Bidirectional polymer-waveguide board-level optical interconnects using CMOS-based transceivers," IEEE Trans. Advanced Packaging **32**, 345–359 (2009).
80. A. W. Fang, H. Park, O. Cohen, R. Jones, M. J. Paniccia, and J. E. Bowers, "Electrically pumped hybrid AlGaInAs-silicon evanescent laser," Opt. Express **14**, 9203–9210 (2006).
81. J. Liu, X. Sun, R. Camacho-Aguilera, L. C. Kimerling, and J. Michel, "A Ge-on-Si laser operating at room temperature," Opt. Lett. **35**, 679–681 (2010).