Abstract—This paper gives a brief historical summary of the development of the field of optical interconnects to silicon integrated circuits. It starts from roots in early optical switching phenomena, proceeds through novel semiconductor and quantum well optical and optoelectronic physics and devices, first proposals for optical interconnects, and optical computing and photonic switching demonstrators, to hybrid integrations of optoelectronic and silicon circuits that may solve basic scaling and other problems for interconnections in future information processing and switching machines.

The idea of using light beams to replace wires now dominates all long-distance communications and is progressively taking over in networks over shorter distances. But should we use light beams at the much shorter distances inside digital computers, possibly connecting directly to the silicon chips, or even for connections on chips? If so, why, and, ultimately, when?

In this short paper, I will try to summarize the history of how we got to the point where we are at now: so that others may more than compensate for the errors of today, and, incidentally, generated some important spinoffs.

The history of optical interconnects is strongly intertwined with the broader field of optics in digital computing and switching. With the benefit of hindsight, aspects of optical computing have been misguided or even naïve; that field, however, provided much of the impetus for key technical innovations today, and, incidentally, generated some important spinoffs along the way that may more than compensate for the errors of the field.

Modern thinking about the use of optics in digital computing dates from the 1960s. The laser was invented. The semiconductor diode laser in its earliest form had been demonstrated. Nonlinear optics was an exciting emerging field. The idea of using lasers or other nonlinear optical devices as the logic devices in computing was examined. The conclusion of these first analyses was that optical logic devices were not a good substitute for transistors in a general-purpose computing machine (see, e.g., [1]). Their biggest problem would be that they would take too much energy. The ideas of optics for communication were being developed also at that time. The notion that the best use of optics inside computers might be for making connections was not seriously considered, perhaps because wires were still capable of delivering the performance needed.

The optics research community continued to be interested in the idea of optical switching, driven by the idea that optics could make switches that would be much faster than any electrical transistor. This particular claim remains valid; nonlinear optics is capable of making logic devices much faster than any electrical device. Much of the work was in the field of optical bistability, which flourished from about the mid 1970s to the late 1980s [2].

The observation of optical bistability in semiconductors [3], [4] increased the interest in the field, because semiconductors offered more nearly practical devices than some of the other systems being investigated (e.g., sodium vapor). Smith [5] published a more optimistic analysis of optical logic devices in which he argued that the use of resonator structures could reduce the operating energy below electronic devices. The nonlinear effects in semiconductors became relatively well understood [6]. These semiconductor bistable devices were very simple, with some (e.g., [4]) requiring only a semiconductor material with plane parallel surfaces. As a result, semiconductors offered the possibility that large parallel arrays of devices might be made; a simple slab of semiconductor might be sufficient. In the early 1980s, it was found that there were particular enhanced optical properties in semiconductor quantum well structures [7], [8]. In particular, quantum wells showed very strong peaks (excitonic peaks) in their absorption spectra at room temperature, and

Invited Paper
relatively easy saturation of those peaks. This led to further interest in semiconductor optical switching devices.

One could now imagine planes of devices communicating with large arrays of light beams from one plane to another, with the information flowing perpendicular to the plane. The numbers that could be connected this way would be much larger than the number of wires of electrical busses. Here we see the emergence of an interconnection advantage as a reason for the use of optics in computing. At that time, however, there was no practical way of getting information in the form of light from electrical circuits at high densities and high speeds, so the discussion was still restricted to interconnecting optical logic devices. There was also early work, using liquid crystal spatial light modulators, to demonstrate elementary optical computers, that exploited the two-dimensional (2-D) parallelism of optics, though the devices available did not offer any serious possibility of making a practical system [10].

The idea of optical interconnection of very large scale integration (VLSI) electronics was proposed and analyzed in a seminal paper by Goodman, et al. in 1984 [11]. This paper pointed out several basic reasons why optical interconnection might be interesting and important, successfully anticipated many subsequent developments, and was arguably the start of the field of optical interconnects proper.

At that time, there still was no good way of getting large numbers of optical outputs from a silicon integrated circuit. Silicon itself, because of its indirect band gap, is fundamentally a very poor light emitter and not a very good light modulator. There has been much subsequent research on making silicon light emitters. Much recent interest (see, e.g., [12]) was stimulated by relatively efficient incoherent emission from porous silicon [13]. Incoherent emission is, however, likely not sufficient for dense, high-speed interconnects, mostly because of the basic optical inefficiencies in focusing incoherent light4 [14]. The use of silicon for efficient optical output devices has remained a very stubborn problem. III–V materials (e.g., GaAs, InGaAs) remain the only viable ones for semiconductor light emission or high-performance modulation. In the early 1980s, the optoelectronic output devices available [III–V (incoherent) light emitting diodes, and III–V edge-emitting lasers] still had relatively high power dissipations, and there was no demonstrated technology for integrating them in large numbers with silicon integrated circuits. The edge-emitting geometry of the lasers was also not attractive for use of any free-space imaging optics; to this day, the packaging of any optics with edge-emitting lasers is a difficult problem.

In 1984, a new optical modulation mechanism, the quantum-confined Stark effect [15], was discovered in III–V semiconductor quantum wells. This effect allowed efficient high-speed optical modulators. The first such modulator (and also the structure used to measure the physical effects) is shown in Fig. 1. Integrated waveguide versions of such modulators are extensively used today for high-speed low-chirp modulation in telecommunication systems (see, e.g., [16]). This effect was important for both optical computing and optical interconnects for three reasons: 1) it allowed low energy devices; 2) the changes in absorption were large enough to make a modulator or switch that would work for light beams propagating perpendicular to the surface, allowing 2-D arrays; 3) it had high enough yields to allow large arrays of devices to be made (e.g., thousands). When used with optical inputs controlling optical outputs, the resulting devices were called self-electrooptic effect devices (SEEDs) [17], [18]. Initial work on using such modulator arrays concentrated on their use as optical logic devices, partly because it was not clear how to integrate them with electronics and also because there was the hope that the parallelism of optical arrays would make them interesting in practical applications even if they were logically simpler than electronic systems.

Simple optical bistability was found to be impractical for making optical logic systems of any substantial size. Simple optically bistable devices are examples of “two-terminal” logic devices, known to the electronic community to be troublesome for large systems [19]. As a result, work on optical computers based on such simple bistability largely died out in the late 1980s. The semiconductor nonlinear absorption effects that were researched for many of these earlier bistable devices did, however, find use in modelockers for short pulse lasers (see, e.g., [20]–[22]).

A variant of the SEED, the symmetric SEED (S-SEED), though bistable, was technically a “three-terminal” device by virtue of being bistable in the ratio of two beam powers5 [23].

4If there is no predictable phase relationship between light emission from different points on a source, there is no way of persuading the light from these different parts to provide the constructive and destructive interference required to focus to a very small spot. Small spots are important because then we can use small photodetectors. Small photodetectors have small capacitance, which means larger voltage swings for a given optical input power (and/or higher speed). Higher voltage swings mean less amplification. Less amplification means less power dissipation, more noise immunity, and less latency. This difficulty of focusing incoherent sources is not merely a technological one; it is protected by the second law of thermodynamics (known in optics as the constant brightness theorem). It is not possible to take the light from two cooler black bodies and combine it with some linear optics so as to heat up a warmer black body without violating the second law of thermodynamics. Attempts to get round the focusing difficulties of incoherent sources tend also to permit this violation of the second law.

5These two-beam devices rely on “time-sequential” gain. Basically, the bias power is turned down to a low value, the device is able to be switched by small absolute powers in this condition, and the bias powers are then turned up to allow the device to be read out at high power, giving an effective signal gain in the switching.
This device allowed relatively substantial processors to be constructed. A team led by Huang [24] at Bell Laboratories demonstrated a simple but complete computer [25] with four arrays of 32 S-SEED devices. This work also led to substantial efforts on architectures and on optics for handling arrays of light beams (see, e.g., [26]). A related team led by Hinton, also at Bell Laboratories, constructed experimental switching machines for telecommunications routing. This work, which overlapped in time with the optical computing work, led to several generations of experimental systems [27], including the largest free-space digital optical system built so far [28], with over 60,000 light beams. This system is shown in Fig. 2. This working system demonstrated the feasibility of dense free-space optical array connections. It should be noted that, during this same time period of the late 1980s to early 1990s, there were several other important efforts and devices in such digital optics. These included the group at Heriot-Watt University (see, e.g., [29]) that had remained active from the earliest semiconductor bistability observations [4], and work by others on other optoelectronic logic devices based on light emission during this time and beyond (see, e.g., [30]–[32]).

The work on the computing and switching systems exposed two issues with digital optical systems. 1) To be efficient in overall performance, the computing and switching systems both appeared to want greater logical complexity than a very simple logic function between the optical inputs and outputs. This lead to the notion of “smart pixels”—devices or units with optical inputs and outputs, but with significant functional complexity between those. 2) Complex optical systems tended to have loss that was not negligibly small, so the available powers to switch the next device in the system were less than originally hoped. This meant that either more optical power or even more sensitive devices were required. The SEEDs were already much more sensitive than other optoelectronic devices available at the time. Hence, some additional gain mechanism was required. A solution to this was to incorporate electronic gain. Both of these issues could be solved by integrating electronic devices between the optical inputs and outputs. One of the first devices to explore this solution was a monolithic integration of GaAs field effect transistors and quantum well modulators and detectors (FET-SEEDs) [33]. This approach was also successfully incorporated into a substantial system demonstration [34]. One of the “smart pixel” units in this system is shown in Fig. 3. These are likely the most complex monolithic optically interconnected systems demonstrated to date.

During this period when much of digital optical computing was moving toward optical interconnects, there was continued analysis of optical interconnects. For example, engineering analysis showed specific energy dissipation benefits of optical interconnects [35]. Other physical analysis [36] showed that one reason for potential energy saving in optical interconnects was that optical devices, being fundamentally quantum devices, could act as efficient impedance transformers between the high impedance of small electronic devices and the low impedance of classical wave propagation. The first experiments on the use of optics for clock distribution were also performed [37].

Another important development during this time was the demonstration of viable vertical-cavity surface-emitting lasers (VCSELS). These had originally been pioneered in the group of Iga [38]. Jewell (in Huang’s optical computing group at Bell Laboratories) and collaborators had been working to make resonant-cavity semiconductor optical switching devices. In an extension of this work, they were able to demonstrate the first electrically pumped VCSEL at room temperature [39]. There has since been much work in turning these into practical devices, especially for low-cost optical fiber connections, where they now dominate such systems. They are also candidate devices for optical interconnects to silicon chips, though they lag behind the quantum well modulators in demonstrations of dense interconnects. (See, e.g., [40] for a critical comparison.
of optical output devices for optical interconnects to silicon chips.) In dense interconnects, low power dissipation will likely be critical. An important development in low-power VCSELs was the demonstration of oxide-confined devices for potentially much lower thresholds [41].

With digital optics and optical interconnects converging, a logical next step was to proceed to interconnecting silicon electronics. The FET-SEED work did manage to demonstrate high-performance integrated systems, but it required a custom GaAs electronic technology. Silicon VLSI technology continued to advance, and the idea of being able to join with the mainstream technology was appealing. It also meant that the optics would then no longer be competing against silicon technology itself (a battle optics had lost many times with general-purpose silicon replacing special purpose analog optical processors), but rather against the metal interconnect technologies. With optics joined to silicon VLSI, as silicon transistors became better, so also would the optical connections.

The monolithic integration of optical emitters with silicon has been, and continues to be, a substantial basic challenge; III–V emitters monolithically integrated to silicon during growth typically have problems with lifetimes because of progressive propagation of the crystal defects normally inherent in such lattice-mismatched growth. Quantum well modulators were quite successfully grown on silicon substrates, however, with apparently good lifetimes [42] (a fact that is perhaps less well known than it ought to be). Despite this success, and subsequent success integrating with actual circuits [43], these workers soon realized that the problem of integration with silicon VLSI was not really solved in a practical sense. The issue is that in such a practical monolithic integration the VLSI process is changed in some ways. Additional high-temperature cleaning steps are needed. The circuits must be made on substrates that are cut off-axis. Considerable care must be taken with extra protection steps to prevent any of the silicon oxide in the circuit from being exposed to gallium metal because gallium makes silicon oxide conducting. Most important from a practical point of view is that the silicon production line will likely not accept the wafer back into the line, e.g., for deposition of final metal levels, if there is gallium on the wafer.

The practical solution to integration, at least in the short to medium term, is therefore some hybrid technique that can be performed on the finished silicon VLSI wafers. Various techniques have been proposed and demonstrated (see, e.g., [40]). Solder bonding has been particularly successful. Large arrays (e.g., thousands) of quantum well modulator/photodiode devices have been successfully bonded to functioning silicon complementary metal–oxide–semiconductor (CMOS) VLSI circuits [44], [45]. Fig. 4 shows one of the first of these bonded circuits as used in photonic switching experiments. More recently, smaller arrays of VCSELs have also been successfully bonded to silicon circuits [46].

During the mid to late 1990s, the scaling arguments for why optical interconnects would be interesting were further clarified. One analysis [47], for example, showed that a likely limit on future dense optical interconnects to silicon CMOS would be the dissipation of receiver circuits, but that, because of the continuing improvement of the silicon transistors themselves, the optical interconnects would likely be able to scale almost to match the growth of the ability of silicon circuits to perform logic. Also during this time period, it was becoming increasingly clear to the silicon semiconductor industry that electrical interconnects were beginning to run into serious scaling limitations. A basic scaling problem had been understood for some time, which is that, as an electrical line is scaled down in all three dimensions, its resistive–capacitive time constant does not change at all (see, e.g., [48] for early electrical discussions). This is, to say the least, undesirable, because it means that, since the transistors get faster as they are scaled down, the wires do not scale to keep up with the transistors.

The analysis of the scaling of electrical connections was extended to include other kinds of electrical lines, with the relatively general conclusion that there is an underlying scaling to simple digital connections at any length scale [49]. This scaling concludes essentially that once an electrically connected architecture runs into speed limitations from the interconnect, and once all the available space is filled with wiring, the performance of the system cannot be improved either by miniaturizing the whole system or making larger; it is necessary to change the interconnect technology. This limit is known as the “architectural aspect ratio” limit [49]. Optical interconnects avoid this problem altogether because they do not have the resistive loss physics that gives rise to this behavior.

In the 1990s, the Semiconductor Industry Association had begun to publish “roadmaps” for performance and technologies for future generations of silicon technology [50]. These roadmaps show substantial problems for interconnects on silicon chips, with no known solution after approximately 2006 for interconnects to keep pace with the desired performance. Interconnects off of the chip, with their greater lengths, are likely a much larger problem. The issue of where optics might make sense as a solution to these problems has recently been analyzed in some detail [51], with the conclusion that optics is a very attractive solution on physical grounds to solve off-chip interconnect problems, and with some possibilities for use on the chips.
also. There is apparently no physical solution other than optics to the problem of dense off-chip interconnect.

The idea of changing the technology used for interconnects to silicon chips in computing and switching systems is a very radical one. It is important to understand exactly what the physical benefits are of doing that. The physical benefits have been codified in [52], and we will not repeat that classification here. Some of the resulting possible practical benefits of optical interconnects are as follows [51].

1) Design Simplification:
   a) Absence of electromagnetic wave phenomena (impedance matching, crosstalk, and inductance difficulties such as inductive voltage drops on pins and wires).
   b) Distance independence of performance of optical interconnects.
   c) Frequency independence of optical interconnects.

2) Architectural Advantages:
   a) Larger synchronous zones.
   b) Architectures with large numbers of long high-speed connections (i.e., avoiding the “architectural aspect ratio” scaling limit of electrical interconnects).
   c) Regular interconnections of large numbers of crossing “wires” (useful, for example, in some switching and signal processing architectures).
   d) Ability to have 2-D interconnects directly out of the area of the chip rather than from the edge.
   e) Avoidance of the necessity of an interconnect hierarchy (once in the form of light, the signal can be sent very long distances without changing its form or amplifying or reshaping it, even on physically thin connections).

3) Timing:
   a) Predictability of the timing of signals.
   b) Precision of the timing of the clock signal.
   c) Removal of timing skew in signals.
   d) Reduction in power and area for clock distribution.

4) Other Physical Benefits:
   a) Reduction of power dissipation in interconnects.
   b) Voltage isolation.
   c) High interconnect density, especially for longer off-chip interconnects.
   d) Possible noncontact, parallel testing of chip operation.
   e) Option of use of short optical pulses for synchronization and improved circuit performance.
   f) Possibility of wavelength division multiplexed interconnects without the use of any electrical multiplexing circuitry.

This short paper has tried to lay out a history of the development of ideas for optical interconnects to silicon chips. It is not a comprehensive review by any means of that field. One recent review is given in [52]. There has been much work in the 1990s on various demonstration systems that we have not reviewed here. The reader is also referred to the recent Special Issue of the PROCEEDINGS OF THE IEEE [53], which contains a much more comprehensive discussion of optical interconnects in digital systems, including optical and device approaches not discussed in detail here (e.g., waveguided systems). A review of proceedings from Optical Computing and Optics in Computing conferences over the last decade will show an increasing enthusiasm for optical interconnects, with this topic now arguably dominating the field.

With the idea of optical interconnects, the prognosis for the use of optics in digital computing and switching is more optimistic and realistic now than it likely has been at any point in the past. There are many challenges, including difficult issues such as packaging, cost, and reliability, and simply getting used to joining silicon circuits with optical links (e.g., receiver circuit issues). There is, however, a growing acceptance and even hope in the mainstream electronic community that optics may be successfully incorporated at some point in the not too distant future. That hope is doubtless shared by the optics community, which has shown an unflagging optimism that the days of optics as a major part of information processing would finally arrive.

ACKNOWLEDGMENT

The author is very pleased to acknowledge the collaboration and support of all of his colleagues at Bell Labs, where much of this work was performed, especially P. J. Anthony, W. Brinkman, D. S. Chemla, A. M. Glass, H. S. Hinton, A. Huang, H. Kogelnik, C. V. Shank, the extraordinarily talented groups of people that they led, and all of the people whom the author had the honor to lead.

REFERENCES

MILLER: OPTICAL INTERCONNECTS TO SILICON


