

## Silicon Photonics Stumbles at the Last Meter

We have fiber to the home, but fiber to the processor is still a problem

By **Anthony F. J. Levi**

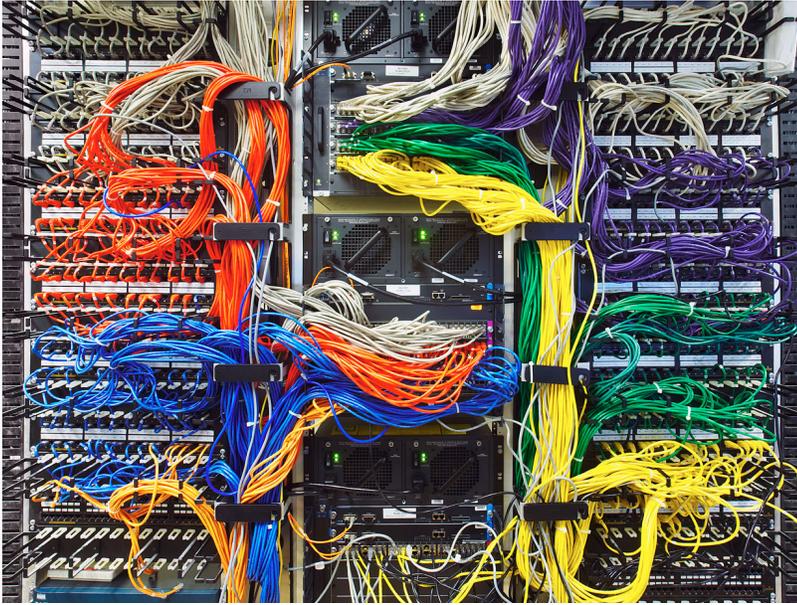


Photo: Getty Images

racks of computer servers inside data centers using optical fibers. And indeed, today, those photonic devices link racks in many modern data centers. But that is where the photons stop. Within a rack, individual server boards are still connected to each other with inexpensive copper wires and high-speed electronics. And, of course, on the boards themselves, it's metal conductors all the way to the processor.

Attempts to push the technology into the servers themselves, to directly feed the processors with fiber optics, have foundered on the rocks of economics. Admittedly, there is an Ethernet optical transceiver market of close to US \$4 billion per year that's set to grow to nearly \$4.5 billion and 50 million components by 2020, according to market research firm [LightCounting](https://www.lightcounting.com/index.cfm) (<https://www.lightcounting.com/index.cfm>). But photonics has never cracked those last few meters between the data-center computer rack and the processor chip.

Nevertheless, the stupendous potential of the technology has kept the dream alive. The technical challenges are still formidable. But new ideas about how data centers could be designed have, at last, offered a plausible path to a photonic revolution that could help tame the tides of big data.

---

## Inside a Photonics Module

Photo: Luxtera

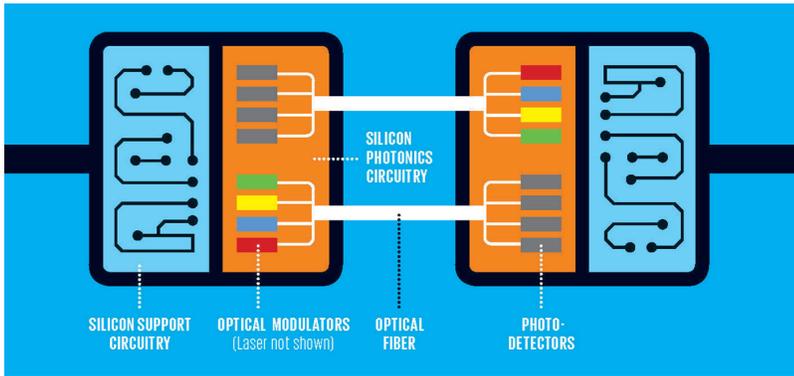
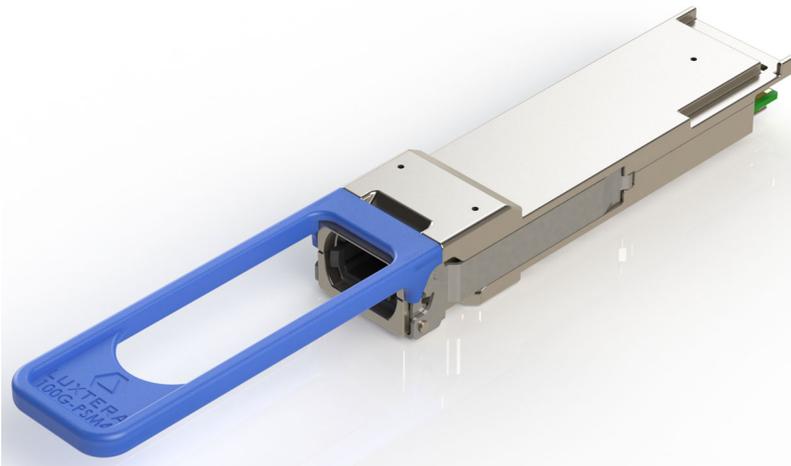
**Plug and Play:** A silicon photonics module converts electronic data to photons and back again. Silicon circuits [light blue] help optical modulators [bottom row, left] encode electronic data into pulses of several colors of light. The light travels through an optical fiber to another module, where photodetectors [gray] turn the light back into electronic bits. These are processed by the silicon circuits and sent on to the appropriate servers.

**If you think** we're on the cusp of a technological revolution today, imagine what it felt like in the mid-1980s. Silicon chips used transistors with micrometer-size features. Fiber-optic systems were zipping trillions of bits per second around the world.

With the combined might of silicon digital logic, optoelectronics, and optical-fiber communication, anything seemed possible.

Engineers envisioned all of these advances continuing and converging to the point where photonics would merge with electronics and eventually replace it. Photonics would move bits not just across countries but inside [data centers \(/computing/hardware/a-view-to-the-cloud\)](#), even inside computers themselves. Fiber optics would move data from chip to chip, they thought. And even those chips would be photonic: Many expected that someday blazingly fast logic chips would operate using photons rather than electrons.

It never got that far, of course. [Companies and governments plowed hundreds of millions of dollars](#) (<http://www.aimphotonics.com/>) into developing new photonic components and systems that link together



**Anytime you access** the Web, stream television, or do nearly anything in today's digital world, you are using data that has flowed through photonic transceiver modules. The job of these transceivers is to convert signals back and forth between electrical and optical. These devices live at each end of the optical fibers that speed data within the data centers of every major cloud service and social media company. The devices plug into switchgear at the top of each server rack, where they convert optical signals to electrical ones for delivery to the group of servers in that rack. The transceivers also convert data from those servers to optical signals for transport to other racks or up through a network of switches and out to the Internet.

Each photonics transceiver module has three main kinds of components: a transmitter containing one or more optical modulators, a receiver containing one or more photodiodes, and CMOS logic chips to encode and decode data. Because ordinary silicon is actually lousy at emitting light, the photons come from a laser that's separate from the silicon chips (though it can be housed in the same package with them). Rather than switch the laser on and off to represent bits, the laser is kept on, and electronic bits are encoded onto the laser light by an optical modulator.

This modulator, the heart of the transmitter, can take a few forms. A particularly nice and simple one is called the Mach-Zehnder modulator. Here, a narrow silicon waveguide channels the laser's light. The guide then splits in two, only to rejoin a few millimeters later. Ordinarily, this diverging and converging wouldn't affect the light output, because both branches of the waveguide are the same length. When they join up, the light waves are still in phase with each other. However, voltage applied to one of the branches has the effect of changing that branch's index of refraction, effectively slowing down or speeding up the light's wave. Consequently, when light waves from the two branches meet up again, they destructively interfere with each other and the signal is suppressed. So, if you vary a voltage on that branch, what you're actually doing is using an electrical signal to modulate an optical one.

The receiver is much simpler; it's basically a photodiode and some supporting circuitry. After traveling through an optical fiber, light signals reach the receiver's germanium or silicon germanium photodiode, which produces a current with each pulse of light.

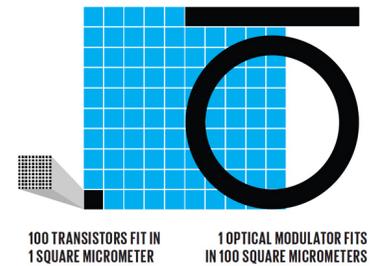
Both the transmitter and receiver are backed up by circuitry that does amplification, packet processing, error correction, buffering, and other tasks to comply with the Gigabit Ethernet standard for optical fiber (<https://ieeexplore.ieee.org/document/737786/>). How much of this is on the same chip as the photonics, or even in the same package, varies according to the vendor, but most of the electronic logic is separate from the photonics.

**With optical components** on silicon integrated circuits becoming increasingly available, you might be tempted to think that the integration of photonics directly into processor chips was inevitable. And indeed, for a time it seemed so. [See "[Linking With Light \(/semiconductors/optoelectronics/linking-with-light\)](https://www.spectrum.com/story/semiconductors/optoelectronics/linking-with-light)," *IEEE Spectrum*, October 2001.]

([/image/MzExOTY0Nw.jpeg](https://www.spectrum.com/story/semiconductors/optoelectronics/linking-with-light))

**Photonics Fail:** Photonics will never be a real option to transport data from one part of a silicon chip to another. A single optical switch, a ring oscillator in this case, performs the same function as an individual transistor, but it takes up 10,000 times as much area.

You see, what had been entirely underestimated, or even ignored, was the growing mismatch between how quickly the minimum size of features on electronic logic chips was shrinking and how limited photonics was in its ability to keep pace. Transistors today are made up of features only a few nanometers in dimension. In 7-nanometer CMOS technology, more than 100 transistors for general-purpose logic can be packed onto every square micrometer of a chip. And that's to say nothing of the maze of complex copper wiring above the transistors. In addition to the billions of transistors on each chip, there are also a dozen or so levels of metal interconnect needed to wire up all those transistors into the registers, multipliers, arithmetic logic units, and more complicated things that make up processor cores and other crucial circuits.



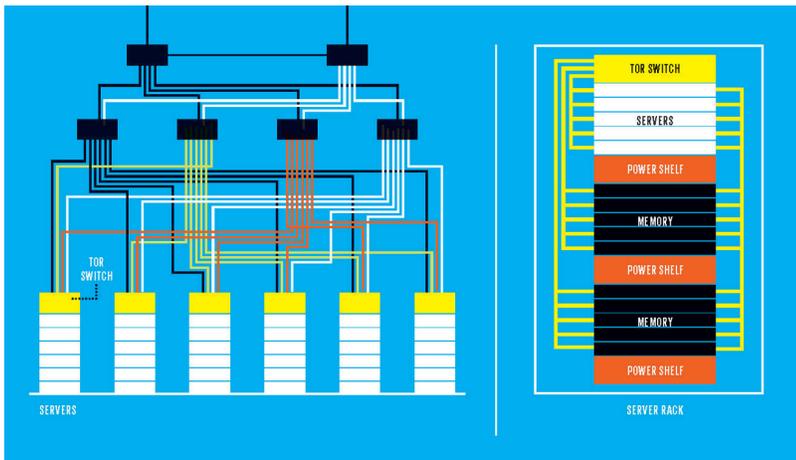
The trouble is that a typical photonic component, such as a modulator, can't be made much smaller than the wavelength of the light it's going to carry, limiting it to about 1 micrometer wide. There is no Moore's Law that can overcome this. It's not a matter of using more and more advanced lithography. It's simply that electrons—having a wavelength on the order of few nanometers—are skinny, and photons are fat.

But, still, couldn't chipmakers just integrate the modulator and accept that the chip will have fewer transistors? After all, a chip can now have billions of them. Nope. The massive amount of system function that each square micrometer of a silicon electronic chip area can deliver makes it very expensive to replace even relatively few transistors with lower-functioning components such as photonics.

Here's the math. Say there are on average 100 transistors per square micrometer. Then a photonic modulator that occupies a relatively small area of 10µm by 10µm is displacing a circuit comprising 10,000 transistors! And recall that a typical photonic modulator acts as a simple switch, turning light on and off. But each individual transistor can act as a switch, turning current on and off. So, roughly speaking, the opportunity cost for this primitive function is 10,000:1 *against* the photonic component because there are at least 10,000 electronic switches available to the system designer for every one photonic modulator. No chipmaker is willing to accept such a high price, even in exchange for the measurable improvements in performance and efficiency you might get by integrating the modulators right onto the processor.

The idea of substituting photonics for electronics on chips encounters other snags, too. For example, there are critical on-chip functions, such as memory, for which photonics has no comparable capability. The upshot is that photons are simply incompatible with basic computer chip functions. And even when they are not, integrating a competing photonic function on the same chip as electronics makes no sense.

## Data-Center Design



**Today (Left):** Photonics slings data through a multitered network in the data center. The link to the Internet is at the top (core) level. Switchgear moves data via optical fibers to top-of-rack (TOR) switches, which sit atop each rack of servers.

**Tomorrow (Right):** Photonics could facilitate a change in data-center architecture. Rack-scale architecture would make data centers more flexible by physically separating computers from their memory resources and connecting them through an optical network.

**That's not to say** photonics can't get a lot closer to processors, memory, and other key chips than it does now. Today, the market for optical interconnects in the data center focuses on systems called top-of-rack (TOR) switches, into which the photonic transceiver modules are plugged. Here at the top of 2-meter tall racks that house server chips, memory, and other resources, fiber optics link the TORs to each other via a separate layer of switches. These switches, in turn, connect to yet another set of switches that form the data center's gateway to the Internet.

The faceplate of a typical TOR, where transceiver modules are plugged in, gives a good idea of just how much data is in motion. Each TOR is connected to one transceiver module, which is in turn connected to two optical fibers (one to transmit and one to receive). Thirty-two modules, each with 40-gigabit-per-second data rates in each direction, can be plugged into a TOR's 45-millimeter-high faceplate, allowing for as many as 2.56 terabits per second to flow between the two racks.

But the flow of data within the rack and inside the servers themselves is still done using copper wires. That's unfortunate, because they are becoming an obstacle to the goal of building faster, more energy-efficient systems. Photonic solutions for this last meter (or two) of interconnect—either to the server or even to the processor itself—represent possibly the best opportunity to develop a truly high-volume optical component market. But before that can happen, there are some serious challenges to overcome in both price and performance.

So-called fiber-to-the-processor schemes are not new. And there are many lessons from past attempts about cost, reliability, power efficiency, and bandwidth density. About 15 years ago, for example, I contributed to the design and construction of an [experimental transceiver that showed very high bandwidth \(https://dx.doi.org/10.1109/JLT.2004.833251\)](https://dx.doi.org/10.1109/JLT.2004.833251). The demonstration sought to link a parallel fiber-optic ribbon, 12 fibers wide, to a processor. Each fiber carried digital signals generated separately by four vertical-cavity surface-emitting lasers (VCSELs)—a type of laser diode that shines out of the surface of a chip and can be produced in greater density than so-called edge-emitting lasers. The four VCSELs directly encoded bits by turning light output on and off, and they each operated at different wavelengths in the same fiber, quadrupling that fiber's capacity using what's called coarse wavelength-division multiplexing. So, with each VCSEL streaming out data at 25 Gb/s, the total bandwidth of the system would be 1.2 Tb/s. The industry standard today for the spacing between neighboring fibers in a 12-fiber-wide array is 0.25 mm, giving a bandwidth density of about 0.4 Tb/s/mm. In other words, in 100 seconds each millimeter could handle as much data as the U.S. Library of Congress's Web Archive team stores in a month.

Data rates even higher than this are needed for fiber-to-the-processor applications today, but it was a good start. So why wasn't this technology adopted? Part of the answer is that this system was neither sufficiently reliable nor practical to manufacture. At the time, it was very difficult to make the needed 48 VCSELs for the transmitter and guarantee that there would be no failures over the transmitter lifetime. In fact, an important lesson was that one laser using many modulators can be engineered to be much more reliable than 48 lasers.

But today, VCSEL performance has improved to the extent that transceivers based on this technology could provide effective short-reach data-center solutions. And those fiber ribbons can be replaced with multicore fiber, which carries the same amount of data by channeling it into several cores embedded within the main fiber. Another recent, positive development is the availability of more complex digital-transmission standards such as [PAM4 \(https://www.edn.com/design/systems-design/4441212/The-fundamentals-of-PAM4\)](https://www.edn.com/design/systems-design/4441212/The-fundamentals-of-PAM4), which boosts data transmission rates because it encodes bits on four intensities of light rather than just two. And research efforts, such as [MIT's Shine \(https://arpa-e.energy.gov/sites/default/files/documents/files/ARPAE\\_ENLITENED\\_Project\\_Descriptions\\_FINAL.pdf\)](https://arpa-e.energy.gov/sites/default/files/documents/files/ARPAE_ENLITENED_Project_Descriptions_FINAL.pdf) program, are working toward bandwidth density in fiber-to-the-processor to demonstration systems with about 17 times what we achieved 15 years ago.

These are all major improvements, but even taken together they are not enough to enable photonics to take the next big leap toward the processor. However, I still think this leap can occur, because of a drive, just now gathering momentum, to change data-center system architecture.

Today processors, memory, and storage make up what's called a server blade, which is housed in a chassis in a rack in the data center. But it need not be so. Instead of placing memory with the server chips, memory could sit separately in the rack or even in a separate rack. [This rack-scale architecture \(https://www.intel.com/content/www/us/en/architecture-and-technology/rack-scale-design-overview.html\)](https://www.intel.com/content/www/us/en/architecture-and-technology/rack-scale-design-overview.html) (RSA) is thought to use computing resources more efficiently, especially for social media companies such as Facebook where the amount of computing and memory required for specific applications grows over time. It also simplifies the task of replacing and managing hardware.

Why would such a configuration help enable greater penetration by photonics? Because exactly that kind of reconfigurability and dynamic allocation of resources could be made possible by a new generation of efficient, inexpensive, multi-terabit-per-second optical switch technology.

**The main obstacle** to the emergence of this data-center remake is the price of components and the cost of their manufacture. Silicon photonics already has one cost advantage, which is that it can leverage existing chip manufacturing, taking advantage of silicon's huge infrastructure and reliability. Nevertheless, silicon and light are not a perfect fit: Apart from their crippling inefficiency at emitting light, silicon components suffer from large optical losses as well. As measured by light in to light out, a typical silicon photonic transceiver experiences greater than a 10-decibel (90 percent) optical loss. This inefficiency does not matter much for short-reach optical interconnects between TOR switches because, at least for now, the silicon's potential cost advantage outweighs that problem.

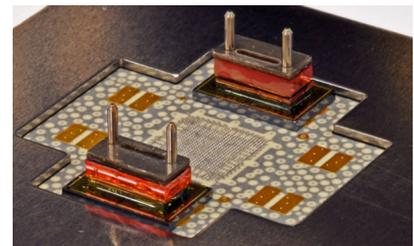


Photo: Anthony F.J. Levi

**Past Perfect:** We've had the technology to bring optical fiber directly to the processor for more than a decade. The author helped conceive this 0.4-terabit-per-second-per-millimeter demonstrator more than 15 years ago.

An important cost in a silicon photonics module is the humble, yet critically important, optical connection. This is both the physical link between the optical fiber and the transmitter or receiver chip as well as the link between fibers. Many hundreds of millions of such fiber-to-fiber connectors must be manufactured each year with extreme precision. To understand just how much precision, note that the diameter of a human hair is typically a little less than the 125- $\mu\text{m}$  diameter of a single-mode silica glass fiber used for optical interconnects. The accuracy with which such single-mode fibers must be aligned in a connector is around 100 nm—about one one-thousandth the diameter of a human hair—or the signal will become too degraded. New and innovative ways to manufacture connectors between fibers and from fiber to transceiver are needed to meet growing customer demand for both precision and low component price. However, very few manufacturing techniques are close to being inexpensive enough.

One way to reduce cost is, of course, to make the chips in the optical module cheaper. Though there are other ways to make these chips, a technique called wafer-scale integration could help. Wafer-scale integration means making photonics on one wafer of silicon, electronics on another, and then attaching the wafers. The paired wafers are then diced up into chips designed to be nearly complete modules. (The laser, which is made from a semiconductor other than silicon, remains separate.) This approach cuts manufacturing costs because it allows for assembly and production in parallel.

Another factor in reducing cost is, of course, volume. Suppose the total optical Gigabit Ethernet market is 50 million transceivers per year and each photonic transceiver chip occupies an area of 25 square millimeters. Assuming a foundry uses 200-mm-diameter wafers to make them and that it achieves a 100 percent yield, then the number of wafers needed is 42,000.

That might sound like a lot, but that figure actually represents less than two weeks of production in a typical foundry. In reality, any given transceiver manufacturer might capture 25 percent of the market and still support only a few days of production. There needs to be a path to higher volume if costs are really going to fall. The only way to make that happen is to figure out how to use photonics below the TOR switch, all the way to the processors inside the servers.

**If silicon photonics** is ever going to make it big in what are otherwise all-electronic systems, there will have to be compelling technical and business reasons for it. The components must solve an important problem and greatly improve the overall system. They must be small, energy efficient, and super-reliable, and they must move data extraordinarily fast.

Today, there is no solution that meets all these requirements, and so electronics will continue to evolve without becoming intimately integrated with photonics. Without significant breakthroughs, fat photons will continue to be excluded from places where skinny electrons dominate system function. However, if photonic components could be reliably manufactured in very high volume and at very low cost, the decades-old vision of fiber-to-the-processor and related architectures could finally become a reality.

We've made a lot of progress in the past 15 years. We have a better understanding of photonic technology and where it can and can't work in the data center. A sustainable multibillion-dollar-per-year commercial market for photonic components has developed. Photonic interconnects have become a critical part of global information infrastructure. However, the insertion of very large numbers of photonic components into the heart of otherwise electronic systems remains just beyond the edge of practicality.

Must it always be so? I think not.

About the Author

[Anthony F.J. Levi](https://sites.usc.edu/alevi/a-f-j-levi/) (<https://sites.usc.edu/alevi/a-f-j-levi/>) is a professor of electrical engineering and physics at the University of Southern California.