The future of packaging with silicon photonics

By Deborah Patterson [Patterson Group]; Isabel De Sousa, Louis-Marie Achard [IBM Canada, Ltd.]

It has been almost a decade since the introduction of the iPhone, a device that so successfully blended sleek hardware with an intuitive user interface that it effectively jump-started a global shift in the way we now communicate, socialize, manage our lives and fundamentally interact. Today, smartphones and countless other devices allow us to capture, create and communicate enormous amounts of content. The explosion in data, storage and information distribution is driving extraordinary growth in internet traffic and cloud services. The **sidebar** entitled, "Trends driving data center growth," provides an appreciation for the incredible increase in data generation and its continued growth through 2020.

To process and manage the unabated growth in data traffic, silicon photonics will be used to define new data center architectures. This article discusses the impact that silicon photonics will have on data center technology trends, and on the next-generation microelectronic packaging developments that address optical-to-electrical interconnection as photon and electron conversion moves to the level of the package and microelectronic (logic) chip.

Data center dynamics

The large-scale restructuring of data centers is one of the most dynamic transformations taking place in information technology. The need to re-architect the data center is being propelled by the staggering surge in shared and stored data along with an increasing demand to effectively interpret the tremendous amounts of content being generated. In addition to the huge growth in data traffic, the infrastructure supporting the Internet of Everything (IoE) will emphasize real-time responsiveness between people and/or objects. The next wave in data processing and data traffic management will require the ability to support cloud computing, cognitive computing and big data analysis along with the necessary speed and capacity to deliver a timely response.

Optics have traditionally been employed to transmit data over long distances because light can carry considerably more information content (bits) at faster speeds. Optical transmission becomes more energy efficient as compared to electronic alternatives when the transmission length and bandwidth increase. As the need for higher data transfer speeds at greater baud rate and lower power levels intensifies, the trend is for optics to move closer to the die. Optoelectronic interconnect is now being designed to interface directly to the processor, application specific integrated circuit (ASIC) or field programmable gate array (FPGA) to support switching, transceiver, signal conditioning, and multiplexer/demultiplexer (Mux/Demux) applications.

Figure 1 shows a forecast for silicon photonics adoption through 2025 with data centers dominating initial growth. Silicon photonics are also being developed to support applications as diverse as high-performance computing and optical sensors.

The data center need for speed and capacity. **Figure 2** illustrates forecasted data center traffic by 2019. One of the more notable trends is that almost three-quarters of all data center traffic will originate from within the data center. The recognition of this statistic, compounded by the enormous increase in data traffic, has significantly altered the approach to data center design. Besides upgrading optical cabling, links and other interconnections, the legacy data center, comprised of many off-the-shelf components, is in the process of a complete overhaul that is leading to significant growth and change in how transmit, receive, and switching functions are handled, especially in terms of next-generation Ethernet speeds. In addition, as 5G ramps, high-speed interconnect between data centers and small cells will also come into play. These roadmaps will fuel multi-fiber waveguide-to-chip interconnect solutions, laser development, and the application of advanced multi-chip packaging within the segment.

The high-end or "Hyperscale" data center is massive in both size and scalability. It provides a single compute architecture

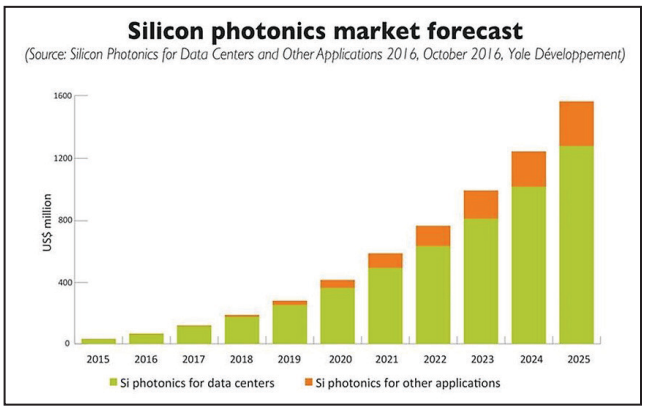

Figure 1: Silicon photonics growth rates will initially be dominated by applications within the data center. SOURCE: Yole Développement, Oct. 2016

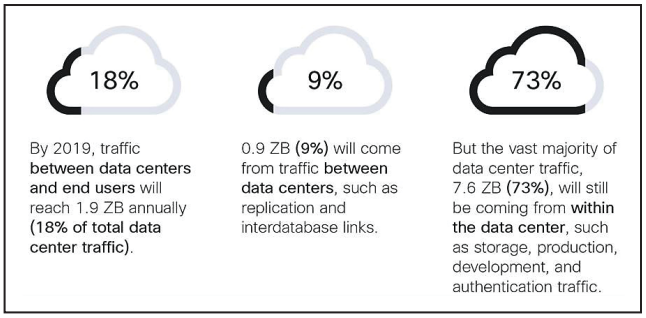

Figure 2: Data center traffic and bit rates show remarkable growth. The vast majority of data center traffic will reside within the data center. SOURCE: Cisco Global Cloud Index, 2014–2019

made up of small individual servers and peripheral functions such as memory, power management, and networking, all woven together through layers of redundant systems. Hyperscale data centers are represented by companies such as Amazon, Microsoft, Google, and Web 2.0 companies like Facebook. They can house over a hundred thousand to a million servers.

Conversely, alternative data center architectures are also under consideration. An example is the build-out of many small data centers linked together through heterogeneous networks. No matter which approach, moving the enormous volume of data within and between data centers will continue to require increased speed and bandwidth with lower latencies and greater power efficiencies. Silicon photonics is positioned to address these fundamental conditions.

Servers will be reconfigured to support higher speeds along with new componentry at 10 gigabit Ethernet (GbE), 25, 40, 100 and 400GbE. Meanwhile, new standards such as 1TbE will eventually be introduced. **Figure 3** illustrates typical data center connections, nomenclature, Ethernet speeds and link distances.

Because each data center server generation lasts 3-5 years and the infrastructure (buildings) typically lasts 3-5 generations, the technology choices incorporated now will become the legacy infrastructure over the next 10-25 years. Therefore, determining which technologies deliver maximum flexibility, capacity and reduction in total system cost-of-ownership is non-trivial and leading-edge development emphasizing concurrent semiconductor and package design is receiving greater visibility.

Optoelectronic integration of transceivers and switches

Because optics can transport more data at significantly lower power than electronic transmission, the intent is to drive optoelectronic conversion as close to the chip and microelectronic packaging level as possible. The data remains in optical form – leveraging high optical densities – until it enters the package and interfaces with the silicon photonic die. At this point, the data is converted from photonic to electronic format to undergo computation, storage, redirection, etc., by conventional logic ICs. Running electrical and optical pathways side-by-side on a micrometer scale is the eventual objective.

Data centers are being "disaggregated" so that compute, storage, memory, networking and power conditioning/distribution can be reorganized and redistributed systemically. Signal routing efficiencies, better thermal management, and densification of connections brought forth by well-designed package integration is expected to improve system upgradability, flexibility, and reliability while lowering cost. Silicon photonics will be especially useful in enabling the high communication bandwidth requirements of disaggregated systems.

Transceivers represent the initial high-volume application for silicon photonics as optics migrate as close as possible to the origin of the data. Outside the data center, optical transceivers are used in transport, enterprise, carrier routing and switching markets. Within the data center, transceivers are located with each server. However, they are mounted at the edge of the board, resulting in large distances between the optical components and the processor chip. The IBM research team in Yorktown, NY has proposed advanced packaging designs where the silicon photonic die can be integrated directly into the processor module, bypassing today's standard transceiver housings. Integrating the transceiver functions within the silicon photonic die or processor module is an area of concentrated activity.

Ethernet switches, on the other hand, currently rely on electronic interconnect but are another targeted area of silicon photonic interest. This is because data center architecture is undergoing a fundamental shift away from traditional three-tiered designs to route information. Power consumption and latency have increased as data traffic volumes have surged. This is due to too many "hops" or handoffs when routing data between source and destination (controlled by routers and switches). "Leaf-spine" network architecture is replacing the older tiered approach. The leaf-spine network is controlled by ASIC switches.

A "leaf" is typically the top-of-rack (ToR) switch that links to all servers within a common tower or rack (shown in **Figure 3**). The next layer of switches is represented by the "spine." The spine is a higher capacity switch (40 or 100 gigabyte per link) that connects to leaf switches (across the server racks), to other spine switches, and to the next level of "Core" switches. This is known as a "flat architecture" and it is implemented through high-capacity ASIC switches. The largest ASICs currently offer 32 x 100GbE ports. ASIC switch capacity and port count dictate the size of the leaf-spine network, e.g., the number of servers that can be linked together. Higher speed ports drive lower bandwidth costs per Gb/s and higher capacity switches drive the total number of links (minimizing the number of data hops, associated latency, and power consumption).

In addition to the electrical domain switching described above, there is significant interest in on-chip switching

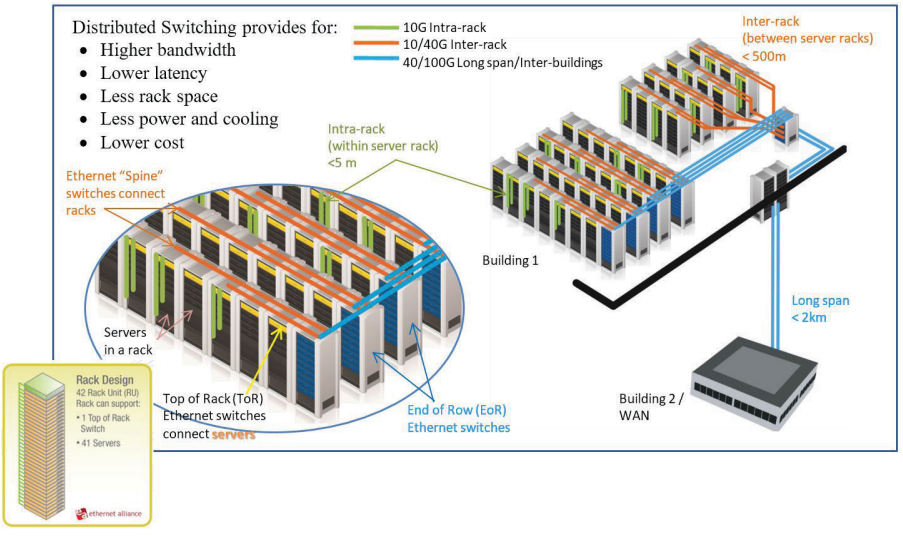

Figure 3: Illustration showing data center connections. The shorter server-to-server connections within the server rack typically require <5m (green) of routing, longer connections between the server racks may require up to 500m (orange), and long span connections within a building or between buildings can measure up to 2km (blue). SOURCE: Image courtesy of Finisar, Patterson Group, IBM, Ethernet Alliance

Trends driving data center growth

The following information is taken from the 2015 Cisco Global Cloud Index [1] and the Cisco Visual Networking Index: Global Mobile Data Traffic Forecast [2] to provide the reader with an appreciation of what is driving today's explosion in data traffic.

The growth in mobile users has had a phenomenal impact on data traffic. The average smartphone user in 2015 consumed more data in less than a minute than a mobile user in 2000 did in an entire month [3]. Mobile video, which represented 55% of data traffic in 2015, is projected to grow to 75% by 2020. The video forecast becomes even more significant when one realizes that the bit rates required for streaming high-definition and ultra-high definition (4K) video are 2x and 9x faster than standard definition video, further driving the need for both capacity and speed. And this does not even include game changers such as virtual, augmented and mixed reality technologies. In addition to the changes in volume and density of video content, there will be a massive migration of stored data to the cloud. Today, almost three-quarters of stored data on client devices still reside on PCs with integrated hard drives. The expectation is that a majority (51%) of stored data will be associated with non-PC devices with always-on wireless connectivity and no hard drives, propelling consumer cloud storage requirements to 1.6GB per month by 2019. Cloud users will store data (which is archival and may or may not require bandwidth) and will pull/push data from/to the cloud (which requires bandwidth).

Data centers manage both cloud traffic and data center traffic. Cloud computing represents a subset of all data center traffic that exists within and between data centers and end users. Cisco's Global Cloud Index forecasts global cloud traffic to quadruple between 2014 and the end of 2019, from 2.1 to 8.6 zettabytes (ZB). Global data center traffic is projected to grow from 3.4 to 10.4ZB (where a zettabyte is equal to a trillion gigabytes). **Figure S1-1** illustrates the concept of 10.4ZB.

The global forecast of devices and connections track to a 10% CAGR. In comparison, they outpace the CAGRs of the global population (1.1%) and the Internet user rate (6.5%). But that's not all. The Internet of Everything (IoE), described by Cisco as the connection of people, processes, data and things, may drive data center and cloud traffic growth to over 500ZB annually by 2019 - almost 50 times that of data center traffic alone (10.4ZB). It is assumed that IoE content and applications will increasingly be stored, managed and propagated through the data center, along with a substantial increase in big data analytics that will continue to mature into indispensable tools of strategic and tactical decision making. Future 5G build-out will drive speed and capacity requirements closer to the user with significant reduction in latency (>100x faster than 3G and roughly twice as fast as 4G), opening the door to interacting in real-time with your environment and people in ways that are only just being imagined.

Figure S1-2 highlight a few examples of the huge increases in devices, their impact on overall Internet and data center traffic and speeds, and the almost four-fold increase in cloud management over the current five-year period. Although the forecast start dates vary from 2014-2016, the story remains consistent. Between the growth of wireless and mobile traffic (which will account for two-thirds of total IP traffic by 2020), social media activity, the progression of video volumes and densities, the ramp of imaging such as embedded vision, virtual/augmented/mixed reality and 3D video, the continued migration to cloud storage, the propagation of sensors feeding the Internet of Everything, and the evolution of big data analytics and cognitive computing, it is clear that the need for speed and capacity growth will continue unabated for a long time to come.

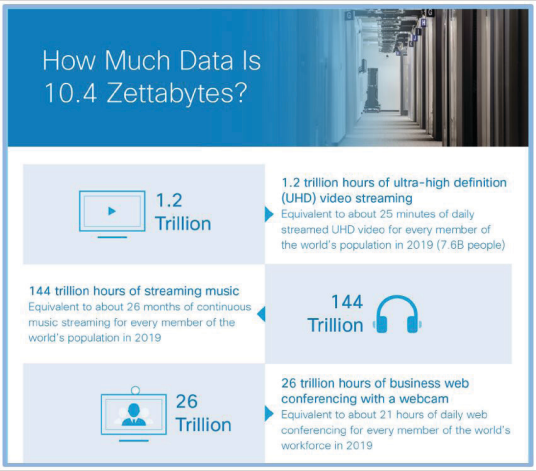

Figure S1-1: Chart illustrating the concept of 10.4ZB. SOURCE: Cisco Global Cloud Index, 2014–2019

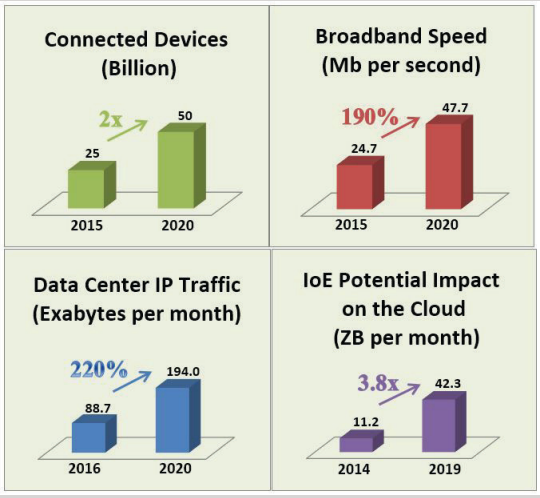

Figure S1-2: Connected devices and their impact on cloud and data center traffic. SOURCE: Data courtesy of Cisco GCI, VNI and IoT White Papers

References

1. Cisco Global Cloud Index, "Cisco Global Cloud Index projects cloud traffic to quadruple by 2019," Oct. 28, 2015; https://newsroom.cisco.com/press-release-content?articleId=1724918
2. Cisco Visual Networking Index (VNI), "The zettabyte era – trends and analysis;" http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html
3. K. Patel, "Cisco visual networking index global data traffic forecast, 2015-2020," SEMICON West, San Francisco, CA, July 12-14, 2016; http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

of the full optical content of fibers without conversion into the electrical domain. Such a fiber optic switching matrix can route tens to hundreds of fibers in, acting as a hub to transmit high-bandwidth optical data while working in tandem with its electronic counterpart.

Advanced packaging with silicon photonic die

In today's data centers, pre-packaged transceivers and other optical modules include discrete, large-pitch components. These traditional "gold boxes" and off-the-shelf components are too expensive for downstream data center proliferation; however, adoption of silicon photonics has the promise of meeting aggressive price targets. Price targets at 100GbE are found between $1-$5 per gigabit for single-mode fiber as compared to $2 per gigabit for multi-mode fiber [1]. It is generally recognized, however, that $1-$2 per gigabit for single-mode fiber is necessary in order to achieve widespread impact in datacom (for example, through the enablement of new architectures). It is also understood that silicon photonics and its packaging will need to leverage the know-how, best manufacturing practices, and massive scalability developed for high-volume CMOS manufacturing and microelectronic assembly. This will bring down cost enough to fuel high-volume integration, which in turn, will contribute to driving costs down along traditional semiconductor growth and cost trajectories. Silicon photonics is projected to offer cost-efficient optical interconnects (I/Os), allow for a reduction in components, and reduce assembly and test steps to decrease manufacturing costs while increasing throughput.

Currently, assembly, test, fiber and laser connections – everything except for the silicon chip(s) – are adding prohibitive cost and slowing down deployment. Much of this is due to low-volume fiber-to-die assembly practices and the impact on overall test and manufacturing flows. In order for the cost of silicon photonic assembly and test to meet end-product, design-for-manufacturing cost targets, clever and disruptive packaging approaches that also address the optical I/O (fiber- or waveguide-to-die) and laser-to-die interconnections are required.

A tale of two supply chains

When assembling a silicon photonic package, two specialized providers are typically employed: the semiconductor assembly and test service (SATS) provider and the optics contract manufacturer. The core competencies of the SATS and fiber optic assembler are quite different and typically do not overlap. The SATS provider specializes in the assembly and test of the electrical (digital and analog) ICs, passive components and electrical interconnect at the substrate or interposer level, whereas the optics contract manufacturer specializes in interconnecting the photonic IC (PIC) to the optical fibers. The optics contract manufacturer manages fiber and laser alignment at sub-micron levels and also provides specialized optical testing at the package and die level. The end customer/OEM must therefore oversee two separate assembly and test process flows as well as define the build and test plans. It is no surprise that assembly and test strategies may include "known good" multi-chip platforms designed within the optoelectronic module.

Optoelectronic multi-chip modules (MCMs) can contain processors, ASICs, FPGAs, memory, passive components and other functional elements plus the PIC. Certain operations can be integrated on the photonic chip [2] while others will be combined in-package (MCM/SiP). Logic and optical functions can include transmitters, receivers, multiplexers/demultiplexers, modulators, splitters, photodetectors, resonators, optical isolators and polarization controls. Non-optical CMOS amplifiers, drivers, or serializer/deserializer (SerDes) functions may also be located within the module. **Figure 4** illustrates various advanced packaging technologies that can be used to integrate silicon photonic (also known as "nanophotonic") digital and analog componentry within an optoelectronic package. These modules often require high-density interconnect and a combination of advanced packaging techniques.

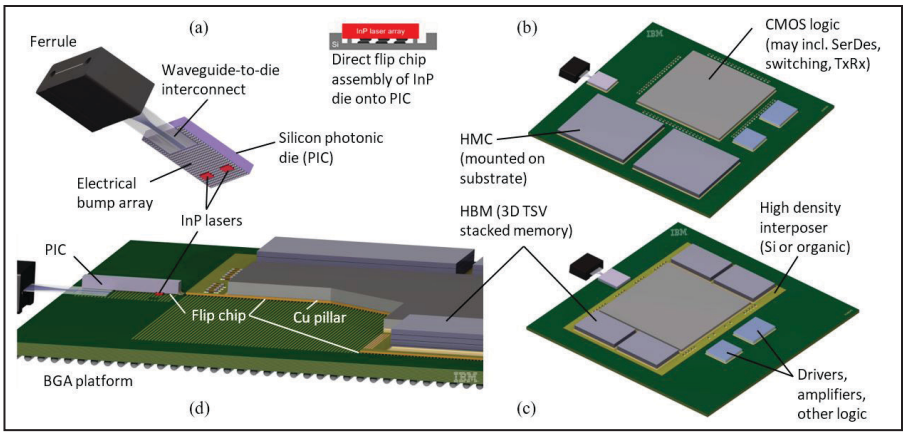

Figure 4: Silicon photonic MCM illustrating (clockwise) a) a photonic integrated circuit (PIC) with flip-chip mounted laser die and edge-coupled waveguide; b) logic, HMC and PIC mounted directly to a BGA substrate; c) logic and HBM mounted on a high-density interposer with copper pillar micro-bumps that enable high-speed connectivity; d) cut-away revealing copper pillar, flip-chip, BGA and optoelectronic interconnections. SOURCE: Image courtesy of IBM Corporation

Figure 4a shows a next-generation silicon photonic die designed to accommodate a waveguide array that will transport the photons into the die where efficient coupling can take place. The die also contains an array of flip-chip bumps to support electrical signals. In this illustration, laser die are attached to the silicon photonic chip through self-aligning flip-chip bumps that provide a low profile and closely coupled chip-to-laser interconnection. Lasers can also be accommodated off-chip through flip-chip or wire bond attachment. **Figure 4b** shows a ball grid array (BGA) substrate with a central logic die such as an ASIC, two hybrid memory cubes (HMC) and associated CMOS devices. A typical HMC is represented by a packaged 4-8 die DRAM stack supported by a controller chip and using 3D TSVs for vertical chip-to-chip connectivity. The packaged HMC has a standard BGA pitch (e.g., 650µm) and is mounted on the substrate along with the other components and PIC. These modules use flip-chip and/or wire bonding for I/O interconnection.

An alternative design is illustrated in **Figures 4c** and **4d**. In this case, the logic die and high-bandwidth memory (HBM) stacks are mounted directly onto a high-density interposer to closely link the logic, memory and PIC. It is desirable to have the PIC closely coupled to the logic die using either approach (b or c) to promote

power efficiencies. The interposer can be inorganic (Si) or it can be a high-density organic laminate material. When used as an interposer, these organic high-density interconnect (HDI) materials are sometimes referred to as "2.1D" to differentiate them from 2.5D silicon interposer structures. HDI substrates support dense wireability and flip-chip interconnect at a lower cost than silicon interposers because they are not manufactured in the foundry. With multiple layers, 10µm vias, and 2/2µm or 3/3µm line and space routability, HDI materials can present a cost-effective alternative to 2.5D approaches. In the case of HBM, where a 55µm bump pitch is used, HDI material presents a viable option. Populated interposers can also be tested before assembly with the PIC and BGA substrate. Substrates and/or interposers may also contain embedded components, low-loss integrated waveguides, or embedded fiber connectivity.

These complex modules are often larger in area, integrate many material types, and must address significant thermal dissipation and signal integrity considerations. Packaging know-how must not be underestimated and deep experience in electrical and thermal modeling, materials characterization, component choice, vetted design rules, comprehensive test, and design for manufacture is essential. In addition, providers need to address package-level reliability testing in consideration of the photonic components, including features such as fiber strain relief [3].

Signal integrity and thermal dissipation pathways are crucial. For example, to support SerDes operation at 28GHz, the silicon photonic die and SerDes interface – integrated into the logic die and placed in close proximity to the PIC – must be optimized. The optical elements need to be calibrated and balanced against the electrical interface as well. Optical systems reduce/eliminate the need for signal conditioning and amplification steps that would be required for systems with long electrical traces. It is also important to mitigate hot spots on the ICs and the thermal drift that some optical components experience.

Silicon photonics programs at their most effective are designed with a system-level integration mindset. Therefore, concurrent package design for the PIC and the electrical and thermal elements within the module are considered together. Development must include a) silicon photonic die design, b) parallel waveguide interconnect technology, c) chip-to-waveguide (fiber) assembly processes, d) overall thermal management, and e) final integration between the logic components and photonic IC using various assembly techniques. The build strategy will depend on the type of advanced packaging employed (and at what level higher density flows are designed into the module), the test complexity, yields for both the optical and electrical process flows, and overall cost tradeoffs. A provider that offers assembly of the electrical/thermal elements plus integration of next-generation chip-to-fiber interconnects would be a welcome addition to the packaging landscape.

Demonstrating silicon photonic "firsts"

IBM has been developing CMOS integrated silicon photonic die and packaging approaches for close to fifteen years. **Figure 5a** shows a photograph of the first CMOS integrated coarse wavelength multiplexed silicon photonics chip. This silicon photonics reference design, demonstrated in 2015, is capable of transmitting and receiving 4 wavelength channels – each operating at 25Gb/s. The chip combines both mixed signal electronic circuits and optical capability on a single die. The die was manufactured through the former IBM Microelectronics Division, now a part of GLOBALFOUNDRIES, whose monolithically integrated silicon photonics technology (called CMOS9WG) was used. It is derived from their 90nm CMOS process, a mature technology node that can facilitate commercialization. The design was simulated and fully verified with a process design kit (PDK) integrated with industry-standard CMOS design tools. The die contains modulators, germanium photodetectors, and ultra-compact wavelength division multiplexers (WDM). **Figure 5b** shows the die with an array of 100µm Pb-free flip-chip bumps on a 200µm pitch. The flip-chip interconnects keep electrical I/O loss to a minimum. **Figure 5c** identifies the transmitter and receiver elements along with their eye diagrams, showing a clean, high-quality response.

In this design (**Figure 5**), the lasers are manufactured separately and connected off-chip with the optical signals coming into the die through the top four laser ports. The chip contains four receive and four transmit channels, each capable of transmitting data at 25Gb/s via independent wavelengths. The four carrier signals (λ) are bundled into 100Gb/s fiber ports through wavelength multiplexing. An extension of this design would allow for up to twelve such fiber ports when using a standard 12-fiber connector, thereby enabling 600Gb/s bi-directional data rates from a single optical transceiver. The intention of the program was to push 100Gb/s data speeds over a range of up to 2 kilometers (1.24 miles) as required by large data center environments [4].

Note that the die in **Figure 5** is designed with only two optical I/Os representing a single input fiber and output fiber. Currently, duplex fiber connections represent the lowest cost and most widely used approach. One fiber in and one out have been sufficient to support current bandwidth needs. Because of the low I/O count, active optical alignment is often employed and involves assembly in a manual or partially automated fashion using specialized tooling. The

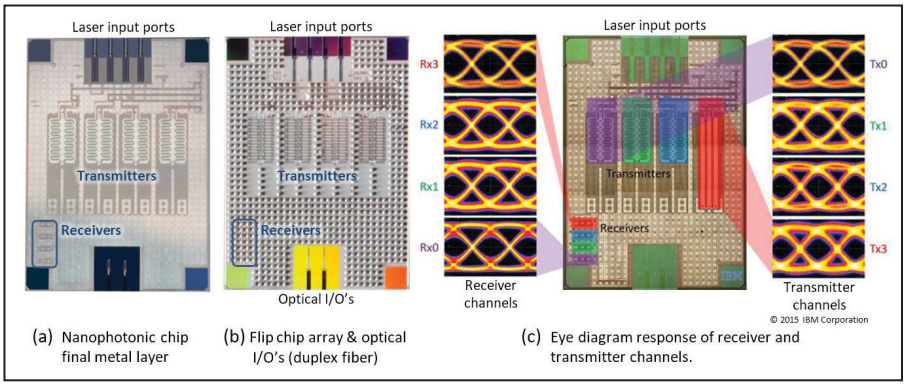

Figure 5: An IBM silicon (nanophotonic) wavelength division multiplexed 4 x 25Gb/s reference design (die) supporting both electrical and optical circuits. Two optical I/Os are seen at the bottom of the die. These connect to optical fibers, one in and one out. SOURCE: Images courtesy of IBM Corporation

result is a lower throughput, limited scalability, and higher cost process as compared to electrical interconnect and manufacturing operations.

Delivering chip-to-fiber assembly know-how to the semiconductor market is a crucial differentiator required to feed turn-key optoelectronic assembly solutions. IBM's silicon photonics program embraces a disruptive approach that leverages existing high-throughput microelectronic tools and techniques to assemble cost-efficient and scalable single-mode optical inputs and outputs that can be deployed in high volume.

Multi-waveguide to chip interconnect

To keep up with data traffic growth, both software and hardware strategies are utilized. Among these are the migration from the duplex fiber connections described in the previous section to multi-fiber or multi-waveguide interconnects (ideally leveraging industry standard 12-fiber connectors). The following sections identify data rate trends, techniques and trade-offs utilized to support these trends, a 12-fiber interconnection approach demonstrating the feasibility of automating fiber-to-chip assembly, and a next-generation multi-waveguide polymer structure created through lithographic patterning for high-throughput assembly.

Riding the wave of data rate progression. In order to carry greater amounts of optical data, the optical port density and optical bandwidth per port must increase, along with higher manufacturing throughput and scalability. Technology development to intercept downstream requirements is being pursued now. The bit rates (speeds) are trending as follows:

- Within the server rack (intra-rack): 10GbE → 25GbE → 40GbE → 100GbE → 400GbE (finalizing standardization);
- Between server racks (inter-rack): 40GbE → 100GbE → 400GbE (finalizing standardization); and
- Within the data center or between buildings (inter-data center or "long haul"): 100GbE → 400GbE → 1TbE/1.6TbE (standardization remains to be determined).

Figure 6 highlights the unabated rise in data center revenue as broken down by data rate through 100GbE.

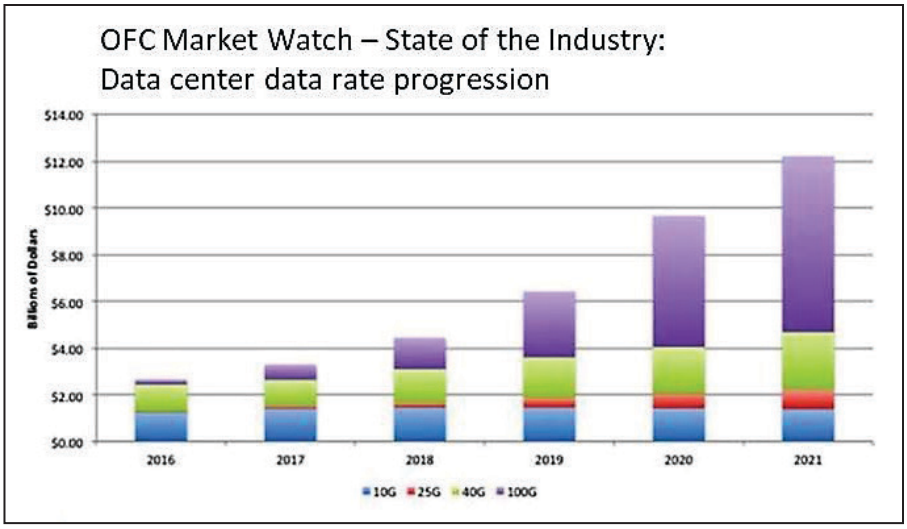

Figure 6: a) Data centers remain a dynamic multi-billion dollar market feeding a progression of high- speed data rates. SOURCE: Data Center Optics, Discerning Analytics presentation, OSA Industry Development Association, OFC May 2016.

400GbE bit rates are now being standardized for 2018 introduction and it's not just for long haul distances. It will also be used at distances below 100 meters to manage intra- and inter-rack data traffic. The **sidebar** entitled, "Techniques to extend data rates," describes various ways to mix and match technologies using next-generation 400Gb/s Ethernet as an example. The number of data lanes (fibers) into and out of the die depends on the application. A single one in/one out port count is used today for fiber links in excess of 500 meters while up to eight (4 transmit/4 receive) fiber connections are used for 100-500 meter core applications such as Parallel Single Mode 4 (PSM-4) in support of 100Gb/s (standard) and 400Gb/s Ethernet (proposed) [5]. Emerging applications such as communication hubs and on-chip fiber switches [6] can require port counts of tens of fibers per chip. 2016 saw a significant shift toward 100GbE components. The development and introduction of complex packaging and optical interconnects to lower 100GbE costs and support next-generation 400GbE is timely.

Adding more chip-to-fiber connections can substantially expand bandwidth and capacity by providing a platform upon which the techniques listed in **Table S2-1** (sidebar) can be deployed. This leap in data will carry forward for several generations to support 100GbE → 400GbE → 1TbE+ roadmap progression. Aggressive cost goals (<$1 per attachment [7]) must be met to justify mass adoption. Therefore, practical, high-volume chip-to-fiber assembly techniques are an area of current concentration.

Bridging legacy infrastructure with next-generation electronics. The 12x1 mechanical transfer (MT) interface is a single-mode fiber standard that has been handed down from the telecom industry. Single-mode fiber has been employed because of its high efficiency in guiding light over long distances with very low loss. Single-mode fiber is preferred over multi-mode fiber and is becoming a de facto standard within several areas of next-generation data center designs.

An automated, self-aligned 12-fiber connection to a silicon photonic die is examined below. The program leverages semiconductor assembly, tools, and process flow experience. The 12 optical I/O solution is attractive because of the vast deployment of 12-fiber connectors, making it the lowest cost for parallel optical connections per fiber port. This type of connector is used in PSM-4 transceivers for 100Gb/s and is planned for 400Gb/s and beyond. Multiple 12-fiber connectors are anticipated to fulfill the optical I/O needs of communication hubs and on-chip fiber switches. Multiple wavelengths may be used in each fiber for CWDM, creating a large requirement on both the optical bandwidth per fiber connection as well as the number of fiber connections.

Figure 7a shows a 12 I/O silicon photonics demonstration die with a flip-chip array providing the electrical

Techniques to extend data rates

Data centers have historically been built using off-the-shelf components. As data center hardware matured, architecture evolved to take advantage of technology improvements at a measured pace. However, the unprecedented data explosion has accelerated data center redesign and is reshaping many technology roadmaps that support this segment.

To ensure higher bandwidth and rapid data delivery across the electrical interface, four general approaches are considered when evaluating system-level trades-offs such as overall efficiency, cost, and return on investment. These are speed (increasing symbols [1] per second or baud rate), modulation, lane count (number of fibers), and transmission technology (often distance dependent). The divergent needs of each end application guarantee that architecture will remain flexible. A granular segmentation in the Ethernet switch market, for example, will provide the data center with customized options it lacked in the past. **Table S2-1** describes various ways to mix and match technologies in support of next-generation 400GbE data rates.

ultra-fast 400GbE infrastructure upgrades.

The general trend is that sophisticated techniques such as PAM-4 modulation, single-mode fiber (SMF) and wavelength multiplexing are becoming the de facto standard at shorter and shorter distances. For example, in the data center environment, SMF has traditionally carried signals over long distances such as 500 meters (inter-rack) to 2km (intra-data center) to 10+km (between data centers). SMF is now employed at 100 meters and is migrating to distances as short as 20 meters in order to provide downstream flexibility for data center upgrades. In fact, Microsoft announced that it will employ SMF at distances of 20 meters and greater. This is an evolutionary departure from traditional multi-mode fiber (MMF) interconnection that supports today's vertical cavity surface emitting lasers (VCSEL).

Trade-offs such as fiber cost become a more dominant factor as the number of links needed at shorter distance increases. Fiber can cost as much as $200-$300 for large switches. Because there are so many more short links (eg., transceivers, switches) incorporated into data centers, the overall capacity cost of fiber becomes significant and its rollout will, therefore, be planned to include subsequent upgrade costs as data center design continues to respond to increased capacity demands.

Feature	Technique	Getting to 400 GbE data rates
Speed	Increase baud rate (data symbols [1] per second) with forward error correction	Baud rate on each lane (fiber) is increased when transitioning from 10GbE to 25GbE or 50GbE per lane
Modulation	Using more bits for each transmission symbol	PAM-4 (used for 50GbE) transmits 2 bits of information per symbol versus NRZ with 1 bit per symbol
Lane Count	Increase copper traces and/or optical fibers	8 lanes at 50GbE (PAM-4) or 16 lanes at 25GbE (NRZ) both produce 400 GbE data rates
Transmission Technology	Tradeoff - copper traces versus optical fibers	Optical fiber can carry signals much farther than copper traces
	Tradeoff - Single Mode Fiber (SMF) versus Multi-Mode Fiber (MMF)	SMF requires more precise transceivers but can carry data over longer distances than MMF
	Tradeoff – number of wavelengths (λ) versus number of optical fibers	8λ at 50GbE on a SMF with PAM-4 modulation will enable 400GbE over a single fiber at distances of 10+km

Table S2-1: Rising data rates leverage current infrastructure while providing a platform for new technology.

The techniques listed in **Table S2-1** will allow for considerable variation. This is necessary because legacy technology at speeds ≤10GbE still represent a significant portion of data center infrastructure. In addition, new usage models in certain IoT/IoE, industrial and automotive segments will leverage slower speeds. This said, there is much excitement and focus on facilitating faster data speeds. The IEEE Ethernet Task Force passed new Ethernet standards for 25GbE and 400GbE with 50GbE and 200GbE under consideration. These multiples will feed today's 100GbE and

Reference

1. Data is sent in "symbols" where a symbol represents a specific state of frequency, amplitude and phase. A symbol can be single or multiple bits of data. In digital communications, the symbol rate is also known as the baud rate.

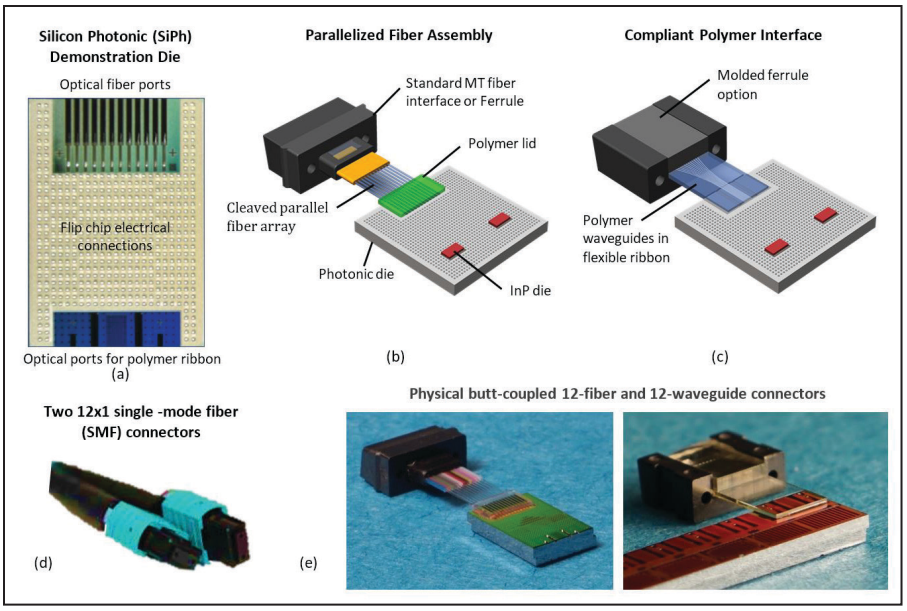

Figure 7: a) IBM silicon photonic die; b) schematic representations of the parallelized fiber and c) compliant polymer ribbon assemblies; d) an industry standard 12x1 single-mode fiber connector; e) photos of the parallelized fiber and compliant polymer structures. SOURCE: Images courtesy of IBM Corporation

connections. The die supports two next-generation photonic-to-electronic interconnect approaches: twelve fiber inputs or "V-grooves" can be seen on the top area of the die (green) and finer waveguide inputs are clustered on the bottom center of the die (blue). The V-groove represents the physical structure where the light is "butt coupled" – that is, where the fibers terminate into on-chip metamaterial optical mode converters. The finer waveguide inputs correspond to adiabatic optical coupling from a polymer waveguide interface to the chip.

The first approach is called the "parallelized fiber assembly" because a fiber optic ribbon emerging from the MT interface (also referred to as a "ferrule") is assembled at once to a photonic chip; this is depicted in **Figure 7b**. The design was specifically developed to utilize high-volume pick-and-place equipment (vacuum pick-tip) that can handle the fiber stub.

The second approach illustrated in **Figure 7c** employs a "compliant polymer" interface. In this design, a polymer ribbon with lithographically defined waveguides and self-alignment structures acts as an intermediary between the chip and the standard fiber connector. The polymer ribbon contains finely spaced embedded waveguides that align to both the ferrule and die. The design results in improved thermo-mechanical reliability as compared to rigid connections such as the direct fiber-to-chip attachment (where fiber pistoning or other chip-package interactions can be an issue).

Both approaches offer flip-chip or wire bond electrical connections and support direct, self-aligning InP laser attach to the silicon photonic die. The technology direction depends upon the application. The parallelized fiber approach employs standard components with known reliability and excellent optical fiber transparency. The compliant polymer ribbon shows much promise in terms of structural reliability, is reflowable, and is compatible with high-volume manufacturing to support more and varied waveguide pathways. **Figure 7d** shows an industry standard 12x1 single-mode fiber connector that would attach to the other end of the MT fiber interface and **Figure 6e** shows the physical assemblies with 12x1 arrays of optical I/Os.

Automating optical microelectronic assembly. CMOS logic and optical transmit and receive functions need to be closely integrated. Many of today's chip-to-fiber attachments use vertical diffractive grating couplers as they are typically the easiest coupling scheme to integrate on a wafer. These are defined by the active alignment of fiber(s) that attach at a semi-vertical angle to the face of the photonic die. However, diffractive grating couplers suffer from limited spectral bandwidth, which is generally insufficient for CWDM applications. In addition, they exhibit unattractive coupling efficiency from fiber-to-chip where the two polarizations of light present in the fiber need to be considered. Finally, they emit light almost vertically from the plane of the chip, which complicates or even negates flip-chip electrical connection and impacts thermal management strategies. With these considerations, in-plane coupling techniques and adiabatic mode converters were pursued instead. These are inherently tolerant to fabrication imperfections, polarization of light and spectral bandwidth.

Two major challenges to mating single-mode optical fibers to silicon photonic waveguides are alignment accuracy and fiber mode conversion [8]:

1. Alignment tolerances between the optical fiber(s) and waveguide couplers embedded within the silicon photonic die should not exceed 1-2µm on-chip for acceptable coupling efficiency. (For comparison, ±10µm placement accuracy is suitable for solder bumps used for electric connections.)
2. There is a large mismatch in energy confinement between an optical fiber and a silicon photonic waveguide. The energy is distributed over a ~10µm spot in the fiber while it forms a sub-micron spot in a typical silicon photonic waveguide. This creates the equivalent of a large impedance mismatch that must be addressed with a suitable optical mode converter, which is analogous to an impedance matching circuit.

Precise self-alignment features are crucial to maximizing assembly tolerances and must be designed into any automated photonic assembly process that seeks to utilize standard electronic manufacturing equipment. Self-alignment structures are designed within the silicon photonic die and optical components (ferrule) to meet the 1-2µm on-chip photonic alignment requirement. Self-alignment techniques must also work within the speeds and tolerance limitations of today's production equipment.

Pick-and-place handling of fiber components in a microelectronic manufacturing process is particularly challenging. Not only are optical fibers typically handled with specialized tools instead of standard vacuum pick-tips, but the placement accuracy of standard equipment is normally on the order of ±10µm (almost an order of magnitude larger than single-

mode optics requirements). In addition, high-speed pick-and-place equipment does not offer the 3D spatial movement required for chip-to-fiber alignment. Equipment is designed to engage the die in-plane. Pressure sensitive movement is only possible in the downward direction after lateral positioning. This means, for example, that you can pick up a parallelized fiber array and lay it down in the V-grooves of the photonic die, but not be able to push the fiber array inward to butt into the waveguide coupler. To counteract this limitation in movement, a platform was designed to transfer a portion of the vertical pick-and-place movement into a horizontal force that successfully butts each fiber termination within the V-groove of its on-chip waveguide coupler [9,10]. The self-alignment features designed into the components and adaptive tooling demonstrated that high-throughput pick-and-place equipment could be applied to photonic fiber assembly, paving the way for large-scale automation and scalability in complexity, yield and manufacturing volume.

The graphic progression in **Figure 8a** shows how vertical force from the pick-and-place equipment is converted to the horizontal force needed to slide the fibers properly into the die. **Figure 8b** shows a 12-fiber planar array assembled onto the silicon photonic die with self-alignment providing ~1µm on-chip final alignment accuracy. **Figure 8c** demonstrates how well the fibers fit into the V-grooves on the silicon photonic chip. **Figure 8d** computes the fiber core-to-waveguide coupler misalignment for 10,000 random structural tolerance combinations to demonstrate the manufacturability of the fiber self-alignment. The results show a 3σ misalignment below ±1.3µm demonstrating tolerances well within required placement accuracies [11].

Compliant polymer ribbon: an elegant approach. Another waveguide-to-chip interconnect method is currently under development. In this approach, a flexible polymer ribbon replaces the cleaved fiber array. The design mechanically decouples the chip's optical interface from the fiber connector by substituting a mechanically compliant polymer ribbon. The ribbon contains an array of mechanical self-alignment structures and polymer waveguides that transmit light from the fiber to the die. This offers several advantages over other known approaches [12]:

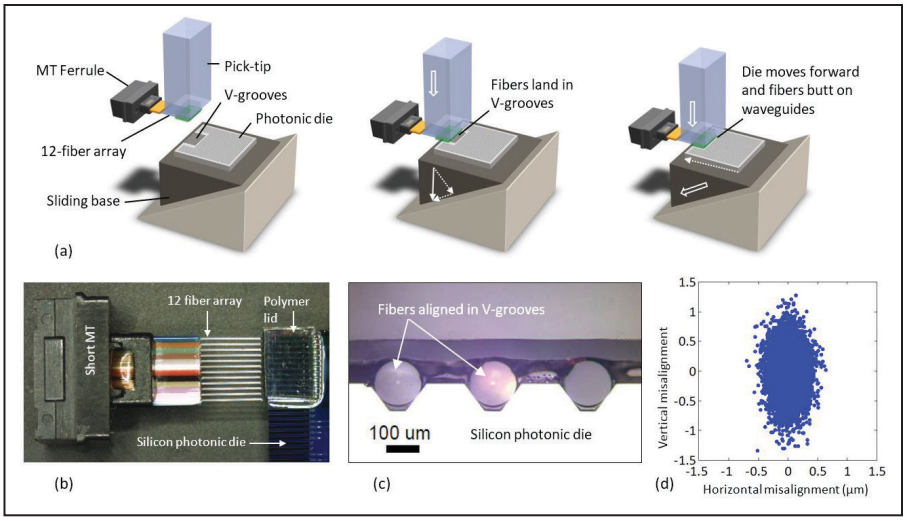

Figure 8: Demonstration of self-alignment of parallelized fiber assembly: a) trigonometric transfer of vertical to horizontal force, pushing the die forward to mate with the fibers; b) top down photograph of a 12-fiber ribbon array attached to a silicon photonic die; c) cross-sectional micrograph showing a close-up of three fibers settled into their corresponding V-grooves within the die; d) Monte Carlo analysis demonstrates manufacturability of fiber self-alignment with a 3σ misalignment below ±1.3µm. SOURCE: Images courtesy of IBM Corporation

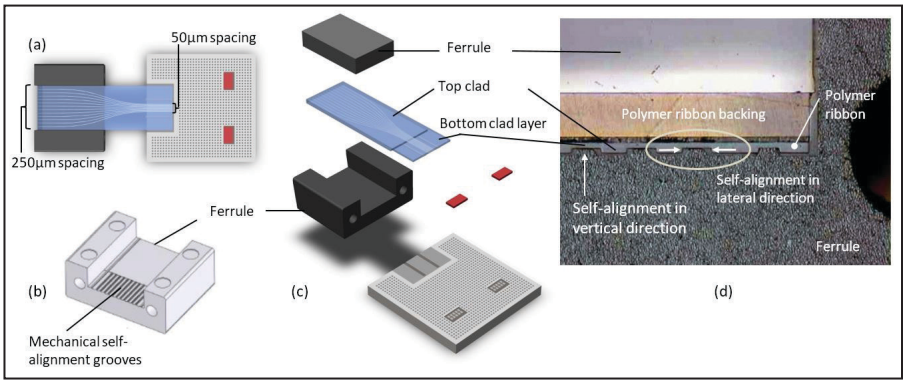

Figure 9: a) Top down view of compliant polymer ribbon aligned with ferrule (no lid) and silicon photonic die; b) ferrule with self-alignment guides for polymer ribbon; c) exploded view of assembly showing two-piece molded ferrule; d) cross section of polymer ribbon aligned to grooves in the molded ferrule. SOURCE: Images courtesy of IBM Corporation

- Cycling strains within the package are separated from the die by the compliancy of the polymer material, improving long-term reliability.
- The material and design supports large optical bandwidth. As compared to diffractive couplers where a 1dB bandwidth of a two-polarization vertical coupler has been reported at ~30nm [13], the polymer ribbon demonstrated a 0.8dB penalty over a 100nm bandwidth and all polarizations [14].
- Integrated waveguides are lithographically patterned within a three-layer polymer stack. Lithography provides for significant flexibility in design to optimize optical mode conversion efficiencies, define self-alignment structures or create other features of interest such as pitch conversions, bent die interfaces or port shuffles.

Figure 9a illustrates a 12-waveguide array, although the number of patterned waveguides can vary. At the fiber connector side, the waveguide spacing is designed at a 250µm standard fiber connector pitch and tapers down to a 50µm pitch at the photonic chip interface. The polymer waveguide is 8µm wide at the chip interface where it mates with the silicon waveguide tapers. As in the parallelized fiber approach, the far end of the polymer interface is compatible with a standard 12x1 MT fiber interface. At the MT fiber connection point, a

custom molded U-shaped ferrule (**Figure 9b**) is used to allow for the alignment and connection to standard metal pins. Ferrule-to-polymer self-alignment structures are used on both the polymer ribbon and the ferrule. **Figure 9d** shows the ferrule's trapezoidal alignment ridges that promote vertical and lateral self-alignment of the polymer ribbon. Rectangular ridges are also patterned on the polymer ribbon to fit into trapezoidal grooves on the silicon die for self-alignment. Each of these features and process steps ensures final alignment accuracy within 1-2µm. When used hand-in-hand with position-tolerant optical coupling approaches, high-throughput assembly with low optical coupling-loss can be achieved.

The ferrule and polymer ribbon are assembled first. Because corresponding self-alignment structures are defined on both components, accurate alignment with low-accuracy, high-throughput assembly equipment is achievable. Like the parallelized fiber, the sub-assembly (ferrule and polymer ribbon) is pick-and-placed onto the silicon photonic die where a UV curable transparent epoxy adhesive is dispensed and then cured once the polymer ribbon is properly self-aligned. Material properties, application, dispense location and quantity, etc., are examined to properly control the bond-line for optimal performance.

The compliant polymer interface has been demonstrated to successfully butt-couple standard single-mode optical fibers to mode-matched polymer waveguides, and adiabatically couple the energy into on-chip silicon photonic waveguides. The results yield wide bandwidth with encouraging peak performance. They also reinforce the objective of adopting high-volume manufacturing processes to chip-to-waveguide assembly.

Summary

In today's data centers, distance dictates the use of copper, single-mode fiber or multi-mode fiber transmission. Silicon photonics enables bandwidths and power efficiencies (energy/bit) that conventional electronic cables cannot compete against. Tomorrow's high-volume data center demands will incorporate advanced packaging techniques that synchronously integrate both electronic and photonic features. Because silicon photonics is a CMOS-compatible process, adopting the microelectronics industry's high-volume manufacturing processes and best practices is realistic. The convergence of two key assembly competencies, advanced multi-chip packaging and chip-to-waveguide interconnect, will contribute to the dynamic growth of this exciting market segment.

Acknowledgements

This article was made possible through the efforts of a dedicated IBM team. Notable contributors include Jean Audet, Tymon Barwicz, Alan Benner, Nicolas Boyer, Gaétan Cossette, Alain Drouin, Paul Fortier, William Green, Patrick Guerin, and Marc Hebert.

References

1. LightCounting webinar, "High-speed Datacenter Optical Interconnects report," Oct. 2016.
2. T. Prickett Morgan, The Channel, "IBM chips the laser light fantastic," www.theregister.com, Dec. 2010; http://www.theregister.co.uk/2010/12/01/ibm_silicon_nanophotonics_cmos/
3. Communication with A. Benner, IBM Poughkeepsie, NY and T. Barwicz, IBM Yorktown Heights, NY. Active program in Yorktown Heights includes fiber-to-module reliability and yield optimization.
4. S. Anthony, "IBM demos first fully integrated monolithic silicon photonics chip," Ars Technica UK, May 2015; http://arstechnica.com/information-technology/2015/05/ibm-demos-first-fully-integrated-monolithic-silicon-photonics-chip/
5. 100G PSM4 Specification (Online); http://psm4.org/100G-PSM4-Specification-2.0.pdf
6. S. Han, et al., "Large-scale silicon photonic switches with movable directional couplers," Optica, Vol. 2, Issue 4, pp. 370-375, 2015.
7. K. Patel, "Cisco visual networking index global data traffic forecast, 2015-2020," SEMICON West, San Francisco, CA, July 12-14, 2016.
8. T. Barwicz, et al., "Automated, self-aligned assembly of 12 fibers per nanophotonic chip with standard microelectronics assembly tooling," Proc. of the IEEE Elect. Comp. and Tech. Conf. (ECTC), San Diego, CA, May 26-29, 2015.
9. T. Barwicz, et al., "Photonic packaging in high-throughput microelectronics assembly lines for cost-efficiency and scalability," Optical Fiber Comm. Conf. and Exh. (OFC), 2015.
10. T. Barwicz, et al., "A novel approach to photonic packaging leveraging existing high-throughput microelectronic facilities," IEEE Jour. of Selected Topics in Quantum Elec., Vol. 22, Issue 6, 2016.
11. T. Barwicz, et al., "Automated, self-aligned assembly of 12 fibers per nanophotonic chip with standard microelectronics assembly tooling," Proc. of the IEEE Elec. Comp. and Tech. Conf. (ECTC), San Diego, CA, May 26-29, 2015.
12. T. Barwicz, et al., "Optical demonstration of a compliant polymer interface between standard fibers and nanophotonic waveguides," Optical Fiber Comm. (OFC) Conf., OSA Tech. Digest (2015), paper Th3F.5.
13. A. Mekis, et al., "A grating-coupler-enabled CMOS photonics platform," IEEE JSTQE 17, pp. 597-608, (2011).
14. T. Barwicz, et al., "A compliant polymer interface with 1.4dB loss between standard fibers and nanophotonic waveguides," 2016 Frontiers in Optics, OSA Tech. Digest (online); Optical Society of America, 2016, paper FTu1D.2.

Biographies

Deborah Patterson received her BS in Electrical Engineering from the U. of California, San Diego and her BS in Biology from Furman U. She is President of the Patterson Group; email deborah@patterson-group.com.

Isabel De Sousa received her MS in Metallurgy from Ecole Polytechnique of Montréal. She is an Advisory Packaging Engineer at IBM Canada, Ltd.

Louis Achard holds an MS in Mechanical Engineering from McGill U. He is an Advisory Packaging Engineer at IBM Canada, Ltd.