Silicon Photonics for Terabit/s communication in Data Centers and Exascale Computers
Stéphane Bernabe, Quentin Wilmart, Kobi Hasharoni, Karim Hassan, Yvain Thonnart, Pierre Tissier, Yohan Desieres, Ségolène Olivier, Tolga Tékin, Bertrand Szelag

To cite this version:

HAL Id: cea-03463711
https://cea.hal.science/cea-03463711
Submitted on 2 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License
Silicon Photonics for Terabit/s communication in Data Centers and Exascale Computers


ABSTRACT

Silicon Photonics Technology using sub micrometer SOI platform, which commercially emerged at the beginning of the century, has now gained market shares in the field of fiber optic interconnects, from Inter-to Intra-Data Center communications. With growing demands in terms of aggregated bandwidth, scalability, transceiver form factor, and cost, Silicon Photonics is expected to play a growing role, especially with the foreseeable need to co-package photonic transceivers with next generation Ethernet switches. This new paradigm will be possible only with an evolution of existing Silicon Photonics manufacturing platforms, in order to solve the challenges of 3D packaging, laser integration, reflow-compatible optical connectors and high efficiency, low footprint modulators. Achieving these challenges may pave the way to Terabit scale communications in Data Centers and High Performance Computing Systems (HPC).

Keywords: Silicon Photonics, Photonic Integrated Circuit, Photonic Module, Transceiver, High speed interconnect, Optical Network on Chip, On-Board Module, System-In-Package, High Performance Computing

1. INTRODUCTION

Using guided light as a way to propagate information is not a new idea, with early experiments starting as early as 1958 at STL’s laboratories [1]. From the early optical fiber field trial at the end of the 1970s to its massive use within Data Center, optical fiber based communications have revolutionized information and communication technologies (ICT), allowing and accompanying the development of Internet, by giving access to incomparable transmission capacity required by massive data transportation [2,3]. Indeed, by exhibiting Bandwidth x link distance products in excess of 100 Gbps.m [4], optical links have supplanted copper based links for any high speed application beyond several meters. Obviously, this trend will increase in the near future, along a roadmap for optical links and related transceiver modules shared by most of the stakeholders in the field.

In this article, we describe the roadmap and examine to what extent Silicon Photonics (SiPh) will help achieving this roadmap, targeting shorter and shorter distances from Intra Data Center Interconnects (DCI) down to intra-chip optical links, with the emergence of Optical Network on Chip (ONoC) approach.

First, we describe the recent evolution of requirements for DCI and ONoCs leading to the current roadmap (section 2). In section 3, we describe the existing Silicon Photonics platforms and their current evolution. Section 3 focuses on the requirements of next generation photonic transceivers, and section 4 describes the various missing building blocks to achieve these requirements, especially laser source integration, electronic-photonic integration, and packaging.
2. OPTICAL INTERCONNECTS EVOLUTION FOR DATACOM AND ROADMAP

In 2019, data traffic inside data centers (DC) accounts for six times of the global internet traffic, reaching 11.6 Zettabytes [5]. Most of this traffic is carried through optical interconnects that count for 60% of the overall links inside data centers [6]. With increasing development of DCs in terms of number and size, the market segment of intra DC photonic transceivers has become a driver for photonic module development. Current data centers are typically using hierarchical architectures (Clos topology), with VCSEL/multimode fiber (MMF) based optical links used for short reach under 300 meters, for example to connect top of rack switch to edge switch. These interconnects are using 40 to 100 Gbps aggregated data rates obtained by using Spatial Division Multiplexing (SDM). For longer links (100m to 2km), spine switch interconnect links using singlemode fibers (SMF) together with Photonic Integrated Circuit (PIC) based transceivers are widespread [7, 8]. The development of next generation data centers, with a jointly increase of the number of hyperscale DCs, is dominated by several trends [9,10,11,12,13]:

- Increase of data center size, leading to an increase of the ratio of SMF links versus MMF
- Increase of data rate used for switch-to-switch interconnects. This is typically obtained by increasing the symbol rate per lane, the number of parallel lines (SDM) or by using more bits into each symbol through advanced modulation formats such as PAM-4.
- Increase of switch aggregated bandwidth, increasing by a factor 2 every two years

As a result, the legacy and mature VCSEL/MMF technology is reducing in terms of overall module ratio, with a growing percentage of row-to-row links being equipped with Silicon Photonics based transceivers, in addition to those already used for long range spine-leaf interconnects. In addition, server to server interconnects will largely migrate to 400 Gbps in the next years, using one of the emerging 400G standards (e.g. DR4, FR4, FR8, LR4,…). To increase the data rate per lane, optical signal will be modulated using PAM-4 format at 28 or 56 GBd, in addition to spatial or wavelength multiplexing (WDM). The complexity of such required transmitter and receiver circuit obviously advantages integrated techniques such as Silicon Photonics. Unlike in the segment of Inter DC interconnects, with lower targeted manufacturing volumes, InP based circuits may difficulty compete with Silicon Photonics, due to lower manufacturing yield, higher manufacturing cost and lower scalability.

The expected increase of Ethernet switch aggregated bandwidth would also lead to an accelerated adoption of PIC based transceivers. The general trend will be to reduce the distance between the photonic transceiver and the switch in order to accommodate higher bandwidth electrical interconnects between them.

All these emerging requirements conduct to the following roadmap for Silicon Photonics modules, which is now commonly shared by most of the stakeholders, from photonic module transceiver manufacturer to data center architects. This roadmap is illustrated by Figure 1:

---

*Figure 1 – Expected photonic transceiver module development roadmap*
Since 2016, several industrial players [14,15] have launched 100Gbps Multiple Source Agreement (MSA) products, such as QSFP28 pluggable modules intended to be installed in rack’s front panels. This module generation is currently evolving to 400 Gbps modules using Double Density QSFP modules [16,17]. As these new modules use the same form factor, it results in a 4 times increase of the overall bandwidth at the front panel. It seems difficult to further increase the aggregated bandwidth, as the physical dimension of optical fiber connectors at the front panel are not expected to evolve.

To mitigate this limitation, front panel pluggable modules are expected to evolve to mid-board micromodules designed to be positioned on the motherboard, close to the Ethernet switch package. Such modules, based on VCSEL arrays, have already been made available (e.g. TE Connectivity’s Coolbit [18], and Avago’s MicroPod [19]. Leveraging Silicon Photonics, Luxtera have also unveiled such modules [20], which integration challenges have been discussed in [21].

With Ethernet switches expected to reach 50 Tbps in 2025, the next step will be System-In-Package co-integration of the Photonic transceiver with the Ethernet switch chip [22,23,24,25]. The resulting device will be made of the switch chip (host chip) sharing the same package (e.g. BGA laminate) as one or several Silicon Photonics chips, providing a bandwidth density of Tbps/mm at the package edge. This approach will be detailed in section 4. The same approach may also be used to address High Performance Computing needs in the case the host chip is a CPU (Central Processing Unit) or a HBM (High Bandwidth Memory).

Finally, in the long term, sub centimeter photonic links should be used in computing systems also, for example to achieve intra-chip communication between cores in manycore computer systems. This kind of architecture has been proven to be an alternative to active interposers by drastically reducing the power consumption of such systems [26].

For all these new classes of module architectures, Silicon Photonics will be a key technology enabler, due to the numerous advantages that have already pointed out [27, 28]: scalability, integration of a wide range of optical functions, ability to achieve performance comparable to discrete devices, and leveraging of CMOS manufacturing and test facilities leading to economy of scale and reduced cost.

3. SILICON PHOTONIC PLATFORMS AND ADD-ONS FOR PERFORMANCE IMPROVEMENT

Silicon Photonics Manufacturing Flow

Early industrial development and commercialization of Silicon Photonics circuits and modules started at the turn of the century with Bookham’s ASOC (application specific optical circuit) platform [29]. This platform, and later the one Kotura Inc. used under licence was relying on micrometric cross section SOI (Silicon on Insulator) based waveguides, demonstrating several functions such as Variable Optical Attenuator, 40 channel multiplexers and 4x25 Gbps C-Band transmitters and receivers based on Silicon Mach Zehnder modulators [30]. However, this version of Silicon Photonics is no more used in industrial products. Nowadays, Photonic device manufacturers and industrial foundries have developed sub micrometric waveguide based platforms, following Luxtera pioneering work [31]. In this approach, 220nm or 310nm thick SOI is used.

Today’s Silicon Photonics landscape is now occupied by circuits and modules manufacturers like Intel, Cisco (who acquired Luxtera in 2019), and Acacia – all but Intel being fabless players - and industrial foundries like Global Foundries [32] TSMC, TowerJazz, AMF and so on. R&D organizations (e.g. imec, CEA Leti or IHP) also provide foundry [33] services or prototyping [34,35], often through Multi Project Wafer (MPW) services. This allows small companies or universities to access Silicon Photonics technology in order to demonstrate proof of concepts or small volume manufacturing. Some other foundries [36,37] have pursued the developments of micrometric waveguide based devices, but their development are out of the scope of this paper.

Whatever the platform is, a generic set of optical functions is now doable using SOI platforms, including routing (using silicon waveguides, splitters, and so on), wavelength management (e.g. multiplexing), polarization management (e.g. polarization separation), light out or in coupling (using optical fiber couplers), modulation, photodetection, attenuation and switching. Light generation through integrated laser source remains a limitation due to the impossibility to use Silicon as a gain material to build laser sources, however some overcoming approaches are examined in section 5.
As the main part of photonic devices required several of these above mention functions to be integrated, for example for photonic transceivers or (switches), the fabrication of Silicon Photonic based Integrated Circuits generally relies on the sequential realization of the following process flow (Figure 2):

![Basic process flow for Silicon Photonics PIC manufacturing](image)

Applying this process flow to SOI wafers leads to the following cross section of the Photonic Circuit:

![Typical cross section of a Silicon Photonics device, as processed at CEA LETI](image)

This typical flow may be somehow modified depending on the foundry; however the macrosteps will globally remain the same.

In parallel to the process development, that is now available in 200mm or 300mm diameter wafer formats, Electronic Design Automation (EDA) software have emerged to help designers assembling photonic building blocks in semiconductor like environments. These EDA tools come with specific features such as Curved shape (non Manhattan) design capabilities, automated tiling generation, specific Design Rule Check (DRC) procedures, and possibly bridges to system level models and photonic device simulations [38,39,40,41]. As a result, photonic Process Design Kits (PDK) are now available from the existing foundries, providing Device Libraries and Design Rule Manuals (DRM) taking into account the used fabrication flow and its maturity level.
Silicon Photonics Building Blocks

Existing Design Kits provide access to a full library of devices able to be integrated into PICs, aiming at routing, multiplexing, modulating and detecting light. An exhaustive description of these building blocks can be found in [42].

- Waveguide and routing devices:

Waveguide and other routing devices are obtained from the patterning of the SOI silicon layer. They are part of the passive device category, which do not required power supply. A typical Silicon Photonics library includes waveguide (rib or strip structure) to route light, Multimode Interferometers (MMI) or Directional Couplers to split optical power, crossings, etc… Figure 4 shows SEM views of such devices.

![Figure 4](image)

Figure 4 – SEM views of routing devices. From the left to the right : rib waveguide (cross section), MMI (top view)

The main characteristics of these devices are propagation losses for waveguides, and insertion losses (IL) for discrete devices. Splitting ratio and uniformity also characterizes Directional couplers and MMIs. Typical values of propagation losses obtained using Leti’s manufacturing platform are reported in the following table, for two different processes [43]. The “annealing” process dramatically reduces the loss by applying an annealing step whose result is to reduce the Si waveguide’s roughness, thus the losses.

<table>
<thead>
<tr>
<th>Waveguide type</th>
<th>Characteristics</th>
<th>Propagation loss (dB/cm), standard process</th>
<th>Propagation loss (dB/cm), with annealing step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strip</td>
<td>Thickness : 300nm Width : 350nm</td>
<td>3.6</td>
<td>&lt;2</td>
</tr>
<tr>
<td>Deep Rib</td>
<td>T=300-65nm W=320nm</td>
<td>3.9</td>
<td>&lt;2.5</td>
</tr>
<tr>
<td>Shallow Rib</td>
<td>T=300-165nm W=400nm</td>
<td>1.3</td>
<td>&lt;0.3</td>
</tr>
</tbody>
</table>

Table 1 – Propagation losses in O-Band of various waveguides processed using CEA LETI’s platform (for a full process flow including Back End Of Line)

- Wavelength management devices: These devices are passive device too, used to multiplex or demultiplex signal of different wavelength, or to act as filters. In Silicon Photonics based circuits, Array Waveguide Multiplexers (AWG), Echelle Grating Multiplexers, and cascaded Mach Zehnder Interferometers are used as mux/demux. Ring filters can also be used as add/drop combiners. A detailed description of these devices can be found in [44]. Figure 5 shows some SEM view of wavelength management devices.
Modulation devices:

In order to modulate light at data rate of 10, 25 or 56 Gbd, active devices such as Mach Zehnder Modulators (MZM) or Ring Resonator Modulators (RRM) can be used. In both case, the devices rely on phase modulation obtained through free-carrier induces index change[45]. In the case of a Mach-Zehnder Modulator, a pn diode structure is formed and superimposed to the optical waveguide of one of the MZM’s arm (Figure 7). As a result, applying a high speed modulated signal to the junction, in reverse bias mode, leads to an amplitude modulation of the optical signal at the output of the Mach Zehnder Interferometer.

![Figure 6 – Cross section of a High Speed Modulation Section used in MZM, and top view (layout) of a Mach Zehder Modulator as designed for SiPh fabrication](image)

The main characteristics of a MZM are Insertion losses, Modulation efficiency (VpiLpi, in V.cm) and bandwidth (in GHz). In reference [46], we reported the following values:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junction Losses (dB/mm)</td>
<td>1.15</td>
</tr>
<tr>
<td>Modulation efficiency (V.cm)</td>
<td>1.75</td>
</tr>
<tr>
<td>Bandwidth (GHz)</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 2 – Typical characteristics of MZM modulators

Photodetectors:

Used to convert optical signals to electrical signals at the receiver side, Ge-photodiodes integrated in a SOI platform have been developed as early as 2009 [47]. This class of end of waveguide Ge-on-Si photodetector requires several process steps. First a cavity is open at the end of the waveguide. Then Ge is epitaxially deposited in the cavity, then Ge is locally doped to form a p-i-n junction. Recently, an alternative architecture has been presented [48], avoiding doping and activation by in achieving ion implantations in Si prior to Germanium epitaxy.

Using this approach and some design optimization, the following characteristics have been measured [49]:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
</table>

under -2V bias, at 1310nm

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark current (nA)</td>
<td>10</td>
</tr>
<tr>
<td>Responsivity (A/W)</td>
<td>0.8</td>
</tr>
<tr>
<td>Bandwidth (GHz)</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 3—Typical characteristics of Ge/Si photodiodes

- **Fiber coupler devices**: an important passive device is the fiber coupler, which is the structure whose role is to couple light to and from the PIC, so as to enable an effective coupling into an optical fiber [50]. Two kinds of structures are used to achieve vertical optical coupling: one-dimensional Vertical Grating Couplers (1D-VGC) and two-dimensional Vertical Grating Couplers (2D-VGC). These structures are shown in Figure 7, left and right, respectively.

![Figure 7](image.png)

The VGC is made of a periodic or quasi periodic patterning of the waveguide, acting as a Bragg reflector that deflects light out of the PIC plane. By tailoring the shape of the Grating Coupler, it is possible to obtain an optical beam with a size and a shape that allows an efficient coupling. For example, coupling losses of 1.35 dB have been reported in [51]. However, this value is for a 1D diffraction structure, which exhibits high Polarization Dependent Loss (PDL). This is not a problem for the extraction of light from the PIC, but for injection into the PIC, polarization independent coupling is required. For this role, 2D vertical grating couplers (2D-VGC) are used. Incident light with an arbitrary polarization state from an optical fiber can be coupled to two waveguides at the output of the 2D-VGC, leading to Insertion loss of 3-4 dB and polarization dependent loss (PDL) <0.5 dB [52].

The VGC grating structure also exhibits limited optical bandwidth (+/-15 nm for 1dB excess loss). To provide efficient coupling over wider spectral ranges, edge couplers are preferred. In this case, light is emitted from the chip facet, at the end of a tapered waveguide. This requires the facet to be etched or polished with an optical quality. The taper at the end of the waveguide is used to enlarge the optical mode in order to match the mode of a lensed fiber (typically 2-3 µm Mode Field Diameter), leading to reduced coupling losses over a wide wavelength range.

**Silicon Photonics PICs**

Using this type of devices either in O-Band or C-Band, a large number of Photonic Integrated Circuits ([Erreur ! Source du renvoi introuvable.]) and packaged modules for Intra Data Center applications have been demonstrated.
This includes Ethernet 100 Gbps transceivers using 4+4 optical ports carrying 25 Gbps OOK-NRZ modulated data streams[53, 54] or WDM based modules at 12.5 Gbps per channel [55] or 25 Gbps per channel [56,57]. More recently, 400 Gbps transceivers have been demonstrated, using 56 GBd PAM-4 modulation combined with 4 lanes SDM, announcing the next generation of Silicon Photonics based transceiver products [58]

For Inter Data Center Interconnects, typically using DP-QPSK modulation formats, 100Gbps coherent transmissions have been demonstrated [59] and are now in production.

**Building Blocks improvements: low loss Si waveguides, and Si-SiN platform**

For next generation circuits in Datacom application, as well as more advanced circuits for emerging applications of Silicon Photonics (optical switches, sensors, quantum photonics), building blocks are continuously being improved, in term of losses characteristics (IL, PDL, propagation losses), speed (photodiode and modulators bandwidth), spectral characteristics (broadband devices), and so on …

For circuits exhibiting increasing size and complexity, reducing optical propagation loss of silicon waveguides is a must [60]. Propagation losses in silicon sub-micron waveguides are mainly due to light scattering on the sidewall roughness of the waveguide. Considering advanced lithographic tools such as immersion lithography in 300mm platform, a lower waveguide roughness can be obtained and losses as low as 1.3 dB/cm have been demonstrated [61 62]. Another approach to reduce losses have been developed jointly by CEA-LETI and CNRS, based on a post-lithography dedicated process to suppress the waveguide roughness [63]. This novel process consists of an annealing step at 850°C under H2 atmosphere causing atomic scale rearrangement to minimize the sidewall surface, resulting in waveguides sidewalls smoothening with no impact on the waveguide width and minor shape alteration. This process enables ultra-low propagation losses for all types of waveguides: 1.1dB/cm and 0.15dB/cm for annealed strip and rib waveguides respectively [64] thus competing with the best propagation losses reported from foundries using immersion lithography patterning. Figure 9a-b shows the effect of this annealing process on shallow rib which is used in Mach Zehnder Modulators (MZM): the propagation losses are reduced from 3.9dB/cm down to 0.6dB/cm. The annealing process preserves the performances of more complex components such as grating fiber coupler or directional coupler and no impact of the annealing on the modulation efficiency of the Mach Zehnder modulators is observed.
In addition to Si waveguide improvements, Silicon nitride layers (SiN) have been recently added in Silicon Photonics stack to enable new devices required for WDM applications [65,66,67,68,69]. SiN exhibits a low thermo-optic coefficient (1.7x10-5K-1): it allows the realization of temperature quasi-insensitive photonic devices. This is particularly interesting considering the (de)-multiplexers of the receiver and transmitter circuits which are possibly at different temperatures depending on their location in a data center707172. Therefore the use of the SiN layer to design (de)-multiplexers avoids the use of a costly temperature controller. As an exemple, we demonstrated very low temperature sensitivity (13 pm/°C) for four channels Echelle grating multiplexers (b and Figure 10e.) The device shows uniform 20nm-space channels, with insertion losses around -2.5dB and cross talk below -30dB. The -1dB bandwidth of the channels is 8nm.

The use of SiN, exhibiting a lower index of refraction than Silicon (1.88), also enables the design of enhanced optical bandwidth vertical grating coupler, which is of interest when considering WDM applications. We designed Si-SiN bilayer grating couplers. Whereas an all-SiN grating coupler has a rather low efficiency (Figure 10g blue line), the bilayer grating directionality is improved to nearly 90%, which translates into a good coupling efficiency and minimum insertion loss of -3 dB as shown in Figure 10f-g (green line). This value is remarkably uniform over the 200mm wafer as shown in Figure 10f, which reflects a good control of the Si-SiN interlayer thickness as well as the SiN low sensitivity to fabrication errors. Considering the 80nm operating bandwidth of a CWDM transceiver, it is critical to minimize the transmission difference between the channels. Therefore, with an insertion loss within the 3-5.5dB range over the O-band, the SiN-Si grating offers an improved solution compared to the all-silicon grating whose transmission shows a 1.9-13dB insertion loss range.

This additional SiN photonic layer is integrated in the photonic process flow with a good flexibility: we used a 600nm thick SiN deposited with Plasma Enhanced Chemical Vapor Deposition (PECVD) at 300°C (see Figure 10a). The integrated SiN waveguides show propagation losses of 0.6dB/cm for single mode strip waveguides and bend losses below 0.02dB/90° bends. Waveguide crossings, which are major components when addressing complex and dense circuits, have been developed, showing 0.08dB insertion loss and -60dB crosstalk between Si and SiN waveguides, and 0.07dB insertion and -60dB crosstalk between SiN waveguides. Transition between Si and SiN waveguides are managed by Si-SiN transitions relying on adiabatic inverse tapers, exhibiting insertion loss <0.2dB over the O-band for the TE mode [73].
Next generation DataCenter: co-packaged transceivers for Tbps switches, using Photonic chiplets

The data center (DC) network, providing connectivity between all servers, is based on several hierarchies of Ethernet switches connected to each other using optical fibers terminated with optical transceivers. Modern DC architectures require that traffic between multiple servers flows using as few switching hierarchies as possible such that latency, power and system cost are minimized. The flow of traffic in the DC is between a large number of compute and storage nodes (i.e., in the east-west direction) implying that each server must be connected to all other servers in its network in a full mesh topology[74]. The most efficient configuration in terms of minimal congestion is the non-blocking Clos network that requires high bandwidth and a large-enough switch radix to avoid congestion in the switch. If the switch radix is not sufficient (e.g., number of I/O ports is too small), packet loss may occur and in order to avoid it, a new switching hierarchy is added to the network increasing latency, power consumption and cost.
The chip I/O interface is based on a serializer – deserializer (SerDes) circuit and the switch bandwidth is the product: number of SerDes × lane speed. Increasing the bandwidth can be realized using either dimension; however, both are limited in their scaling. In the former case, the number of SerDes that can be allocated in the package is limited due to board breakout constraints as the number of BGA pins that need to be connected to metal traces on the board cannot exceed the line-space design rules. At 100Gb/s per lane, these constraints are becoming very strict making lane breakout from the switch a formidable task. On the other hand, the SerDes data rate will probably not exceed 53GBd due to losses and using amplitude modulation (PAM4), the highest useable data speed will be 100Gbps (progress to 100GBd is very unlikely given the loss associated with board material). Historically, switch vendors have increased both radix and data speed as shown in Figure 11.

![Figure 11. Ethernet switch historical progress.](http://www.ieee802.org/3/ch/1117_3/CFI_03_1117.pdf)

High capacity switches with a bandwidth of 51Tbps are expected to begin deployment in 2024. In order to support the switch, 64 optical transceivers will be required operating at 53GBd PAM4 on both the host and line sides. An 800Gbps transceiver will consume 15-20W (1.0 – 1.3kW for the system); adding ~900W for the switch ASIC, the power consumption is about 2kW. Given the total power consumption of a modern DC, the switch power dissipation accounts to about 1-2% of the total power and seems to be negligible. However, within a single rack, the power allocated for the switch is limited and must be addressed. Similarly, assuming an ambitious cost target of $0.75/Gb/s, the total budget of the optical transceivers is more than 5x the cost of the switch itself. In addition to these power and cost constraints, the complexity of routing 512 differential traces at 100Gb/s across the board is too high and can be carried out only for short distances within the ASIC package. It seems thus that a new approach is needed to enable these high bandwidth switches.

At the Nyquist frequency, 26.56GHz, the loss of a ~15” trace can be as high as 25-30dB, even when using a (costly) high frequency board material (e.g. Megtron 7). The outcome is that the SerDes on the switch I/O port must be able to compensate for the high signal loss. Thus, a single 100Gb/s SerDes consumes 700mW (using 7nm CMOS, ~550W with 5nm CMOS); most of the power consumption arises from the equalizer required to enhance the signal, and with 512 interfaces, the power required just for chip I/O is about 350W – out of a total chip power of ~900W. This high-power consumption of chip I/O implies that less power is available for logic processing thereby limiting the switch performance.

These problems are grave to the extent that switch vendors are seeking alternative technologies that will enable building next generation devices efficiently. The problem is even more severe when considering the 102Tb/s switches that will follow. One approach for enabling these large switch systems is to decouple chip I/O from the switch ASIC and move the interface functionalities - digital, analog and optical, to custom chiplets that will handle them. The optical transceivers are thus integrated with the chiplet and the entire system is co-packaged on a common multi-chip module (MCM), such that all chips are in close vicinity of each other. Shortening the trace length from MR to XSR (350mm to ~5mm) has the immediate benefit that the SerDes power consumption drops significantly, from 700mW to about 150mW as the lossy metal trace is replaced by a 2-3dB optical fiber link [75]. Additional power saving is achieved if the SerDes are replaced with a parallel interface between switch and chiplets.

Practically, chip I/O is handled by a chiplet with the photonic interconnect assembled either on its surface or adjacent to it. The optical interconnect is based on large, 2D silicon photonics (SiPh) matrices that have enough channels to support the chip radix requirements. SiPh is the natural technology for this task based on its scalability to large matrices, and small footprint and cost. By integrating a large number of transmit and receive cells in a single SOI chip and by integrating the analog drivers PHY functions, significant cost savings can be achieved with respect to individual, SiPh-based transceivers.
Fiber optic bundles connect the SiPh arrays to the switch chassis front panel; an array of passive fiber connectors is used to route the fibers to their recipients in the DC network. With a 51Tb/s switch bandwidth, a single chiplet should be able to handle 12.8Tbps using 128 transmit and receive lanes each running at 100Gbps. Four I/O-specific chiplets would be needed on the MCM to support one switch. The chiplet may be a purely analog device consisting of 128-element arrays of drivers and trans-impedance amplifiers (TIA); or it may be a mixed signal ASIC with both analog and digital functionalities that off-load I/O specific features from the switch silicon (e.g., FEC, MAC).

A similar approach has been demonstrated by Rockley Photonics [76 77] demonstrating a full co-packaged system based on a small Ethernet switch with an integrated optical interconnect. Ayer Labs demonstrated a chiplet solution with integrated optical interface that can be used to interconnect two ASICs. Both modules make use of external laser source fiber coupled to the device and base their photonic interface on SiPh optical interconnect. An MCM approach is used in both examples to assemble the ASIC (switch/PHY, respectively) as well as integrated fiber coupling solutions. A slightly different co-package approach was demonstrated by IBM and Finisar [78] with a VCSEL-based chip scale optical module integrated on an MCM. While the assembly effort is similar to the previous examples, the use case is different as VCSEL-based optical links are limited to sub 100m reach. Unlike the SiP based devices that can span the entire DC network, up to 2000m; this device is useful for middle-of-row or end-of-row application.

A schematic view of the suggested co-packaged design is shown in Figure 12; a 51Tbps switch ASIC is assembled on an MCM with four chiplets each handling 12.8Tbps of chip I/O. The number of chiplets could be doubled to improve the yield on the optical interconnect. The SiPh based transmit and receive modules are assembled on the chiplet or can be mounted on a small substrate next to it. The electrical connection between the switch and chiplets can be a serial one at 100Gbps based on reduced power SerDes (XSR reach) or on a parallel bus. The latter solution has an even lower power consumption at the expense of highly dense interconnect on the package.

![Figure 12. Schematic layout of a 51Tbsp co-packaged system with I/O-specific optical chiplets](image)

In terms of signal integrity, this solution does not require the high frequency traces to pass through the package core thereby improving performance as reflections are minimized. The layout of each chiplet is shown in Figure 13; the two-dimensional SiPh arrays, one for transmit and the other receive, are flipchip assembled on top of the chiplet ASIC. The analog modulator driver array as well as the TIA arrays are located directly below the photonic matrices such that transmission line effects are minimal since the trace length is the size of the solder bump, < 100µm. The chiplet is flipchip assembled onto the MCM using solder bumps on its top surface. Traffic from the switch ASIC is routed on the MCM and to the chiplet via the solder bump interface. In this design, the high frequency 100Gbps signals are confined to the package and will not be routed to the system linecard. Fiber coupling is handled by the fiber interface unit which may be either an edge coupled device or a vertically coupled device using grating couplers on the SiPh chips. For a 12.8Tbps chiplet using WDM there are 32 transmit and 32 receive fibers with four wavelength each, 20nm spacing. Alignment of the fibers is carried out using v-groove arrays and microlenses to assist in mode size conversion to the fiber. In a first approach of such approach, laser sources will be kept outside of the package for temperature range and temperature sensitivity reasons, however some strategies for integrating laser sources should be addressed in the long term.
An example of such a 2D SiPh array is shown in Figure 14, built as part of the H2020-funded project L3MATRIX [79]. A 64 element SiPh array was fabricated in two groups of 32 elements; each cell has a MZ modulator and an integrated DFB laser. The lasers are on a CWDM grid and every 4 adjacent cells are multiplexed together using an Echelle grating located in the area between the two groups of transmitters. The SOI chip was designed for flipchip assembly on a chiplet.

The main rational behind decoupling of chip I/O from the switch is to save power via reducing the SerDes complexity. One advantage of the fact that the chiplet is a CMOS chip is that it can be designed as a mixed signal chip and to off-load any I/O-related functionality from the switch to the chiplet. Thus, the FEC and MAC blocks, as well load balancing (Rx) and packet reassembly (Tx) can be integrated in the chiplet along with the analog blocks – modulator drivers and TIAs. In addition, signal monitoring functionality may be added as well as redundant channels to improve reliability. By moving these functions from the switch to the chiplet, additional power savings are obtained as well as freeing silicon area in the switch. The SerDes power drops from 350W to about 77W; using four or eight chiplets to drive the optical interconnect, equivalent to 64 transceivers - 20W each, allows for additional power saving as the unified circuits of the PHY chip have a lower power consumption. Similarly, the transceiver cost is greatly reduced, by a factor of ~2 as a result of the combined overhead and unification of laser sources and driver electronics.

**Optical Network On Chip : rationale , state of the art and requirements**

While the present and near future of silicon photonics in computing systems is on the tight integration of boards and modules with photonics in the data center, the trend towards deeper and deeper integration continues further. New high performance computing applications for big data, graph analysis and artificial intelligence push for more and more cores and dedicated accelerators densely integrated with huge amounts of memory with fast access times. Moving towards exascale high-performance computing systems is indeed not feasible by simply scaling-out the current systems with more nodes and an efficient data-center network. Computation and energy efficiency needs to be increased with more locality and sharing of memory and communication resources within integrated many-core computing nodes.

Nevertheless, increasingly complex CMOS technologies create big challenges on modularity and scalability of many-core systems. Single large-scale monolithic multiprocessors become more and more costly to develop because of fabrication yields. Even though core redundancy allows to bypass non-functional cores, the chip level communication
infrastructure cannot tolerate data link failures due to fabrication defects [80]. To overcome this issue, a clear trend has appeared to partition the design in smaller dies with yield management and known-good-die sorting. These smaller dies become “chiplets” that can be integrated with 3D-stacking technology on large-scale interposers in mature process nodes. This advanced packaging strategy of partitioned subsystems allows for modular system composition and high scalability [81], with a high fabrication yield improvement. Indeed, a big opportunity for chiplets and interposers is the ability to couple the different technology platforms between chiplets and with the interposer.

For this chiplet-based approach to be efficient, though, a high pressure is put on die-to-die communication at interposer level, whether for core to memory, offloading to dedicated accelerators and GPUs, or core-to-core coherence traffic. In all communication schemes and for any number of chiplets, the requirements are a high bandwidth density and low latency within a low power consumption envelope. To achieve these goals, the interposer technology is a key choice to perform in view of the communication and scalability needs [82].

Passive interposers with metal transmission lines on organic or silicon substrate have been developed in industry for a few years [83]. They represent the direct evolution of board-level traces to the interposer for a scale-in integration, but the large free area of the interposer allows for many parallel lines with relaxed throughputs and aggressive pitch where board-level communication would quickly rely on data serialization. Nevertheless, this technology does not allow for multiplexing or switching, and is limited to perimeter-bound communication. It is well suited to point-to-point communication between two dies, but suffers from very poor scalability, which cannot be extended much even when re-introducing serialization.

The idea of moving to active CMOS interposers brings more flexibility with the ability to buffer the signals across long distances, and route, arbitrate and switch the communications using networks on chip (NoC). The first demonstrations of NoC on CMOS interposers with are only emerging [84] from research labs, yet with a strong interest from industry [85]. Full digital synchronous operation of the NoC, however, brings constraints on signal propagation, clock distribution and pipelining, which limit throughput and dramatically increase latency for large-scale interposers.

Consequently, with increasing maturity of silicon photonics platforms and compatibility with face-to-face die hybridization, it is now possible to consider the implementation of optical network-on-chip (ONoC) topologies on silicon photonic interposers to go beyond the scalability limitations of active CMOS interposers. Indeed, with efficient optical modulators and switches and with low-loss optical waveguides, silicon photonics could allow building complex routing topologies at large scale with the latencies of electromagnetic wave propagation.

First demonstrations of optical communication for computing architectures were done in 2016, with an optical link between a dual-core RISC-V processor and a 1MB memory [86]. But this architecture included a single optical link between the processor and the memory, with no optical routing scheme. Nevertheless, the initial proposals for ONoC architectures were done as early as 2007 [87]. These initial works, both at architecture level and for demonstrators, have considered a monolithic integration of silicon photonics with CMOS. On the process side, the use of SOI technologies both for some CMOS nodes and for silicon photonics has drawn efforts to make these technologies compatible, which is a cost effective solution for small-scale architectures. Nevertheless, as Moore’s law draws high-performance computing towards more aggressive nodes, the dependence of a monolithic integration to SOI appears as a clear limitation. Besides, regarding architectures, modularity and yield management have taken time to settle as correlated requirements of large-scale manycore systems. Opportunely, 3D integration technologies have developed in parallel with silicon photonics, and ONoC architecture proposals have also evolved with this perspective.

Initial large-scale ONoC architectures were inspired by their electronic counterparts. The 2D-mesh topology, which has been favored in NoC for its scalability by tiling of identical blocks for high-performance computing was present in several proposals [88,89,90,91]. Nevertheless, electrical and optical NoCs have a fundamental difference regarding routing: packet-switching has become common in electrical NoCs, where successive routers on the data path are switched with in-band routing information contained in the control preamble of the data, while optical switches require out-of-band circuit-switching sent before the data to configure the whole path. For this reason, to avoid the many hops incurred during traversal of a 2D-mesh grid, shallower topologies have been favored to limit the amount of routing. Crossbar topologies, sometimes expanded to Clos networks with fixed connections in the input or output stage for given wavelengths, present a single point of switch required between source and destination. These topologies have been derived into several variants, either centralized or distributed: the most frequent topologies use a dedicated link per source, which is tapped by all destinations for Single-Writer-Multiple-Reader (SWMR) topologies [92,93], or conversely a dedicated link per destination, accessible by all potential sources for Multiple-Writer-Single-Reader (MWSR) topologies [94]. By sharing a link for Multiple-Writer-Multiple-Reader topologies, the amount of resources is reduced, with additional contention to the accessed medium [95, 96, 97,98], which may be solved by using reserved wavelengths multiplexed in a waveguide. To lay out these shallow topologies while being compatible with a 2D tiling of computation cores, there has been a consensus towards ring-
based waveguide bundles arranged in a serpentine layout [99,100]. This however comes with long waveguides passing across all tiles.

In all proposed optical topologies, the key requirement for scalability is on the optical power budget along the optical path, starting from the laser source to the electro-optical photodetector on the receiver side. This optical power budget faces the accumulation of optical losses all along the optical path, starting from light couplers in the case of an external light source, then propagation losses inside the waveguides, possible attenuations in crossings, and, more importantly insertion losses within the optical modulators and switches, both for a given wavelength and due to inter-wavelength attenuation. Given a typical MWMR topology, with a series of modulators from different sources followed by a series of filters to different destination, the optical budget evolves as described in figure below, with one of the filters routing light from the shared waveguide to a dedicated waveguide connected to the photodetector.

![Figure 15. Optical Link Budget in a photonic ONoC](image)

This Figure 14 shows some implications on the performance requirements of the photonic devices for large ONoC topologies. Obviously, the long optical paths required in such topologies advocate for ultra-low loss waveguides with attenuations around 0.2dB and below. Besides, the losses in the switching filters, when routing from a waveguide to another, should be minimized, and as said before, the topology itself should avoid multiple switches to limit those losses. But additionally, most of the losses for large-scale topologies would come from the passing modulators and filters, which may scale proportionally to the number of electronic tiles connected to the ONoC. The passing insertion losses of these devices should be made minimal, while providing the ability to switch them quickly when the communication needs to be established.

Moreover, along with power budget, the close integration of optical devices with computation cores in an ONoC architecture brings several additional constraints on the optical devices: on performance, on temperature robustness, and on area & form factor.

Regarding performance, the applicative traffic patterns inside the ONoC are more distributed than Datacom & Telecom transceivers. Bandwidth aggregation is not so important, and occurs primarily at the module interface, not internally: on most applications, core bandwidth requirements are around 1 byte per flop, i.e. in the 10-50Gbps range. Clustered computation may increase these needs at the electro-optical interfaces, but wavelength multiplexing relieves these constraints. Therefore, transmitters and receivers performance does not need to be pushed to extreme values and should be established as a power tradeoff between amortization of static optical costs and increasing SerDes costs to stabilize around 10Gbps.

Regarding temperature robustness, dense integration of photonic devices in close proximity to intensive computation logic creates a hard thermal environment, with temperatures ranging from ambient to 80°C and more, subject to local hotspots and millisecond-range variations. For this reason, either athermal broadband devices or actively stabilized photonic devices are required. Several solutions have been developed for thermal control of resonant devices [101]. In this latter case, actual thermal isolation of the devices is sought for, in order to increase the thermal efficiency of the control on the device resonance. This can be achieved using back-side cavity etching under the devices to isolate.

Finally, regarding area and form-factor, dense integration in a computing architecture requires the actual footprint of all the electro-optical control to be minimal with respect to the computation resources. Besides for a localized control,
form-factors of the optical devices should be limited, as well as groups of optical devices consistently controlled to interface between a processing cluster and the ONoC. The compute chiplets may actually have different IP distribution models, integrating photonic drivers or leaving the photonic control to dedicated chiplets. In this case, transceiver chiplets should have a near-square aspect ratio. For this reason, compact modulators and filters such as microring resonators are favored.

In terms of integration, the photonic interposer and the chiplets will exchange information at high datarate, both between the processing clusters and the transceiver chiplets and between the electro-optical control in the transceiver chiplets and the photonic devices on the interposer. Therefore, face-to-face assembly of the chiplets and the interposer using microbumps is favored for high-speed serialized electrical signals, removing the need for Through-Silicon Vias (TSV) in the chiplets. Nevertheless, the photonic interposer itself needs to include TSV, not so much for external communication, which may be extended optically, but for power delivery to the processing and transceiver chiplets, which needs to pass through the interposer from back-side bumps. The corresponding power TSV and low-bandwidth TSV for configuration and service signals have however quite different requirements from the short-term optical modules with serialized data through the TSV. The resulting conceptual assembly for ONoC interface to compute chiplets is shown in Figure 16. The details of this 3D integration will be presented in the next section.

![Figure 16: Conceptual assembly of a photonic ONoC interface relying on a Silicon Photonic Interposer withs TSVs.](image)

5. REMAINING CHALLENGES FOR NEXT GENERATION MODULES AND PHOTONIC INTERPOSER

Section 4 highlighted two emerging applications of Silicon Photonics Circuits: co-packaged transceivers and Optical Network On chip for manycore computers used in exascale systems. Both application emphasized future needs for Photonic Integration which are not already existing in current foundries, in order to increase bandwidth density, complexity, and insertion in a Semiconductor manufacturing and packaging flow. To address these needs, at least four key challenges need to be solved: i/ 3D integration, in order to make PICs benefit from dense TSV based electrical interconnects, thus reducing the Interposer or chiplet footprint; ii/ laser integration, enabling the photon source to be integrated in the same chip as the photonic functions; iii/ fiber packaging, with new techniques definitely needed to reduce cost and packaging throughput, while moving to semiconductor assembly processes; iv/ high efficiency modulators, with enhanced bandwidth/footprint/voltage bias function of merit.

**3D Integration:**

Advanced packaging allows for new architectures and better signal transmission, leveraging so called “3D Integration” methods already widely used in CMOS processes [102]. In such applications, standard wire-bonding techniques are replaced with denser interconnects with TSVs (Through Silicon Vias) and micro bumps (also known as copper pillars, with typical pitches of 50µm on the chip front side). The lower inductances and shorter electrical paths compared to standard wire bonding improve signal integrity and throughput between the CMOS chips and the motherboard. Moreover, 3D interconnects can be integrated on the whole surface of the silicon die, while wire bonding is limited to the periphery, significantly reducing the number of available I/O. As explained in section 4, this 3D integration approach should be
applied to Photonic Integrated Circuits, enabling Photonic Interposers approaches for ONoCs. Also, for photonic chiplets used in co-packaged transceiver, 3D technique could help increasing the module density and avoid complex and bandwidth limited wire bonding interconnects. One of the key components of such structures is the TSV, connecting the front and back sides of the interposer. The two main types are the “via middle” (TSVM) approach and the “via last” (TSVL) approach. TSVMs are processed after FEOL and before BEOL, while TSVL are done after BEOL. The main difference between the two are filling and aspect ratio. TSVM exhibit higher aspect ratios (typically 1:10) and are fully filled with Cu, while TSVL are larger and filled with a Cu liner only. Because of their smaller diameter and higher aspect ratio, TSVM are preferred in applications where I/O density is crucial, such as in HPC or photonic interposers.

In the past years, silicon photonic interposers have been demonstrated by A*STAR IME [103]. TSVM process has been used, with dimensions 20µm by 100 µm (AR 1:5), on 200mm SOI wafers. The architecture is presented in Figure 17. This study was a proof of concept for Photonics/TSV integration, showing suitability of TSV architectures to connect standard photonic building blocks, such as modulators or photodetectors.

**Figure 17: Cross-section of the photonic interposer architecture from IME [103], system overview for imec [104] and CEA-Leti’s Starac architecture**

IMEC proposes another structure for the photonic interposer, with only one chip processed with TSVM (10 x 100 µm) [104]. The TSV can handle RF signals up to 50 GHz. In this paper, the approach to the coupling between optic fibers and the interposer is different from the previous design. The fibers are laterally coupled with a SiN/Si transition, allowing for a much larger bandwidth than vertical grating couplers with similar insertion losses.

Coefficients of thermal expansion (CTE) mismatch between copper and silicon is significant ($a_{Cu} = 17 \cdot 10^{-6}K^{-1}, a_{Si} = 2.6 \cdot 10^{-6}K^{-1}$), eventually causing residual stress in the SOI layer. This stress has to be investigated in photonic interposer applications, as some photonic building blocks may be sensitive to mechanical stress (e.g. resonator filters). Reference [105] defines optical Keep-Out Zone to be respected by photonic designer in order to avoid stress-induced in photonic structures.

**Figure 18: TSV mid AR 10 (left) and TSV Last AR 5 (right) – CEA Leti**

Moreover, the interposer is prone to bowing, which can interfere with the stacking of the chips on the front side or with the application of solder balls on the back side of the interposer. The deformation of the chip is caused to the thinness of the chip (100 µm or less), the presence of a thick BOX (Buried Oxide) layer and a process on both sides of the interposer. In order to secure the process, one must keep the bow under control [106] on a temperature interval ranging from operating temperature of the final product (generally 25°C or close) up to the reflow temperature of the solder joints, typically above 250°C.
**Laser Integration:**

In this section, we highlight some recent achievements on the integration of a laser source on Silicon for Datacom applications, focusing on hybrid III-V on Silicon devices that offer a high integration level compared to flip-chip [107]. A large number of academic and industrial actors have been involved in the development of hybrid III-V on silicon laser sources [109,110,111,112,113,114,115,116,117]. Among these developments, one can observe three main types of optical coupling methods between the III-V Gain region and the silicon photonic circuits, namely butt-coupled, evanescently coupled, and supermode coupling [118] (also referred as adiabatic tapering) hybrid cavities. Although limited by alignment and mode-matching constraints at a packaging level, butt-coupled hybrid lasers have shown equivalent performances compared to evanescent/adiabatic hybrid ones. In order to solve these integration constraints and increase the scalability of hybrid PICs, the co-integration of III-V gain medium and silicon PICs through a monolithic integration appears as the most efficient way to simultaneously improve the performances and reduce the cost of optical transceivers. Figure 19 depicts two types of monolithic integration cross-sections based on evanescent and adiabatic optical coupling.

![Figure 19 Schematic view of hybrid III-V on silicon waveguides based on (a) evanescent coupling adapted from Erreur ! Source du renvoi introuvable, and (b, c) supermode control through adiabatic tapering adapted from Erreur ! Source du renvoi introuvable.](image)

In both cases, the process flow involves first a conventional sequence of SOI patterning in order to define passive devices, modulators, and eventually germanium photodiodes. Next, the III-V epitaxies are produced separately on either InP or GaAs based wafers with a particular attention on the surface defect density to remain compliant with wafer to wafer bonding specifications. The evanescent coupling shown in Figure 19(a) relies on a direct bonding of III-V on silicon, which benefit from the intrinsically flat surface of the SOI layer for the hybridization together with air gaps to ease the extraction of gasses produced during the thermal annealing performed after the bonding [119]. This hybridization was formerly developed by UCSB and Intel [120,121] and next pursued by several groups. Alternatively, the III-V and silicon materials can be spaced by a thin bonding layer of about ~100 nm, in order to encapsulate the overall photonic integrated circuit, made of either silicon dioxide or BCB. The space between the two waveguide induces a degeneracy lift of the optical modes, which can be handled to realize adiabatic mode transformation [122]. This feature allows an improvement of de confinement on the MQW region, which increases the accessible modal gain compared to the evanescent coupling for an equivalent III-V stack, at the expense of a complex design method [123].

From a fabrication point of view, the evanescent hybrid lasers requires mandatory III-V tips at both extremities of the gain section in order to transfer the part of light from the III-V to the silicon, preventing undesired intra-cavity reflections, as shown in 0.
While the critical dimensions for such optical tips (~500 nm) are accessible with any DUV lithography, the large III-V stack height must be tapered in several section to prevent from undesired fragility and to limit the backward reflections.

On the other hand, adiabatic couplers patterned onto the SOI layer makes the use of III-V tips optional (depending on the SOI thickness and III-V epitaxy), facilitating the III-V patterning of the Gain, MQW, and Mesa layers, which can nevertheless suffer from sidewall slope during the etching, as shown in 0More recently, CMOS-compatible process for the patterning and metallization of III-V on silicon was demonstrated (Figure 20(d)), using only dry etching steps and low access resistance BEOL based on titanium contacts, tungsten plugs, and aluminum pads [127,128].

From the two types of optical coupling discussed above, different laser cavities were proposed, starting historically by Fabry-Perot cavities. Next, single frequency lasers were successfully demonstrated, based on either Distributed Bragg Reflectors (DBR) [129,130], ring/racetrack cavities [131], tunable lasers using ring filters [132,133], or Distributed Feedback (DFB) lasers [134,135,136], as shown in 0

For intra-datacenter communications, such single mode hybrid III-V on silicon laser exhibiting more than 3dBm output power and high sidemode suppression ratio (>50dB) coupled to a single mode fiber in the O-band is of practical interest. Since the patterning of the quarter-wave shifted DFB is realized on the SOI layer with a high accuracy, such cavity can be declined to deploy wavelength division multiplexing (WDM) links [137]. Co-integration with high-speed silicon Mach-Zehnder modulator was subsequently demonstrated with adiabatic coupling [138]. Similarly, evanescent hybrid lasers were co-integrated for producing 100G CWDM4 transceivers [139]. To the best of our knowledge, state-of-the-art performances
of single frequency hybrid lasers in term of threshold current (<20mA), optical power (>25mW), and thermally behavior (up to 80°C) over 120nm discrete wavelength range was also demonstrated by Intel Corp. for next generation 400G transmitters[140].

At this stage, direct modulation of the optical signal (NRZ, PAM4, PAM8...) appears as a straightforward encoding solution for short-reach links up to 800Gbps. Nevertheless, the cost reduction offered by silicon photonics may open some opportunities for high-speed coherent transceivers for Exascale computing [141,142]. In this regard, widely tunable hybrid lasers for the emitter and for the local oscillator at the receiver side, already developed in the C-band [143,144], might be of great interest in the future. For coherent transceivers, it is worth noting that phase and amplitude noises of hybrid DFB lasers can be largely improved by an appropriate design cavity [145].

In summary, heterogeneous integration of III-V on silicon have been widely developed for more than a decade. The motivations that arose initially from telecom domain in the C-band have quickly shift to O-band for answering the rapidly growing needs of datacenters. The performances have been enhanced years after years thank to an improvement of both laser designs and fabrication process, bringing hybrid III-V on silicon PICs onto the market. The reader is encouraged to pursue the short review proposed here by taking into account the thermal environment and packaging which impacts directly the performances of the lasers [146,147]. From the integration point of view, one can note a strong incentive on the micro-transfer printing processing as an alternative of the molecular bonding process discussed above [148,149]. Recently, hybrid III-V on silicon lasers with a reduced footprint based on either microring [150] or membrane lasers with lateral current injection [151] have been demonstrated, which opens bright prospects for next generation high density interconnect, for short-reach and ultra-short reach, depending on the emitted power, direct modulation bandwidth, and thermal behavior.

Fiber Packaging:

Early generation of Silicon Photonics modules, e.g. from Luxtera, currently uses legacy fiber coupling techniques that have been developed for previous generation of Integrated Optics devices [152, 153]. Typically, an optical fiber ribbon, inserted into a v-groove array, is actively aligned with submicronic accuracy to a linear array of Vertical Grating Couplers, then fixed using UV curing epoxy. The main difference with non SiPh modules is the required angle (typically 8°) of the optical fiber axis to the surface of the Photonic chip. This makes low footprint of photonic modules difficult to achieve, however it can be reduced to 5 to 8 millimeters by using bend tolerant fiber. This technique has been pushed to extremely high channel number density by using Multicore Fiber (MCF), that could be aligned to VGC 2D arrays [154].

This approach suffers from several drawbacks when low cost (<1 USD) assembly is targeted, as it is expected for next generation transceivers. First, in new applications, the overall height of the coupling element should be kept as thin as possible, to enable thermal heat sink assembly, and board stacking in blades. Furthermore, it is mandatory to rely on an assembly technique that withstand further temperature sensitive steps, like solder assembly on a motherboard. This should limit the use of epoxies and fiber pigtails. Also, fiber pigtails are difficult to manage by using a standard pick and place equipment, required by high assembly throughput. This three limitations (assembly cost, solder reflow compatibility and compatibility with semiconductor packaging equipment) lead to the development of new solutions to perform fiber connector assembly.

As a result, several alternative approaches have been recently developed and applied to Silicon Photonics circuits instead of using standard fiber to grating active alignment. The a first approach relying on interposing microoptics element, such as microlenses array, between the VGC and the fiber optic connector [155]. A second one is using passive alignment of optical fiber at the end of the PIC coupler (edge coupling configuration) [156], and a third one relies on evanescent coupling between the PIC waveguides and a polymer flex embedding optical waveguides [157]. The following figures summarizes the three coupling scenarios:
The microoptics approach aims at enlarging the output beam size from the PIC, leading to relaxed alignment tolerances between the PIC and the fiber optic connector. For example, a mode enlarged to a diameter of 80 µm and collimated at the output of the microlens should lead to axial tolerances of 19 µm at 1dB at 1310nm when using lenses on the top of a VGC, and at the tip of the fiber. This approach enables the insertion of an optical isolator or a reflector to get low profile connector system. In this configuration, angle tolerances should be kept below 0.2°, and the absolute position of the lens towards the VGC and the optical fiber core below 1 µm. This range of tolerance can be obtained by using self-alignment approaches, as described in [158].

The in-plane edge butt coupling technique requires 1-2 µm tolerance between the output mode of the fiber and the PIC coupler for a output mode size diameter of 8µm. By using metamaterial based edge couplers, Barwicz et al. [159] demonstrated coupling losses of 0.7 dB (TE mode) 1.4 dB (TM mode) between the PIC and an optical fiber. This metamaterial edge coupler, acting as a Spot Size Converter (SSC) enables the use of self-alignment techniques such as etching V-grooves in the PIC substrate to allow self positioning of a flat cleaved fiber at the end of the taper. This configuration is much more compatible to mass production, and more scalable than coupling of smaller mode size to lensed fibers.

The same team from IBM demonstrated another technique [160] involving evanescent coupling of the optical beam from the PIC to an optical flexboard with embedded waveguides. This compliant flexboard is patterned with mechanical microstructures which allow the flex to be passively self-aligned to the PIC, provided that the optical chip contains complementary etched grooves. At the other end of the flex, other microstructures enable an MT connector interface to be accurately aligned, allowing an MT connectorized fiber ribbon to be connected to the flex. 0.8 dB coupling penalty have been demonstrated over 100 nm bandwidth, for both polarization states, after UV epoxy curing of the various parts.

Other approaches are emerging to set new standards in term of optical coupling. Among them, Ayar Labs is using angle polished gradient index fibers to reduce the thickness of the module and PETRA is using optical pins. This works are detailed in references [161] and [162].

High efficiency modulators:
A last key device to optimize for next generation applications are modulators. With length in excess to 500µm, standard Si Mach Zehnder modulators fails to cope with emerging requirements in term of footprint and efficiency. The most mature device to replace MZM are Ring Resonator Modulators. Their intrinsic small dimensions (<30µm), plus the possibility to cascade them to build WDM transmitters lead to bandwidth density of 1 Tbps/mm2 [163]. However, using a WDM approach and ring resonator require thermal tuning, adding penalties in term of energy consumption in a complete system.

Other new type of modulators could be considered in the future.

SiGe-based Electro – absorption modulators, based on the modulation of optical absorption under applied electric field [164] are certainly the most promising in term of CMOS compatibility, footprint and consumption, with typical length < 100µm, consumption <100fJ/bit and electrical bandwidth up to 50GHz for the C band [165,166,167,168]. Ge/SiGe
multiple quantum wells (MQWs) have been introduced to sharpen the band edge absorption spectrum, in order to improve the ratio between extinction and insertion losses. Ge/SiGe MQWs can also be used to tune the spectral position of the band edge around the C band. Encouraging approaches to reach the O band have been proposed using SiGe/SiGe quantum wells or strained quantum wells [169] but remain a challenge. Hybrid IIIV-on-silicon modulators, demonstrated until now only in the C band [170,171], could also be tuned in the O band or used to reach higher electrical bandwidth in the C band, at the cost of a more complex fabrication process.

Hybrid IIIV-on Silicon capacitive Mach Zehnder modulators have been introduced recently with reported efficiency in the range $V_L \sim 0.05-0.1\text{Vcm}$, $V_L \alpha$ products of $\sim 1-2\text{dB.V}$ [172,173], and expected consumption around 10-100fJ/bit. High mobility and low effective mass of IIIV allows enhancing both the efficiency and the bandwidth of Si capacitive modulators. Further, quaternary alloys (InGaAsP) allow addressing either the C or O band. Electrical bandwidth of only few GHz have been however reported for such efficient designs, because of too high RC constants. Electrical bandwidth >30GHz have been reported using the lower capacitance of depletion regime (at the cost of reduced modulator efficiency) [174]. In the near future, high electrical bandwidth might be reached without losing efficiency by combining carrier depletion and Franz Keldish effect [175,176,177].

Unlike Silicon-based phase shifters, phase shifter lengths <200µm could be achieved without using microring resonators, ie, keeping a broadband behavior and avoiding power-demanding heaters to trim the resonant wavelength.

Hybrid Organic-inorganic modulators are based on the Pockels effect in organic materials. Performances have risen during the last few years due to improvement of organic materials performances [178,179] and stability [180], as well as the use of advanced waveguide designs, like slot waveguides [181,182] or plasmonic waveguides [183,184]. $V_L \alpha$ down to $\sim 0.03\text{Vcm}$ and $V_L \alpha$ product of 1.2dB.V have been demonstrated for Silicon-organic hybrid (SOH) modulators using organic material with record Pockels activity ($r_{33} \sim 390\text{pm/V}$) [185]. Using such devices, a phase shift of $\pi/3$ should be reached along a 50µm-long phase shifter driven at 2V with only 0.2dB of insertion loss. These losses, due partly to the slot waveguide roughness, might be even decreased using state-of-the-art nanofabrication facilities. The viability of the SOH modulators has been further demonstrated by generating high quality on-off-keying signals at 40 Gbit/s with Q factors in excess of 8 and consumption <1fJ/bit. Line rates (symbols rates) of 128 Gbit/s (64 GBd) using quadrature-phase-shift-keying (QPSK) modulation and of 160 Gbit/s (40 GBd) using 16-state quadrature-amplitude-modulation (16QAM) have been also obtained using a custom ceramic printed circuit board [186]. POH modulators, based on plasmonic waveguides, have even smaller access resistance and higher optical confinement than SOH, and overcome SOH performances [187]. Record efficiencies (0.013V.cm), data rate transmission (up to 400Gbit/s) and ultra-low consumption (0.07 fJ/bit -0.3 fJ/bit) [188] have been recently demonstrated, at the cost of high (~ 10-15 dB) insertion losses in gold-based plasmonic waveguides. Low-loss and CMOS-compatible plasmonic materials remains the biggest issue to make POH modulators realistic candidates for datacom applications. Finally, as far as nonlinear organic materials is concerned, challenges remain like preservation of high Pockels coefficients ($rij$) during device operation. Encouraging results have been recently reported [189], with high $r_{33}$ value (up to 290pm/V) maintained for over 500 h at 85°C. Other challenges are mostly related to integration of such high performance material at wafer-scale in a CMOS foundry. Toxicity management, processing, poling, and encapsulation have still to be investigated.
6. CONCLUSION

Recently, Silicon Photonics Technology has been adopted to build high speed (100Gbps, then 400Gbps) transceivers modules addressing optical interconnects in Data Centers, and also for inter Data Centers links. Intrinsic advantages of this technology in term of : co-integration with CMOS electronics, scalability, and the versatility of the wafer level process flow, enabling materials such as InP, Ge or SiN to be added to the existing platforms should enable Silicon Photonics to be the key enabler for achieving two major challenges of the coming years to be taken. The first challenge will be first to allow multi terabit switches or CMOS digital chips to benefit from optical I/Os , by directly co-packaging the host chip with the photonic transceiver. The second challenge will be to develop optical intra-chip communications, paving the way to next generation manycore computers.

To achieve these two challenges, some progress are still needed in four fields : to develop 3D packaging /Photonique convergence in order to get Silicon Photonic Interposers ready ; to get integrated process flows for laser diode integration; to develop novel optical connectors ; and to introduce high efficiency, low footprint modulators.

ACKNOWLEDGEMENTS

“This work was supported (in part) by the following projects : French National program “Programme d’investissement d’avenir, IRT Nanoelec, n° ANR-10-AIRT-05”

COSMICC project
PICTURE project
MASSTART project

The authors would like to thanks Dr. Daived Fowler for providing valuable information on Vertical Grating Couplers.

REFERENCES
1 Hecht J., City of light, The story of fiber optics, Oxford University Press, 340 pages, 2004
4 Po Dong, Young-Kai Chen, Guang-Hua Duan and David T. Neilson, Silicon photonic devices and integrated circuits, Nanophotonics, 3, 215-228, 2014
7 T. Liljeborg, Silicon photonics and the future of optical connectivity in the data center in 2017 IEEE Optical Interconnects Conference (OI), 2017
12 A.V. Krishnamoorthy, Optical Interconnect Packaging for the Cloud, ECTC Special Session, 2016
14 http://www.luxtera.com/100g-qsfp28-modules/
15 https://www.intel.fr/content/www/fr/fr/architecture-and-technology/silicon-photonics/optical-transceiver-100g-psm4-qsfp28-brief.html
16 http://www.qsfp-dd.com/
18 End-to-End Communications with Advanced Fiber optic Technologies, white paper, available online: https://www.te.com/content/dam/te-com/custom/styles/everyconnectioncounts/assets/docs/data-communications-coolbite-te-25g-active-optics-white-paper.pdf
23 Rockley Photonics, Sailing through the data deluge with pervasive optical connectivity, available online: https://rockley photonics.com/wp-content/uploads/2019/02/Rockley-Photonics-Sailing-through-the-Data-Deluge.pdf
24 P. Meade et al., TeraPHY: A High-density Electronic-Photonic Chiplet for Optical I/O from a Multi-Chip Module, paper M4D.7, Proceeding OFC 2019
27 Roel Baets, 20 years of silicon photonics: lessons learned, lessons to be learned, 2018 INTERNATIONAL CONFERENCE ON OPTICAL MEMS AND NANOPHOTONICS (OMN) . In International Conference on Optical MEMS and Nanophotonics p.213-214, 2018


40. Mentor Graphics, Silicon photonics verification: challenges and solutions (white paper)


46. B. Szelag et al., Optimization of 64Gbps O-band thin-rib PN junction. Mach-Zehnder Modulator fabricated on a 200mm silicon photonics platform, SSDM 2018.

47. L. Vivien et al., « 42 GHz p.i.n germanium photodetector integrated in a silicon-on-insulator waveguide », Optics Express, vol. 17, n° 8, p. 6252-6257, 2009.


49. H. Zegmout et al., High speed integrated waveguide lateral Si/Ge/Si photodiodes with optimized transit time, accepted for publication, Proceedings of the SPIE, 2020.


54. 100G Parallel Single Mode Data Center Connectivity, Product Brief, Intel, 2016.


https://www.intel.fr/content/www/fr/fr/architecture-and-technology/silicon-photonics/optical-transceiver-100g-cwdm4-qsfp28-brief.html


Yamaoka, Suguru, et al., "239.3-Gbit/s Net Rate PAM-4 Transmission using Directly Modulated Membrane Lasers on High-Thermal-Conductivity SiC." ECOC 2019

D. Guckenberger et al., « Advantages of CMOS photonics for future transceiver applications », présenté à European Conference on Optical Communication, ECOC, 2010


Takemura et al., High Density Optical and Electrical Interfaces for Chip-scale Silicon Photonic Receiver, ICEPT 2017 Proceedings TE2-1.

Thonnart op. cit.


Sanghyeon Kim, Member, IEEE, Younghyun Kim, Member, IEEE, Yoojin Ban, Marianna Pantouvaki, and Joris Van Campenhout, "Simulation Study of a Monolithic III-V/Si V-Groove Carrier Depletion Optical Phase Shifter", IEEE journal of quantum electronics, 56(2), 6300208 (2020).

Qiang Li, Chong Pei Ho, Shinnichi Takagi, and Mitsuru Takenaka, “Optical Phase Modulators based on Reverse biased III-V/Si Hybrid Metal-OxideSemiconductor Capacitors”, accepted for publication in IEEE photonics technology letters, 2020.


