Monolithic silicon-photonic platforms in state-of-the-art CMOS SOI processes [Invited]

VLADIMIR STOJANOVIĆ,1,2* RAJEEV J. RAM,2,3 MILOS POPOVIĆ,2,4 SEN LIN,1 SAJJAD MOAZENI,1 MARK WADE,2 CHEN SUN,2 LUCA ALLOATTI,3 AMIR ATABAKI,3 FABIO PAVANELLO,5 NANDISH MEHTA,1 AND PAVAN BHARGAVA1,2

1Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94709, USA
2Ayar Labs, 6460 Hollis St, Ste A, Emeryville, CA 94608, USA
3Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
4Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215, USA
5Department of Electrical, Computer, and Energy Engineering, University of Colorado - Boulder, 425 UCB 1B55, Boulder, CO 80309, USA

Abstract: Integrating photonics with advanced electronics leverages transistor performance, process fidelity and package integration, to enable a new class of systems-on-a-chip for a variety of applications ranging from computing and communications to sensing and imaging. Monolithic silicon photonics is a promising solution to meet the energy efficiency, sensitivity, and cost requirements of these applications. In this review paper, we take a comprehensive view of the performance of the silicon-photonic technologies developed to date for photonic interconnect applications. We also present the latest performance and results of our “zero-change” silicon photonics platforms in 45 nm and 32 nm SOI CMOS. The results indicate that the 45 nm and 32 nm processes provide a "sweet-spot" for adding photonic capability and enhancing integrated system applications beyond the Moore-scaling, while being able to offload major communication tasks from more deeply-scaled compute and memory chips without complicated 3D integration approaches.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

OCIS codes: (250.3140) Integrated optoelectronic circuits; (250.5300) Photonic integrated circuits; (200.4650) Optical interconnects.

References and links


40. M. Iqbal, M. A. Gleeson, B. Spaugh, F. Tybor, W. G. Gunn, M. Hochberg, T. Bachr-Jones, R. C. Bailey, and...
1. Introduction

Three decades ago the work of Soref and Bennett [1] signaled the dawn of silicon photonics. To many, this meant that finally, optics would realize the same economies of scale that silicon-based microelectronics (especially CMOS) has enjoyed for decades. In this review paper, we take a look at the development trajectory of the silicon-photonic technology and the state-of-the-art in the capability of silicon-photonic processes available today, in the context of the photonic interconnects as the flagship application for this technology.

Being able to create passive photonic devices in silicon, as well as affect the index of refraction through some current or voltage controlled mechanism are the necessary steps towards creating optical coupling, guiding and modulating devices for photonic interconnects. However, the other key pieces of technology are the photodetector and the approach for integration with electronics, which determine the effectiveness of photon-electron conversion, and ultimately the energy cost, speed, and bandwidth-density and integration cost of the overall solution.

Indeed, the first commercial high-volume process, developed by Luxtera, attempted to address all of the issues above at the same time, by integrating the photonic devices in a then state-of-the-art 130nm silicon-on-insulator (SOI) CMOS process [2]. To yield good photonic device performance the process had to be modified with epitaxial Ge step for efficient photodetectors, as well as Si body partial-etch for passive and active waveguide structures. Small parasitic capacitances between transistors and devices were realized through monolithic integration, enabling energy-efficient, high-bandwidth transmitter and receiver components. However, the process customizations and economic investment that led to having the improved photonic device performance, also prevented the technology from following the CMOS scaling trends of shrinking the device features every two years, and hence improving the transistor and system speed and energy-efficiency. Furthermore, since interconnect speeds are scaling at an even faster rate of 4x every two years, this meant that the technology would soon fail to deliver the speeds required in new interconnect standards. For example, it has been challenging for this technology to achieve 25 Gb/s modulation even into relatively small photonic loads such as ring-based optical modulators [3]. To make a major impact, every process technology has to be qualified and available for high-volume production, and every additional process step complicates and slows down this process, further preventing the technology from following the mainstream scaling trends. Similarly, IBM’s monolithic photonic platform [4], which was implemented in a more advanced 90 nm node, took several years to qualify and achieve high-volume and availability due to process customizations.

The issues with limited transistor performance and process qualification/availability recently have taken the manufacturers in a different direction. Both STMicroelectronics and TSMC have demonstrated hybrid integration of CMOS with Luxtera’s photonic technology implemented in standalone photonic SOI wafers [5,6]. This approach decouples the transistor process development from the photonic process development and is seemingly very attractive since it allows the latest node CMOS circuitry to be used in conjunction with optimized photonic devices. However, this arrangement suffers from several issues which significantly limit its effectiveness to a narrow range of applications. First, the micro-bumps used to connect the chip with transistors...
to the chip with photonic devices have limited scaling pitch (to about 40-50 µm) and parasitic capacitances larger than 20 fF, which significantly impacts the speed and energy-efficiency of photonic interconnects. Second, this connectivity arrangement limits the integration scenarios of photonics to 100G pluggable transceivers applications [5]. To enable larger density and quality of electrical connections to the transceiver circuit chip, such as those needed in 400G pluggable and mid-board optics scenarios, the photonic process has to be modified to add through-silicon-via (TSV) technology, further complicating the process and qualification. Alas, this multi-chip stacked solution is cumbersome for highly-integrated optics-in-package scenarios.

Photonic interconnects can achieve high volumes and remain the technology of choice for future system connectivity applications, if they can help address the bandwidth density and energy-cost limits of electrical I/O on large system-on-chip (SoC) chips such as switches, graphics and multi-core processors (GPUs and CPUs) or field-programmable gate arrays (FPGAs). The integration and packaging approach have to enable both a low-energy, high bandwidth-density connection from the large SoC to the photonic transceiver chip, and a photonic connection out of the transceiver chip. The only way to achieve this is if: 1) the transceiver chip is integrated as close as possible (preferably in the same package and on the same interposer) to these large SoCs; 2) photonic interconnects are monolithically integrated with transistors that enable the highest performance in mixed-signal transceiver applications while not further complicating the packaging and process development. With this in mind, our team has worked to create a photonic technology platform in high-volume 45 nm and 32 nm SOI process nodes, creating photonic devices in the native processes with no required process modifications. The advanced features of these processes opened new degrees of freedom in photonic device design that mitigated some of the inherent process limitations and also enabled tight electronic-photonic design optimization. This approach led to record breaking energy-efficient high-speed transmitters as well as the highest degree of electronic-photonic integration demonstrated in the world’s first microprocessor with photonic I/O [7].

In this paper, we summarize and compare the results of these technology platforms, demonstrating the potential of the monolithic integration technologies, and in particular our “zero-change” 45 nm/32 nm SOI.

2. Survey of existing SOI platforms

In this section, we summarize the performance of the state-of-the-art silicon photonic process technology platforms and discuss the advantages of monolithic integration in advanced high-performance CMOS processes for meeting the needs of future optical interconnects.

High-performance integrated systems demand advanced CMOS technologies with high $f_T$ (frequency at which transistor current gain is unity) and $f_{max}$ (frequency at which transistor power gain is unity). These are the performance metrics of transistors representing analog circuit’s speed and sensitivity, and logic speed. Figure 1 shows the trend of $f_T$ for NMOS devices in IBM/GlobalFoundries technology nodes, which is representative of the performance in other similar foundries and process nodes. Notice that $f_T$ has peaked in 45 nm and 32 nm CMOS nodes, due to the change of focus for more scaled-down nodes on logic energy and area density optimization for memory and logic chips, rather than the speed and performance of analog and mixed-signal circuits. Since photonic interconnects are primarily based on mixed-signal transceiver circuitry, these transistor metrics directly impact the link performance metrics such as speed, sensitivity and energy efficiency. For photonic interconnects to be attractive alternative to electrical short-to-long-range (chip-to-chip to backplane) I/O of large SoC chips, they have to provide a sub-1 pJ/b 25-50 Gb/s links with low-energy electrical connection to the SoC and aggregate throughputs larger than 10 Tb/s. In addition to these performance metrics, for such large volume applications, it is key that photonic interconnects are manufactured in a high-volume, state-of-the-art 300mm foundry.
From this perspective, non-monolithic platforms are expected to achieve high energy efficiencies and receiver sensitivities for high-speed optical transceivers due to the flexibility to choose the best performing electronics process independent of the photonics process. A performance summary of the latest non-monolithic silicon photonic technologies is shown in Table. 1. Despite the advantage of optimizing the electronics and photonics separately, these platforms still consume >1 pJ/\text{b} modulator driver energies with >50 \mu A receiver sensitivity, which clearly does not satisfy the electrical and optical power budget of future optical interconnects. The main reason are the additional parasitic inductance and capacitance of wire-bonds or micro-bumps (Cu-pillars) interconnecting electronic and photonic chips. This extra capacitance ranging from 20 fF to 100 fF degrades transmitter’s energy efficiency and also imposes stringent gain-bandwidth constraint for the receiver design leading to degraded receiver sensitivity.

Aside from the packaging of photonics with mixed-signal transceiver circuits, the final packaging with the SoC chip is important for the overall photonic interconnect performance since it determines the quality of the electrical link between the SoC and the photonic transceiver. Current non-monolithic platforms require wire-bonds to connect the photonic transceivers to the package, which degrades the electrical link channel between the SoC and the electronic transceiver chip in the photonic interconnect module. Flip-chip packaging capability is required for high-performance applications such as 400G optical transceivers, mid-board modules and co-packaging with the SoC. Solving this problem demands the development of silicon photonics platform with TSVs [5].

Monolithic silicon photonic integration can minimize both the parasitic capacitance of the interconnection between optical transceiver electronic and photonic devices (now implemented on the same die), and the transceiver chip and the package substrate or the interposer through flip-chip packaging. However, a major challenge in monolithic integration is that process optimizations for photonics and electronics cannot be performed independently of each other. As such, the transistors in monolithic platforms tend to derive from older CMOS processes, where transistor properties are not so sensitive to fabrication changes for photonics. For instance, adding epitaxial Ge to the process requires front-end process modifications that can more easily be tolerated in old CMOS nodes above 90 nm (Table. 2). Such front-end process modifications are significantly more challenging in more advanced process nodes with higher performance transistors. Furthermore, old CMOS processes do not have enough lithography precision for building high quality ring-resonators with good coupling and relative resonant wavelength control required for dense wavelength division multiplexed (DWDM) applications, and consequently transmitters use Mach-Zehnder modulators (MZM) which are much less area and energy efficient.
Our solution to the above mentioned problems is to use unmodified high-volume 45 and 32 nm SOI CMOS technologies, which have the highest $f_{\text{r}}/f_{\text{max}}$ demonstrated, and achieve the needed photonic performance by utilizing the advanced lithography and new process features, coupled with device and circuits co-optimization. We call this approach “zero-change”, as we are not changing the native CMOS fabrication steps. These nodes are the latest partially depleted SOI (PDSOI) processes, which provide thick-enough crystalline silicon (c-Si) body layer to build low-parasitic electrical signaling to the host SoC.

Table 1. Summary and comparison of non-monolithic silicon photonic platforms.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability</td>
<td>Prototyping/Research</td>
<td>Prototyping/Research</td>
<td>High-volume</td>
<td>High-volume</td>
</tr>
<tr>
<td>Integration Method</td>
<td>Wire-bound</td>
<td>Wire-bound</td>
<td>Wafer-bound</td>
<td>Wafer-bound</td>
</tr>
<tr>
<td>Photonics process node</td>
<td>40 nm SOI</td>
<td>130 nm SOI CMOS</td>
<td>130 nm SOI CMOS</td>
<td>55 nm Be/360/65 nm CMOS</td>
</tr>
<tr>
<td>Waveguide material</td>
<td>SiO2/Si</td>
<td>SiO2/Si</td>
<td>SiO2/Si</td>
<td>SiO2/Si</td>
</tr>
<tr>
<td>System Performances</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transmitter Data-rate</td>
<td>50 Gb/s</td>
<td>25 Gb/s</td>
<td>25 Gb/s</td>
<td>25 Gb/s</td>
</tr>
<tr>
<td>Extinction Ratio (Insertion Loss)</td>
<td>2 dB</td>
<td>3 dB</td>
<td>2.5 dB</td>
<td>2.5 dB</td>
</tr>
<tr>
<td>Receiver Energy</td>
<td>0.38 pJ/(b)</td>
<td>0.68 (pJ/(b)</td>
<td>N/R</td>
<td>N/R</td>
</tr>
<tr>
<td>Receiver Sensitivity</td>
<td>79 A/W</td>
<td>72 A/W</td>
<td>N/R</td>
<td>N/R</td>
</tr>
<tr>
<td>Receiver Energy</td>
<td>0.38 pJ/(b)</td>
<td>0.68 (pJ/(b)</td>
<td>N/R</td>
<td>N/R</td>
</tr>
</tbody>
</table>

NR = Not Reported
* High-volume assumes a 300mm foundry
† Modulator and driver energy efficiency

Table 2. Summary and comparison of monolithic silicon photonic platforms.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability</td>
<td>High-volume</td>
<td>High-volume</td>
<td>High-volume</td>
<td>High-volume</td>
</tr>
<tr>
<td>CMOS node</td>
<td>45 nm SOI CMOS</td>
<td>130 nm SOI CMOS</td>
<td>90 nm SOI CMOS</td>
<td>55 nm Be/360/65 nm CMOS</td>
</tr>
<tr>
<td>Waveguide material</td>
<td>SiO2/Si</td>
<td>SiO2/Si</td>
<td>SiO2/Si</td>
<td>SiO2/Si</td>
</tr>
<tr>
<td>Photonics Performances</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modulator Bandwidth</td>
<td>13 GHz</td>
<td>13 GHz</td>
<td>13 GHz</td>
<td>13 GHz</td>
</tr>
<tr>
<td>Modulator Loss</td>
<td>1.3 dB/(\mu)m</td>
<td>1.3 dB/(\mu)m</td>
<td>1.3 dB/(\mu)m</td>
<td>1.3 dB/(\mu)m</td>
</tr>
<tr>
<td>Modulator Device</td>
<td>Ring-resonator</td>
<td>Ring-resonator</td>
<td>MEMS</td>
<td>MEMS</td>
</tr>
<tr>
<td>Photonic Performances</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receiver Sensitivity</td>
<td>79 A/W</td>
<td>72 A/W</td>
<td>N/R</td>
<td>N/R</td>
</tr>
<tr>
<td>Receiver Energy</td>
<td>0.38 pJ/(b)</td>
<td>0.68 (pJ/(b)</td>
<td>N/R</td>
<td>N/R</td>
</tr>
</tbody>
</table>

NR = Not Reported
* Modulator and driver energy efficiency
† Full receiver energy including samplers and digital circuits
3. Overview of the “zero-change” SOI platforms

We implement our “zero-change” silicon photonic platform in an unmodified commercial 45 nm CMOS SOI process [9]. All photonic devices are designed to conform to the purely-electrical foundry design flow [19], without any modifications to the native process Fig. 2. Optical devices are implemented in the sub-100 nm thick high-index crystalline silicon (c-Si) layer, normally used as the transistor body. We avoid using the first six metal layers and block foundry’s filler cells on top of photonics structures to prevent metallic optical losses. However, we place our custom high-density filler cells around the photonic blocks to satisfy minimum density rules for these layers.

Photonic design layouts are generated automatically via our SKILL based CAD tool in Cadence Virtuoso [20]. This allows seamless integration of photonic and electronic designs in a single environment in addition to fixing design-rules violations by using Manhattan discretization, boolean and sizing operations. This tool is also utilized for photonic auto-routing and making larger photonic integrated circuits (PIC) layouts.

Since the buried oxide (BOX) layer is not thick enough to optically isolate the c-Si waveguide core from the silicon substrate, we have to remove the silicon substrate to reduce the waveguide optical loss. Substrate removal is done in a single post-processing step on the flip-chip die-attached chips [19]. The flip-chip underfill keeps the released die mechanically stable and dissipates the heat even under thermal stress tests. Moreover, thermo-optic heater structures needed in many applications such as resonance tuning or phase matching achieve high efficiencies in this type of packaging scheme through the removal of thermally conductive silicon substrate. This step does not affect the transistor performance, and all existing foundry IP, timing libraries, and simulation models remain valid [7,19]. Waveguide loss of approximately 3dB/cm is achieved after this step. In addition, the flip-chip packaging is favorable for high-performance electronics due to better power delivery, pin counts, and signal integrity of the I/O pins. Light is coupled to the chip via vertical grating couplers. Couplers have been fabricated by patterning c-Si and polysilicon layers. Active devices including microring modulators and photodiodes are implemented using existing source/drain and well implant doping levels and the available SiGe in this process.

Figure 3(a) presents the timeline of platform development for this technology, utilizing the available multi-project wafer runs and without the explicit foundry support. Owing to the maturity of the high-volume 45 nm SOI process, but constrained by the multi-project run availability and turnaround times, we were able to go from device test-chips to a fully-functional processor with photonic I/O in less than four years, on limited research grant funds. Commercially available CMOS technologies normally have much faster turnaround time, which expedites device development and development of new systems. In translating the learning experiences
from this platform into the 32nm SOI platform we have already shrunk the development cycles significantly. Further optimization and acceleration will be possible with tighter foundry support and coordination.

The demonstrated processor with photonic I/O using the 45 nm “zero-change” platform showcases the power of this technology. Ultra-power-efficient ring-resonator based silicon photonic links, with millions of transistors and hundreds of photonic devices fabricated on the same chip, are aimed to improve processor-memory link bandwidth [7]. This SoC, Fig. 3(b), has a dual-core RISC-V processor [21], 1 MB SRAM based cache memory, and DWDM optical I/Os illustrated in Fig. 3(c). Figure 3(d) shows the key photonic devices of an optical link implemented in this technology. This work achieved the highest level of integration scale and system complexity among the state-of-the-art electronic-photonic systems.

3.1. “Zero-change” platform in 32nm SOI CMOS

We have extended our “zero-change” monolithic integration approach to a more advanced 32 nm technology node to further improve the speed of electronics and performance of the device platform by exploiting new process features such as channel SiGe with higher Ge concentration available in these technologies [22, 23]. This process node features high-k/metal gates (HKMG) with the minimum gate length of 25 nm and 33% logic speed improvement over 45 nm node [10].
Due to a similar silicon body thickness to the 45 nm process, photonic designs are directly imported into this technology node. In a single multi-project wafer run, we were able to recently demonstrate 12 Gb/s transceivers in the standard telecom O-band in this platform. As with the 45 nm platform, we expect that future development of photonics that are able to exploit the improved lithography and expanded material selection will allow for continued improvement of the 32 nm platform.

4. Photonic devices

4.1. Passive devices

Waveguides are built in the sub-100 nm thick c-Si body layer by blocking all the transistor body dopants, to lower the optical loss. In the 45 nm platform, the measured loss is 3.7 dB at 1280 nm and 4.6 dB at 1550 nm [19]. The extracted intrinsic quality factors of 227 k and 112 k were obtained for 1280 nm and 1550 nm undoped rings with 7 µm radius, respectively. These high Qs are made possible by advanced processing that offers very small line-edge roughness. Advanced photo lithography and patterning in this process also allow ring-resonators with radii as small as 5 µm with small bending loss.

Vertical grating couplers have been used in “zero-change” platforms to couple the light from on-chip waveguides to optical fibers. Our grating couplers are implemented using both the c-Si and transistor gate poly-Si layers. Since we can pattern two silicon layers independent of each other, we have more degrees of freedom for design and optimization compared to other custom silicon photonic processes that use a partial silicon etch for uni-directional grating couplers. This allows us to achieve 1.5 dB loss (including the taper), and 78 nm 1 dB bandwidth around 1320 nm wavelength [24–26] Fig. 4. The measured pigtailed optically packaged couplers also achieved 2.5 dB loss [26].

Fig. 4. (a) 3D layout of a unidirectional grating coupler, (b) Optical transmission at 10.5 degree vertical angle.

4.2. Active devices

Microring-modulators have been realized by placing interleaved p and n junctions along the ring cavity. This technique utilizes the fine lithography advantage of the deeply-scaled 45nm process, in order to enable efficient modulation through high junction capacitance density, in the absence of the partial etch or customized doping to form other types of junctions. Resonance wavelength can be modulated by changing the carrier density in the depletion regions of interdigitated junctions via carrier-plasma effect in silicon [1,27]. A variety of different p and n doping profiles can be implemented by combining available implants for transistor well and source/drain dopings that set various threshold voltage options. Cathode and anode segments are all connected via
spoke-shaped metal contacts in the center of the ring in order to avoid proximity of metal to the optical mode. These 5 µm-radius active microrings achieved intrinsic Q-factors of 18 k and up to 10 k loaded Q-factors with 3.2 THz free spectral range (FSR) in the telecom O-band [28–30]. Measured resonance wavelength shift efficiency is 20 pm/V in the depletion region (reverse bias). The resonator has an embedded silicided c-Si heater structure for thermal tuning of the resonance required to compensate for thermal and process variations. The ring heater resistance is 500 Ω with a high thermal tuning efficiency (3.8 µW/GHz).

Segmented ring-resonator can also be configured as an optical digital-to-analog converter (ODAC) [31]. One can control the amount of resonance shift by independently controlling individual interleaved junction segments. Figure 5(a) shows the 3D rendered layout of a spoked ring-modulator with separate anode contacts. Our analysis showed improved linearity of this structure over conventional method of controlling the resonance shift by using electrical DACs to control the applied voltage on PN junctions. This device is used to perform 40 Gb/s PAM-4 transmission, and can also be used in other systems such as optical arbitrary waveform generators. Figure 5(b) shows measured optical transmission of a WDM transmitter row with 11 channels. Due to the high lithographic precision and film thickness control of 45/32 nm processes, the measured resonances are fabricated in order as designed, with channel-to-channel resonance variation less than half of the channel spacing, across WDM row length of 1.5 mm.

Both 45 nm and 32 nm CMOS nodes feature epitaxial SiGe materials to improve the performance of PMOS devices, Fig. 6(a). Embedded SiGe (eSiGe) with Ge% concentration around 20% has been used in the source/drain regions of PMOS transistors to apply compressive stress since the 45 nm technology node [32]. In order to compensate for the low Ge% concentration and minimize the PD parasitic capacitance, we built resonant PDs by forming PIN junctions in the ring resonator’s cavity as shown in Fig. 6(b).

Resonated eSiGe detectors in 45 nm platform showed the responsivity of 0.55 A/W and 0.5 A/W at 1180 nm and 1270 nm wavelengths, respectively with −4 V bias voltage [26, 33]. This PD has the best-in-class dark current of 20 pA and the electro-optical 3 dB bandwidth of 5 GHz limited by the RC of the junctions.

Two types of resonant SiGe-based PDs are implemented in the 32 nm process using the two variants of epitaxial SiGe available in this process. Photonic structures in this process are still at an early stage of development and PDs are implemented using unoptimized microring resonators with a loaded Q-factor of 6.5 k (intrinsic Q > 15 k). Q’s are expected to improve 2 – 3× by design optimization. PDs using eSiGe layer achieved 0.06 A/W responsivity at 1310 nm.
This technology node also features another epitaxially grown SiGe layer with a higher Ge% concentration (approximately 40%), which leads to higher responsivity. This SiGe epi layer is used for PMOS channels (cSiGe) to reduce the threshold voltage ($V_{TH}$) after introducing metal gate to the process [10]. Measurements showed that cSiGe-based resonant PDs have an improved responsivity of 0.13 A/W at −8 V bias. The responsivity of both types of SiGe PDs will improve with the improvement of Q of microrings through the reduction of optical loss. These devices exhibit a >12.5 GHz 3 dB bandwidth (measured via a 13.5 GHz VNA) and 150 nA dark current.

We have also extended the operation of modulators and PDs beyond the O-band. We have redesigned the microring spoked modulators for operation in 1550 nm (C-band) and have demonstrated 25 Gb/s modulation [34]. The loaded Q-factor of 13 k of these modulators shows that the sub-100 nm thickness of the silicon device layer is not a limiting factor of these platforms (32 nm and 45 nm PD-SOI nodes) to implement compact and high performance devices for wavelengths longer than the O-band (most silicon photonic platforms have a silicon thickness of 200-250 nm). Also, in addition to the SiGe PDs, defect-based resonant PDs have been demonstrated covering the optical telecommunication O to L bands [35]. These PDs work based on the absorption enabled by the defect states in the transistor gate polysilicon layer [36]. Figure 6(c) is the micrograph of this design and Fig. 6(d) depicts the cross-section of the absorption region. This PD achieved 0.15 A/W responsivity with 10 GHz bandwidth at −15 V bias.

Table 3 summarizes the performance of our photonic devices implemented in “zero-change” platforms.

5. System-level demonstrations

Machine learning workloads are increasingly driving the convergence of high-performance computing and data-center architectures, requiring high bandwidth density and low-energy interconnects right from the SoC sockets. Limited socket substrate area and power budget, as well as the large number of connections, require sub-1 pJ/b, >1 Tb/s/mm² and low-cost <0.1 $/Gb/s photonic interconnect solutions. Even the near-term next generation pluggable 400G interconnects require sub-5 pJ/b and <1 $/Gb/s for intra- and inter-rack communications, which
Table 3. Photonic devices performance summary.

<table>
<thead>
<tr>
<th>CMOS Technology</th>
<th>45nm SOI</th>
<th>32nm SOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength</td>
<td>1550</td>
<td>1310</td>
</tr>
<tr>
<td>Waveguide Loss</td>
<td>4.6 dB/cm</td>
<td>3.7 dB/cm</td>
</tr>
<tr>
<td>Grating Coupler Loss</td>
<td>10 dB*</td>
<td>1.5 dB</td>
</tr>
<tr>
<td>Grating Coupler 1-dB Bandwidth</td>
<td>N/A</td>
<td>78 nm</td>
</tr>
<tr>
<td>Modulation Speed+</td>
<td>25 Gb/s</td>
<td>40 Gb/s</td>
</tr>
<tr>
<td>Photodetector Responsivity</td>
<td>0.15 A/W</td>
<td>0.5 A/W</td>
</tr>
<tr>
<td>Photodetector Bandwidth</td>
<td>10 GHz</td>
<td>5 GHz</td>
</tr>
</tbody>
</table>

*Without electrical equalization, +Bidirectional couplers.

are the specifications that will be hard to achieve with traditional MZM-based photonic structures. The portfolio of advanced ring-resonator based photonic devices closely integrated with fastest transistors in “zero-change” 45nm/32nm SOI platforms can be utilized for these aims. Here, we summarize our latest system-level demonstrations of optical transceivers and DWDM links in this technology.

5.1. Optical NRZ and PAM4 transmitters

Microring resonator (MRM) based optical transmitters can meet the needs of next generation optical interconnects. While MZMs with high-enough extinction ratio (ER) are millimeter-sized devices with large capacitance and high-insertion loss (IL), which leads to high energy consumption, limited modulation rate and large footprints, ring-resonators have very compact footprints suitable for high-speed and energy-efficient optical transmitters. Moreover, microring modulators enable DWDM for large-scale, high-bandwidth density integrated systems with bandwidth densities in the $\text{Tb/s/mm}^2$ range. Our experimental results (Tables 1 and 2) show two orders of magnitude improvement in energy efficiency compared to industry workhorse MZM-based transmitters by using MRM-based transmitters in “zero-change” monolithic platforms.

Figure 7(a) shows the die photo of a 40 Gb/s NRZ transmitter in 45 nm SOI CMOS. The modulator and driver stage consume only 40 fJ/b and the total energy consumption and area including a serializer and a dedicated high-speed clock source per transmitter are 0.7 pJ/b and 0.03 mm$^2$, respectively. Figure 7(b) presents the area and energy breakdown; while the high-speed clock source (digital PLL) dominates the area and energy, it can be shared between multiple transmitter blocks and macros (typically driving at least up to 16 transmitters on a single DWDM macro). Also, notice that due to the small footprint of MRMs, photonics takes only 15% of the total area.

As resonant devices, microring modulators are subject to the fundamental trade-off between optical bandwidth and optical modulation amplitude (OMA). In order to support 40 Gb/s NRZ data stream the MRM’s Q-factor is set to 5500 by tuning the drop-port coupling. To improve the OMA, we have used high swing drivers (2.4 V$_{pp}$) with differential driver and an AC coupler. In doing so, we increase the depletion width of the junctions, which consequently introduces a larger resonance shift and improves the OMA. During operation, the modulator sees voltages of −0.5 V and −2.9 V. The modulator is always reverse-biased to keep enough electric field in the depletion region to sweep-out the generated carriers, enabling fast modulation. Figure 7(c) shows the measured NRZ eye diagrams at 40 Gb/s with the ER of 3 dB and IL of 4.7 dB. Thanks to the monolithic integration and device-circuit co-optimization, this work has achieved higher bandwidth density (1.3 Tb/s/mm$^2$ at 40 Gb/s assuming a dedicated PLL per transmitter) and an improved energy efficiency compared to other MRM-based transmitters with electronics and photonics on separate dies. This design also has the highest data-rate, energy-efficiency, and bandwidth-density compared to prior works in monolithic photonic platforms. At these line rates, further modulator optimization is possible through customized higher-doping concentrations [37].

To achieve data-rates higher than the modulator bandwidth, higher order pulse amplitude modulation (PAM) schemes can be exploited to mitigate the fundamental energy-bandwidth
trade-off at the system level. We recently demonstrated a 40 Gb/s PAM-4 transmitter based on the optical DAC design described in Section 4.2 in 45 nm “zero-change” platform [31]. This PAM-4 transmitter directly converts the digital data into optical levels with programmable look up table (LUT) to linearize the ring’s Lorentzian response. We achieved 685 fJ/b total transmitter energy efficiency (42 fJ/b modulator and driver energies) with an area bandwidth density of 0.67 Tb/s/mm². This MRM operates on 0 and −1.5 V voltages leading to 3 dB ER and 5.5 dB IL (eye-diagram shown in Fig. 7(d)) at 1285nm laser wavelength. ER and IL can be improved further by critically coupling the ring via adjusting the bus waveguide and ring gap.

Due to the limited transistor speed scaling in advanced processes, and relative increase in transistor and interconnect parasitics, link speeds in excess of 50 Gsymbol/s will be highly energy inefficient, while links above 100 Gsymbol/s will be very difficult to realize. Utilizing the microring-based DWDM technology with large bandwidth density allows for line rates per wavelength to stay in the energy-efficient regime of around 25 Gsymbol/s, while growing the overall system throughput by adding more wavelengths.

Despite multiple advantages of MRMs, they have not been used commercially in photonic interconnect applications due to the need for active tuning to counteract process and thermal variations. Additionally, thermal tuning is essential in DWDM chip-to-chip optical I/O to find and lock each resonance to an optical carrier. We have addressed this issue for both NRZ [26, 38] and higher order PAM modulations [31] in our platform. These bit-statistical tuners decouple tracking of optical one- and zero-levels to realize non-dc-balanced data transmission, an “eye-max”-locking controller, and ring self-heating cancellation without the need for a high-speed sensing frontend. The tuner consumed 18 fJ/b at 40 Gb/s in the logic. At maximum power, the heaters output 5.4 mW and can compensate for wavelength or frequency offsets of approximately 8.2 nm or 1.53 THz (a temperature range of 140 K) at a tuning efficiency of 3.53 µW/GHz. While the chip to chip and wafer to wafer variations are expected to be larger than the DWDM channel spacing, the variation is still well within the heater tuning range. Since this mismatch results in the rotation of the ring order of the whole bank around the periodically repeating FSR, it can also be handled

Fig. 7. 40 Gb/s NRZ and PAM4 transmitters results: (a) Micrograph of the 40 Gb/s NRZ transmitter, (b) Total area and energy breakdown for 40 Gb/s NRZ transmitter, (c) NRZ eye-diagram, (d) PAM4 eye-diagram.
5.2. Optical receiver

Receiver sensitivity directly impacts the laser power budget required in an optical link. Hence, improving receiver’s sensitivity in an energy-efficient way is essential to lower laser energy and packaging costs and increase link margin tolerances. The \( f_r \) of the transistors and parasitic capacitance of interconnects between electronics and photonics are both critical factors in determining optical receiver’s sensitivity and energy. The “zero-change” platforms provide the highest available \( f_r \) with smallest parasitic capacitances among all available silicon photonic technologies.

A high-sensitivity, fully-differential optical receiver for high-density photonic interconnects is demonstrated in 45 nm “zero-change” platform [40]. To realize fully-differential operation, a 3-dB power splitter and SiGe photodetector are integrated with the receiver. Each PD is connected to a trans-impedance amplifier (TIA) followed by two pre-amps and samplers that operate on interleaved clocks to provide double-data rate (DDR). The DDR output is retimed, deserialized and fed into the digital backend for on-chip bit error rate (BER) measurements. This receiver improves sensitivity further by suppressing common-mode and supply noise through fully-differential operation. The receiver achieved BER < \( 10^{-12} \) at 12 Gb/s with input sensitivity of 8.6 \( \mu A_{pp} \) while consuming 4.3 mW (0.36 pJ/bit). In combination with PDs with 0.5 A/W responsivity, receiver’s optical sensitivity would be about \(-20 \, \text{dBm}\). Despite the sub-10 GHz bandwidth of PDs in 45 nm “zero-change”, higher data-rates are also achievable by using equalization techniques [41]. These additional circuits will trade-off the energy efficiency and sensitivity with higher data-rates.

5.3. Optical WDM link

Microring based optical transceivers can be used to build DWDM links to achieve Tb/s aggregate bandwidths over a single fiber. We have demonstrated stand alone DWDM capabilities with 11-channel within 3.2 THz FSR at 1180 nm wavelength with 5 Gb/s data-rate per channel [38].

A 4.1×DWDM link with 8 Gb/s channel data-rate has been also demonstrated recently in 45 nm “zero-change” platform using a 4.1 bench-top CW laser source constructed using 4 DFB lasers, a star coupler, and an SOA (to amplify power after the star coupler back to >7 dBm per CW wavelength) [26]. The channel spacing between wavelengths is nominally 2 nm, though it does not need to be exact owing to the high wavelength tunability of the microring transceivers; the link will tolerate any channel spacing down to approximately 0.5 nm. In the 4.1×8 Gb/s configuration with both transmit and receive-side thermal tuning heaters set at half-strength, the link consumes an end-to-end power of 109 mW, or 3.4 pJ/bit, including clocking and serializer/deserializer circuitry with 7 dBm of laser power per wavelength.

6. Platform capabilities and future applications

For the last decade, optical interconnects have been the primary application target for the silicon-photonic platforms. Even in this domain, further advances are possible by utilizing more advanced device concepts implementable in the advanced 32nm/45nm SOI process platforms, with fast and low-cost development cycles. Furthermore, the position of the 32nm and 45nm SOI platforms as planar process nodes with fastest transistor and analog/mixed-signal performance, adds the possibility for inexpensive process customizations such as photonic structure doping optimizations that would further improve the device performance for a number of applications, even beyond photonic interconnects. With inherent high connection density between transistors and photonics, these platforms are also suitable for emerging photonic applications such as phase-arrays [42] for lidar and free-space optical communication, as well as molecular sensing arrays [43,44]. Beyond classical applications, integrated photonic platforms hold a great promise
for quantum communication and computing \([45,46]\), as well as low-energy interconnects from cryogenic environments.

7. Conclusion

With adoption by major semiconductor foundries, prospects for silicon-photonics are bright. A combination of a slow-down in transistor performance scaling beyond the 32 nm process node and successful photonic device and transceiver demonstrations with state-of-the-art energy, bandwidth-density and cost performance in the 45 nm platform, point to these process nodes as potential strongholds for a variety of integrated electronic-photonic systems-on-a-chip. These fully-integrated electronic-photonic SoCs will be able to offload the communication work from co-packaged more-deeply-scaled compute and memory SoCs, as well as perform other complex sensing and communication functions. Furthermore, the advanced lithographic capability and some limited process customizations of these process nodes will likely enable an even more powerful class of photonic devices and electronic-photonic systems to address the future applications.

Acknowledgments

The authors would like to thank the corresponding research groups members at UC Berkeley, MIT and Boston University and the Berkeley Wireless Research Center (BWRC) staff. The authors would also like to thank the Trusted Foundries, MOSIS and the IBM/GlobalFoundries staff involved in handling the design submissions. The views expressed are those of the authors and do not reflect the official policy or position of the DoD or the US Government.

Funding

National Science Foundation (NSF) (ECCS award 1611296); Defense Advanced Research Projects Agency (DARPA) (POEM award HR0011-11-C-0100).