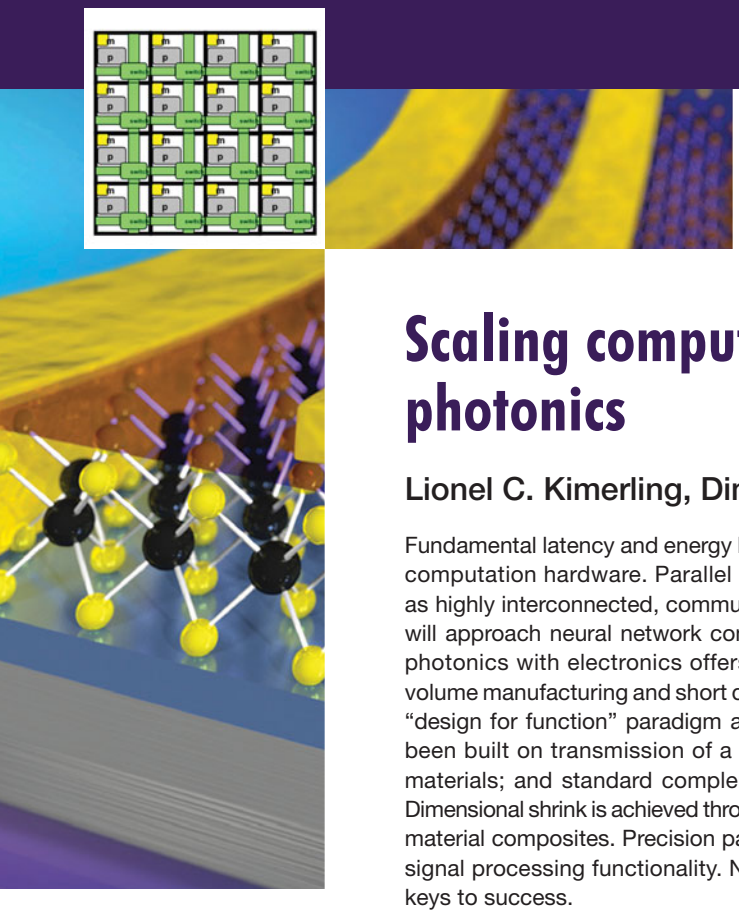

Scaling computation with silicon photonics

Lionel C. Kimerling, Dim-Lee Kwong, and Kazumi Wada

Fundamental latency and energy limitations are driving major changes in communication and computation hardware. Parallel multicore and multiprocessor architectures are emerging as highly interconnected, communication-centric computation tools that at high node count will approach neural network connectivity and complexity. Monolithically integrated silicon photonics with electronics offers a promising platform for scaling functionality with high volume manufacturing and short design cycle times. The system parameters for this emerging "design for function" paradigm are cost, energy, and bandwidth density. The platform has been built on transmission of a $\lambda \approx 1550$ nm photon wavelength; Si, Ge, Si_3N_4, and SiO_2 materials; and standard complementary metal oxide semiconductor foundry processing. Dimensional shrink is achieved through strong signal confinement in high refractive index contrast material composites. Precision pattern transfer has enabled both photonic interconnect and signal processing functionality. New materials, process integration, and packaging are the keys to success.

Scaling computation at constant cost, energy, and footprint

The deployment of new information technology must meet both performance and scalability specifications. Information hardware must conform to a common platform to leverage a common research, development, and manufacturing base. System architecture scaling has migrated from central processor units (CPUs) with ever-faster clock frequencies to increasing numbers of processors working in parallel with reduced clock frequencies. This fundamental architectural change resulted from the approach of power density limits as transistor switching voltages reached their scaling asymptote near 0.5 V. Parallel architectures scale linearly in power density with processor count, while continuation of clock frequency scaling at constant voltage would scale as the second or even third power of frequency ($P = ½\ CV^2f$; P = power, C = capacitance, V = voltage, and f = frequency). Multiple processors working in parallel place an emphasis on fast and efficient interprocessor communication. Parallel, communication-centric architectures should ideally possess all-to-all connectivity with zero latency and energy. Software for massively parallel systems must adapt to minimal data movement or to detailed specification of data location and transport in the lines of code. Two vectors of convergence are active in creating scalable solutions for the next two decades: (1) electronic-photonic convergence to natively integrate communication into the computation process and (2) hardware-software-architecture convergence to provide intelligent, dynamic provisioning of energy-consuming computational resources. Microphotonic integration on the silicon platform embraces both vectors by utilizing the tools of silicon microelectronics and by supporting optimized complexity with novel electronic-photonic partitioning of functionality.[1] This article presents the authors' vision of the grand challenges for new materials, process integration, and foundry manufacturing platforms for scaling computation functionality.

Scaling computation architecture: A connectivity case study

The physics of information processing and transport dictates an ultimate tradeoff between clock/communication frequency and parallelism.[2] While aggregate bandwidth must continue to scale at all levels of the interconnection hierarchy, the data rates of each channel must reach an asymptotic limit defined by the dissipated power density. Monolithic waveguide integration of optoelectronic devices with transport media, and high index contrast materials, which tightly confine light in dimensions of wavelength/refractive-index and allow for

Lionel C. Kimerling, MIT Microphotonics Center, Massachusetts Institute of Technology, USA; lckim@mit.edu
Dim-Lee Kwong, Institute of Microelectronics, Agency for Science, Technology and Research, Singapore; kwongdl@ime.a-star.edu.sg
Kazumi Wada, Department of Materials Engineering; University of Tokyo, Japan; kwada@material.t.u-tokyo.ac.jp
DOI: 10.1557/mrs.2014.165

micron-dimensioned layouts, provide a path for scaling to small photonic devices coupled with low parasitic capacitance to electronic drivers and receivers.

Microprocessor performance is gated by two factors: communication bandwidth and programmability. Bandwidth is needed for low latency signal processing among the compute nodes without delays associated with contention and routing. As the number of cores (processors), N, grows exponentially with time, the desired communication bandwidth must grow as N^2. **Figure 1**a shows a mesh interconnection solution, where data are transported by routing point-to-point between nearest-neighbor cores. The mesh architecture enables hundreds of cores to communicate with inexpensive local communication, before contention (blocking of a path while it is being used) and distance-dependent routing energy and latency become limits. However, the mesh introduces heterogeneity that increases programming complexity by requiring locality in the lines of code for point-to-point coordination of thousands of processes. Broadcast communication from one core to all other cores can relieve the need for locality in programming, but it is slow and power inefficient on an electrical mesh. A broadcast photonic network that allocates a unique wavelength of light to each sending core and an array of wavelength-selective receivers at each core (Figure 1b) can provide high bandwidth, low latency, and power-efficient communication among cores with no routing requirement by the hardware or data locality in the lines of code. This parallel architecture is known as ATAC: all-to-all computing.[3,4]

The ATAC architecture utilizes multichannel photonic waveguides that relieve both the bandwidth and loss limitations of the metal wire bus used in multicore microprocessors today. By allowing direct core-to-core communication, ATAC reduces the frequency of the memory calls implicit in the traditional bus architecture. Each off-chip memory read can consume 500 pJ of energy and 250 clock cycles, while a photonic network transfer consumes <3 pJ and <3 cycles. ATAC requires no routing, utilizes resources only when needed, and enables each core to directly communicate with any other core. A view of extended global memory can be efficiently enabled with an ATAC photonic network and distributed cache memory. Communication on the ATAC network is explicit and contention-free, encouraging the programmer to use it often with reduced overhead.

The ATAC architecture is likely to be used generically in complementary metal oxide semiconductor (CMOS)-integrated photonic circuits.[3] Each core (node) is assigned a unique wavelength, and an optical power bus distributes all wavelengths around the chip. A sending core selects its wavelength through an optical filter (ring resonator), encodes it with data using an optical modulator, and launches it on an optical data waveguide for distribution to all other cores. If power-gated, monolithic lasers can be placed in each core, and additional advantages (~50x) in energy efficiency can result. Each core has a filter/photodetector bank to independently drop each wavelength channel, to convert its data to electrical current, and to condition and process it as an electrical signal. This level of network complexity would be impossible with the discrete photonic components used in fiber optics, but waveguide integration provides the necessary small size and power efficiency.

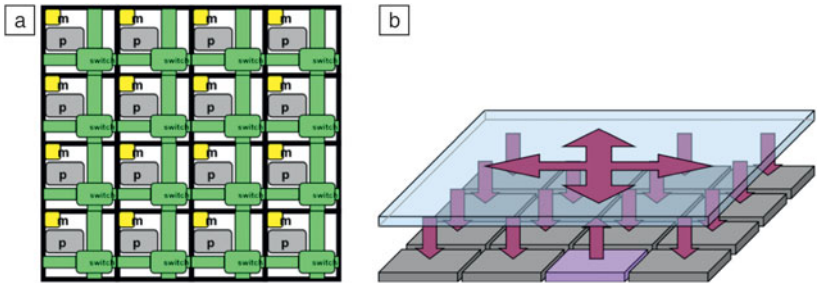

Figure 1. Scalable communication for multicore microprocessors. (a) Electrical mesh routing network for a 16 core microprocessor; (b) photonic all-to-all broadcast network.[3,4] Note: m, local cache memory; p, processor; switch, to launch and route near-neighbor, core-core communication.

Materials and devices for monolithic silicon microphotonics

Waveguide materials for high density optical interconnection

Two determining constraints for electronic and photonic integration on silicon are waveguides capable of (1) small bend radii that do not radiate transmitted light and (2) high transparency transmission. In silicon microphotonics, the silicon-on-insulator (SOI) platform has been widely employed, referred to here as the high index contrast (HiDex) platform, see **Figure 2**.[5] A single mode Si waveguide, clad by SiO_2 (n_r = 1.5), confines light in small dimensions due to the high refractive index of silicon (n_r = 3.5). For the telecommunication standard wavelength of 1550 nm, the single TE (transverse electric) mode silicon waveguide has dimensions of 200–250 nm thick and 400 nm wide for the HiDex platform. Doped-SiO_2 fibers and planar lightwave circuits, here a low index contrast (LoDex) platform, where the typical index difference between core and cladding (Δn) is <0.1, possess a much larger 2–8 μm core thickness. HiDex materials typically have a larger temperature dependence of the refractive index (thermo-optic coefficient, TOC), which leads to a large fluctuation in channel wavelength in the wavelength division multiplexing (WDM) optical filters. The interchannel wavelength difference is 100 GHz (i.e., 0.8 nm at the 1550 nm range), which will soon reduce to 50 GHz. Thus, the fluctuation limits the maximum number of wavelength channels, because larger channel spacing is required to avoid signal overlap. In computing applications, the chip and package temperatures can fluctuate in operation from room temperature up to ~70°C.[6]

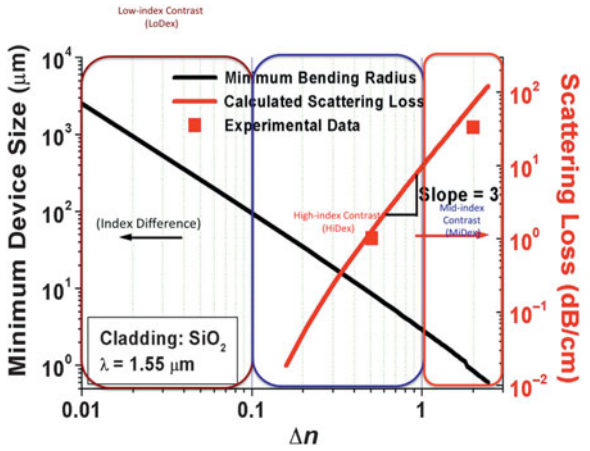

Figure 2. Dimensional scaling for silicon photonics is dependent on the index difference (Δn) between the core and clad materials. SiO_2 is the clad material for silicon photonics. The optical device footprint can be scaled smaller as the index contrast Δn increases; but device fabrication tolerances, such as waveguide sidewall roughness, become more demanding.

For high bandwidth communication, the HiDex silicon waveguide enables phase modulation by the injected carrier—plasma-induced change in index in one arm of a microphotonic Mach-Zhender interferometer (MZI). Therefore, the choice of material platform is application based: HiDex for active functions (photodetector, modulator, and light emitter) and LoDex for passive functions (such as WDM-MUX/DEMUX, wavelength division multiplexor/demultiplexor).

To accommodate small bend radii for waveguide integration on a chip of 1–2 cm^2 area, the mid-index contrast (MiDex) materials platform (Figure 2) using silicon nitride (Si_3N_4) or silicon oxynitride (SiON) can offer advantages. **Table I** compares the behavior of thermally stable, ring resonators composed of Si_3N_4 and Si with a compensating polymer (negative TOC) cladding.[7] The composite materials waveguide structures give zero thermo-optic response. The smaller TOC of Si_3N_4 allows stronger confinement Γ in the waveguide core, and the effective index (n_{eff}) and minimum ring bend radii (R_b) are similar: Si_3N_4 (R_b = 5.5 μm) and Si (R_b = 3.5 μm). The stability and smaller intrinsic scattering loss (Figure 2) favor MiDex for on-chip dense wavelength division multiplexing (DWDM) applications in silicon photonics. Possibly the most significant advantage of the Si_3N_4 MiDex platform is its superior power handling capacity: two photon absorption at 1550 nm is negligible, and optical nonlinearity is more than one order of magnitude smaller than for Si. For example, scaling of the ATAC multicore DWDM application, described in the section on scaling computation architecture, is ultimately limited by the optical power in the signal bus waveguide: determined by the number of optical channels, the number of drop ports (N-1 receiver ports for each wavelength), and the drop power for each port (required for signal integrity at the employed data rate). If N is the number of nodes receiving the broadcast, then the required optical power in the waveguide is N^2 x power/drop, which can exceed 1 W (~1 GW/cm^2 for Si) as the system architecture approaches >1000 nodes.

Table I. Polymer clad composite waveguide materials systems for thermally stable optical filters: zero resonant wavelength change between 0–70°C for Si and Si_3N_4 core materials.

Material	n	Γ	n_{eff}	n_g	R_b(μm)
Si	3.48	0.58	1.78	3.68	3.5
Si_3N_4	2.05	0.9	1.75	2.2	5.5

Note: Core refractive index (n), optical mode confinement factor (Γ), effective index (n_{eff}), group index (n_g), and minimum bend radius (R_b).[7]

Device and process integration

The targeted implementation of silicon photonics is monolithic integration of optical circuits with electronic circuits in a "CMOS-like" standard process flow, as shown schematically in **Figure 3**.[8] The historical path to monolithic integration began with discrete breadboards, then hybrid packages and chips, and finally monolithic integration. Current commercial products feature monolithic modulator-waveguide-detector integration with hybrid flip-chip bonding of the laser and electronics. The laser is the most difficult device to monolithically integrate on a silicon chip,[9] and III–V light emitters have been hybrid-integrated with transparent adhesives or wafer bonding in the back-end-of-line (BEOL).[10,11] Monolithic electronic-photonic integration in silicon microphotonics is expected to enable scaling of cost, manufacturing, energy, and bandwidth.

Silicon CMOS process technology is one of the most elaborate human-made resources in history. It provides the most accurate yet cost-effective processing for monolithic integration of electronics and photonics on a chip. The fundamental challenge for CMOS process technology for Si microphotonics is dimensional tolerance, because wavelength-dependent optical devices are more dimensionally sensitive than transistors. Fabrication errors in waveguide width/thickness as small as 1 nm can be consequential for on-chip DWDM. The absolute process tolerance of MiDex should be larger because of weaker confinement of light modes than in HiDex. The drawback of Si_3N_4 materials deposited by chemical vapor deposition is residual hydrogen in the film, leading to strong N–H bond absorption near 1520 nm,[12] which is at the middle of the optical communication wavelength range. Annealing at 1100°C or higher is necessary to remove hydrogen from SiN_x layers, and these temperatures are not tolerable in the CMOS BEOL processing. However, it should be noted that Si_3N_4 deposited by PVD is a compatible process that introduces no hydrogen.

Multiplexing for high aggregate bandwidth

Scaling interconnection to the highest bandwidth density will necessarily utilize DWDM,

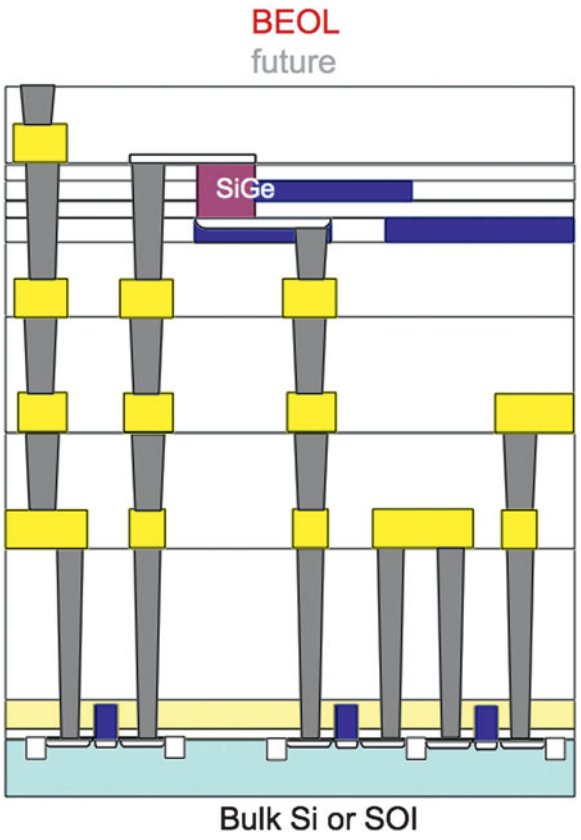

Figure 3. Schematic layout for monolithic integration of silicon photonic interconnection into the back-end-of-line (BEOL) complementary metal oxide semiconductor interconnect stack. This architecture is consistent with the size and function of photonics, and it can potentially simplify fabrication by eliminating interconnection layers.[8] For the schematic structures, yellow represents the intralayer metal interconnects, gray is the interlayer metal interconnects, dark blue is the silicon transistor and optical waveguide elements, and light blue is the silicon wafer of the SOI layer. Note: SOI, silicon on insulator.

where 50 GHz (0.4 nm) channel spacing is typical. An additional track is quadrature optical phase shift keying (QPSK) with electronic digital signal processors using silicon MZI modulators with advanced CMOS technology (Cisco Systems announced a 100 G CFP module in the Optical Fiber Conference [OFC] 2013, and Acacia Communications announced in OFC 2014). QSPK provides higher spectral efficiency in terms of bits/Hz by modulating the phase of the light, with specified phase delays, as well as the traditional light intensity (OOK, on-off keying). The signal for the same modulation frequency can be multiplexed into time-domain and phase-domain encoded channels.

The challenge of on-chip and on-module dimensional precision and thermal stability for WDM and DWDM has been noted. Athermalization using a composite silicon-polymer waveguide can achieve 0.5 pm/°C resonant wavelength stability that is capable of 220 wavelength channels in the communication C-band near 1550 nm. The partial delocalization of the optical mode allows a higher power transmission limit for the waveguide, and the channel number achievable by athermalization can be more than 10 times that normally handled in Si microphotonics. The addition of a thin, photosensitive As_2S_3 chalcogenide glass layer between the negative TOC polymer and positive TOC silicon, as in **Figure 4**,[7] adds a resonance trimming capability to the filter, since an index change is optically equivalent to a dimensional change.[7,13]

Silicon modulators

Electro-optic modulators based on electro-refractive (ER) and electro-absorptive (EA) effects have been reported. Germanium EA devices are described later. Ring resonator and MZI ER modulators have been demonstrated. Si rings showed a higher bandwidth density and smaller energy/bit.[14] The drawback is the previously described sensitivity to thermal fluctuations. Currently, thermal heater control of the ring modulators is employed to stabilize the operation and to precisely tune the modulation wavelength. The drawback of this approach is the added power penalty.

Monolithic optical isolators

The layer-by-layer processing that is inherent to the Si CMOS platform enables buffer-assisted deposition for enhanced phase stability and device functionality. A good application of this principle is the optical isolator shown in **Figure 5**.[15] The magneto-optical cerium-doped yttrium-iron-garnet (YIG) material provides a Faraday-effect-induced non-reciprocal transmission medium. The waveguide presents different indices of refraction (resonant wavelengths) to launched and reflected signals, and the resonator acts as an optical diode. A layer of YIG-on-Si stabilizes Ce:YIG against phase separation and preserves its optical non-reciprocity. The optical resonator structure gives field concentration and enhanced effective interaction length, and Ce:YIG on one side of the

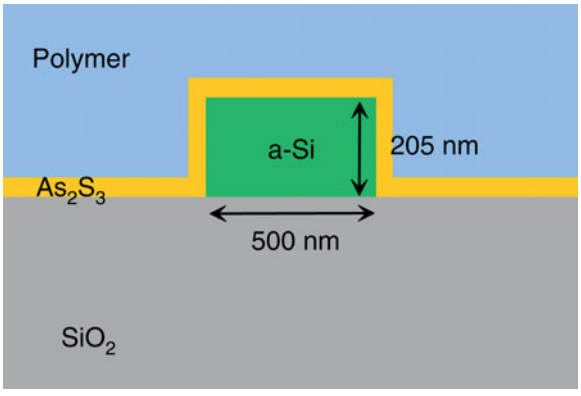

Figure 4. Athermal and trimmable design for silicon HiDex waveguides. Engineering mode sharing between the core (+ thermo-optic) and the clad (– thermo-optic) results in thermal stability for add-drop filters of 0.5 pm/K. The thin As_2S_3 layer provides a compatible photosensitive index for precision resonant wavelength trimming.[7,13] Note: a-Si, amorphous silicon.

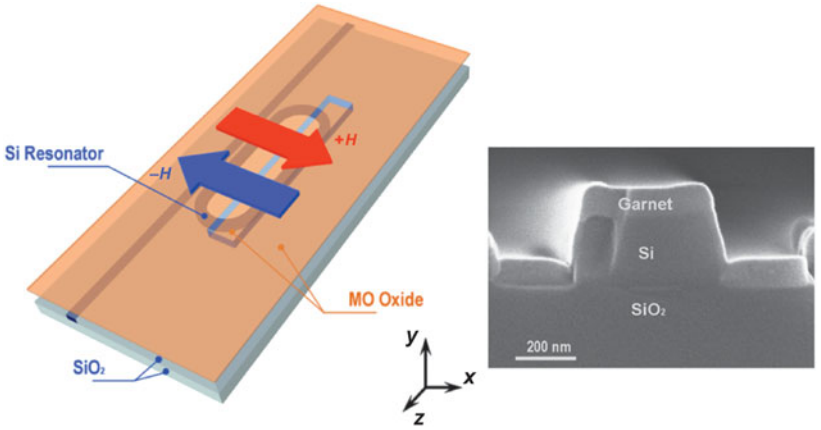

Figure 5. Schematic of monolithic optical isolator based on a Ce:YIG/YIG/Si composite waveguide structure embedded in an optical resonator device.[15] The direction of the applied magnetic field, *H*, determines the effective index of the composite waveguide. The micrograph depicts the magneto-optic, Ce:YIG garnet layer deposited on a silicon waveguide. Note: YIG, yttrium-iron-garnet; MO, metal oxide.

"racetrack" resonator discriminates against the direction of optical signal transmission. The result is a device with the smallest footprint and a high performance ~20 dB isolation at ~1550 nm.

The seamless ability to design and process multilayer composite structures adds unique value to the Si photonics platform that is not accessible elsewhere. The examples given here for athermal optical filters, trimmable resonant wavelengths, and optical isolation are only the beginning of the exploitation of optical performance scaling with composite materials structures.

Monolithic germanium photodetectors and modulators

Germanium growth without a thick graded buffer layer and dislocation-free Ge in selective area epitaxy on Si are essential to waveguide-integrated optical devices.[16] Because of the built-in biaxial tensile strain in Ge-on-Si following growth, the absorption edge is red-shifted beyond 1600 nm,[17] which is beneficial for detection of C+L bands (wavelength channels allocated for standard wavelength division multiplexed communication [C-band] and extended longer wavelength communication [L-band]) of optical fiber communication. Selective area epitaxy with SiO_2 masking allows submicron wide Ge stripes to be directly integrated on the waveguide. For an optical layer in the BEOL interconnect stack, Ge on a non-crystalline substrate such as SiO_2 or amorphous Si must be developed. There have been two approaches reported: polycrystalline Ge melting/solidification growth on SiO_2 in the front-end-of-line[18] and Ge on amorphous SiO_2 passivation layers and/or amorphous Si waveguides.[19] Both approaches are quite promising for monolithic integration.

The Franz-Keldysh (FK) GeSi electro-absorption modulator (EAM) has been demonstrated[14] for operation at λ =1550 nm. The dilute alloying of Ge with Si moves the absorption edge of as-grown, tensile-strained SiGe to 1550 nm. The FK EA modulator should provide the lowest energy/bit performance because of its reverse bias field-only operation. The major issue in device design is the reduction of the operating voltage to CMOS power supply levels of <1 V. A compressive-strained Ge-based EAM for wavelengths shorter than 1550 nm based on the quantum confined Stark effect (QCSE), which is an electric field induced energy shift in the absorption edge in semiconductor quantum well structures, has been proposed[20] and demonstrated.[21] Operation of this QCSE modulator near 1260 nm (the communication O-band) would have significant value.

Monolithic germanium lasers

Germanium is an indirect semiconductor with a Γ-L valley separation of 140 meV. Photons are most efficiently generated at the Γ-point (direct gap position) where no change in wave vector is required. As the energy of the conduction minimum at the direct gap position is lowered and approaches the minimum for the band structure at L (indirect gap position), photon emission for lasing becomes faster and the preferred pathway for the recovery of the forward-biased injected carrier density to approach equilibrium. This small energy separation together with tensile strain-engineering allows for electrons to populate in the Γ valley under *n*-type doping. This is schematically shown in **Figure 6**. Approximately 0.2% tensile-strain is typically present in Ge-on-Si epilayers due to the thermal expansion mismatch between Ge and Si. This level of strain will shrink the Γ-L valley separation to 110 meV and allow net positive gain when doped at a donor doping concentration $N_D > 5 \times 10^{19}$ cm^3.

Based on this concept, Ge lasing has been demonstrated by optical pumping[22] and electrical injection.[23] The current challenge is to reduce the lasing threshold current.[24]

High volume manufacturing on the CMOS platform

Over the last decades, silicon photonics technology has reached new heights in terms of device performance and levels of integration. It has garnered vast interest due to its tremendous market potential for applications in high-performance computing, optical interconnects and long haul communications for various distances (e.g., ranging from a few meters to hundreds or thousands of kilometers), and photonic sensors. Recent acquisitions of silicon photonics companies and public announcements from major semiconductor players indicate an inflexion point for commercialization in the coming years.[25–27] At this critical juncture, concerted efforts are required to develop technology and process know-how for silicon photonics fabrication in a yield-oriented manufacturing environment with upside potential to meet

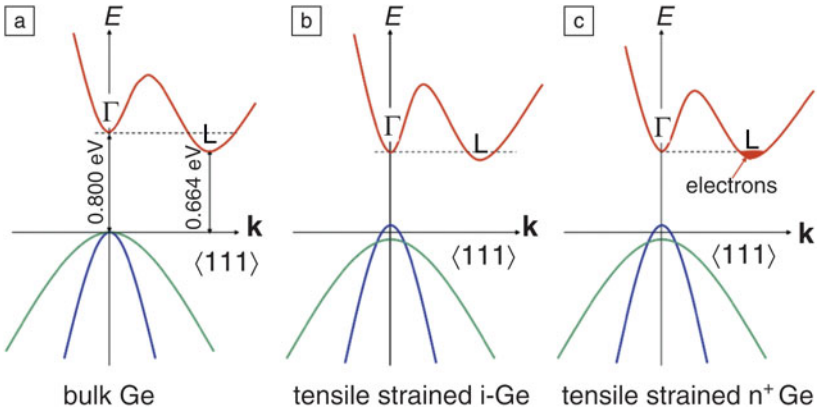

Figure 6. Band structure engineering for Ge lasing at 1550 nm: (a) bulk Ge: the conduction band minimum is at the indirect gap wave vector (**k**) position (L-point); (b) epitaxial Ge-on-Si under 0.25% tensile strain: the energy gap (*E*) decreases as the conduction band minimum for the direct gap (Γ-point) is lowered and the valence band degeneracy is lifted with light hole maximum increasing; and (c) quasi-Fermi level equilibration with *n*-doping: the energy of the highest filled conduction band states at the L-point is raised by doping with donors.

volume demand so as to establish a robust and cost-viable pathway for production in commercial foundries. One of the examples is the silicon photonics technology platform that the Institute of Microelectronics (IME) in Singapore developed and transferred to GlobalFoundries in Singapore (GFS) using their 200 mm, 0.18 μm CMOS foundry line. This breakthrough leverages the low-cost, high-yield processing capabilities that are available in a commercial manufacturing facility.[28,29]

Figure 7a shows a schematic of this platform, in which a SOI with 220 nm top Si/3 μm buried oxide (BOX)/high resistivity base is used. The Si passive fabrication process involves two partial etches: one for a 70 nm grating etch-depth and another for a 90 nm slab thickness in a Si rib waveguide. After the two partial etches, a third Si etch to the BOX is used to complete the passive device formation. The surface grating coupling scheme allows in-line wafer-level probing of the optical device performance immediately after passive device and circuit formation, as well as after active device formation, as shown in Figure 7c. For active devices, multiple Si implantations are conducted prior to Ge selective epitaxy, after which an n^{++} Ge implantation is performed. The standard BEOL process consisting of tungsten plugs and aluminum metal is used for a two-level metallization. Figure 7b shows an example of the Ge-photodetector portion after metal-1 patterning. A TiN heater is integrated between the two metal levels for thermal-optical tuning and modulation. An undoped oxide is used as the inter-layer dielectric and the inter-metal dielectric. A silicon nitride top layer is used for passivation and can be optionally removed from areas where efficient surface coupling is required.

Ge-Process qualification

In CMOS integrated circuit (IC) foundries, the use of pure germanium for fabrication is considered "new," thereby requiring careful introduction into existing manufacturing lines in order to minimize any undesirable impact on the momentum of ongoing production lots. For this effort, we successfully introduced Ge-integration with the existing Si-process line, while preserving consistent performance and yield. High-speed Ge photodetectors (>20 GHz) with low dark current (~11 nA) and high responsivity (~1.06 A/W) at 1550 nm were verified. Tight distributions in electrical device characteristics were achieved, as illustrated in **Figure 8**, indicating excellent process-control uniformity in the CMOS foundry.

Transfer evaluation results and (manufacturing) process readiness

Process modules were individually qualified during the technology transfer to ensure that the photonic device structures met the specifications in terms of dimension control and uniformity. For instance, for in-line process control, a post-etch final inspection critical dimension of 247.9 +/– 4.93 nm

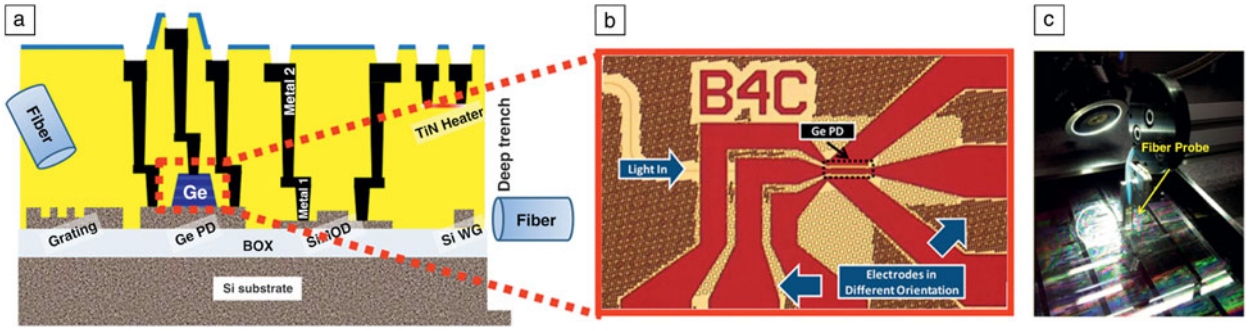

Figure 7. (a) Schematic of silicon photonics platform that was set up in GlobalFoundries of Singapore's 0.18 μm complementary metal oxide semiconductor foundry line. (b) Optical image of an 8 × 25 μm Ge photodetector (PD) after metal 1. (c) Wafer-level optical tester used for photonic device characterization.[27] BOX, buried oxide layer; MOD, modulator; WG, waveguide.

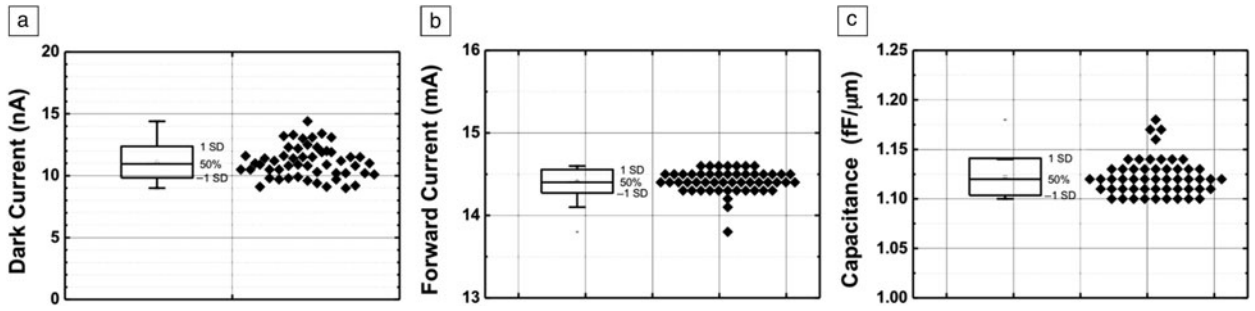

Figure 8. Statistical plots for (a) dark current, (b) forward current, and (c) normalized capacitance (at |1V|) for an 8 × 25 μm² Ge photodetector showing excellent uniformity.[27] (SD = standard deviation)

(mean +/– 1σ) with good process controllability, for example, a C_P (process capability index C_{PK}) of 2.03 (1.89) is typically obtained from a 0.18 μm CMOS wafer (for a 250 nm nominal critical dimension [CD] target). This again gives an indication of the strict process controls achievable from the manufacturing facility.

From wafer-level optical measurements, low-loss passives (e.g., waveguides [CD ~500 nm] and bends [with a radius of 5 μm and a CD of ~500 nm]) were achieved. The Si-channel waveguide loss was ~1.8 +/– 0.2 dB/cm, Si-rib waveguide loss was ~1.0 +/–0.2 dB/cm, and Si-bending loss was ~0.014 +/–0.002 dB. These results were also verified at the die-level through edge-coupling from a lensed fiber and a 180 nm nanotaper with typical coupling loss <3 dB/facet.

In addition, MZI *p–n* junction-based Si modulators, a TiN heater for thermal-optics, and various efficient fiber-to-waveguide coupling schemes have also been established. **Figure 9** shows an example of the MZI modulator, in which 4 mm Si MZI modulators exhibited typical figure-of-merit, the modulator efficiency, $V_\pi L$ of ~2.3 V.cm (V_π = applied voltage and L = device length for a 180° phase change) with phase shift insertion loss and device capacitance of 0.9 dB/mm and <0.4 pF/mm at –1 V, respectively. The loss and speed performance can be optimized via doping engineering, allowing further improvement of $V_\pi L$ for lower power operation and also higher speed,[30] in which an excellent 50 Gb/s modulating performance was obtained.

Both lateral and surface couplings of light are facilitated by edge couplers and gratings, respectively, in this platform. For lateral coupling, a ~120 μm deep Si trench with a smooth oxide sidewall facet (above the Si trench) eliminates the need for sidewall polishing. A suspended oxide coupler with a broadband wavelength operating range (>100 nm) and large fiber-to-chip alignment tolerance have been fabricated on the same platform. In summary, this enablement of an accessible commercial foundry caters to the increasing demands in silicon photonics prototyping and mass production needs.

Si-photonics and electronics integration platform/2.5D TSI

To date, silicon photonics technology has advanced and matured considerably with regard to the development of fundamental

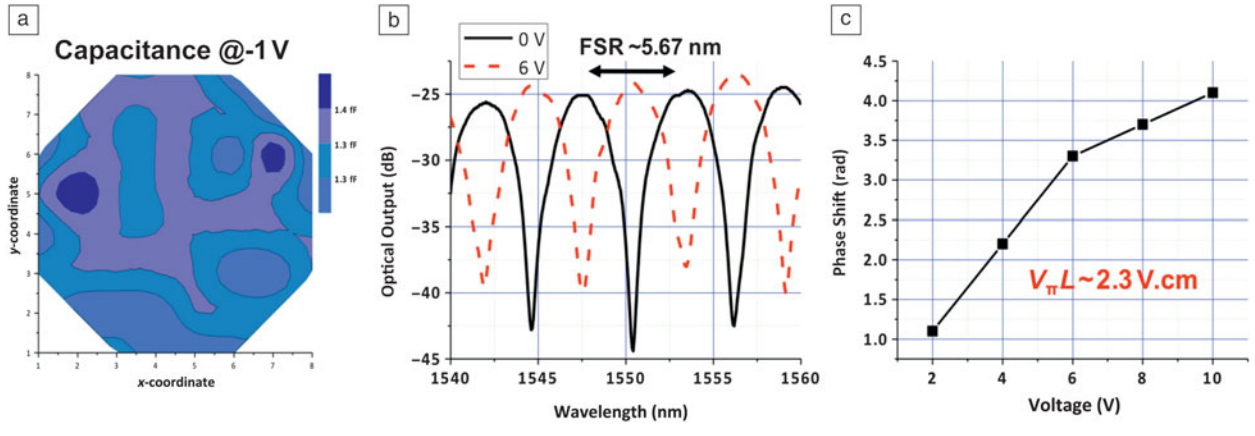

Figure 9. Efficient silicon Mach-Zhender interferometer modulator with low insertion loss and capacitance. (a) The spatial variation in the junction capacitance controlling the phase modulation is remarkably uniform; (b) the signal extinction dependence on wavelength and applied junction bias shows the behavior expected for an ideal Mach-Zhender interferometer; and (c) the optical phase shift is linear with applied bias voltage (V) for V < 6V with high efficiency indicated by a low $V_\pi L$. Note: FSR, free spectral range; V_π and L, applied voltage and device length for 180 degree phase change (maximum signal extinction), respectively.[30]

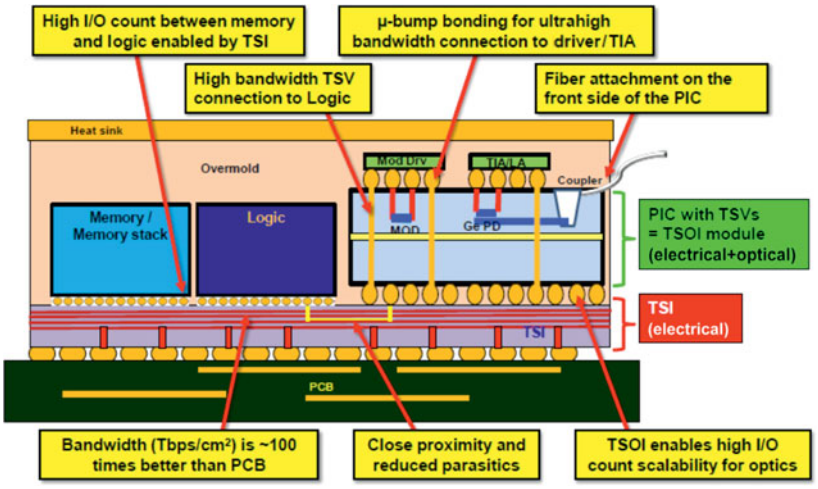

Figure 10. Heterogeneous electronic and photonic integrated circuit integration based on the 2.5D through-silicon interposer (TSI) platform. This class of structures is the favored hybrid integration scheme preceding full monolithic integration. Note: TIA, transimpedance amplifier; TSV, through-silicon via; I/O, input/putput; PCB, print-circuit board; PD, photodetector; LA, limiting amplifier; MOD, modulator; MOD Drv, modulator driver; PIC, photonic integrated circuit; TSOI, through-silicon-optical-interposer.[7,34]

devices and establishment of commercialization pathways. With this, the next vital milestone is the convergence of electronics and photonics. In addition to the approach through monolithic integration,[31,32] heterogeneous electronic-photonic integrated circuits integration presents a more versatile alternative. It provides a path that harnesses the best performing electronics and photonics circuits, since each circuit type can be optimized using its own technologies and material systems without having to sacrifice their respective performance for the large-chip area difference, process thermal-budget incompatibility, and difficulty in managing large surface topographic features. 2.5D through-silicon interposer (TSI) is a technology platform that offers versatile heterogeneous integration capabilities for expanding the scope of applications with the mission of increasing functionalities (e.g., logic, memory, analog/mix-signal, radio frequency, microelectromechanical systems) and performance (e.g., shorter distance and finer bumping/solder features). In this platform, a chip area is allocated to host multiple functional guest chips (e.g., electronic or photonic integrated circuit, memories). This allocated chip provides the needed electrical metallization layers to connect all guest chips laterally and also to connect vias through silicon vertically to the bottom print-circuit-board. While the benefits of 2.5D TSI architecture have been actively explored in the last few years for memory and logics integration for attaining higher performance and lower power,[32,33] the integration of photonic functions is a critical strategy to meet many important performance targets needed for handling large data, lower power budget, and cost.

Figure 10 illustrates one of the implementation schemes. It has memory and logic on one side and optical I/O (input/output) on the other, complete with electrical connection via the bottom Si-interposer with both redistribution layers (RDLs) and through-silicon-vias (TSV). The driver and amplifier chips are flip-chip bonded on to a silicon photonic chip, which also has TSV connecting the electrical signals to the common (i.e., bottom) interposer for communicating with the memory/logic units. Thus, different from the monolithic approach, 2.5D TSI/3D IC technology is considered to be a more practical and versatile approach to integrate with silicon photonics. The TSI serves as a host to connect various functional chips and allows the photonics chips to be brought much closer to logic and memory chips, enabling higher I/O count scalability (e.g., with fine RDL pitch <2 μm and fine micro-bump pitch in the range of 10–40 μm) for higher data bandwidth and lower energy consumption (e.g., <pJ/bit) at lower cost.

Summary

As photonics has penetrated computation system design to shorter links, photonic components have become a larger fraction of the total system interconnection, and issues of system performance, manufacturing yield, and reliability are being reevaluated. Questions of chip design, redundancy strategy, assembly and packaging, modulation format, and the ability to support scaling of system density and input/output (I/O) bandwidth with low latency and energy have become fundamental. Data centers and high-performance computer installations today can feature as many as 20 million optical transceivers. Within the next five years, photonic interconnection for the PCIe processor to memory communication at the board level will be widespread. These high component counts far exceed current data communication and telecommunication demands. Silicon photonics is the unique platform that will deliver high volume manufacturing and cost reduction through monolithic integration. Silicon microphotonics today, as silicon microelectronics during the last 40 years, has the potential to satisfy these system constraints and provide 40 more years of performance scaling.

Acknowledgments

D.-L. Kwong would like to acknowledge funding from Exploit Technologies Pte Ltd (ETPL) A*STAR under "Silicon Photonics Commercialization Flagship Grant ETPL/11-R15FSH-001," the collaborative efforts of GlobalFoundries Fab 3/5, Singapore, as well as P. Dong and Y.K. Chen from Bell Labs, Alcatel-Lucent in IME's silicon photonics technology platform transfer, and M. Hochberg, T.B. Jones, and the team in the development of the optical wafer-level test set-up at IME.

References

1. The Microphotonics Center at Massachusetts Institute of Technology, "Communications Technology Roadmap, CTR I (2005), CTR II (2009), CTR III Scaling and Energy (2010); Scaling Copper (2011); Short Reach Optical Interconnection (2013);" http://mph-roadmap.mit.edu/.

2. A. Agarwal, *IEEE Trans. Parallel Distrib. Syst.* **2**, 4 (October 1991).
3. G. Kurian, J.E. Miller, J. Psota, J. Eastep, J. Liu, J. Michel, L.C. Kimerling, A. Agarwal, *Proc. Parallel Architectures and Compilation Techniques, PACT'10*, 477 (2010).
4. G. Kurian, S. Chen, C.H.O. Chen, J.E. Miller, J. Michel, W. Lan, D.A. Antoniadis, P. Li-Shiuan, L. Kimerling, V. Stojanovic, A. Agarwal, *Parallel & Distributed Processing Symposium (IPDPS), 2012 IEEE 26th International*, (2012), pp. 1117–1130.
5. K.K. Lee, D.R. Lim, A. Agarwal, K. Wada, L.C. Kimerling, *Mater. Res. Soc. Symp. Proc.* **637**, K. Wada, T.F. Krauss, P. Wiltzius, K. Asakawa, E.L. Thomas, Eds. (Materials Research Society, Warrendale, PA, 2001), p. E3.4.1.
6. H.F. Hamann, A. Weger, J.A. Lacey, Z. Hu, P. Bose, E. Cohen, J. Wakil, *IEEE J. Solid-State Circuits* **42**, 56 (2007).
7. V. Raghunathan, W.N. Ye, J.J. Hu, T. Izuhara, J. Michel, L.C. Kimerling, *Opt. Express* **18**, 17631 (2010).
8. L.C. Kimerling, J. Michel, *ECS Trans.* **41** (7), 3 (2011).
9. R.A. Soref, J.P. Lorenzo, *IEEE J. Quantum Electron.* **22**, 873 (1986).
10. J. Van Campenhout, P. Rojo-Romeo, D. Van Thourhout, C. Seassal, P. Regreny, L. Di Cioccio, J.M. Fedeli, C. Lagahe, R. Baets, *Opt. Express* **15**, 6744 (2007).
11. H. Park, A.W. Fang, S. Kodama, J.E. Bowers, *Opt. Express* **13**, 9460 (2005).
12. D.J.W. Klunder, F.S. Tan, T. van der Veen, H.F. Bulthuis, G. Sengo, B. Docter, H.J.W.M. Hokstra, A. Driessen, *J. Lightwave Technol.* **21**, 4 (2003).
13. S. Grillanda, V. Raghunathan, V. Singh, F. Morichetti, J. Michel, L. Kimerling, A. Melloni, A. Agarwal, *Opt. Lett.* **38**, 24 (2013).
14. J. Liu, M. Beals, A. Pomerene, S. Bernardis, R. Sun, J. Cheng, L.C. Kimerling, J. Michel, *Nat. Photonics* **2**, 7 (2008).
15. L. Bi, J. Hu, P. Jiang, D.H. Kim, G.F. Dionne, L.C. Kimerling, C.A. Ross, *Nat. Photonics* **5**, 12 (2011).
16. H.C. Luan, D.R. Lim, K.K. Lee, K.M. Chen, J.G. Sandland, K. Wada, L.C. Kimerling, *Appl. Phys. Lett.* **75**, 2909 (1999).
17. Y. Ishikawa, K. Wada, D.D. Cannon, H.-C. Luan, L.C. Kimerling, *Appl. Phys. Lett.* **82**, 2044 (2003).
18. A. Solomon, X. Fengnian, S.W. Bedell, Y. Zhang, T. Topura, P.M. Rice, Y.A. Vlasov, *Opt. Express* **18**, 4986 (2010).
19. K.A. McComber, X. Duan, J.-F. Liu, J. Michel, L.C. Kimerling, *Adv. Funct. Mater.* **22**, 1048 (2012).
20. Y.H. Kuo, Y.K. Lee, Y. Ge, S. Ren, J.E. Roth, T.I. Kamins, D.A.B. Miller, J.S. Harris, *Nature* **437**, 1334 (2005).
21. P. Chaisakul, D. Marris-Morini, M.-S. Rouifed, G. Isella, D. Chrastina, J. Frigerio, X. Le Roux, S. Edmond, J.-R. Coudevylle, L. Vivien, *Opt. Express* **20**, 3219 (2012).
22. J.F. Liu, X. Sun, L.C. Kimerling, J. Michel, *Opt. Lett.* **34**, 1738 (2009).
23. R.E. Camacho-Aguilera, Y. Cai, N. Patel, J.T. Bessette, M. Romagnoli, L.C. Kimerling, J. Michel, *Opt. Express* **20**, 11316 (2012).
24. Y. Cai, Z. Han, X. Wang, R.E. Camacho-Aguilera, L.C. Kimerling, J. Michel, J.-F. Liu, *IEEE J. Sel. Top. Quantum Electron.* **19**, 1901009 (2013).
25. Luxtera, "Molex Purchases Luxtera's Silicon Photonics-Based Active Optical Cable (AOC) Business; Partners on Future AOC Development," available at http://www.luxtera.com/20110111226/molex-purchases-luxtera's-silicon-photonics-based-active-optical-cable-aoc-business-partners-on-future-aoc-development.html. News release. January 11, 2011.
26. Cisco, "Cisco Completes Acquisition of Lightwire," available at http://www.cisco.com/web/about/ac49/ac0/ac1/ac259/lightwire.html. News release. March 19, 2012.
27. Mellanox Technologies, Ltd. "Mellanox Technologies, Ltd. Completes Acquisition of Kotura, Inc.," available at http://www.mellanox.com/page/press_release_item?id=1096. News release. August 15, 2013.
28. A.E.J. Lim, T.Y. Liow, J.F. Song, C. Li, Q. Fang, X. G. Tu, N. Duan, K.K. Chen, R. Poh, C. Peng, B.W. Mun, M.N. Islam, J.S. Park, C. Subbu, G.Q. Lo, in *Optical Fiber Communication Conference and Exhibition/National Fiber Optic Engineers Conference (OFC/NFOEC)*, 2014.
29. Institute of Microelectronics, Alcatel-Lucent, and GlobalFoundries, "Collaboration between IME/A*STAR, GFS and Alcatel-Lucent to Bring Advanced Photonics Chips to Market," available at http://www.ime.a-star.edu.sg/files/news/20120303202216005O.pdf. News release, November 28, 2011.
30. X.G. Tu, T.Y. Liow, J.F. Song, X.S. Luo, Q. Fang, M.B. Yu, G.Q. Lo, *Opt. Express* **21** (10), 12776 (2013).
31. S. Sahni, A. Ayazi, Y.M. Chi, A. Dahl, P. De Dobbelaere, S. Gloeckner, K.Y. Hon, S. Hovey, Y. Liang, M. Mack, G. Masini, A. Mekis, M. Peterson, T. Pinguet, J. Schramm, M. Sharp, P. Sun, R. Timpe, L. Verslegers, *2013 IEEE 10th International Conference on Group-IV Photonics* (IEEE, Seoul).
32. A. Solomon, S. Shank, W. Green, M. Khater, E. Kiewra, C. Reinholm, S. Kamlapurkar, A. Rylyakov, C. Schow, F. Horst, H. Pan, T. Topuria, P. Rice, D.M. Gill, J. Rosenberg, T. Barwicz, M. Yang, J. Proesel, J. Hofrichter, B. Offrein, X.X. Gu, W. Haensch, J. Ellis-Monaghan, Y. Vlasov, *Tech. Dig. Int. Electron Dev. Mtg.* (2012).
33. J.Y. Xie, "Interposer Integration through Chip on Wafer on Substrate (CoWoS) Process" (Semicon West, San Francisco, CA.) 2012; http://www.semiconwest.org/sites/semiconwest.org/files/docs/John%20Xie_Altera%20Corporation.pdf.
34. Institute of Microelectronics, "2.5D Silicon Interposer Multi-Project Wafer;" https://www.a-star.edu.sg/ime/SERVICES/25d_silicon_interposer_multi_project_wafer.aspx. □