Ultralow-power silicon photonic interconnect for high-performance computing systems

Guoliang Li[a*], Xuezhe Zheng[a], Jon Lexau[b], Ying Luo[a], Hiren Thacker[a], Thierry Pinguet[c], Po Dong[d], Dazeng Feng[d], Shirong Liao[d], Roshanak Shafiiha[d], Mehdi Asghari[d], Jin Yao[a], Jing Shi[a], Ivan N. Shubin[a], Dinesh Patil[b], Frankie Liu[b], Kannan Raj[a], Ron Ho[b], John E. Cunningham[a], Ashok V. Krishnamoorthy[a]

[a]Sun Microsystems Physical Sciences Center, San Diego, CA 92121, USA;
[b]Sun Laboratories, Menlo Park, CA 94025, USA;
[c]Luxtera Inc., Carlsbad, CA 92009, USA;
[d]Kotura Inc., Monterey Park, CA 91754, USA
*guoliang.li@sun.com

(Invited Paper)

ABSTRACT

The Ultra-performance Nanophotonic Intrachip Communication (UNIC) project aims to achieve unprecedented high-density, low-power, large-bandwidth, and low-latency optical interconnect for highly compact supercomputer systems. This project, which has started in 2008, sets extremely aggressive goals on power consumptions and footprints for optical devices and the integrated VLSI circuits. In this paper we will discuss our challenges and present some of our first-year achievements, including a 320 fJ/bit hybrid-bonded optical transmitter and a 690 fJ/bit hybrid-bonded optical receiver. The optical transmitter was made of a Si microring modulator flip-chip bonded to a 90nm CMOS driver with digital clocking. With only 1.6mW power consumption measured from the power supply voltages and currents, the transmitter exhibits a wide open eye with extinction ratio >7dB at 5Gb/s. The receiver was made of a Ge waveguide detector flip-chip bonded to a 90nm CMOS digitally clocked receiver circuit. With 3.45mW power consumption, the integrated receiver demonstrated -18.9dBm sensitivity at 5Gb/s for a BER of 10^{-12}. In addition, we will discuss our Mux/Demux strategy and present our devices with small footprints and low tuning energy.

Keywords: Optical interconnect, silicon photonics, high performance computing (HPC), optical transmitter, optical receiver, wavelength-division multiplexing (WDM), Si ring resonator, hybrid bonding

1. INTRODUCTION

In the recent several decades, computing technology has made tremendous progress in the benchmark performance, such as gigaupdates per second (GUPS) and global fast Fourier transforms (FFT). The progress can be mostly attributed to the following three factors: 1) faster transistor, 2) more advanced design, and 3) more transistors in the system. These factors have resulted in several orders of magnitude increase in the computer clock frequency, and led to the modern computer processor chips with multi-core, multi-thread, and heavy on-chip cache. Behind the above progress was the ever-advancing manufacturing technology [1]. By shrinking the transistor size, not only the transistors are made faster, but also more transistors are packed into a single processor chip.

However, nowadays the manufacturing technology is approaching the physical limits; device size scaling alone can no longer provide exciting improvement to the computing performance. The slowdown of device scaling not only saturates the clock frequency, but also makes the further increase of transistor count very difficult. In order to significantly increase the transistor count, one has to either increase the processor chip size, or use multiple processor chips in the same system. Unfortunately, significantly sizing up the chip is economically prohibitive due to unacceptably lower yield and higher manufacturing cost. On the other hand, using multiple processor chips encounters another daunting challenge - the inter-chip communication. The traditional way is to assemble multiple, often separately packaged, processor chips

on a printed circuit board (PCB), and use wire-bonding and solder balls to connect the chip I/O pins to the PCB. Since such I/O pins and paths are bulky, power-hungry SerDes (serializer/deserializer) circuits often have to be used to increase communication bandwidth over a limited number of I/O, which may result in orders of magnitude higher power consumption than on-chip wires. Since the inter-chip high-speed signals have to travel through the lossy transmission lines on PCB, the communication distance is limited to a few inches for 10 Gb/s signal, and the distance limit drops fast when the signal speed gets higher, unless some power-hungry equalization circuit is used.

A better way for co-packaging multiple chips is to use 3D stacking combined with through-silicon vias (TSVs), which can significantly reduce the inter-chip communication distance and allow much higher I/O density [2-4]. Power-hungry SerDes and equalization circuits can be avoided. Miniaturized multi-chip packaging using this approach is achieving remarkable success in the wireless IC industry. However, the tough thermal management in 3D stacking limits its capability for co-packaging large number of high-power chips (such as processor chips).

Another recently emerged approach is to use proximity coupling, such as capacitive coupling [5]. This approach is based on the fact that high-speed signal can travel through closely placed tiny metal pitches (with high capacitance), and suffer very little signal loss and distortion. This effectively extends on-chip wires across multiple chips, and enables high-density, high-speed and low-power interconnects together with chip reworkability (since no soldering is involved). Using this approach, many processor chips can be placed in a 2D-grid configuration, and the adjacent chips can be interconnected with overlapping facing-down "bridge" chips. However, in spite of so many advantages, a large multichip system using proximity coupling would suffer from long message latency [6]. This is because 1) standard on-chip wires propagate at only 5% to 10% the speed of light [7], and 2) communication between far away chips must go through multiple other chips. This long message latency will penalize the system performance. In addition, the long travel distance results in large capacitance of on-chip wires, which adds power consumption and limits signal speed.

It is now widely recognized that optical interconnect could be the ultimate solution to removing the inter-chip communication bottleneck for high performance computing (HPC) systems using many processor chips [8-11]. True speed-of-light communication and point-to-point communication network can effectively reduce message latency; while wavelength-division multiplexing (WDM) can offer unmatched bisection bandwidth. Among different optical solutions, Si photonic links are especially attractive owing to 1) compatibility with CMOS fabrication process, and 2) compact and low-loss Si optical waveguide due to high refractive-index contrast between Si and SiO_2. With closely integrated electronics, Si photonic links can potentially offer huge advantages in achieving small footprint, low power, and low cost. Furthermore, optical proximity communication allows multiple chips to communicate without soldering, which enables reworkability for high system yield [12].

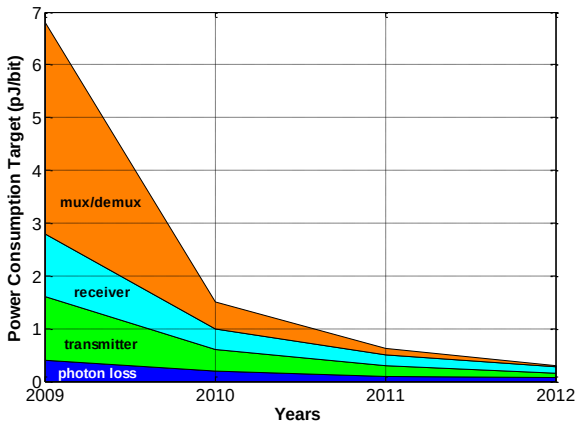


Figure 1. Power consumption targets in the UNIC project. Different colored areas represent portions for different components; the top curve represents the power consumption of the entire photonic link (including the transmitter and receiver circuits), which scales down to 300 fJ/bit in year 2012.

In 2008, DARPA announced its funding to the Ultra-performance Nanophotonic Intrachip Communication (UNIC) project. This project aims to achieve unprecedented high-density, low-power, large-bandwidth, and low-latency optical interconnect for highly compact supercomputer systems. It sets extremely aggressive goals on power consumptions and

footprints for optical devices and the integrated VLSI circuits. Figure 1 shows the power consumption targets for this project. The power for the entire link, including the VLSI circuits in the transmitter and the receiver, exponentially scales down to 300 fJ/bit over 4 years. Compared with the tens of pJ/bit power consumption in a state-of-the-art optical transceiver, this represents a ~100X improvement. The different colored areas in Figure 1 represent power portions for different components. The "photon loss" portion represents the consumption of laser power in the link (external laser is assumed in this project). For example, in 2012, each wavelength channel will have 1mW laser power at the beginning of the link with a data rate of 15 Gb/s, hence the photon loss will be 67 fJ/bit. This part of power consumption becomes significant after other power consumptions are aggressively scaled down. To maintain a closed link with minimum photon loss, it requires excellent performance of all the components in the link, including low loss, high extinction ratio, low jitter, and high receiver sensitivity.

2. INTEGRATION OF ELECTRONICS AND PHOTONICS

One of the most attractive attributes of Si photonics is that it can be monolithically integrated with VLSI electronic circuits that are needed in a photonic link. These circuits may include modulator driver, transimpedance amplifier, control/tuning circuits, and some other signal conditioning circuits. Such monolithic integration can potentially save manufacturing cost, reduce electrical parasitics, and provide more system design flexibility. Monolithically integrated electronic and photonic circuits using slightly modified mainstream CMOS fabrication foundries have been developed by several companies [13]. A drawback of monolithic integration is that it may not have the best-in-breed electronics and photonics since they would prefer different manufacturing platforms. For example, state-of-the-art CMOS circuits are now popularly built on bulk Si substrates; while photonic devices normally require SOI substrates for optical waveguides. Finding an innovative technique to monolithically integrate the best-in-breed electronics and photonics is a challenge that is being actively tackled [14].

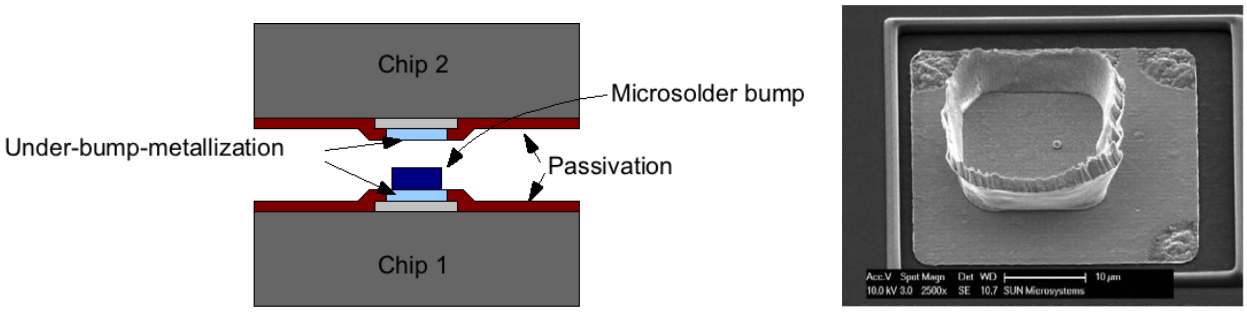


Figure 2. (a) Illustration of flip-chip hybrid-bonding. (b) SEM of a microsolder bump on top of a UBM pad.

In the UNIC project, Sun team collaborates with multiple partners for developing different devices required in the link. Each partner uses a different fabrication platform. Therefore monolithic integration of photonics and electronics in this project is not practical. In order to still demonstrate ultralow-power and high-performance Si photonic transmitters and receivers, we have developed a flip-chip bonding technique with very low electrical parasitics. Figure 2(a) illustrates such bonding technique. To bond a photonic chip with a VLSI circuit chip, both chips are first processed to add under-bump-metallization (UBM) on their respective bonding pads. The UBMs serve as both strong adhesion layer and diffusion barrier, and they enable low-resistance contact from microsolder bump to the aluminum pads. The microsolder bumps are fabricated on the UBM pads of one of the two chips. Figure 2(b) shows an SEM picture of such a microsolder bump. It is a metal alloy with ~18 μm footprint and a crown shape of several micrometers tall. The two post-processed chips are then flip-chip bonded together by thermocompression, which results in extremely small chip-to-chip separation. During this process the crown-shaped microsolder is compressed and embedded into the UBM pad of the opposing chip, resulting in extremely small resistance. We have measured <0.1 Ω resistance for each bonding site. The parasitic capacitance due to bonding depends on the pad size. It is estimated to be 20-30 fF with the current 30 μm pads.

As the footprint of the microsolder bumps can be scaled down to several micrometers, the bonding pad size can potentially be scaled down to ~10 μm range, and the resulted capacitance can be ~10X smaller.

To facilitate the tests, the hybrid-bonded chips are die-attached and wire-bonded to a PCB board for power, control, and high speed digital I/O connections. A photo of the complete assembly is shown in Figure 3. In this picture, the top bigger chip is a photonic chip, which is facing down and bonded to the bottom VLSI chip. The bottom surface of the VLSI chip is attached to the PCB using epoxy, and the VLSI chip is electrically connected to the PCB by wire-bonding. The photonic chip could have modulator or photodetector devices from one of Sun's partners. The VLSI chip, which is fabricated using TSMC 90nm CMOS technology, contains many copies of modulator driver and receiver circuits custom-designed for each partner's devices. It also has on-chip digital buffers and clock distribution. The modulator driver and receiver circuits can each be activated individually by a software control accessed via a graphical user interface on a computer. The computer software talks to the PCB through a JTAG interface.

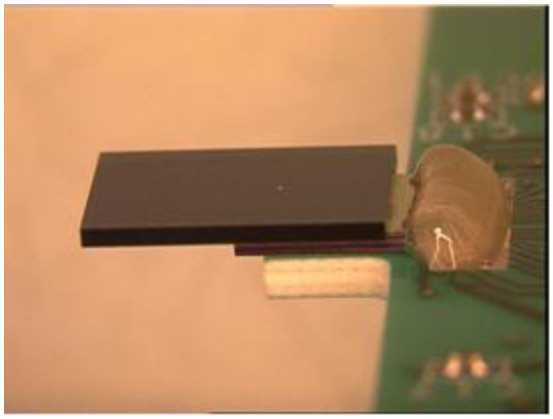


Figure 3. Photo of a hybrid-bonded photonic chip and VLSI chip wire-bonded to a test PCB board.

3. THE OPTICAL TRANSMITTER

3.1 Transmitter power consumption

An optical transmitter normally consists of an optical modulator and a modulator driver. Its power consumption includes the power dissipated in both parts. The following equation represents a complete consideration for the transmitter power consumption:

$$P_{TX} = P_{static}^{ckt} + P_{dynamic}^{ckt} + P_{bias}^{\mathrm{mod}} + P_{dynamic}^{\mathrm{mod}} + P_{tuning}^{\mathrm{mod}} \quad (1)$$

The first two terms are the power dissipated inside the driver circuit. This is the dominant part of the power consumption in almost all of the transmitters with external laser. It requires aggressive circuit designs and careful margining to minimize the power dissipated inside the driver circuit. In some cases, pre-emphasis circuit is used in order to reach higher modulation data rate [15], which may result in significant extra power consumption.

The 3rd term in equation (1) is due to the modulator biasing with non-zero voltage, which often causes DC current flow. The modulator bias voltage is the DC part of the voltage applied across the modulator. For example, when a high-speed voltage swing from 0V to 2V is applied across the modulator, the DC bias is 1V. Some modulators may require additional bias voltage for optimized performance. The bias current may be caused by the modulator diode structure, such as leakage current in a reverse-biased PN (or PIN) diode, or diffusion current in a forward-biased PIN diode. It can also be caused by photocurrent in an electroabsorption modulator. This part of power consumption can be pretty significant in some cases. For example, in case of a modulator with forward-biased diode, if the bias voltage is set at 0.5 V with a DC current flow of 2 mA, it results in 1 mW power consumption; in the case of an electroabsorption modulator, if the bias voltage is set at 2 V with a photocurrent of 0.5 mA, it also results in 1 mW power consumption. This term

alone can be 100 fJ/bit at 10Gb/s data rate, which must be taken into account. On the other hand, biasing a modulator with a reverse-biased PN diode based on plasma dispersion effect normally causes very small leakage current, and this part of power consumption can be ignored.

The 4th term in equation (1) is due to the ON/OFF switching of the modulator. This part of power consumption was often estimated using $1/4CV^2$ (assuming PRBS data), where C is the diode junction capacitance; but a modulator often has some parasitics that need to be taken into account (see section 3.2). To minimize switching energy, it is desirable to have an optical modulator with very low capacitance. Mach-Zehnder modulators normally have large capacitance, and may even have a 50 Ω termination in traveling-wave devices, which results in large switching energy. Modulators with forward-biased PIN junction [15] have relatively high switching current (or high equivalent capacitance), so its switching energy is also high. There are some promising modulator candidates with potentially small switching energy, including reverse-biased Si microdisk modulators (reported 85fJ/bit switching) [16], Franz-Keldysh effect GeSi modulators (reported 50fJ/bit switching) [17], GeSi quantum well modulators [18], evanescently coupled III-V modulators [19] and reverse-biased Si ring modulators used in this work.

The 5th term in equation (1) has also been often overlooked. A modulator may not work with specified performance at all laser wavelength channels. In fact, the performance of most modulators deteriorates very fast as the laser wavelength goes away from the optimum point. A resonator modulator may need to tune the resonant wavelength, and an electroabsorption modulator may need to tune its absorption spectrum. The tuning challenge for resonator type of devices has been discussed thoroughly in [8].

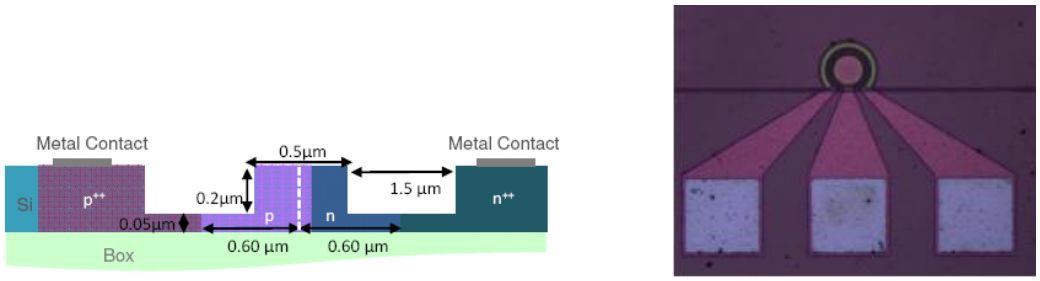


Figure 4. (a) Cross-sectional diagram of the ring waveguide. (b) Top-view photo of the microring modulator.

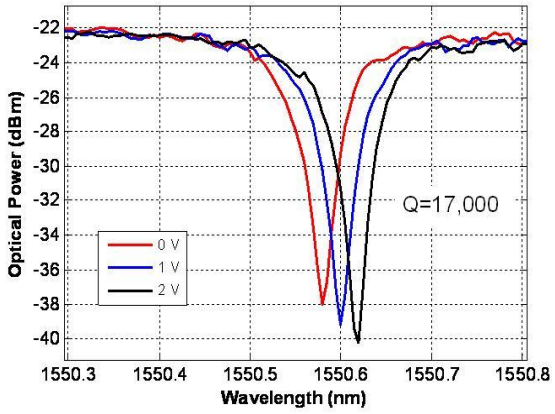


Figure 5. Resonance spectrum at different voltages (reverse bias to the PN junction) of the ring modulator.

3.2 The Si ring modulator

In this work we used a Si ring modulator chip designed and fabricated at Kotura, Inc [20]. Figure 4(a) shows a cross-sectional diagram of the phase modulation waveguide, and Figure 4(b) shows a top-view photo of the device. The radius

of the ring modulator was 15 μm. DC characterization was performed on die using electrical probe and lensed fibers. Figure 5 shows the resonance spectrum at different reverse bias voltages of 0V, 1V and 2V respectively. The measurement results indicate a quality (Q) factor of ~17000 for the ring resonator, and a voltage-induced wavelength shift of ~18pm/V. When operated around 1550.62nm wavelength and with a voltage swing from 0V to 2V, this ring modulator can achieve >10dB DC extinction ratio and ~2 dB insertion loss at on-state.

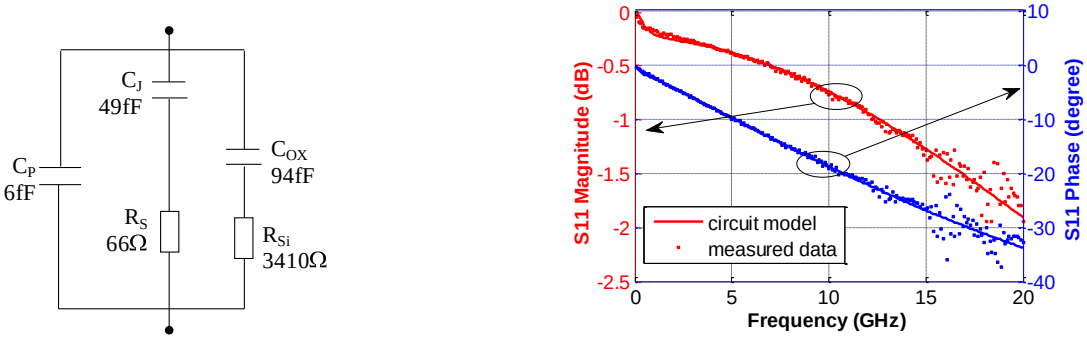


Figure 6. (a) Small-signal circuit model for reverse-biased ring modulators built on SOI substrate. The marked parameter values are for 0V bias. (b) Curve-fitting of the measured S11 at 0V using the circuit model in (a).

The high-speed behavior of the ring modulator has been carefully studied using a circuit model extracted by curve-fitting the measured S11 data. In this circuit model shown in Figure 6(a), C_P represents the capacitance between the electrodes (mostly due to the contact pads) through the top dielectrics and the air, C_J denotes the capacitance in the reverse-biased diode junction, R_S denotes the diode series resistance, C_{OX} denotes the capacitance through the dielectric and Si layers, and R_{Si} is the resistance in Si layer. The parameter values in Figure 6(a) were extracted at 0V, indicating a junction capacitance of 49fF. The curve fitting result is shown in Figure 6(b). The same data analysis at 1V and 2V bias indicated a slightly smaller junction capacitance, consistent with our expectation. Using the extracted circuit model, we calculated a switching energy of the ring modulator at ~100 fJ/bit when modulated by a pseudo-random data with a 2V swing at 5Gb/s (note that C_{OX} can be partially charged and discharged during the modulation). This switching energy can be readily lowered by reducing the ring modulator size.

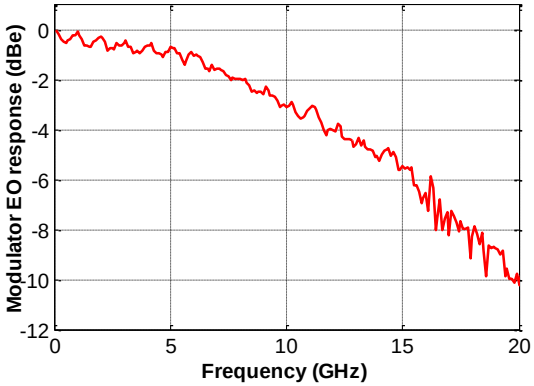


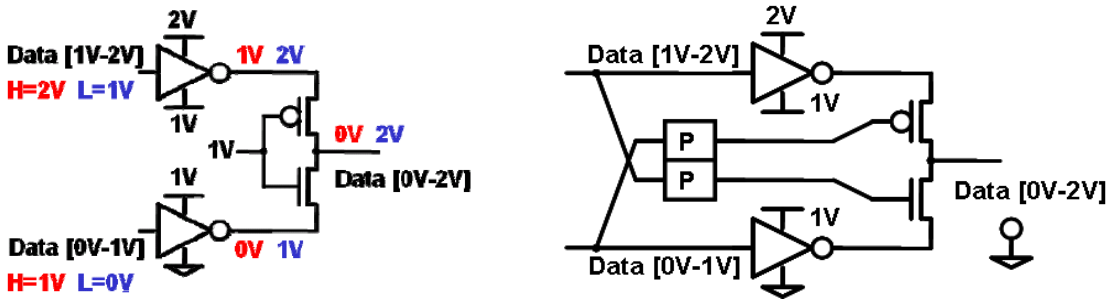
Figure 7. Measured frequency response for the 15 μm ring modulator biased at 0V.

Small-signal modulation bandwidth has also been measured for the ring modulator using a microwave network analyzer and a reference detector with known frequency response. The measured result in Figure 7 indicates a 3dB bandwidth of 10 GHz at 0V, with the wavelength set at ~1550.6nm in the resonance spectrum shown in Figure 5. The modulation bandwidth of the device is subject to both the RC limit, which is ~24GHz estimated from the circuit model (with a 50 Ω

source), and the photon lifetime limit, which is ~11.4GHz based on the measured quality factor. Clearly, the photon lifetime represents the primary limit to the modulation bandwidth for this device.

3.3 Modulator driver circuit

Our modulator driver was fabricated using TSMC 90nm CMOS technology [21]. One challenge for transmitter circuits is that modulators often require a modulation voltage of 2 V or higher, while the CMOS technology process only supports 1 V transistors with high speed. To solve this problem, a "cascode" driver design can be used, as illustrated in Figure 8(a). The data signal is fed into two inverters (with different voltage references) whose outputs in turn drive the rails of a final inverter with a fixed 1V input. At "HI' data value, the upper inverter outputs "LO" at 1V while the lower inverter outputs "LO" at 0V, in this case the final inverter's 1V input is "HI", thus it outputs "LO" at 0V; At "LO" data value, the upper inverter outputs "HI" at 2V while the lower inverter outputs "HI" at 1V, in this case the final inverter's 1V input is "LO", thus it outputs "HI" at 2V. Using this approach, a voltage swing from 0V to 2V can be achieved with only 1V across each transistor.



(a) (b)

Figure 8. (a) Schematics of a cascode modulator driver design. (b) Improved cascode modulator driver to avoid temporary overstress to transistors. "P" stands for pulse driver.

One problem with the circuit in Figure 8(a) is that if the two input inverters are fast, the final inverter may not activate until its sources have fully transitioned. This can cause a full 2V swing being temporarily applied to the transistors of the final inverter. For example, at stable "HI" data value, the final inverter outputs 0V; when the data transitions to "LO", the upper inverter may act fast and output 2V. Before the final inverter activates, its top transistor sees 2V across its source and drain. This temporary 2V stress may cause damage to the transistors. A solution is to replace the fixed 1V input of the final inverter with pulse drivers, which can selectively activate the transistors early during the data transition, as shown in Figure 8(b) (see [22] for more detailed discussions). This design was used for our modulator driver, together with a digital clock circuit. The whole circuit takes only 650 μm^2 of chip area [21].

3.4 Transmitter test results

After the ring modulator and the VLSI driver were flip-chip bonded together and attached to the PCB, we tested the high-speed performance of the transmitter assembly. With the on-chip digital buffers and the clock distribution turned on, the VLSI chip generates ~1W heat, which may cause difficulty to the high-speed testing since ring resonators can be very sensitive to temperature change. To combat this problem, the photonic chip was closely contacted with a copper heat sink. The resulted optical output eye of the transmitter is shown in Figure 9, tested with 5Gb/s PRBS31 data. The double falling edge is due to a timing error related to non-optimal clock loading, which can be easily corrected with timing and layout optimizations in the subsequent tapeout. The optical eye is wide open, with >7dB extinction ratio. Using an EDFA (to compensate for the large coupling loss between the fiber and the ring modulator bus waveguide) and a reference receiver, we have achieved stable error-free transmission at 5Gb/s for over 2 hours without active resonator bias tuning or temperature control, indicating a bit-error-rate $<10^{-13}$. The total transmitter power consumption was measured as 1.6mW by reading the power supply voltages and currents. This corresponds to 320 fJ/bit for the entire

transmitter including the ring modulator and its CMOS driver, as well as any excess power required to drive circuit parasitics including internal wiring and flip-chip pads.

It should be noted that our measured transmitter power consumption didn't include any power that may be needed for tuning the ring resonant wavelength. Ring resonator tuning is a challenging problem that needs innovative solutions [8]. We will report our experimental results on ring tuning in Section 5.

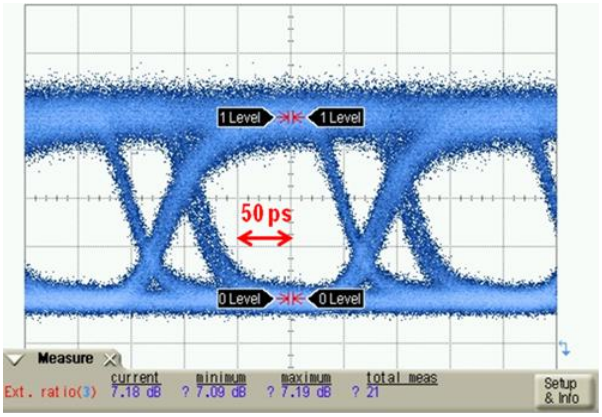
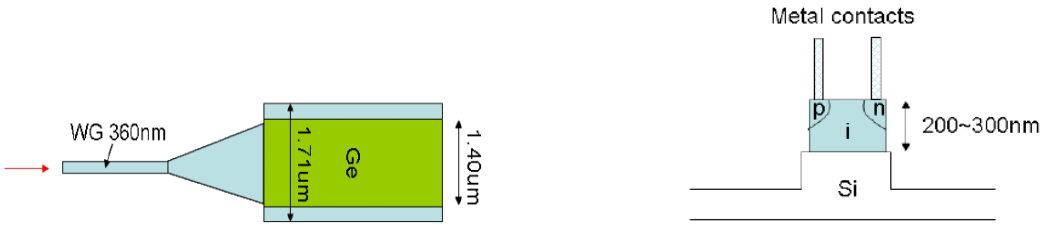


Figure 9. Measured eye-diagram of the integrated transmitter tested with 5Gb/s PRBS31 data, showing an extinction ratio of >7dB.

4. THE OPTICAL RECEIVER

4.1 The Ge photodetector

An optical receiver needs a photodetector (PD) at its front end to convert the received optical signal to electrical current signal. SiGe alloy has been long recognized as one of the best materials for optical detection in the near infrared wavelength range. It is particularly attractive in Si photonics since it is relatively easy to be monolithically integrated on the SOI substrates. Since its absorption strength starts to roll off at longer wavelengths (>1550nm), normally only pure Ge material is considered for C-band photodetection. In order to make high quality Ge PD on SOI substrate, a critical problem to be solved is the 4% lattice mismatch between the Ge and Si crystals. When growing Ge film on top of the Si surface, the large lattice mismatch can cause significant defects, particularly threading dislocations, which can damage the PD performance. In addition, when integrating Ge PDs with electronic ICs, the high temperature (600-700 ^{o}C) growth of Ge film can also degrade the transistor performance. In spite of these challenges, tremendous progress has been made on Ge PD devicess in recent years [23-29], along with great improvement on Ge-on-SOI material growth technique.

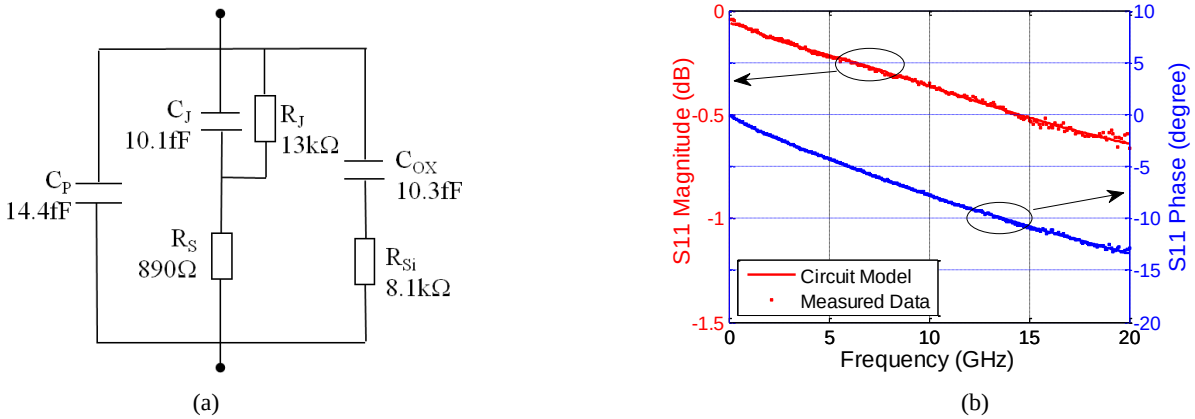


(a) (b)

Figure 10. Schematics of Ge waveguide photodetector: (a) top view, (b) cross section.

In this work we used a Ge PD chip designed at Luxtera and fabricated with Luxtera-Freescale's Ge-enabled CMOS-Photonic process [25]. Figure 10 shows the schematics of the Ge waveguide PD design. A 200~300nm thick Ge film is grown on top of the Si waveguide. The edges of the Ge strip are doped to form a p-i-n diode and for contacts. Light coming in from the Si waveguide is evanescently coupled into the Ge region and converted to photo current. To make the Ge fabrication compatible with CMOS process, the Ge film was epitaxially grown using a reduced pressure chemical vapor deposition (RPCVD) technique at 350 ^{o}C [25]. The single-step low-temperature Ge growth minimizes the thermal impact on the transistors, but it may result in more defects in the Ge film, which can lead to larger dark current.

To achieve high receiver sensitivity with low-power circuit, the PD needs to have high photo responsivity and low dark current. Our characterizations show that the Ge waveguide PDs on Luxtera's chip have a high responsivity of ~0.7 A/W at 1550 nm, and their DC responsivity is insensitive to the applied reverse bias voltage. To lower dark current, we chose a low 0.5V reverse bias for the Ge PD used in our integrated receiver. At 0.5V bias, the Ge PDs have a dark current of 1-3 μA, and a 3-dB electrical bandwidth of 7-10 GHz measured in a 50Ω system.



(a) (b)

Figure 11. (a) Small-signal circuit model for a Ge PD at a reverse bias of 0.5V. (b) Curve-fitting of the measured S11 using the circuit model in (a).

To understand the details of the high-speed behavior of the Ge PD, and to help the receiver circuit design, we have measured the device S11 and extracted a circuit model using curve fitting. The Ge PD circuit model shown in Figure 11(a) is very similar to the modulator circuit model shown in Figure 6(a), except that here we add one extra resistance R_J to model the junction leakage path. The parameter values in Figure 11(a) were extracted at 0.5V, and the curve fitting result is shown in Figure 11(b). The tested PD has a very small junction capacitance of 10 fF and a reasonable parasitic capacitance of 14.4 fF. A small PD capacitance is critical for the receiver circuit to achieve low power and high speed. The extracted circuit model also indicates that the Ge PD has a pretty large series resistance. But this is just a minor problem since it doesn't propose a limit to the device speed. Using the circuit model, we can estimate that the RC limited PD bandwidth is ~16 GHz in a 50Ω system. Clearly, some other effect (such as carrier transit time) is the primary limit to the device speed.

4.2 VLSI receiver circuit

The goal of the receiver circuit is to take the small photocurrent from a PD and output a correct (error-free) data symbol sequence. It generally consists of amplification and decision circuits. The amplification circuit may include a transimpedance amplifier (TIA) and a post amplifier in order to provide a transimpedance gain of ~10kΩ and an output voltage swing of a few 100mV; the decision circuit may include a clock-and-data-recovery (CDR) circuit and a slicer.

With a target receiver sensitivity of -15 dBm of modulated light with extinction ratio of >5dB, and a PD responsivity of >0.6A/W, we can expect a photocurrent swing of >20μA peak-to-peak from the Ge PD. We use a 3-stage amplification circuit to convert this 20μA current swing to a 200mV voltage swing. The 3-stage amplification circuit consists of a TIA

(converting current to voltage), a linear transconductance stage (converting voltage to current), and another TIA, as shown in Figure 12 [21]. Each stage burns about 0.3 mA of current. The total current is about 1 mA across all three stages (0.2 pJ/bit for a 5 Gbps data rate) and has a total bandwidth exceeding 4 GHz. Over all three stages the total gain is about 10kΩ, resulting in 200 mV of output for a minimum 20 μA input current. The simulated input-referred noise current is 0.8 μA, including the shot noise from the PD diode.

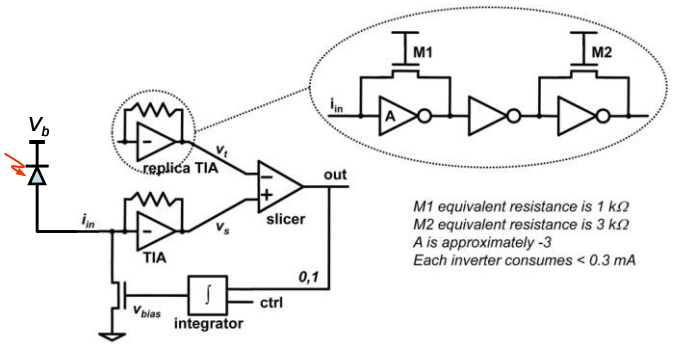


Figure 12. Schematics of the receiver design. The TIA uses n-transistors in triode as feedback resistors. The replica TIA sets the reference voltage of the slicer at Vdd/2, and the Vbias is controlled so that the TIA output is also centered around Vdd/2 [21].

The 3-stage TIA is followed by a slicer, for which we use a simple and power-efficient sense amplifier, as shown in Figure 12. The slicer needs a reference voltage to determine whether the data value is a "0" or a "1". In modern optical links, this reference voltage is obtained by long-running averaging of the TIA's output, which is approximately halfway voltage between a digital "0" and "1" if the data is DC-balanced. For this purpose, most modern optical links enforce DC balance through schemes such as 8B10B encoding [30]. However, such coding imposes a 25% overhead, which is equivalent to 25% lower data rate and 25% higher energy per bit. Furthermore, a traditional optical receiver has to include a power-hungry full CDR circuit in order to control the slicer timing.

Our receiver design solves these two issues using novel low-cost approaches by recognizing that our optical receiver is designed for inter-chip communications, which is in a small system. Every chip in the small system can share the same global clock source; clock phase differences between any two chips in the system are bounded and only changes slowly. Based on this concept, our VLSI chip with multiple receivers (and transmitters) was designed with a distributed global digital clock, with clock phase adjusted for each receiver. This avoids the full CDR circuit and significantly cuts the power consumption. To solve the reference voltage issue for the slicer, we use a replica TIA to fix the reference voltage at Vdd/2, as shown in Figure 12, and control the TIA input level to make its output centered around Vdd/2. The control of TIA input is done by shunting a transistor to the TIA input to "steal" away a proper amount of DC current, while the amount of "stolen" current is in turn controlled by the proper Vbias of the transistor. Due to signal and environmental variations, this Vbias needs to be adjusted periodically. Since it is a small system, we can use a synchronized global calibration session to periodically adjust Vbias for all the receivers in the system at the same time. This shared global sense of time makes such calibration feasible and inexpensive: it takes only ~100 cycles of calibration time in every 100,000 cycles of data transmission, which is a much lower overhead than the 8B10B encoding.

4.3 Receiver test results

After the Ge PD and the VLSI receiver circuit were flip-chip bonded together and attached to the PCB, we tested the high-speed performance of the receiver assembly. In the test, a pattern generator was used to generate 5Gb/s PRBS31 data which drives a reference lightwave transmitter. The transmitter output, with an extinction ratio of >10 dB and controlled by an optical attenuator, was coupled to the Ge PD through a lensed fiber and a grating coupler. An external clock from the pattern generator was fed to the chip to control the slicer timing. To send the high-speed signal off-chip, the digital output of the slicer was converted to CML signal by high-speed I/O pads. The high-speed output signal had to

travel through short wire-bonds, several inches of PCB traces, a few RF connectors and RF cables to reach the error detector on a BER tester. To achieve optimized performance, we manually adjusted the delay between the clock and the data at the slicer input. The Vbias at the TIA input, as shown in Figure 12, was automatically adjusted through a programmed calibration sequence.

With 18 μA of average photocurrent from the Ge PD (corresponding to an optical power of -16 dBm with detector responsivity of 0.7 A/W), we obtained error-free operation for more than 6 hours, indicating a BER better than 10^{-14}. The receiver BER was tested at different input power levels. Figure 13 shows the measurement results, indicating a receiver sensitivity of about -18.9 dBm for a BER of 10^{-12}. With 0.7 A/W detector responsivity, an -18.9 dBm optical signal would generate 9 μA of average photocurrent, which implies that the input referred noise of the TIA is ~1.3 μA for a Q of 7. The total receiver power consumption was measured as 3.45mW (excluded the power consumed by the digital data buffers and the on-chip clock distribution) by reading the power supply voltages and currents. This corresponds to 690 fJ/bit for the entire receiver including Ge PD and the CMOS receiver circuits, as well as any excess power required to drive circuit parasitics including internal wiring and flip-chip pads.

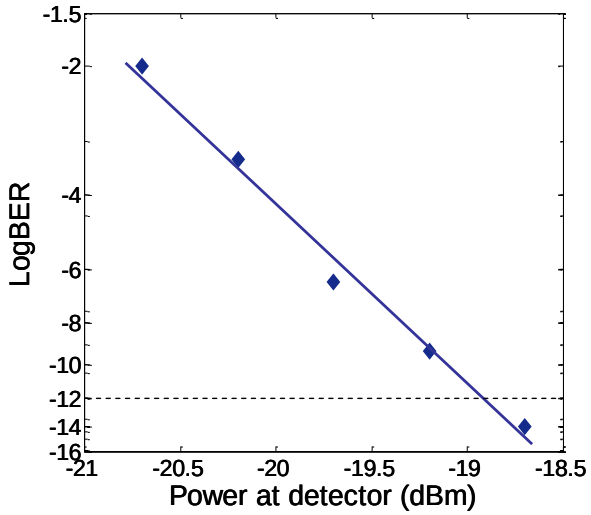


Figure 13. Receiver sensitivity measurement results tested with 5Gb/s PRBS31 data.

5. MUX/DEMUX DEVICES

Si photonic interconnects require silicon real estate for its signal routing using silicon waveguides. Although special design enables off-chip network routing [8], most of the networks require significant on-chip routing [9-11]. Wavelength division multiplexing (WDM) effectively reduces the number of interconnect waveguides, and consequently improve the integration density. Various approaches have been reported for wavelength channel multiplexing and de-multiplexing on silicon platform, including array waveguide grating (AWG) [31], Echelle grating [32], MZI based interleaver [33], as well as cascaded ring add/drop filters [34-35]. Multiplexers based on AWG, Echelle grating devices and interleavers are typically large in size, not desirable for area sensitive intra-chip applications. Ring resonator based add/drop filters using high index contrast Si waveguide, on the other hand, have the potential to make very compact Mux/Demux with desirable optical performance. Ultra-compact single ring add/drop filter has been demonstrated using 3um radius ring resonator [36-37]. High order ring resonators with multiple coupled rings were used to improve the pass band and channel isolation [38-39]. To make multi-channel multiplexers, multiple add/drop filters were cascaded using rings slightly different in size [40].

One critical hurdle for ring resonator based WDM filters to overcome before practical application is its center wavelength accuracy, as well as the accurate channel spacing for multi-channel multiplexers or demultiplexers. Due to manufacturing tolerances, the effective index of the Si waveguide varies due to silicon layer thickness variation on SOI substrate, waveguide width variation, etch depth variation for ridge waveguide structure, as well as the residual stress on

substrate. Each factor increases the required tuning range [8]. In addition, ambient temperature change also affects the waveguide effective index. All these variations cause significant wavelength shift for ring resonator based WDM filters. Although special fabrication techniques have been reported effective in achieving accurate channel spacing for multi-channel devices [41], tuning would still be required to align the filter center wavelengths with the pre-selected wavelength channels.

In this work, we used a 4-channel CMOS WDM multiplexer/demultiplexer using cascaded identical single ring resonators with integrated thermal tuner [42]. To have good channel isolation, we chose channel spacing of 200GHz (or 1.6nm) for the 1x4 multiplexer/demultiplexer. We also chose ring radius of 12μm to have big enough free spectral range (FSR) to accommodate 4 channels at 1.6nm spacing. Expecting low penalty for 10Gbps data transmission, we aim for drop filter with larger than 0.4nm 3dB bandwidth. The pass band is determined by the loaded Q of the device, which is dominated by the coupling coefficient *K*. We designed a gap of 325nm from the bus waveguide to ring to achieve a *K* of about 0.15. Four add/drop filters are cascaded on the same bus waveguide to form a 4 channel multiplexer/demultiplexer. Instead of using rings slightly different in size to achieve filters with different center wavelengths, we used identical rings with integrated thermal tuning. For better efficiency, we integrated the heater directly to the ring waveguide by doping part of the ring slab as doped resistor, as inset SEM picture shown in Fig. 4.

The 1x4 WDM multiplexer/demultiplexer with integrated thermal tuning was fabricated using Luxtera-Freescale 130nm SOI CMOS platform, as shown in Fig. 14. The total area of a single ring resonator including the tuning resistors was 26x40 microns. The pitch for the ring devices was 500μm. Grating couplers at 250μm pitch were used for optical I/O, coupling to fiber array. Injecting current to the doped resistors through the tuning pads, we can heat up the ring waveguide, and in turn change the index of the waveguides to shift the filter center wavelength until it's aligned with the target wavelength channel. The integrated thermal tuning resistor showed resistance about 200Ω. Fairly uniform and linear tuning response was obtained with efficiency of about 90 pm/mW.

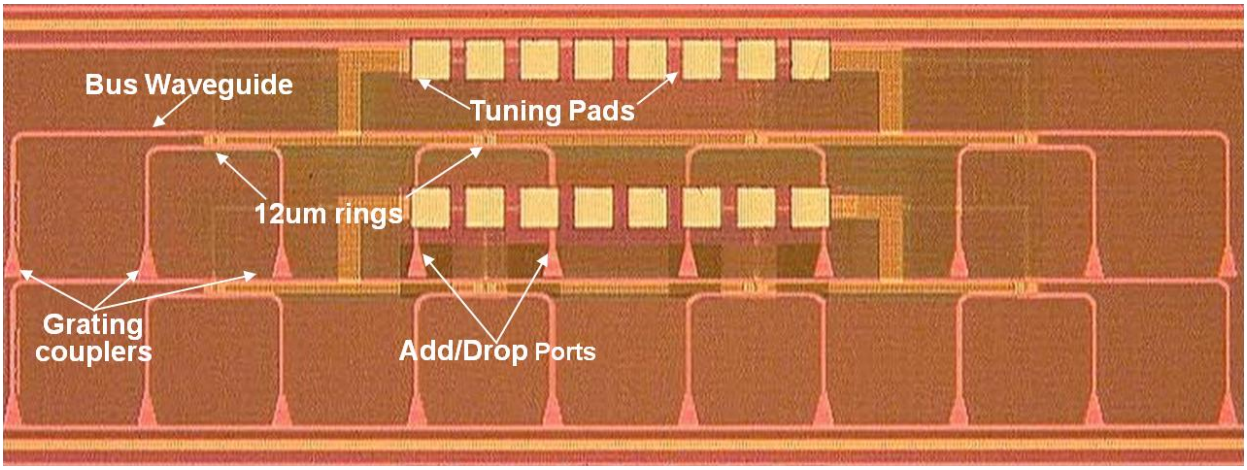


Figure 14. Fabricated 1x4 multiplexer/demultiplexer by cascading 4 ring add/drop filters with integrated doped resistor thermal tuner using FreeScale 130nm SOI CMOS process. Grating couplers are used for optical I/Os.

The measured 4-channel multiplexer/demultiplexer optical performance is shown in Fig. 15. The channels were tuned to ITU WDM grids at 200GHz spacing with center wavelength of 1554.13nm, 1555.75nm, 1557.36nm, and 1558.98nm respectively, by heating up the rings with different power through the doped silicon resistor. Uniform performances, larger than 0.4nm 3dB passband, less than 1dB insertion loss, and better than 16dB channel isolation, were achieved across all four channels. More test results can be found in [43].

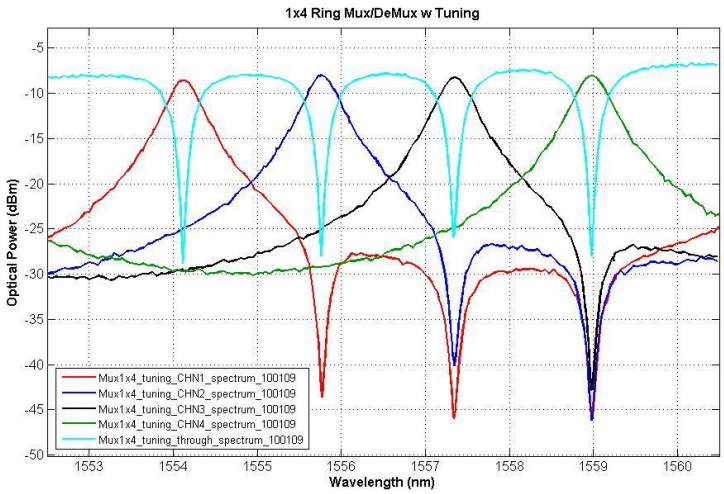


Figure 15 The through spectrum (light blue) and drop spectrums (red, blue, black, and green) of the CMOS 4-channel WDM multiplexer/demultiplexer thermally tuned to WDM ITU wavelength channels 1554.13nm, 1555.75nm, 1557.36nm, and 1558.98nm.

6. CONCLUSIONS

Si photonics is a promising solution for the inter-chip communication in a high-performance computing system with many processor chips, owing to its potential advantages of low latency, high bandwidth, high density, and low power consumption. In this paper, we have reported our flip-chip bonding technique with micro-solder bumps, which can integrate photonic and electronic devices with extremely low electrical parasitic. With this technique, we have demonstrated ultralow-power transmitter and receiver. The transmitter includes a small (15μm radius) Si microring modulator flip-chip bonded to a 90nm CMOS driver. With only 1.6mW power consumption, the integrated and digitally clocked transmitter operates at 5Gb/s with wide open eye and >7dB extinction ratio, and has achieved error-free transmission for over 2 hours without active resonator bias tuning or temperature control. We have shown that for the inter-chip communication in a small system, the receiver can avoid the power-hungry full CDR circuit and instead use a distributed global clock. Our receiver includes a Ge PD flip-chip bonded to a 90nm CMOS chip with amplification and decision circuits. With 3.45mW power consumption, the integrated receiver demonstrated -18.9dBm sensitivity at 5Gb/s for a BER of 10^{-12}. Finally we have reported a compact CMOS 1x4 tunable Mux/Demux using cascaded ring resonators with 12μm radius and with integrated thermal tuners. We measured an insertion loss of < 1dB, a channel isolation of >16dB for a channel spacing of 200GHz, and a uniform 3dB pass band >0.4nm across all 4 channels. We demonstrated accurate channel alignment to WDM ITU grid wavelengths using integrated silicon heaters with a tuning efficiency of 90 pm/mW. This is a significant step towards an optical link with a few 100fJ/bit power, which is critical for future high-performance computing systems.

ACKNOWLEDGEMENT

This material is based upon work supported, in part, by DARPA under Agreement No. HR0011-08-09-000. The authors thank Dr. Jag Shah of DARPA MTO for his inspiration and support of this program. The views, opinions, and/or findings contained in this article/presentation are those of the author/presenter and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

REFERENCES

[1] M. Horowitz, E. Alon, D. Patil, S. Naffziger, R. Kumar, and K. Bernstein, "Scaling, power, and the future of CMOS," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2005.

[2] B. Black, D. Nelson, C. Webb, and N. Samra, "3D Processing Technology and Its Impact on iA32 Microprocessors," In *Proceedings of the 2004 IEEE International Conference on Computer Design*, pp.316-318, Oct. 2004.

[3] A.W. Topol et al., "Three-dimensional integrated circuits," *IBM Journal of Research and Development*, no. 4-5, pp. 494–506, 2006.

[4] P.S. Andry, C. Tsang, E. Sprogis, C. Patel, S.L. Wright, B.C. Webb, L.P. Buchwalter, D. Manzer, R. Horton, R. Polastre and J. Knickerbocker, "A CMOS-compatible Process for Fabricating Electrical Through-vias in Silicon," In *Proceedings of the 2006 IEEE Electronic Components and Technology Conference*, pp831-837, 2006.

[5] R. Drost, R. Hopkins, R. Ho, and I. Sutherland, "Proximity communication," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9, pp. 1529-1536, Sept. 2004.

[6] R. Drost, C. Forrest, B. Guenin, R. Ho, A. V. Krishnamoorthy, D. Cohen, J. E. Cunningham, B. Tourancheau, A. Zingher, A. Chow, G. Lauterbach, and I. Sutherland, "Challenges in building a flat-bandwidth memory hierarchy for a large-scale computer with proximity communication," in *Proc. 13th Symp. High Perform. Interconnects*, Aug. 2005, pp. 13–22.

[7] R. Ho, K. W. Mai, and M. A. Horowitz, "The future of wires," *Proc. IEEE*, vol. 89, pp. 490–504, Apr. 2001.

[8] A.V. Krishnamoorthy, R. Ho, X. Zheng, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, and J.E. Cunningham, "Computer systems based on silicon photonic interconnects," *Proceedings of the IEEE*, 97 (7) 1337-1361 (2009).

[9] A. Shacham et al., "On the Design of a Photonic Network-on-Chip," *Proceedings of the First International Symposium on Networks-on-Chip* (NOCS), 2007.

[10] D. Vantrease et al., "Corona: System Implications of Emerging Nanophotonic Technology," *ISCA '08*, pp. 153-164, June 2008.

[11] C. Batten et al., "Building manycore processor-to-DRAM network with monolithic silicon photonics," *HOTI* 2008, pp.21-30, Aug. 2008.

[12] X. Zheng, J. Lexau, J. Bergey, J. Cunningham, R. Ho, R. Drost, and A. V. Krishnamoorthy, "Optical proximity communication using reflective mirrors," *Opt. Express*, vol. 16, no. 19, pp. 15 052–15 058, Sep. 2008.

[13] These companies include Luxtera, Lightwires, and Intel, etc. Please see their websites for more details.

[14] C. W. Holzwarth, J. S. Orcutt, H. Li, M. A. Popovic, V. Stojanovic, J. L. Hoyt, R. J. Ram, and H. I. Smith, "Localized substrate removal technique enabling strong-confinement microphotonics in bulk Si CMOS processes," in *Conference on Lasers and Electro-Optics* (Optical Society of America, 2008), paper CThKK5.

[15] S. Manipatruni, Q. Xu, B. Schmidt, J. Shakya, and M. Lipson, "High speed carrier injection 18Gb/s silicon micro-ring electro-optic modulator," in *20th Annual Meeting of the IEEE LEOS* (Institute of Electrical and Electronics Engineers, New York, 2007), pp.537-538.

[16] M.R. Watts, D.C. Trotter, R.W. Young, and A.L. Lentine, "Ultralow power silicom microdisk modulators and switches," in *IEEE Conference on Group IV Photonics* (Institute of Electrical and Electronics Engineers, New York, 2008), pp.4-6.

[17] J. Liu, M. Beals, A. Pomerene, S. Bernardis, R. Sun, J. Chen,L.C. Kimerling, and J. Michel, "Ultralow energy, Integrated GeSi electroabsorption modulators on SOI," in *IEEE Conference on Group IV Photonics* (Institute of Electrical and Electronics Engineers, New York, 2008), pp.10-12.

[18] J. E. Roth,O. Fidaner, R.K. Schaevitz, Y. Kuo, T.I. Kamins, J.S. Harris, and D.A.B. Miller, "Optical modulator on silicon employing germanium quantum wells," *Optics Express*, 15 (9), 5851-5859 (2007).

[19] Y. Kuo, H. Chen, and J. E. Bowers, "High speed hybrid silicon evanescent electroabsorption modulator," *Optics Express*, 16 (13), 9936-9941 (2008).

[20] P. Dong, S. Liao, D. Feng, H. Liang, D. Zheng, R. Shafiiha, C.-C. Kung, W. Qian, G. Li, X. Zheng, A. V. Krishnamoorthy, and M. Asghari, "Low Vpp, ultralow-energy, compact, high-speed silicon electro-optic modulator," *Optics Express*, 17, 22484 (2009).

[21] R. Ho, J. Lexau, F. Liu, D. Patil, R. Hopkins, E. Alon, N. Pinckney, P. Amberg, X. Zheng, J. Cunningham, and A. Krishnamoorthy, "Circuits for silicon photonics on a macrochip," in *IEEE Asian Solid-State Circuits Conference*, Nov. 2009, pp. 17–20.

[22] S. Palermo et al., "High-speed transmitters in 90nm CMOS for high-density optical interconnects," *ESSCIRC*, pp. 508-511, September 2006.

[23] T. Yin, R. Cohen, M. M. Morse, G. Sarid, Y. Chetrit, D. Rubin, and M. J. Paniccia, "31GHz Ge n-i-p waveguide photodetectors on Silicon-on-Insulator substrate," *Optics Express*, 15, 13965-13971 (2007).
[24] D. Ahn, C. Hong, J. Liu, W. Giziewicz, M. Beals, L. C., Kimerling, J. Michel, J. Chen and F. X. Kärtner, "High performance, waveguide integrated Ge photodetectors," *Optics Express*, 15, 3916–3921 (2007).
[25] Masini, G., Capellini, G.,Witzens, J. & Gunn, C., "A four-channel, 10 Gbps monolithic optical receiver in 130 nm CMOS with integrated Ge waveguide photodetectors," post deadline paper PDP31, in *Conference on Optical Fiber Communication* (2007).
[26] L. Chen, and M. Lipson, "Ultra-low capacitance and high speed germanium photodetectors on silicon," *Optics Express*, 17 (10) 7901-7906 (2009).
[27] L. Vivien, D. Marris-Morini, J. Mangeney, P. Crozat, E. Cassan, S. Laval, J-M. Fédéli, JF Damlencourt et Y.Lecunff, "42 GHz waveguide Germanium-on-silicon vertical PIN photodetector," *Proceedings of the 5th International Conference on Group IV Photonics*, 185-187 (2008)
[28] D. Suh, S. Kim, J. Joo, G. Kim, and I. G. Kim, "35 GHz Ge p-i-n photodetectors implemented using RPCVD," Proceedings of the 5th International Conference on Group IV Photonics, 191-193 (2008).
[29] D. Feng, S. Liao, P. Dong, N. Feng, H. Liang, D. Zheng, C. Kung, J. Fong, R. Shafiiha, J. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "High-speed Ge photodetector monolithically integrated with large cross-section silicon-on-insulator waveguide", to be published in *Appl. Phys. Lett.*, Dec. 28, 2009.
[30] A. Widmer and P. Franaszek, "A DC-balanced, partioned-block 8B/10B transmission code," *IBM J. Res. Develop.*, vol. 27, no. 5, pp. 441–451, Aug. 1983.
[31] K. Sasaki, A. Motegi, F. Ohno and T. Baba, "Si Photonic Wire AWG of 70 × 60 μm^2 Size ," *IEEE Conference on Lasers & Electro-Optics* (CLEO), pp.478-479, Aug. 2005.
[32] D. Feng, W. Qian, H. Liang, C. Kung, J. Fong, B. J. Luff, and M. Asghari, "Novel Fabrication Tolerant Flat-Top Demultiplexers Based on Etched Diffraction Gratings in SOI," *IEEE Conference on Group IV Photonics*, pp.386-388, Sept. 2008.
[33] J. Song, Q. Fang, S. H. Tao, M. B. Yu, G. Q. Lo, and D. L. Kwong, "Proposed silicon wire interleaver structure," *Optics Express*, 16 (11) 7849-7859 (2008).
[34] B. E. Little, S. T. Chu, H. A. Haus, J. Foresi, and J.-P. Laine, "Microring resonator channel dropping filters," *IEEE Journal of Lightwave Tech.*, vol. 15, no. 6, pp. 998-1005, June 1997.
[35] S. T. Chu, B. E. Little, W. Pan, T. Kaneko, S. Sato, and Y. Kokubun, "An Eight-Channel Add–Drop Filter Using Vertically Coupled Microring Resonators over a Cross Grid," *IEEE Photon. Tech. Lett.*, 11 (6) 691-193 (1999).
[36] B. E. Little, J. S. Foresi, G. Steinmeyer, E. R. Thoen, S. T. Chu, H. A. Haus, E. P. Ippen, L. C. Kimerling, and W. Greene, "Ultra-compact Si-SiO_2 microring resonator optical channel dropping filters," *IEEE. Photonics Tech. Letters*, vol. 10, no. 4, pp. 549-551, April 1998.
[37] M. R. Watts, W. A. Zortman, D. C. Trotter, G. N. Nielson, D. L. Luck, and R. W. Young, "Adiabatic Resonant Microrings (ARMs) with Directly Integrated Thermal Microphotonics," *IEEE LEOS Conference on Quantum electronics and Laser Science* (CLEO/QELS), pp.1-2, June 2009.
[38] B. E. Little, S. T. Chu, P. P. Absil, J. V. Hryniewicz, F. G. Johnson, F. Seiferth, D. Gill, V. Van, O. King, and M. Trakalo, "Very High-Order Microring Resonator Filters for WDM applications," *IEEE. Photonics Tech. Letters*, vol. 16, no. 10, pp. 2263-65, October 2004.
[39] M. A. Popovic, T. Barwicz, M. R. Watts, P. T. Rakich, L. Socci, E. P. Ippen, F. X. K. and H. I. Smith, "Multistage high-order microring-resonator filters with relaxed tolerances for high through-port extinction," *IEEE Conference on Lasers & Electro-Optics* (CLEO), pp266-268, Aug. 2005.
[40] S. Xiao, M. H. Khan, H. Shen and M. Qi, "Multiple-channel silicon micro-resonator based filters for WDM applications," *Optics Express*, 15 (12) 7489-7498 (2007).
[41] C. W. Holzwarth, T. Barwicz, M. A. Popović, P. T. Rakich, E. P. Ippen, F. X. Kärtner, and Henry I. Smith, "Accurate resonant frequency spacing of microring filters without post fabrication trimming," *J. Vac. Sci. Technol.* B 24 (6) 3244-3247 (2006).
[42] O. Schwelb, "Transmission, Group Delay, and Dispersion in Single-Ring Optical Resonators and Add/Drop Filters—A Tutorial Overview," *Journal of Lightwave Technology*, 22 (5) 1380-1394 (2004).
[43] X. Zheng, I. Shubin, G. Li, T. Pinguet, A. Mekis, J. Yao, H. Thacker, Y. Luo, J. Costa, K. Raj, J. E. Cunningham, and A. V. Krishnamoorthy, "A tunable 1x4 silicon CMOS photonic wavelength multiplexer/demultiplexer for dense optical interconnects," submitted to *Optics Express* for publication.