Design of a 10 GHz Silicon Modulator Based on a 0.25 μm CMOS Process - A Silicon Photonic Approach

D.W. Zheng*, D.Z. Feng, G. Gutierrez, and T. Smith
Kotura, Inc, 2630 Corporate Place, Monterey Park, CA 91754
* dZheng@kotura.com, www.kotura.com

ABSTRACT

We report the design of a 10 GHz non-return-to-zero (NRZ) silicon modulator based upon 0.25-μm CMOS/BiCMOS processes. The basic optical component is a ridge waveguide slightly-doped with P and N impurities, which forms a reverse-biased P/N junction. The diode typically operates between reverse and zero biases, so as to change the number of free carriers overlapping with the optical mode and consequently modulate the phase of the light. This type of phase shifters form the arms of a push-pull Mach-Zehnder interferometer to realize amplitude modulation.

Keywords: Waveguide modulators, Photonic integrated circuits, Si photonics

1. INTRODUCTION

As Moore's law drives the transistor density and chip size larger, it is widely accepted that interconnect will be the bottle-neck for computers. As the clock frequency goes up, copper interconnect based system is facing augmented interconnect latency, signal loss, and cross-talk coming from electromagnetic interference. It is broadly believed that optical interconnect will come into play to replace copper at certain stage, which will start from the board level BUS signal and eventually get into the CPU itself. However, as for when this will happen there is still a split view, and it is generally estimated to happen between 2008 and 2012 [1].

Si photonics has been thought as a natural choice for intra-chip optical interconnect because it shares the same material and fabrication facilities main stream electronics – Complementary-Metal-Oxide-Semiconductor (CMOS) or Bipolar devices use. In recognition of the enormous size of the potential market and its impact on revolutionizing computing technology, optical interconnect based upon Si photonics becomes a glamorous subject to study, reflected by the large amount of publications in the past year or so [1-7], in particular there were several papers made to "The letters to Nature" [2-5]. Nevertheless, considering on-chip or inter-chip optical interconnect will most likely not go into production within 5 yrs or so, there have not been many commercial companies working in this field. Any revolutionary technology needs to find its appropriate entry point into the market to serve the society, on the other hand, enough commercialization is needed to generate the incentive for investment into the research and development of the technology. Now it comes the opportunities from both market and technology perspectives. On the market side, the emerging 10 Gbps transceiver market for datacom applications might open the door for optical to replace electrical interconnect at short reach. Luxtera, Inc [Carlsbad, CA] web site has a very good white paper [6], which conducted a detailed comparison of the electrical and optical interconnect at 10 Gbps node. The conclusion is quite convincing that the balance leans towards the optical interconnect solution due to a smaller total cost of ownership. On the technology side, great progress has been made on 10 Gbps Si modulators, which is an important ingredient of Si transceivers [7]. This marriage of the right market with the progress in Si photonics technology would probably change the forecast picture and bring Si photonics based optical transceivers/optical connectors into the market much earlier than people expected earlier.

As for Si modulators demonstrated larger than 1 Gbps transmission, there are mainly two types published in the public domain: MOS based [7] and ring resonator based [4]. As a Si photonics company, when Kotura started evaluating different technology approaches, we analyzed from the following aspects: 1) device performance, 2) Process tolerance and sensitivity, 3) upgradability (showing a clear path to higher performance in the future, 4) cost of the final product, 5) cost of development. Intel team first demonstrated the Gigahertz modulation [3] then pushed it to 10 Gbps transmission all based upon a MOS phase shifter[7]. These pioneering works are really encouraging and the idea of a split waveguide (there was a thin gate oxide residing inside a ridge waveguide) overlapping with a Metal-Oxide-Semiconductor (MOS)

Silicon Photonics, edited by Joel A. Kubby, Graham T. Reed, Proc. of SPIE Vol. 6125,
61250E, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.644942

transistor concept really depart away from all traditional modulator concepts. Ref [3] pointed out two ways to improve the device efficiency: reducing the waveguide dimension and gate oxide thickness. From the equivalent circuit point of view, the electrical signal drives a MOS capacitor plus a series resistance. As the MOS channel length shrinks towards 65 nm and below, gate dielectric thickness drops, capacitance per unit area from the gate oxide actually increases, considering the shrinkage of waveguide lateral dimension, the capacitance drops relatively slower. The capacitance per millimeter for the improved structure reported in Ref. [7] is still fairly large 7.65 pF/mm. The biggest issue of this is that it could dramatically slow down the electromagnetic signal on the metal signal lines, which makes traveling wave design very difficult as you can realize later.

Ring resonator reported in Ref. [4] demonstrated 1.5 Gbps RZ transmission. There are many concerns with this approach. A) this kind of structure needs a tight control of the gap between the ring and the straight waveguide, this is not straightforward even with the current 130nm equipment. The process tolerance is small. B) The stress in the waveguide caused by the cladding layers used for electrical isolation and passivation purpose is difficult to control. C) Temperature could change the effective index of Si, and shift the resonate wavelength of the ring. For low cost optical transceivers, it is preferred to avoid temperature stabilization, which usually involves an expensive thermal electric cooler (TEC) to compensate the change of chip temperature. In this particular case, it mainly comes from the environment temperature change, because this chip consumes very little amount of power. D) This structure utilizes a forward-biased p-*I*-n diode, whose rise and fall time is only about 10 ns. It took advantage of the fact that optical response completes much earlier before electrical signal reaches its steady state. To avoid light touching the highly-doped p and n region, the distance between p^+ and n^+ region has to be more than 2 μm. The speed of carrier injection and extraction limits the speed of optical response intrinsically. Carrier lifetime limits the extraction speed, the distance between the p^+ and n^+ sets the carrier injection time given a fixed drift velocity. It is not easy to see a clear path to the higher speed of 40 or 100 Gbps. However, if some innovative approaches could be figured out to overcome afore-mentioned issues, then ring resonator could be a great candidate for short range communications due to its extremely compact size and therefore low chip cost.

After conducting these analysis, we decided to take a different route to implement the amplitude modulation. Since free carrier absorption seemed to be the only effect in pure silicon, which was practical enough to change both the real and imaginary part of the index [8], we also went along this direction. As well known that speed of a *p-n* diode working at forward bias is limited by the carrier lifetime, so instead of injecting carriers into the waveguide region, we figured out we could put them there at no-bias condition and remove them using an extracting electrical field. By doing this, we formed a p/n junction inside a ridge waveguide, which works between zero and a reverse bias. It is well known that a diode working at reverse bias is only limited by the carrier mobility, therefore could be potentially very fast.

In this paper, we will demonstrate the basic design concept of a Mach Zehnder (MZ) amplitude modulator working in push-pull mode. Each MZ arm consists of a phase shifter where a revere-biased diode resides inside a ridge waveguide. We will demonstrate that 10 Gbps non-return-to-zero (NRZ) transmission could be achieved. More importantly, we will show that there is a clear path to 40 Gbps or even higher modulation by shrinking the waveguide dimensions.

2. CONSIDERATION ON THE DEVICE AND PROCESS INTEGRATION

As pointed out in the introduction, at 10 Gbps transmission, optical is favored in comparison to electrical signals over shielded cables [6]. However, on the optical side, there are still a variety of technologies to compete with Si photonics technology. For example, GaAs or InP directly modulated laser transceivers. Therefore, Si photonics has to bring extra value to stay competitive. One natural thought would be to integrate driver electronics and other control (logic) electronics onto the same chip where optical modulator is built on. Therefore, process and device integration has to be considered at the very beginning. For low cost Si electronics, either CMOS or BiCMOS process has to be adopted. Cost will be minimized if minimal extra process is introduced into the wafer flow.

Silicon-on-Insulator (SOI) (100) wafer was selected to be the starting substrate. The buried oxide was selected to be 1 μm for optical confinement reason. Doping levels in the SOI and substrate layers on the starting wafer were selected to be p-type 200 Ω·cm or larger, which would not affect the doping profile of either CMOS or bipolar transistors.

In a typical modern CMOS process, there is only one Si etch step, which is the shallow trench isolation (STI), the depth of this trench is related to the technology node, and generally gets shallower as the CMOS gate length shrinks. Ridge waveguide was chosen to guide light as all other research groups, because of the small difference in effective index between the two polarizations. If STI etch is used to form the ridge waveguide, then no extra Si etch is needed. In principle, it is all right to do a separate Si etch in depth different from the STI etch depth, however, it could cause either trench filling problem if it is much deeper than the STI etch depth, or Chemical Mechanical Polishing (CMP) problem if it is too shallower than the STI etch. There are in general very small amount of Si consumption in a typical CMOS or BiCMOS process, only gate oxidation and oxidation after shallow trench etch consumes a little. Therefore, the starting SOI thickness and the depth of STI etch defines the geometry of the ridge waveguide.

We designed our Si modulator using a 0.25-μm BiCMOS process from a commercial 8-inch SOI foundry. The cutoff frequency (f_T)of the SiGe hetero-junction bipolar transistor (HBT) is 120 GHz for NPN transistors and it could provide more than 300 mA of current using push-pull emitter-coupled logic (ECL). The reason to select BiCMOS over CMOS comes from the current driven requirement of our device. As will be clarified later, we could also choose 0.18-μm BiCMOS or 0.13-μm CMOS, which is actually simpler from design perspective and perform better. However, the cost per custom run at the foundry is almost 3 times more at 0.13 μm node. As a first step of proof of concept, we adopted 0.25-μm BiCMOS. To avoid infringing the intellectual property of our foundry and not losing generality, we present here a sets of parameters slightly different from the actual values used in the foundry, but they are very close. The thickness of the starting SOI layer was 1.5μm, the depth of the STI etch is 0.6 μm, the width of the waveguide is the same as its height.

3. DEVICE PHYSICS

The shape of the optical mode was modeled using a home-written software, and its TE mode is displayed as Fig.1. The effective index of TE mode was 3.45, optical group index was calculated to be 3.6, all at 1550 nm. Highly conductive poly-silicon residing on top of the ridge moves actual Tungsten vias away from the optical mode to reduce the loss.

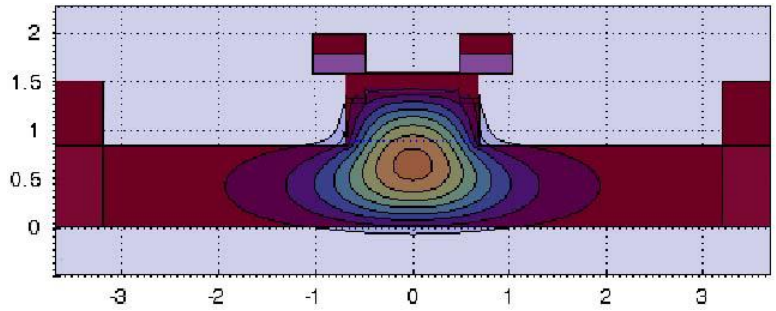

Figure 1. Mode shape of the waveguide

A cross-section of the waveguide with doping profile and electrical contacts is shown as Fig. 2. There are mainly 3 regions. The top of the ridge is doped to about $5.5x10^{16}$ cm^{-3} N-type, the bottom of the ridge is doped to about $3.5x10^{16}$ cm^{-3} P-type. Metallurgical junction between the P and N type dopants is 0.6 μm from the ridge top. P^{++} ($> 5x10^{18}$ cm^{-3}) doping locates about 4 μm away from the P-type dopant and provides a low resistance path to the anode metal contact.

A series of electrical simulations of this device structure were done using ATLAS simulation tool (Silvaco data systems, Santa Clara, CA) including DC, transient and AC small signal analysis. Shapes of the depletion region under 0 and –4 V static bias are exhibited as Fig.3 A and B respectively. Since the phase change caused by electron and hole concentration change is almost linear based upon experimental observations [9] and the classical Drude's model, logarithm of electron plus hole concentration contours are plotted in both figures. It is expected to see that depletion region width increased with reverse bias. It is obvious that light traveling inside the waveguide shall experience a phase shift if the static bias swings from 0 to –4V, which can be calculated based on the overlapping of optical mode with electron and hole concentration distribution, and it is illustrated in Fig.4. The major challenge here is that usually the mesh of the device simulation tool and optical tool needs to be aligned so that accuracy is not lost during the process of interpolation.

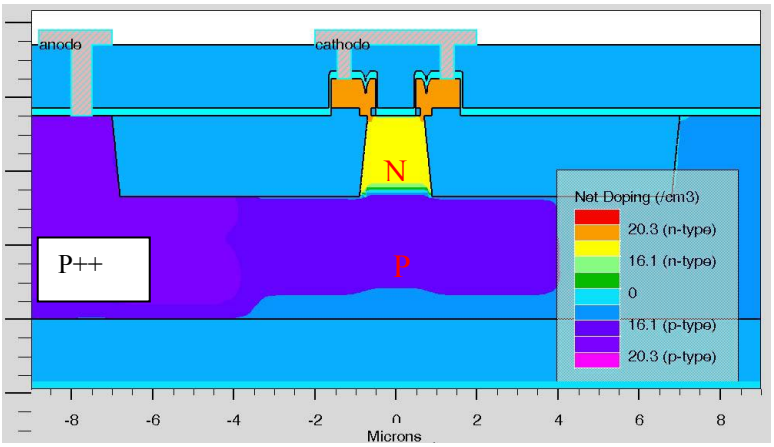

Fig.2 A cross-sectional schematic showing the doping profile and metal contacts.

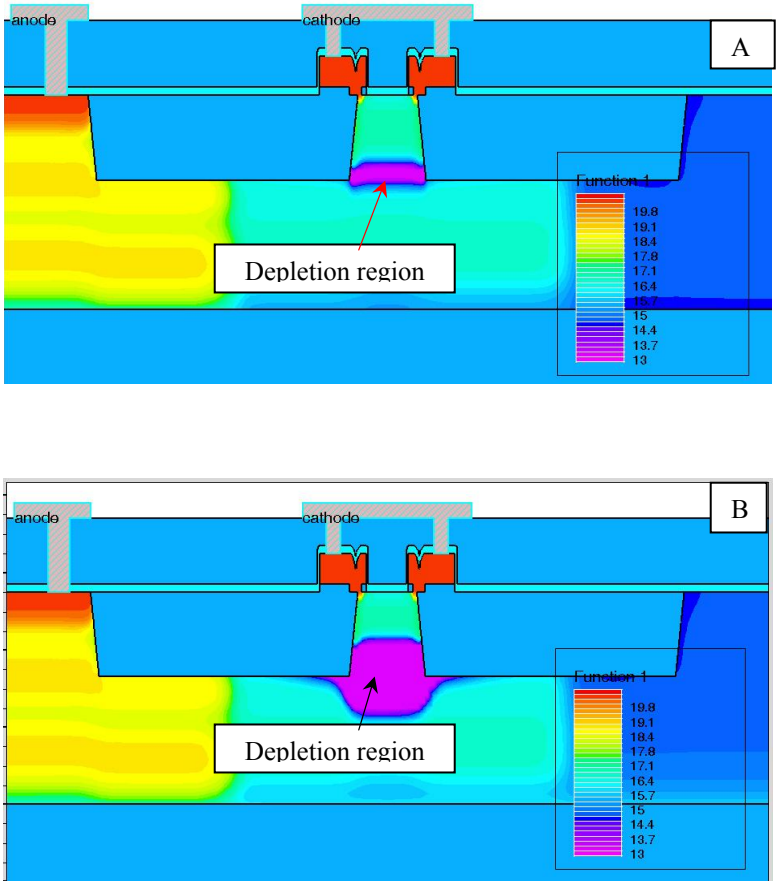

Fig.3 Contour of Log10(electron + hole concentration) showing the depletion region at A) 0 V bias, B) –4v bias.

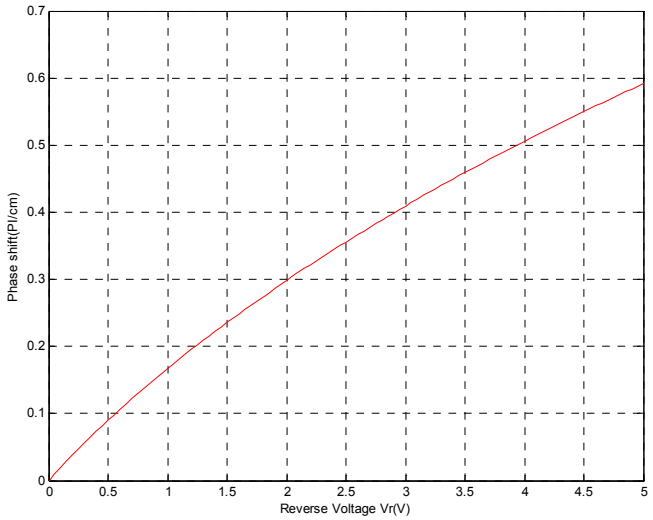

Fig.4 Phase shift per-centimeter-long waveguide versus the reverse bias voltage of the diode.

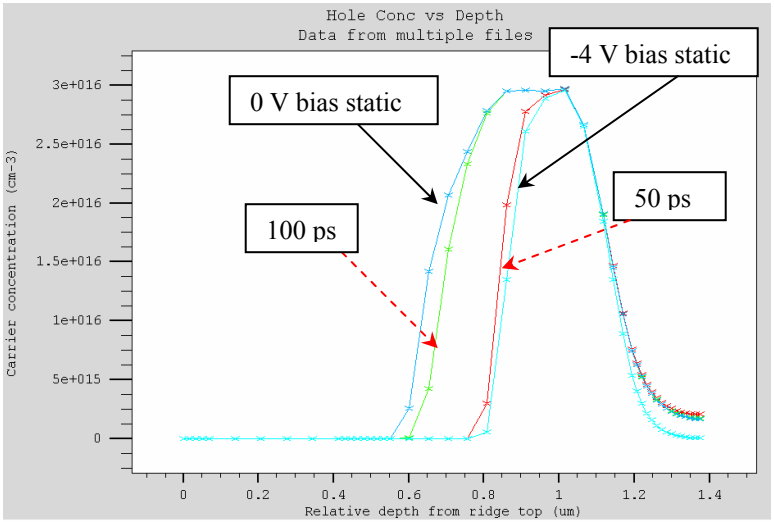

Fig.5 Hole concentration along the vertical bi-sectional line of the waveguide.

Fig.4 shows that under steady state, a centimeter long waveguide could introduce roughly 0.5 π phase change, for a push-pull MZ architecture, this shall be sufficient to reach π phase difference between two MZ arms. However, during high speed signal transmission, when a switch from logic "0" to "1" and back to "0" is needed, transient analysis has to be done to see whether carriers could move fast enough in response to the voltage change at the terminals. Anode was assumed to be grounded and only the cathode voltage was swept. This assumption did not reflect the situation when the electrodes was used as transmission lines carrying traveling waves. Anode along the transmission line may not be the true ground, and the transmission line was only grounded to the ground signal of the package at both ends but no elsewhere along the line. To exam the carrier response, a transient signal was applied to the cathode, voltage ramped up in 8 ps to 4 V, hold until 50ps, and ramp down to 0 V in 8 ps, the period was 100 ps. Fig. 5 plots the hole concentration along x=0, which is the vertical bi-sectional line of the waveguide. As labeled in the plot, two out-bound curves correspond to the static 0 and 4V hole concentration profiles, two in-bound curves are the values at 50 ps and 100 ps

transient time. We can see that the dynamic response is slightly different from the static carrier distribution. So during high speed transmission when the carriers move in and out of the waveguide, induced phase is slightly smaller than what is given by Fig.4 if a simple "0-1-0-1…" logic signal is applied at the terminal. To analyze its effect on the signal transmission eye diagram is too involved to be included in this context. As expected, similar phenomena could be found for electrons, which are found at the top of the ridge waveguide.

In order to investigate the signal degradation as it travels along the electrode, an equivalent circuit model has to be setup for the reverse-biased diode. For simplification, the diode was modeled as a capacitor in series with a resistor. The latter reflects the loss mechanism associated with the *p* and *n* doped area outside the deletion region, the former refers to the junction capacitance associated with the depletion region capacitance under an external bias. Simple estimation leads to the diode resistance value of 2.0 Ω/cm, and the capacitance value of 8.28 pF/cm at zero bias. Obviously this capacitance drops with increased reverse bias, and their relationship could be well characterized by the classical derivation of [13]:

$$C_d(V) = C_j(0) \cdot (1 - V/\phi_0)^{-m} \qquad (1)$$

Where $C_d(V)$ is the diode capacitance at an external bias of voltage V, ϕ_0 is the diode built-in potential, $C_j(0)$ is the zero bias junction capacitance. *m* is an exponential fitting parameter, and it is 0.5 for an abrupt junction.

4. LAYOUT AND THE TRAVELING WAVE DESIGN

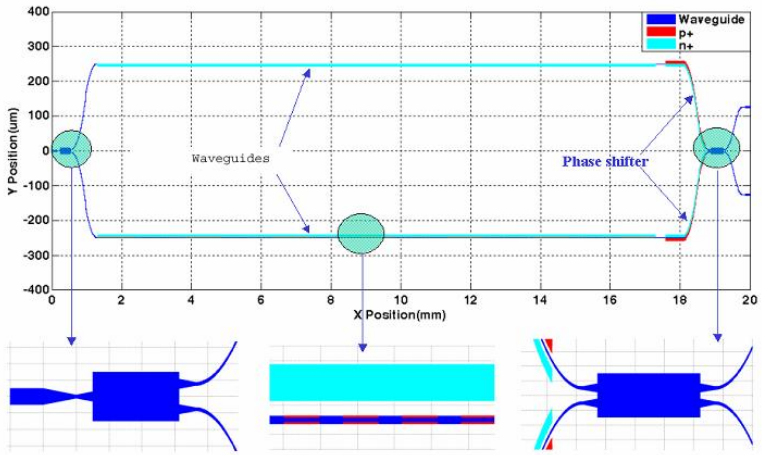

Fig.6 Layout of the MZ interferometer based amplitude traveling wave modulator.

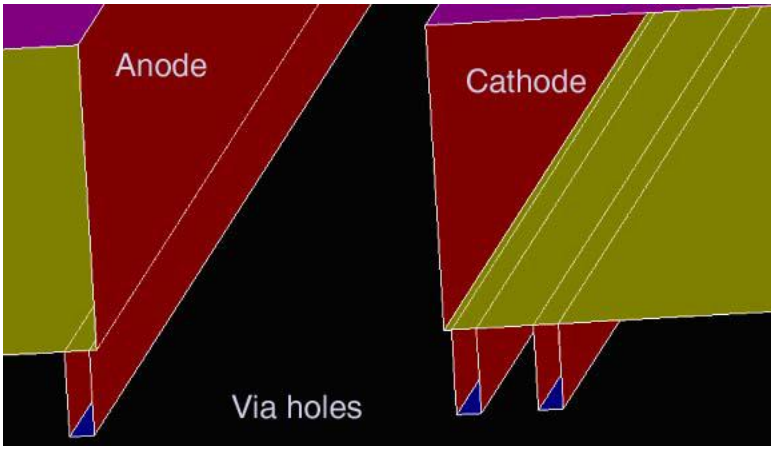

Fig. 7 3D view of the metal transmission line.

Layout of the proposed MZ based amplitude is displayed as Fig.6. Each arm has a phase shifter based upon the basic device structure described in the last section. Near the output, a small section of the forward-biased *p-I-n* diode is used to adjust small phase deviation between the two arms due to fabrication. Dimensions used in the plot as well as the overall layout configuration is mainly for illustration purpose and does not represent the most optimized configuration.

Traveling wave (TW) design has been widely used for both electro-optic or electro-absorption modulators [10-12]. In TW design, the electrodes (in our case anode and cathode) are treated as transmission lines, the group velocity of electrical bias signal and the group velocity of optical signal travels are at the same speed. [10] In this kind of design, electrode capacitance as well as the capacitance of the diode is distributed and does not limit the modulator speed. In order to have a feeling of the problem, different transmission line geometries were simulated using IE3d software from Zeland software (Fremont, CA). The total metal thickness was 6 μm, all interlayer dielectrics (ILD) and passivation dielectrics were included in the simulation. A 3-dimensional (3D) view of the metal structure is presented as Fig. 7, where the contribution from the vias were also included. Two hundred micron long and 1-mm long transmission lines were modeled to obtain the electromagnetic wave effective index and loss parameters. Two scenarios are listed in Table 1 and 2, where the gap of the two electrodes were varied. Electrical parameters have frequency dependence. Listed value is for 5 GHz. Re(Z_c) stands for the real part of the transmission line characteristics impedance.

Table 1. Transmission line effective index and loss information

Case No.	Anode-cathode gap (μm)	Anode width (μm)	Cathode width (μm)	Re (Z_c) * (Ω/mm)	Effective index (N_{eff})	Propagation loss (dB/mm)
A	7.0	100	50	43.1	2.344	0.147
B	20.0	100	50	61.6	2.365	0.132

Table 2. RLCG equivalent circuit parameters

Case No.	R@ 0.1 GHz (Ω/mm)	R@ 5GHz (Ω/mm)	L (nH/mm)	C (pF/mm)	G(Siemens/mm)
A	0.384	0.562	0.337	0.181	4.82×10^{-4}
B	0.102	0.355	0.487	0.128	4.00×10^{-4}

It can be seen that the effective index of the unloaded transmission line is smaller than that of the optical wave, therefore traveling wave design is feasible.

According to Ref.[11], simple equivalent circuit model could be used to conduct speed matching between electrical and optical waves. If the transmission line of case A is connect to a 1 mm long diode, then the effective index of the electrical signal travels along the transmission line N_L could be estimated as:

$$N_L = c \cdot \sqrt{L \cdot (C_t + C_d)} \qquad (2)$$

where *C* is the speed of light in vacuum. L is the transmission line inductance per unit length. C_t is the transmission line capacitance per unit length, and C_d is the diode capacitance per unit length. Here we use the 2 V reverse bias value of C_d (0.576 pF/mm) for illustration purpose because it is half of the voltage swing. It is noteworthy to point out that this is a non-linear circuit due to voltage dependent capacitance values. N_L could be easily calculated to be 4.79, much higher than the optical group index. Therefore a segmented design is needed[11].

Unloaded (not connected to diode) transmission line of length L1 is put adjacent to a loaded transmission line (connected to the diode) of L2 in length as illustrated in Fig.6. To match electrical and optical waves in the time domain, the following equation has to be satisfied:

$$\frac{L_1}{(c/N_{ul})}+\frac{L_2}{(c/N_L)}=\frac{(L_1+L_2)}{(c/N_{op})} \quad (3)$$

Where N_{ul} is the group index of the unloaded transmission line, N_{op} is the optical group index, and c the speed of light in vacuum.

To reduce the reflected electrical wave at the end of the transmission line, a terminating resistor equivalent to the real part of the loaded transmission line is used, a value of 25 Ω was adopted. Performance of the modulator was evaluated using a 64-bit pseudo random digital data string, per bit length was 100 ps in the time domain. The eye diagrams of the RF signal at the input and output terminal, and the optical signal at the detector was plotted as Fig.8.

5. DISCUSSION AND THE PATH TO THE NEXT GENERATION DEVICE

Total capacitance of the loaded transmission line of this design is 10.1 pF/cm for a zero bias diode including the transmission line, which represents the highest capacitive load the driver sees. If we recall the capacitance per cm for the Intel device discussed in the introductory part, C=76.5 pF/cm, very long waveguide has to be employed to achieve π/2 phase shift using a traveling wave design.

Radio frequency (RF) simulation showed that the resistance per millimeter at 0.1 GHz is much smaller for the large gap transmission line (case B) than that of the small gap. The DC resistivity for the Al used as the electrode in this design is 3.0 μΩ·cm, simple calculation would shall that at 0.1 GHz skin effect is actually negligible comparing to the metal thickness, but the calculated resistance value seems to be higher than the number estimated using simple ohm's law. We speculate that extra loss is caused by the electrical current seeing the strong magnetic field when the two electrodes are put really close. This extra loss mechanism is balanced by a smaller inductance value, which governs a smaller delay or less effective index at the loaded transmission part. Equivalently, we reduced the total length required. The penalty for this choice is a lower line impedance, which leads to a higher DC power consumption.

It shall be clear at this stage that shrinking the waveguide dimension proportionally shall improve the device performance, and the advantages are manifold. First, the overlap between the optical mode and the depletion region will be improved; Second, the capacitance of the reverse-biased junction will be dropped, which means the slow-down of the electrical drive signal at loaded section will be reduced, and accordingly the length of the unloaded section could be abridged. This shall lead to a shorter device. Alternatively, since the capacitance is reduced, we could move the metal electrodes further apart and increase the mutual inductance, which will lead to a higher equivalent line impedance. This shall lessen the DC power consumption of the device. In a traveling wave design, the modulator speed is primarily limited by the intrinsic speed of the diode, which is how fast we could move the carriers in and out of the diode. As device gets smaller, a faster modulator is expected.

Certainly we like to point out that the afore-mentioned shrinkage has a limitation. After the core size is trimmed down to a certain number (0.5 μm is a good reference number), optical mode start to leak more to the cladding dielectrics around. To avoid introducing extra optical loss, within certain radius from the depletion region, no high level doping could be introduced, this will limit the intrinsic speed of the device.

7. CONCLUSION

We illustrated the design concept of a 10 Gbps optical modulator based upon SOI ridge waveguide interacting with a reverse-biased *p-n* diode. Various aspects of the design including process and driver circuit integration, device physics, traveling wave and transmission line design were discussed. The main purpose of this article is to stimulate some discussion among the technology community so that this design concept could be validated against experiments. Kotura

also welcome joint development proposals leading to the higher performance Si photonics optical interconnect technology.

ACKNOWLEDGMENT

Thanks are given to Drs. G. Rakuljic and A. Keywitsch for stimulating discussions during the incubation stage of this project.

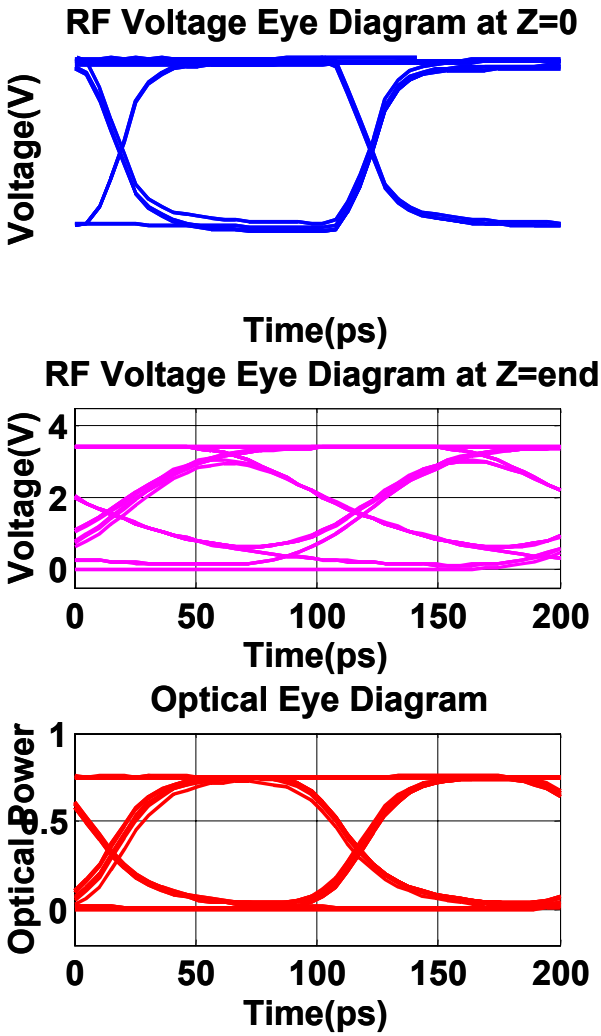

Fig. 8 Electrical and optical eye diagram showing the performance of the modulator, 64-bit pseudo random signal input.

REFERENCES

1 J. Bautista, "The Potential Benefits of Photonics to the Compute Platform", Integrated Photonics Research & Applications (San Diego, CA 2005).

2 H. Rong, R. Jones, A. Liu, O. Cohen, D. Hak, A. Fang, and M. Paniccia, "A continuous-wave Raman silicon laser", Nature 433, 725(2005).

3 A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia, "A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor", Nature 427, 615(2004).
4 Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator", Nature 325 (2005).
5 Y.-H. Kuo, Y.K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D.A.B. Miller, and J.S. Harris, "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon", Nature 437, 1334(2005).
6 www.luxtera.com/technology_wp.htm
7 L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U.D. Keil, and T. Franck, "High speed silicon Mach-Zehnder modulator", Optics Express 13, 3129(2005).
8 R.A. Soref and B.R. Bennett, "Electrooptical effects in silicon,", IEEE J. Quantum Electron. 23, 123(1987).
9 R.A. Soref, B.R. Bennett, "Kramers-Kronig analysis of electro-optical switching in silicon", Proc. SPIE 704, 32(1986).
10 R. Spickermann, S.R. Sakamoto, and N. Dagli, "In traveling wave modulators which velocity to match",
11 G.L. Li, T.G. B. Mason, P.K. L. Yu, "Analysis of Segmented Traveling-Wave Optical Modulators", J. Lightwave Tech. 22, 1789(2004).
12 G.L. Li, c.K. Sun, S.A. Pappert, W.X. Chen, and P.K.L. Yu, "Ultrahigh-speed Traveling-Wave Electroabsorption Modulators-Design and Analysis", IEEE trans Microwave Theory and Techniques 47, 1177(1999).
13 G. Massobrio and P. Antognetti, Semiconductor Device Modeling with Spice, (McGraw-Hill, Inc., 1993), Cap.1.