CMOS Enabled Silicon Photonics for Data Center Packet Switching

L. Chen, Y. Tang, J. E. Bowers, L. Theogarajan
University of California, Santa Barbara

***Abstract*— A low-power high data rate optical packet switch integrating hybrid silicon Mach-Zehnder Interferometer (MZI) based optical switch and 0.13μm CMOS IC for data center switching application is described. A novel truly differential transimpedance amplifier (TIA), with high-bandwidth power supply rejection ratio (PSRR), that serves as the front-end for a 2Gbps packet-header receiver is presented. The TIA achieves 55.7dBΩ and 1.97GHz f_{3dB} while consuming 3.3mW from 1.2V supply excluding 50Ω output driver. We also demonstrate a low-power optical switch driver enabled by a compact supply-regulated driver that features variable output voltage and power-boost for fast rise times. The CMOS driver chip is wirebonded to photonic IC, achieving$\leq$5ns switching time while consuming $250\mu W$.**

***Index Terms*— Silicon photonics, data warehouses, data center, truly differential transimpedance, capacitive switch driver.**

I. Introduction

The growth in data centers traffic today is demanding more energy efficient way of computing and moving data across servers. Recent advances in silicon photonics open up exciting opportunities to solve the bandwidth limitation both on-chip and off-chip wire interconnects. On-chip photonic interconnects have been proposed to link multi-core systems to create a single logical compute node with terascale processing capability, in an energy efficient manner [1]. Within the data center, packet switching network is the preferred way of connecting compute servers because of low latency and high network utilization. Optical packet switching becomes an attractive solution as the demand for bandwidth increases. Traditional electronic network switches approach power consumption limits as data-rates continue to scale to higher. While optoelectronics and MEMS optical network switches have been demonstrated before as solutions, the high costs and power consumption are barriers to mass commercialization. The key to overcoming this barrier is to integrate both active and passive photonic devices with CMOS and take advantage of the economy of scale of mature silicon manufacturing processes. Thus lowering the cost while simultaneously enhancing the reliability of optoelectronic systems. In this paper, we present a unique optical packet switch based on ultra low power hybrid silicon MZI switching elements. We describe a novel truly differential TIA used for the header packet receiver frontend, and a low-power adjustable output, capacitive switch driver circuit that includes a low dropout voltage regulator. Finally we demonstrate the operation of a MZI optical switch driven with the switch driver chip, validating the operation of the switch. The rest of this paper is organized as follow. In section II, we briefly introduce the optical packet switching system, in section III, we describe in detail the circuit implementation. Measurement results are presented in section IV and followed by conclusion in section V.

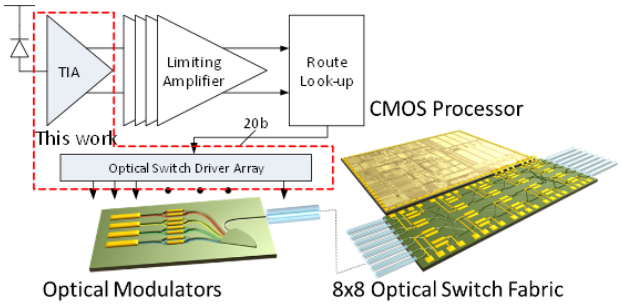


Fig. 1. Envisioned Optical packet switch fabric system. (Fig. Courtesy of M. Heck)

II. Optical Packet Switch

Figure 1 shows a conceptual rendering of our proposed 8x8 optical packet switch. On the transmitter end, a hybrid silicon modulator converts electrical signal into optical packets. Data can be encoded in different wavelengths to increase the bandwidth of the system. Recently, a 50Gbps Wavelength multiplexed (WDM) system has been demonstrated [2]. The switch operates as follows, the encoded packets are sent to switch where the CMOS processor extracts the header information and electronically configures the switch fabric to allow the data packet to reach its intended destination. Since the header information can be encoded at a much lower data rate, due to the packet-switching nature, we chose to encode it at 2.5Gbps. The switch fabric will be built monolithically using 2x2 switching elements, arranged in a Benes network configuration, with all the necessary waveguides integrated on the same chip. The 2x2 MZI switching elements operate in carrier depletion mode and consumes $\approx$100 μW each [3]. A switch fabric with 400Gbps bisection bandwidth could consume as little as 2mW, which results in a power efficiency of 5 fJ/bit. The switch element can be modeled as a reversed bias diode, which presents a capacitive load of 1pF to the driver. For low latency, the switch fabric needs to be configured within a few nanoseconds, for our

design, the switching time is chosen to be under 5ns.

III. Circuit Implementations

A. Truly Differential Transimpedance Amplifier

The TIA converts the received optical signal into electronic signal. For highly integrated systems differential operation is preferred because high supply/ground noise rejection on-chip and high sensitivity are desired. Since photodiodes are single-ended, to make a truly differential TIA, typically a replica TIA is used along with a DC feedback loop to equalize the DC level of the differential output. The added feedback circuitries introduces imbalance to the system as it presents extra load to the one branch of the TIA and not the other, resulting in limited power supply rejection at high frequency and added circuit noise. We propose a novel TIA architecture based on current splitting principle that does not require DC offset correction and achieve high supply rejection up to the bandwidth of the amplifier. Fig. 2 shows the transistor level schematic of the TIA. The first stage consists of equal size transistors $M_1 - M_4$ with nominally the same transconductance gm because they carry the same current. The input current I_{ph} from the photodiode is sensed by the drain of M_2 and the gate of M_1 and results in a voltage change $V_{in} = Z_{in} \cdot I_{ph}$ where $Z_{in} \approx 1/(gm_2 + gm_4)$ is the input impedance of the TIA. The change in M_1's gate voltage results in current $I_1 = gm_1 \cdot V_{in} = gm_1 \cdot I_{ph}/(gm_2 + gm_4)$. Simultaneously, the current in M_4 is reduced by the same amount since it is applied to the source i.e. $I_s = -gm_4 \cdot V_{in} = -gm_4 \cdot I_{ph}/(gm_2 + gm_4)$. The cascode pair M_3 and M_4 relay the resulting currents to the load, R_L, and the differential output voltages $V_{outp} = gm_4 \cdot I_{ph} \cdot R_L/(gm_2 + gm_4)$ and $V_{outn} = -gm_1 \cdot I_{ph} \cdot R_L/(gm_2 + gm_4)$, and since $gm_1 = gm_2 = gm_3 = gm_4$, $V_{odiff} = V_{outp} - V_{outn} = I_{ph} \cdot R_L$. The lowpass filter R_{LPF} and C_{LPF} extracts the DC from the output voltage, and self-bias the cascode pair $M_3 - M_4$. The extracted DC is forwarded to differential pairs $M_5 - M_6$ and $M_7 - M_8$ which remove DC content from V_{outp} and V_{outn} resulting in truly differential output. Negative miller capacitor C_M reduces the loading of M_5 and M_8 on the drains of M_3 and M_4. Superior power supply noise rejection is a key feature of this novel architecture. Asymmetry in a circuit topology leads to finite PSRR. The only asymmetry in this topology is the gate-drain connection of M_2, therefore to analyze the supply sensitivity, we focus on the core TIA circuit. Qualitatively we can treat supply noise as perturbation δV_{dd} resulting in current δI_{dd} across the load R_L and will be coupled to the gate of the cascode nodes. This in turn will be coupled to the gate of M_1 via the mirroring action of M_2, M_1, mitigating the change in V_{dd}. However, this is only effective at low enough frequencies (up to the BW of the TIA) where the difference in currents in the two branches due to capacitive mismatches can be ignored.

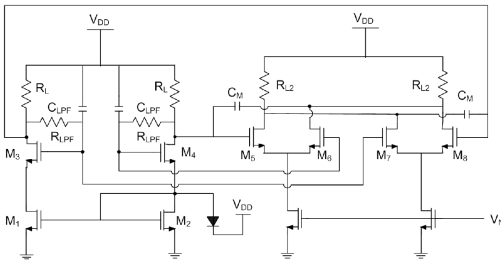


Fig. 2. Novel truly differential TIA based on current splitting frontend.

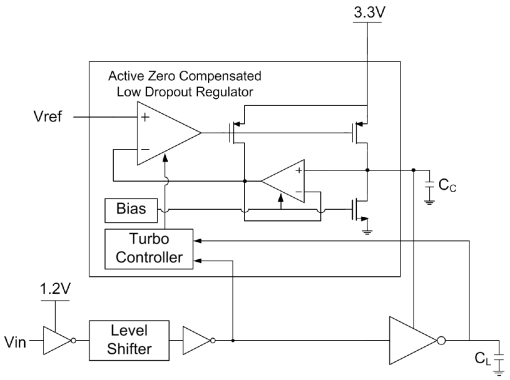


Fig. 3. Switch driver system schematic.

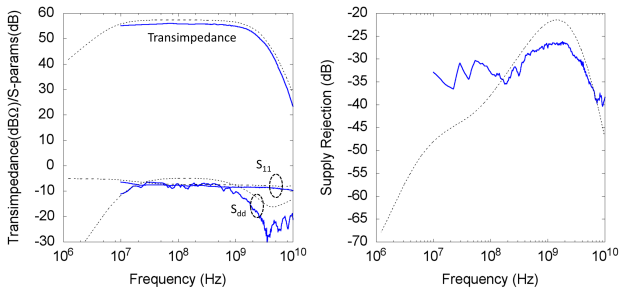
B. Switch Driver

Figure 3 shows the block level schematic of the switch driver circuit. Maximum drive voltage for the optical switch is near 3V, thus 3.3V I/O devices are used for the switch driver. The input, V_{in}, is a 1.2V digital signal and is converted to a 3.3V digital signal through a cross-coupled level shifter. Since the switch fabric is built using many instances of the elemental 2x2 switch element, multiple drivers in an array are susceptible to interference from neighboring switching activities. To mitigate interference, the drivers power supply was regulated by a low-dropout regulator (LDO). Furthermore, the LDO allows the output voltage to be set to a reference voltage, which can be determined in an offline calibration step. Calibration helps mitigate process variation in the fabrication of the optical switches by optimizing individual drive voltages to get maximum extinction ratio. The LDO incorporates a turbo-mode, where the op-amp's bandwidth is temporarily increased for a brief (10ns) period of time during the rising edge, resulting in a fast rise-time. While this technique saves power it complicates the design of the LDO. Typically, a two-stage amplifier, such as the one present in the LDO, is compensated using a lead-lag compensator by introducing a zero near the unity gain frequency of the amplifier. Turbo-mode dynamically changes the unity gain frequency of the amplifier, thus the zero location also needs to be moved accordingly, else the stability of the op-amp will be compromised. We employ active zero compensation (AZC) [4] to overcome this problem. In our

approach, the zero location is made proportional to the bias current, thus making sure that the regulator stable under all switching conditions. Another added benefit of the AZC scheme is that the large compensation capacitor C_c normally used in passive compensation, can be placed at the output of the regulator, providing a more stable supply. The LDO consists of AZC error amplifier and a $64\mu m$ wide PMOS output device and an on-chip 4.6pF dual metal-insulator-metal compensation/decoupling capacitor.

IV. Measurements

Three-port S-parameters of the TIA was measured and converted to transimpedance Zt. Fig. 4a shows the measured (solid) and simulated (dotted) Zt, input return loss S_{11} and balanced output return loss S_{dd}, which are in excellent agreement. Zt is 55.7 dBΩ and f_{3dB} is 1.97 GHz, sufficient for 2.5Gbps operation. The low frequency cutoff seen in the simulated result is due to the lowpass filter used to extract the DC content and can be sized to reduce the cutoff frequency. PSRR is characterized with 3-port S-parameters with the input applied at the supply node. Fig. 4b shows the measured and simulated PSRR. The measured result, at frequencies below 100 MHz, shows a deviation from simulated result, we suspect this is due to a calibration error of the output ports resulting in an imbalanced measurement. The electrical



(a) (b)

Fig. 4. (a) Measured versus simulated differential transimpedance. (b) Measured versus simulated PSRR of the TIA.

characterization of the switch driver circuit was done with a packaged chip. The output is measured with a high-speed sampling scope and 10:1 attenuated active probe. Fig. 5 shows the measured output waveform at several output voltages and measured versus simulated 10-90% rise time as function of V_{ref}. For optical measurement, the driver IC is wire bonded to the optical switch. Optical input/output at 1550nm are coupled to the switch via lensed fiber, and output is fed into a passive photodetector attached to a sampling scope. Fig. 6 shows the switching with input at port 1 and output at port 3. The 20-80% switching time is 3.8ns. The driver consumes $< 150\mu W$ of power while the optical switch consumes 100 μW.

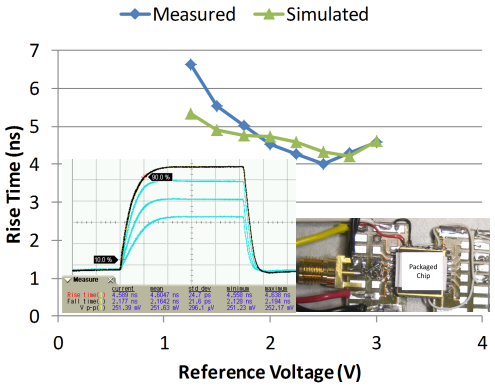


Fig. 5. Measured switch driver output rise time versus reference voltage.

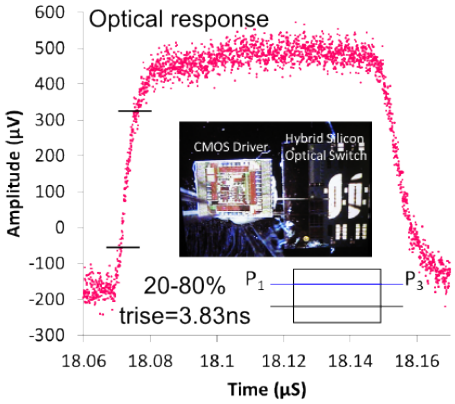


Fig. 6. Measured optical output waveform.

V. Conclusion

Optical packet switch with CMOS electronics integrated for data center switching is presented. A novel truly differential TIA architecture that achieves high BW PSRR is presented. The integrated CMOS driver IC with photonic IC shows $< 5ns$ switching time while consuming less than $< 250\mu W$ total.

Acknowledgment

The authors thank E. Kintzer for useful suggestions and MOSIS, in particular Wes Hansford, for funding the fabrication of the chip. This work was supported in part by Intelligence Advanced Research Project Activity under contract W911QX-11-C-0112.

References

[1] Kannan Raj, John E. Cunningham, Ron Ho et al., ""Macrochip" computer systems enabled by silicon photonic interconnects," *Proc. of the SPIE.*, vol. 7607, pp. 760702-760702-16, February 2000.

[2] "The 50G Silicon Photonics Link, white paper, Intel, July 2010. [Online].

[3] H.-W. Chen, "High-Speed Hybrid Silicon Mach-Zehnder Modulator and Tunable Microwave Filter, Ph.D. Dissertation," University of California, Santa Barbara, March, 2011.

[4] L. S. Theogarajan, "A Low Power Fully Implantable 15-Channel Retinal Stimulator Chip," *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 10, 2008, pp. 2322-2337.