A Monolithic Electronic-Photonic Platform in Conventional Silicon Photonics Processes

Philippe Arsenault and Wei Shi*
Department of Electrical and Computer Engineering, COPL, Université Laval, Québec, Canada
*wei.shi@gel.ulaval.ca

***Abstract*—We present the design methodology to make any active silicon photonics node, a monolithic platform. The approach is based on the usage of the waveguide and complementary dopants to create MOSFETs and digital gates allowing the inclusion of basic on-chip processing and specialized analog circuitry.**

***Index Terms*—Silicon photonics, Monolithic integrated circuit, electronic-photonic co-design, CMOS integrated circuits**

I. Introduction

State-of-the-art silicon monolithic electronic-silicon photonics processes are a monumental advancement in the co-design of electronic and photonic circuits. For high speed applications, sharing the same substrate has its advantages: no interconnection needed reducing the parasitic load, integrated simulation environment and simpler mechanical assembly of the integrated circuit (IC) into more complex systems. Although at the moment, even the best monolithic processes struggle to compete with recent CMOS nodes in terms of maximum frequency of operation for digital signal processing or the driving power needed for optical modulators. For lower speed applications that do not need these higher speeds or driving power, monolithic nodes seem to be an enticing solution for its ease of use, embedded testing capabilities and easier physical assembly. Nevertheless, since monolithic nodes are usually built upon an existing CMOS node and modified to fit the less strenuous fabrication of the photonics, they tend to exceed the price per area of similar CMOS processes [1].

Integrating CMOS capabilities in a zero-change silicon photonics node, although possible, is relatively new [1], [2]. It grants researchers and industry access to a monolithic node without the high entry price at the cost of poorer electrical performances than equivalent CMOS and has no foundry supplied electronic libraries. Nevertheless, it has recently been used in power monitoring and signal multiplexing applications [3], [4].

In this paper, we aim to reduce the gap between commercial monolithic nodes and a enhanced silicon photonics node by showcasing a digital core, proposing improvements on the transistor design and discussing the limitations of the current implementation and tooling.

The authors want to acknowledge the financial support of the NSREC and Alireza Geravand, Sammy Noel-Parisé and Francesco Zanetto for useful discussions.

II. Implementation

A digital core can be used for many thing, as it's logic is dictated by behavioral code like *Verilog*. One proposed application for such a core would be to handle the DC bias setting points of the photonics structure. Instead of having roughly two external pads for each structure, the core proposes a fixed number of pads (10), changing the connection paradigm. It's 16-bits word can either read (0) or write (1) onto one of the eight structures (3-bits address) a 12-bits value. The length of the word, the width of the address and the precision are all fixed by the designer at the behavioral level. The Fig.1.A shows the proposed architecture along with the supporting mixed or analog circuits that can be made of the same transistors. The Fig.1.B shows how the core is built inside, as well as its inputs and outputs in relation with Fig.1.C where all the connections are routed to pads instead of supporting circuits. The placement and routing of the 5800 transistors of this core were made automatically by the open-source tool OpenROAD [5], with the appropriate technology files and gate drawings, created using the python GDS drawing tool GDSFactory. Fig.1.D shows a close-up of a NMOS transistor, an essential building block of the gates. Compared to the transistors used in analog circuits, digital transistors aim to be the smallest and the fastest they can while having the best transconductance gain. As such, a modeling of its parameters that affect both the maximum frequency of operation and the gain will be presented in the next section.

III. Fundamentals and Principle

The small signal gain (g_m) and the cut-off frequency (F_t) of the transistor are both important in digital and analog circuits. A higher small signal gain means that the analog circuitry requires less transistors for the same amplification factor, reducing the circuits size and parasitics. The increase of the small signal gain also increases the current output strength of digital gates, helping in the charge and discharge of capacitive loads, such as gates or PN junctions. The other parameter, the cut-off frequency of the transistor, is a measurement of how fast a single transistor can switch. Improving the cut-off frequency of the transistor opens up the design area for analog and digital circuitry as it represents the upper limit at which a circuit can run in a parasitic-less environment. The saturation region small signal gain is the measurement of how much the drain current (I_D) changes in relation to the gate-to-source

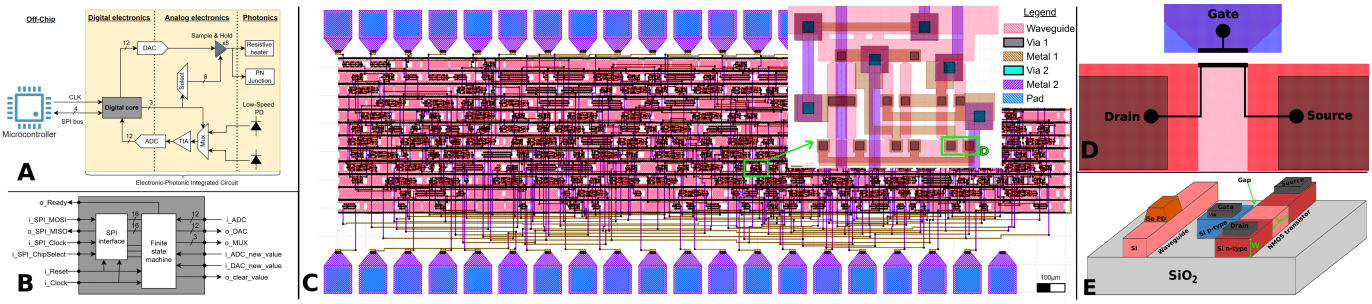

Fig. 1. A. The proposed digital core schematic, B. The digital core in a photonics system, C. Layout of the digital Core, the inlet shows the detail of a singular digital gate, D. A zoomed view of a single NMOS transistor and E. The 3D-view of a NMOS transistor beside a waveguide. The waveguide has a germanium photodetector on top.

voltage (V_{GS}) in a specific condition of the drain-to-source voltage (v_{DS}).

$$g_m = \left.\frac{\delta I_D}{\delta V_{GS}}\right|_{v_{DS}} = \mu_x C_{ox} \frac{W}{L}(V_{GS} - V_T)(1 + \lambda v_{DS}) \quad (1)$$

In 1, μ_x is the mobility of the charges, C_{ox} is the oxide capacitance per unit area, W and L are the size of the gate, V_T is the threshold voltage and λ is the channel-length modulation parameter. The length of the transistor is controllable by the distance between two tubs of the same type of doping (Fig.1.D, in red) and the length of the gate while the width is fixed by the waveguide layer height (see the W and L on Fig.1.E). Depending on the foundry, the minimal L distance can vary and is subject to mask alignment precision but should be used for maximum g_m value. As for maximizing C_{ox}, reducing the gap between the channel and the gate to the foundry's limit is the way to go, as one cannot easily change the material used between silicon waveguides.

The second metric, the cut-off frequency F_T, is defined by the small-signal gain over its gate-source (C_{GS}) and gate-drain capacitance (C_{GD}).

$$F_t = \frac{g_m}{2\pi(C_{GS} + C_{GD})} = \frac{\mu_x(V_{GS} - V_t)}{2\pi(\frac{2}{3}L^2 + 2L_{ov}L)}, \quad (2)$$

where L_{ov} is the overlap length between the gate and the source or drain area. Together with the diminution of the L value, one could theoretically increase the maximum frequency of operation up to 25 GHz and the g_m up to 4 μA/V as shown in Fig.2 and Fig.3. Both g_m and F_t are subject to carrier velocity saturation at shorter L values.

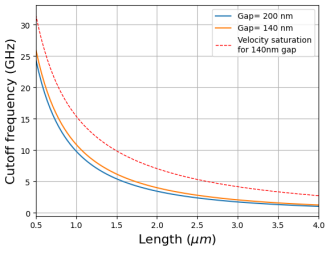

Fig. 2. F_t of a NMOS at V_{gs} = 5V

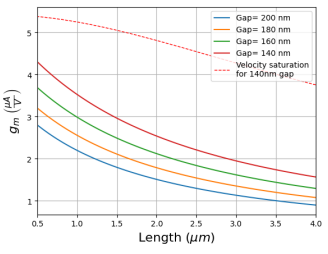

Fig. 3. g_m of a NMOS at V_{gs} = 5V

IV. Conclusion

Advancements toward a fully electronics-photonics monolithic node were demonstrated by improving the fundamental building block, the MOSFET, and by building of a digital electronics library. This library was then used with open-source tooling to produce a digital core with SPI capabilities showing that the implementation of a digital library was successful, albeit some restrictions and barriers are still in place. First is a file we were not able to provide to the automated router is the timing library of the gates, which can either be obtained through parasitic extraction simulations or constructed via a measurement campaign Nevertheless, not having the exact gate timing influences the maximum clock frequency of the circuit, not its logics, as the circuit will still be capable to compute at a lower speed. Second challenge encountered was the routing density. As photonic nodes are not geared toward complex and dense metallic routing (i.e. limited number of metal levels, vias do not stack, wide metals only), the automated router was not finding a solution to the maze problem when the density of the circuit was over 70%. Hence, the core size is determined not only by the gate dimensions but also by the metal routing rules specified by the foundry. To continue, the next steps of this endeavor are varied. The supporting components in Fig.1.A are yet to be built, in part because they are mixed or analog circuits and they need a reliable simulation environment supported by measurement of an array of sizes of L and gap values. Once done, the timing library should be generated and given to the automated router in order to complete a monolithic electronic-photonic platform in a conventional silicon photonic process.

References

[1] S. Shekhar *et al.*, "Roadmapping the next generation of silicon photonics," *Nat Commun*, vol. 15, no. 1, p. 751, 2024.

[2] F. Zanetto *et al.*, "Unconventional monolithic electronics in a conventional silicon photonics platform," *IEEE Transactions on Electron Devices*, vol. 70, no. 10, pp. 4993–4998, 2023.

[3] M. Crico *et al.*, "Monolithic transimpedance amplifier for on-chip light monitoring in pure silicon photonics," *IEEE Photonics Technology Letters*, vol. 36, no. 15, pp. 965–968, 2024.

[4] F. Zanetto *et al.*, "Time-Multiplexed Control of Programmable Silicon Photonic Circuits Enabled by Monolithic CMOS Electronics," *Laser & Photonics Reviews*, vol. 17, no. 11, p. 2300124, 2023.

[5] T. Ajayi *et al.*, "Openroad: Toward a self-driving, open-source digital layout implementation tool chain," *Proc. GOMACTECH*, pp. 1105–1110, 2019.