Systematic Simulation-Based Predictive Synthesis of Integrated Optical Interconnect

Ian O'Connor, Member, IEEE, Faress Tissafi-Drissi, Frédéric Gaffiot, Joni Dambre, Member, IEEE, Michiel De Wilde, Joris Van Campenhout, Student Member, IEEE, Dries Van Thourhout, Member, IEEE, Jan Van Campenhout, Member, IEEE, and Dirk Stroobandt, Member, IEEE

Abstract—Integrated optical interconnect has been identified by the ITRS as a potential solution to overcome predicted interconnect limitations in future systems-on-chip. However, the multiphysics nature of the design problem and the lack of a mature integrated photonic technology have contributed to severe difficulties in assessing its suitability. This paper describes a systematic, fully automated synthesis method for integrated microsource-based optical interconnect capable of optimally sizing the interface circuits based on system specifications, CMOS technology data, and optical device characteristics. The simulation-based nature of the design method means that its results are relatively accurate, even though the generation of each data point requires only 5 min on a 1.3-GHz processor. This method has been used to extract typical performance metrics (delay, power, interconnect density) for optical interconnect of length 2.5–20 mm in three predictive technologies at 65-, 45-, and 32-nm gate length.

Index Terms—Integrated optical interconnect, multidomain design methods, synthesis.

I. INTRODUCTION

THE emergence of very high performance systems-on-chip (SoC) is necessary to achieve future required application performance in terms of resolution (audio, video, and computing) and CPU power/total MIPS (real-time encoding-decoding, data encryption-decryption). The shift to distributed multiprocessor architectures is the recognized route to such performance and, therefore, requires organized high-speed communication between processors. Metallic interconnect will be highly inefficient in this role due to unachievable trade-offs between design parameters (the main limitations due to metallic interconnects are inter-line crosstalk, latency, global throughput, connectivity, and power consumption).

It is believed that the concept of integrated optical interconnect is a potential technological solution to alleviate some of these issues involved in exchanging data between cores in SoC architectures. However, the multiphysics nature of the design problem and the lack of a mature integrated photonic technology have contributed to severe difficulties in assessing the suitability of integrated optical interconnect for on-chip data transport. In turn, these difficulties have rendered unclear the research directions for the integrated photonic device community. The objective of optical interconnect assessment should thus be two-fold: first, to explore the performance capabilities of optical links with existing technological constraints; and second, to provide feedback to be taken into account in component and integration oriented research.

To enable the analysis previously described, it is necessary to define and quantify the specification set for the various subcomponents in an integrated optical link. This requires the development of models and tools to design and optimize optical interconnect in a range of technologies and under a range of specification conditions, in order to evaluate the properties of the technologies developed. Ultimately, the objective is to compare them to the performance of current and future electrical interconnect alternatives in terms of important performance metrics (mainly area, data rate, and power). A contiguous aim is to determine the specifications for which the proposed technologies would outperform traditional interconnect.

Optical-electrical interconnection comparison is not new: several authors have already published comparative studies between optical and electrical on-chip interconnect technologies [1]–[3]. Most of those consider expected technological evolutions to provide a roadmap of interconnect performance. This is done either by using analytical models, or by a simulation-based approach. Analytical models are usually based on ITRS [4] or other expected future CMOS technology parameters and rough estimates of optical interconnect performance. Simulation models are either self-constructed, based on ITRS parameters, or using publicly available Predictive Technology Models (PTM) for gate lengths ranging down to 32 nm.

In this paper, as it was intended to be able to determine device-level specifications, it was necessary to generate accurate optical link performance estimates that can be traced back to individual component parameters. Part of the optical link consists of analog CMOS circuitry. The design thereof can be done based on analytical models, but it is very sensitive to MOSFET parasitics so it is crucial that these parameters are available. However, the data presented in the ITRS is intended for digital design applications. This implies that estimates on transistor parasitics

are very rough or in some cases not even available. Hence, for an accurate evaluation, it is best to use technologies for which simulation models are available.

For these reasons, we have developed a simulation-based approach for optical interconnect synthesis. By way of illustration, we apply our approach to predictive technology models with gate lengths of 65, 45, and 32 nm. In contrast to the model-based approaches, our choice implies that we presently cannot stretch our analysis beyond 32-nm gate lengths according to the availability of models. However, according to the ITRS, a 32-nm gate length technology (which is the smallest available transistor model on the PTM website) would be taken into production for high-end applications in 2008, and this is only the beginning of its lifetime. Therefore, we claim that our analysis does stretch sufficiently far into the future. Also, by extrapolating observed trends in performance metric evolution over technology nodes, some insight into the possible results for future technology nodes can be obtained.

The models and software developed were, therefore, made to be sufficiently generic in order to enable the exploration of: 1) performance capabilities of optical links with existing technological constraints; and 2) the impact of improvements in device performance on interface circuit sizing and overall link performance. This set of data allows tradeoffs to be understood, as well as their relevance in the context of overall link performance metrics.

This paper first describes some technological and device aspects in Section II. The generation of optical link parameters based on link specifications is based on analog synthesis techniques, and the individual interface circuit synthesis methods are described in Section III. To enable complete optical link simulation, a number of behavioral models were developed, and this work is outlined in Section IV, although details are outside the scope of this paper. Section V concentrates on describing the complete optical link synthesis method exploiting the simulation work and individual circuit synthesis methods, as well as the details of the investigation program. Extraction methods concerning interconnect density analyses are described in Section VI. To generate performance metric data for considered technologies and optical link lengths, the simulation-based quantitative assessment of optical interconnects at the physical link level is described in Section VII. Through identification of the impact of individual device parameter variations on particular performance metrics, an exploration of these variations is shown in Section VIII.

II. TECHNOLOGY AND DEVICE ASPECTS

Various technological solutions exist for using optics with a standard CMOS SoC [5]. Our approach is generic, in the sense that it can be applied to any type of integrated photonic technology, as long as a behavioral model is available for each component (see Section IV). In this paper, we have supposed a microsource-based link using heterogeneous integration [6] to achieve an above-IC optical transport layer (see Fig. 1). A CMOS driver circuit modulates the current flowing through the microsource laser, and thus the intensity of light emission.

The microsource is coupled to a passive waveguide structure and provides a signal to an optical receiver (or possibly several, as in the case of a broadcast function). At the receiver, the high-speed optical signal is converted back to an electrical digital (rail to rail) signal and subsequently distributed by a local electrical interconnect network, represented by an output capacitance (fanout) that the link can drive. Similarly, the overall input capacitance presented to the link by the digital signal generator is part of the specification set for the optical link.

To form a planar optical waveguide, the most common CMOS-compatible materials used are polymers Si3N4 and silicon. In a first approach, we assumed silicon as the core and SiO2 as the cladding material. Si/SiO2 structures are compatible with conventional silicon technology and are transparent for 1.3–1.55 µm wavelengths. Such waveguides also benefit from a high relative refractive index difference Δ ≈ (n1 − n2) / 2n12 between the core (n1 ≈ 3.5 for Si) and claddings (n2 ≈ 1.5 for SiO2) and enable the realization of compact optical circuits, with bend radius of the order of a few micrometers [7] and low pitch to crosstalk ratio. To avoid modal dispersion, improve coupling efficiency, and reduce loss, single mode conditions are applied to the waveguide dimensions. For a wavelength of 1.55 µm in Si/SiO2, the waveguide dimensions are around 0.5µm x 0.22 µm (width x height).

Si3N4 waveguides are also contenders for optical interconnect. Since the index difference is lower, the propagation delay is also lower, but the penalty for this is higher crosstalk and bend loss, which implies a requirement for higher pitch and limited use of bends. The waveguide dimensions for these waveguides are around 0.4 µm x 0.8 µm (width x height). Data for both types of waveguide is given in the optical link simulation Section IV.

Integrated III-V sources can achieve >10-GHz 20% total efficiency (including coupling to waveguide) when biased above a 40 µA threshold current, and have a footprint of some 10 x 10 µm2. [8]. Similarly, 50 µm2 InP detectors exhibit 24 GHz 70% responsivity and low (< 10 fF) capacitance [9]. Heterogeneous integration/bonding of such devices on CMOS has also been proven [10].
III. SYNTHESIZABLE MODEL LIBRARY

Manual design of physically heterogeneous systems results in long design cycles and increasingly apparent bottlenecks in the overall design process [11]. This explains the growing awareness in industry that the advent of heterogeneous synthesis and optimization tools is a necessary step to increase design productivity by assisting or even automating the design process. The fundamental goal of heterogeneous synthesis is to quickly generate a first-time correct sized schematic from a set of specifications. This is critical since the heterogeneous design problem is typically under constrained with many degrees of freedom and with many interdependent (and often conflicting) performance requirements to be taken into account across physical domains. For this paper, this type of approach is essential since the objective is to explore the design of optical links subject to many sets of performance requirements and to many sets of technological and device-level constraints.

For this paper, we used our synthesis framework RunII [12], which enables the specification and automated design of heterogeneous systems, using a hierarchical approach to partition the system at a behavioral description level, and domain-specific simulators to design individual blocks. It exploits the concept of synthesizable AMS IP [13], which extends the concept of digital and software IP to the analog and heterogeneous domain. In essence, all information necessary for the design of a given block is formalized in a technology independent way such that a generic framework (such as RunII) can exploit the various facets of the AMS IP information to arrive at a sized system schematic satisfying all performance requirements in the target technology.

To carry out a systematic design analysis of optical interconnects, a set of design requirements and synthesizable models must be established. This section will detail the various synthesizable models developed (their final integration into a complete optical link synthesis method will be described in Section V).

A. Receiver Circuit Synthesizable Model

The classical structure for a receiver circuit is shown in Fig. 2: a transimpedance amplifier (TIA) converts the photocurrent of a few microamperes into a voltage of a few millivolts; a comparator generates a rail-to-rail signal; and a data recovery circuit eliminates jitter from the restored signal.

Of these, the TIA is arguably the most critical component, since it has to cope with a generally large photodiode capacitance situated at its input. For the comparator, scaling figures were used from a reference circuit manually designed at the 90-nm gate length node, and the input capacitance estimated for TIA synthesis.

It should be noted that photodetector capacitances in advanced devices can be around 10 fF or less— at this level, it is likely that the input capacitance of the TIA itself will play a large part in the determination of bandwidth. Bandwidth/power ratio maximization can be achieved in the following several ways:

- parametric optimization: for a given transimpedance structure, find the combination of component parameters necessary for maximum bandwidth;
- structural modification: for a given preamplifier architecture, make structural modifications, usually by adding elements such as inductors for shunt peaking [14] or capacitors as artificial loads or feedback [15];
- architectural exploration: use complex architectures such as bootstrap or common-gate input stages [16].

The basic transimpedance amplifier structure in a typical configuration is shown in Fig. 3, with the internal structure of the amplifier used in the resistive feedback configuration [17]. In
the following analysis, $C_v$ represents the aggregated capacitance at the TIA input node (including the photodetector capacitance $C_d$), $C_v$ represents the aggregated capacitance at the TIA output node (including the load capacitance $C_l$), $C_m$ represents the aggregated Miller capacitance between the TIA input and output nodes, $R_f$ represents the feedback resistance, $R_{in}$ represents the fast inverter output resistance, and $A_v$ represents the fast inverter voltage gain. The bandwidth/power ratio of this structure can be maximized by using small signal analysis and mapping of the individual component values to a filter approximation of Butterworth type, which gives

$$Z_{id} = \frac{R_0 - R_f A_v}{1 + A_v}$$

$$\omega_0 = \frac{1}{R_0 C_d} \sqrt{\frac{1}{M_f (M_v + M_m + M_x M_m)}}$$

$$Q = \sqrt{\frac{M_f M_v}{M_m (1 + M_m) (1 + M_x)}} \frac{1}{1 + A_v}$$

where the multiplying factors $M_f = R_f/R_{in}$, $M_v = C_v/C_{in}$, and $M_m = C_m/C_{in}$ are introduced, normalizing all expressions to the time constant $\tau = R_{in} C_{in}$. By rearranging these equations, it is then possible to develop a synthesis procedure which, from desired transimpedance performance criteria ($Z_{id}$, bandwidth, and $Q$) and operating conditions ($C_{ij}, C_d$) generates component values for the feedback resistance $R_f$ and the voltage amplifier ($A_v$ and $R_{in}$).

Taking into consideration their physical realization, amplifiers with requirements for low gain and high output resistance (high $R_{in}/A_v$ ratio) are the easiest to build, and also require the least quiescent current and area. Fig. 4 shows a plot of this quantity against the TIA specifications (bandwidth and transimpedance gain) for $C_v = C_d = 400 \text{ fF}$ and $C_v = C_l = 45 \text{ fF}$.

Approximate equations for the small signal characteristics and bias conditions of the circuit allow a first cut sizing of the amplifier. The solution is then fine tuned by a local numerical optimization method driven by a sum of weighted cost functions to achieve the desired values of $A_v$ and $R_{in}$, using simulation for exact results [18]. Parasitic capacitances are extracted from the resulting transistor sizes and are used to update the values of $C_v$ and $C_y$. The transimpedance amplifier sizing and amplifier optimization steps are subsequently run again with continually updating values of $A_v$, $R_{in}$, $C_v$, and $C_y$ until the process converges.

Using this methodology and predictive BSIM3v3 and BSIM4 models for technology nodes from 65 nm down to 32 nm, we generated design parameters for 1-THz TIA transimpedance amplifiers to evaluate the evolution in critical characteristics with technology node. Fig. 5 shows the results of transistor level simulation of fully generated photo-receiver circuits at each technology node. According to traditional “shrink” predictions, which consider the effect of applying a unitless scale factor of $1/x$ to the geometry of MOS transistors, the quiescent power, and device area should decrease by a factor of $1/x^2$. Between the considered technology nodes 65-32 nm, $x = 4$, which is approximately verified through the sizing procedure (quiescent power scales by $1/4.4$, gate area scales by $1/3$). This methodology also allows us to find a particular specification to a given tolerance, as shown in Fig. 6. This shows the active area and power of the generated TIA for bandwidths of 1-5 GHz (with $Z_{id} = 1 \text{k} \Omega$ and $Q = 1/\sqrt{2}$ in the 45-nm technology).

### B. Driver Circuit Synthesizable Model

The basic current modulation configuration of the source driver circuit is shown in Fig. 7. While more advanced driver schemes exist (such as voltage-pulsed drivers [19]), the current modulation circuit is sufficient to evaluate the characteristics of the overall link in a systematic way. The source is biased above its threshold current by $M_I$ to eliminate turn-on delays, and as the bias current value is the main contributing factor to emitter power, reducing the source threshold current is a primary device research objective (figures of the order of 50 $\mu$A appear achievable [8]). Device $M_I$ modulates the current flowing through the source, and consequently the output optical power injected into the waveguide. The reduced dimensions of the
microsource (with respect to commercial VCSELs) translates to lower parasitic capacitance and good bandwidth performance. This circuit is sized simply by numerical bisection using the required values for the modulation current $i_{m0}$ and the bias current $I_B$ (a unique solution exists).

IV. SIMULATION

To enable complete link simulation in an EDA framework, it is necessary to develop behavioral models for the optoelectronics devices and passive waveguides. For all behavioral models, the choice of an appropriate level of description is prerequisite to developing and using the models in the required context. Essentially, the description level falls into one of two categories: functional modeling or structural modeling. A functional model will describe the behavior of a device according to its specifications and behavioral equations, without defining the structure of the device. A structural model will describe the behavior of a device according to its internal structure and physical parameters, without necessarily satisfying the specification criteria (which do not have to be formalized in this approach).

Ideally, both functional and structural models should exist for the devices considered. However, this is not absolutely necessary and careful consideration was given to choosing the appropriate description level for each device. Since the source behavior is arguably the most complex and likely to exhibit nonlinear behavior (thermal roll-off, temperature changes) important to complete link simulation, it was decided to model this element at a structural level. The waveguide and detector were modeled at a functional level.

The models were all implemented in the OVI-96 Verilog-A subset of Verilog-AMS, an extension of the IEEE 1364–1995 Verilog hardware description language. This extension is an industry standard for analog simulation model description and can be simulated with a number of general-purpose circuit simulators (we use Spectre). This way the optical and photonic devices can be simulated together with the interface circuitry and with the rest of the optical link given adequate simulation models. This enables interesting optimization strategies (e.g., joint power optimization) and the analysis of link performance sensitivity to various parameter variations as well as temperature changes.

Detailed description of these models [20] is outside the scope of this paper, but the device parameters for optical interconnect varied in this analysis are shown in Table I. with minimum
and maximum values defining the limits of the parameter variation. These limits are based on discussions with experts in the field [21], and on data available in the literature on sources [8], waveguides [7], and detectors [9]. The values in bold italics represent the (pessimistic) nominal values.

This enabling design technology was essential to carry out simulation in an EDA environment of the complete optical link, associating these behavioral models with transistor-level circuit schematics. The nonlinear behavior of the microsource laser was modeled (enabling visualization of physical limits) and converges systematically in Spectre. Achieving complete link simulation was a necessary step to enable subsequent simulation-based link synthesis (using interface circuit design variables) over a range of target technologies and specification sets to extract link performance data. The iterative optimization step is facilitated by the low simulation time required for the complete link (a few seconds for ten data bits on a 1.3-GHz processor).

V. POINT-TO-POINT LINK SYNTHESIS

The objective of our paper was to carry out transistor-level sizing of the receiver and of the driver circuits according to complete link specifications. The optical link under consideration is represented by Fig. 8 (the CMOS structures used at transistor level are fixed in this synthesis procedure).

A. Synthesis Procedure

The synthesis approach implemented consists of creating scenarios allowing the specification of each model, evaluation, and design methods, and of communication between the different blocks using synthesizable AMS IP. A device library containing the synthesizable models of each device in the optical link based on the UML language was developed to allow the modeling of this hierarchical synthesis problem. The procedure used to automatically synthesize an optical point-to-point link is shown in Fig. 9.

The process starts by defining the photodetector characteristics and the required data rate. Using the method described in Section III-A, the transistor-level schematic for the transimpedance amplifier is automatically generated and linked to a manually scaled comparator circuit.

The value of the root mean square (rms) noise power $i_n$ is extracted from simulation of the schematic and updated for each synthesis loop using the Morikuni formula [19] in the transimpedance amplifier noise calculations

$$i_n^2 = \left(2g(I_{gate} + I_{dark}) + \frac{4kT}{R_f}\right) \frac{C}{4D} + 4kT \frac{C^2}{16\pi^2DE \frac{(2\pi C_F)^2}{g_m}} \tag{4}$$

where

$$C = 1 + g_m R_f$$
$$D = R_c (C_x + C_y) + R_f (C_x + C_m) + g_m R_f R_e C_m$$
$$E = R_f R_e \left[(C_x + C_y) C_m + C_x C_y\right]. \tag{5}$$

For a given bit error rate (BER) specification and noise signal associated with the photodetector and transimpedance circuit, we can then calculate the minimum optical signal power $\Delta P_o$ required by the receiver to operate at the given error probability

$$\Delta P_o = i_n \times SNR \tag{6}$$

where

$$SNR = \left(\text{erfc}^{-1}(2\text{BER})\right)^2. \tag{7}$$

Here SNR represents the linear signal-to-noise ratio (absolute value, not in decibels). BER, defined as the rate of error occurrences, is one of the main criteria in evaluating the performance
of digital transmission systems. In our analyses, we fixed BER at $10^{-18}$ bits$^{-1}$ (this corresponds to 1 error/3.17 years for a single link at 10-Gb/s communication, or 1 error/18 days for a 64-bit data bus at the same data rate.)

The value of the power which needs to be emitted by the laser source is evaluated from the calculated value of the minimum optical power at the receiver, and from the power losses induced by the waveguide structure (length and intrinsic loss, number of bends and loss/90° bend) and coupling. These figures depend to a large extent on the materials used.

The final sizing step is to calculate the driver and associated bias and buffer circuits using the emitted power value and the source characteristics in conjunction with the method described in Section III-B. This then enables the simulation of the complete optical link, using transistor-level schematics for the interface circuits and behavioral models for the photonic devices. From the simulation results, the performance criteria can be extracted.

Using this approach, the synthesis problem is considered to be complete, such that no constraint partitioning is required. In fact, the constraints are derived directly from system specifications, and thus constraint exploration is achieved directly by the user, as shown in the investigation strategy detailed in Section VII.

### B. Metrics

In order to be able to evaluate and optimize link performance criteria correctly, a clear definition of the performance metrics is required.

First, the aim is to establish the overall power dissipation for an optical link at a given data rate and BER. The calculation is essentially conditioned by the receiver as explained before, since the BER defines the lower limit for the received optical power. This lower limit can then be used to calculate the required power coupled into waveguides by optical sources, the required detector efficiency (including optical coupling) and acceptable transmission losses. Power can then be estimated from source bias current and photo-receiver front-end design methodologies.

For interconnect density aspects, source and detector sizes must be taken into account, while the width, pitch, and required bend radius of waveguides is fundamental to estimating the size of the photonic layer. On the circuit layer, the additional surface due to optical interconnect is in the driver and receiver circuits, as well as the depassivated link to the photonic layer. The circuit layout problem is compounded by the necessity of using clean supply lines (i.e., separate from digital supplies) to reduce noise (for BER).

The data rate is essentially governed by the bandwidth of the photo-receiver: high modulation speed at the source is generally more easily attainable than similar detection speed at the receiver. This is essentially due to the photodiode parasitic capacitance at the input of the transimpedance amplifier.

The limitations of this analysis are the following:

- PTM models do not take noise into account particularly well, which means that no real noise analysis can be carried out. However, in the sizing process, this problem was circumvented by using the Morikuni formula to estimate noise at block level.
- No automatic layout generation tools were used as it is not in general possible to achieve optimal layout for high-speed analog circuits. Parasitic capacitances were, therefore, extracted from layout estimations rather than from real layouts.

### C. Specifications

Table II shows the sets of specifications used for analysis and interface circuit sizing. Three predictive technologies were considered for this analysis: PTM models [23] for 65-, 45-, and 32-nm gate length technology nodes.

### VI. Interconnect Density Analyses

In this paper, we supposed that only one photonic routing layer would be available above IC. The impact of this hypothesis is that interconnect links can only be defined in a single routing direction (this does not exclude the occasional use of bends to avoid routing obstacles) due to unacceptable crosstalk (around 10% power loss) for straight waveguide crossings, and unacceptable area penalties (around $10 \times 10^3$) for optimized crossings (1% power loss at 1.55-$\mu$m wavelength) [24].

Direct optical link bundles (see Fig. 10) were considered to be used for all inter-processor communication, except between physically adjacent processors (electrical buses used in this case). Normally of course, a design engineer would not create a fully connected architecture, but this is still a useful example to explore and quantify achievable interconnect density.

A number of parameters were defined in order to carry out interconnect density analyses, and are summarized in Table III.

If we assume that the overall chip is square, then the total chip size $= c_s^2$. If each IP block is also square then the IP block size $= c_b^2/p_c^2$. Since the coupling from sources to waveguides and from waveguides to detectors is electromagnetic, we used the
Routing scheme shown in Fig. 10(a) to prevent any waveguide routing below an optoelectronic device and avoid in this way any potential crosstalk issues. If \( s_x \neq d_x \), then using minimum device spacing will result in data skew and should be evaluated or avoided by nonoptimal spacing of the smaller device fixing \( s'_x = d'_x = \max(s_x, d_x) \). This approach will also result in a delay penalty due to longer waveguide lengths.

For a given number \( p_x \) of IP blocks, we can then calculate the following:

- the number of optical routing channels necessary for total connectivity (excluding adjacent IP block communication which will be carried out by electrical interconnect);
- the maximum number of links in each bundle, based on waveguide pitch and optoelectronic device dimensions.

### A. Number of Optical Bundles

The number of necessary optical routing channels can be calculated by examining the connection scheme between IP blocks and determining the number of channels at the plane where the maximum connection density occurs. A number of cases can be considered in order to derive a general model for the number of necessary channels, as shown in Fig. 10(b). In fact, two basic cases can be extracted: for an even number of IP blocks, a single dense plane exists, whereas for an odd number of IP blocks, two dense planes exist.

In general\(^3\)

\[
\begin{align*}
n_b = p_x (p_x - 2) - \sum_{i=2}^{p_x-1} \sum_{j=2}^{p_x-1} p_b - 2i. \tag{8}
\end{align*}
\]

This is shown in graphical form in Fig. 11.

### B. Number of Optical Links Per Bundle

The total number of optical links in a row depends on the side of the total chip \( c \), the number of rows (equal to \( p_y \) if the chip is square and the IP blocks are also square), the pitch between waveguides \( w_y \), the source and detector pitches \( (s_x, s_y, d_x, d_y) \), and waveguide bend radius.

The maximum average number of aligned sources or detectors over the width of an IP block, \( n_{sd} \), is calculated as

\[
\begin{align*}
n_{sd} = \left\lfloor \frac{s_x}{p_x} \cdot \frac{s_y}{p_y} \cdot \frac{d_x}{p_x} \cdot \frac{d_y}{p_y} \right\rfloor. \tag{9}
\end{align*}
\]

The maximum number of links in a row, \( n_l \), is calculated as

\[
\begin{align*}
n_l = \left\lfloor \max(s_x, d_x) + 2w_b + n_{sd}w_y \right\rfloor n_{sd}. \tag{10}
\end{align*}
\]

The maximum number of bidirectional links in a bundle \( n_{lb} \) is calculated as

\[
\begin{align*}
n_{lb} = \left\lfloor \frac{\left\lfloor \frac{n_l}{n_y} \right\rfloor}{2} \right\rfloor. \tag{11}
\end{align*}
\]

Fig. 12 shows a plot of the maximum achievable number of bi-directional links in a bundle for varying number of IP blocks in a row and varying source/detector sizes using this routing scheme. It can be seen for example that 64-bit, "crosstalk-free" optical communication buses between 8 IP blocks in a single
row can only be achieved for source/detector sizes of around $10 \times 10 \, \mu m^2$ and waveguide pitch of 1.1 $\mu m$.

C. Via Stack Calculations

In order to calculate the area penalty (at the CMOS level) of the via stack for passing a relatively high level of current (mainly driver modulation current to the source), we calculate the via matrix area $A$, defined by host technology characteristics and the maximum current passing through the stack. The area required by the via matrix is $A = N(a + b)^2$, where $A$ represents area ($\mu m^2$), $N$ represents the number of contacts in the matrix, and $a$ and $b$ represent minimum contact width ($\mu m$) and spacing ($\mu m$), respectively.

The overall surface area of a given metal, $N a^2$, is proportional to the maximum current $I_{\text{max}}$, following $N = k_m I_{\text{max}} / a^2$, where $k_m$ is a constant associated with the via metal. Rewriting the equation for $A$:

$$A = k_m I_{\text{max}} \left(1 + \frac{b(2a + b)}{a^2}\right). \quad (12)$$

Based on a set of industrial 130-nm gate length design rules for which damascene copper is used for all metal layers, we can empirically determine that for the upper-level vias (worst case) $k_m = 0.166 \, \mu m^2 / mA$.

Even if $a$ and $b$ scale with shrinking process rules, it is trivial to show that $a$ does not scale with the shrinking factor, as long as $a$ and $b$ retain the same ratio to each other. In the aforementioned 130-nm gate length process, $a = 0.36 \, \mu m$ and $b = 0.54 \, \mu m$ for via. Hence:

$$A (\mu m^2) = 2.913 \times I_{\text{max}} (mA). \quad (13)$$

For example, if $I_{\text{max}}$ is situated at around 5 mA, the total via stack area as seen at the circuit level is 14.56 $\mu m^2$.

D. CMOS Interface Circuit Area Calculations

In order to validate the area work, the final step is to estimate the real circuit area necessary from gate area sizes. In the absence of satisfactory high-speed analog layout synthesis tools, our approach was to define a skeleton layout based on worst case transistor sizes and derive technology independent layout rules from this. An important consideration in this work is the actual size of the via stack and its position in each interface circuit layout. The resulting estimated layout was used both to calculate circuit area and also to estimate parasitic capacitances on all circuit nodes.

Following transistor-level synthesis of the interface circuits under worst case conditions (65-nm technology for 20-nm link length, 15-mA source threshold current and 10% efficiency, 2.7-dB/cm waveguide loss, 100-IF photodiode capacitance and 50% responsivity), the total gate area of the interface circuits was calculated. The total circuit area was estimated from this using the skeleton layout template shown in Fig. 13 (including via stack area) and the lambda rules described in the Appendix.

For the receiver circuit, a similar analysis to that of the driver circuit was carried out using the skeleton layout shown in Fig. 14 to enable prediction of actual circuit size from transistor gate dimensions.

The total area was estimated using this method as 55.5 $\mu m^2$. This worst case figure for overall CMOS area is considerably less than the area required for the smallest active device. As such, the active devices remain the limiting factor for area concerns rather than interface circuits.

VII. INVESTIGATION PROGRAM RESULTS

A. Gate Area Analysis for Varying Technologies

The link sizing method described in Section V was applied according to the specifications for the PTM 65-, 45-, and 32-nm technologies. Fig. 15 shows the results in terms of gate area (i.e., transistor channel dimensions only), extracted as the sum of all transistor gate channel areas $W \cdot L$. These results show that the gate area metric approximately verifies the scaling law

A skeleton layout defines relative positions of transistors, without associating any information concerning actual transistor sizes.
(A_{32} \text{ nm} \approx A_{45} \text{ nm} \cdot s^2 \approx A_{65} \text{ nm} \cdot s^2 \cdot s^2), \text{ where } s \text{ is equal to 0.7 (scaling factor between technology generations).}

B. Delay Analysis

The link sizing method described in Section V was applied according to the specifications for the PTM 65-, 45-, and 32-nm technologies. The 50% propagation delay was extracted from simulation, as well as a point of reference in the form of the intrinsic waveguide delay, using \( t_{\text{TOF}} = 13.3 \text{ ps/mm} \) for the Si/SiO\textsubscript{2} waveguides\textsuperscript{5} Fig. 16 shows the delay results for varying link lengths.

It can be seen that: 1) the circuit delay (i.e., the difference between the total delay and the intrinsic waveguide delay) decreases with smaller gate lengths and 2) the same quantity also decreases with longer interconnect. This latter effect is due to higher modulation current \( I_m \), required to compensate higher overall waveguide loss, being able to drive the source capacitance faster.

\textsuperscript{5}TOF: Time of Flight.

C. Power Analysis

The link sizing method described in Section V was applied according to the specifications for the PTM 65-, 45-, and 32-nm technologies. The average static power was extracted from transient simulations using

\[ P = \frac{I_{\text{source}} + I_{\text{source1}}}{2} \cdot V_p + \frac{I_{\text{det}} + I_{\text{det1}}}{2} \cdot V_d + \frac{I_{\text{cest}} + I_{\text{cest1}}}{2} \cdot V_{dd} \]  

where \( I_{\text{source}}, I_{\text{det}} \) and \( I_{\text{cest}} \) represent the currents flowing through the source, detector, and circuit voltage supplies of \( V_p, V_d \), and \( V_{dd} \), respectively. Fig. 17 shows the average static power results for varying link lengths.

The average switching energy is calculated from rising and falling edge transitions (extracted from simulations as the integral of supply currents in edge transitions). Fig. 18 shows the average switching energy results for varying link lengths.

Using static power and switching energy information, the total power can be calculated (see Fig. 19).
Fig. 18. Average switching energy (in picojoules) versus interconnect length for PTM 65-, 45-, and 32-nm technologies.

Fig. 19. Total power (in milliwatts) versus interconnect length for PTM 65-, 45-, and 32-nm technologies.

Fig. 20. Total power dissipation (W) for varying number of interconnected IP blocks.

The figures for aggregated maximum number of links per optical bundle and associated bundle lengths (including additional length due to device staggering) were then used in conjunction with the total power dissipation figures to calculate the total power in one row of interconnected IP blocks, for varying number of IP blocks (from 2 to 10) (see Fig. 20). It can be seen that the total power decreases with a rising number of IP blocks. This is due to the fact that, while there are more links, each link is shorter. Power increases exponentially with length and is dominated by the source power. For a 10 × 10 matrix of IP blocks, and for the most long-term technology node (32-nm printed gate length), this analysis gives a total power of 33.3 W (source threshold currents of 1.5 mA). Compared to the figure given by the ITRS for maximum power at this technology node (167 W), this works out as equal to 20% of total power.

These results show the capacity of the method to synthesize optical links subject to technological specifications, both CMOS and optical. The generation of each data point requires approximately 5 min on a 1.3-GHz processor.

VIII. DESIGN SPACE EXPLORATION

By carrying out variations on the photonic device characteristic set (within the bounds established in Table I) and rerunning the investigation program, it is possible to establish the impact of individual device characteristics on link performance. This section presents some conclusions from this analysis, as an illustration of the type of feedback our approach can give to photonic device engineers.

A. Improvement of Waveguide Materials

Overall link delay (see Fig. 21) can be reduced by about a factor of 2 for 20-mm interconnect length by shifting to lower index waveguide materials. Interestingly, it can be seen from the figure that the waveguide delay increase for longer lengths is compensated for by faster modulation with higher currents, leading in some cases to a decrease in interconnect delay with length. It should be possible, in this configuration, to optimize the system such that overall delay is constant for all interconnect length. This would enable synchronous data communication and reduce dependence on data recovery circuits.

B. Reduction of Source Threshold Current

Source threshold current has a significant impact on static power (see Fig. 22). Approximately 5 mW can be saved by
Fig. 22. Average static power (mW) for varying interconnect length and source threshold currents: PTM45 nm.

Fig. 23. Schematic λ-rule layout used for inverter active area and standard cell area approximations.

bringing the threshold current down to 150 μA from 1.5 mA. While the additional reduction down to 15 μA appears negligible in the context of overall link static power, it should be noted that, when multiplied by the total number of links on the chip, even a small individual power saving can represent a significant reduction at the chip scale.

IX. CONCLUSION

This paper has described a methodology and toolset for the systematic design space exploration of integrated optical interconnect. In particular, it allows the automated design of optical links, resulting in a significant reduction in design time and, more importantly in this context, the capacity to carry out a detailed analysis of the available design space. This technology is particularly useful for repetitive design of fixed optical link structures subject to varying design constraints, technology parameters, and performance requirements. Further, as designer variability is removed from the overall design process, the resulting set of designs is likely to be relatively coherent and this facilitates interpretation of data and extrapolation to future design scenarios.

We have illustrated the direct application of our approach for link synthesis and technology performance characterization by analyzing optical link performance for a single set of photonic components and three CMOS technology generations. We have also hinted at the ways our toolset can be applied to generate useful feedback from system designers to component designers.

APPENDIX

\textbf{λ-RULE SET FOR LAYOUT CALCULATIONS}

In a λ-rule layout, all distances and design rules are expressed in so-called λ-units. This means that the layout grid unit is called λ, and it is set to half the minimal gate length (λ = L/2). All distances are expressed as a number of grid units. If the λ-based design rules are not made too tight, a λ-rule layout can be ported across several technology generations. In this paper, the specific design rules used were taken from an existing standard cell library and reexpressed as a function of λ.

For determining the finger widths, we have simplified our calculations by making all fingers equally wide. Hence, we are disregarding finger width restrictions from the layout grid, since we expect this will have very little impact on the simulation results.

\textbf{TABLE IV SUMMARY OF PSEUDO-LAYOUT PARAMETERS AND EQUATIONS TO DERIVE CMOS CELL AREA ESTIMATES}

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Interpretation</th>
<th>Value or equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_{\text{cell}} ) (λ)</td>
<td>Standard cell height</td>
<td>69</td>
</tr>
<tr>
<td>( w_{\text{active}}(λ) )</td>
<td>Maximal total active width, i.e., maximal available cell height for MOSFET diffusion regions</td>
<td>47</td>
</tr>
<tr>
<td>( w_{\text{PMOS}}(λ) )</td>
<td>MOSFET gate width, expressed as a multiple of λ.</td>
<td>Set by designer</td>
</tr>
<tr>
<td>( w_{\text{NMOS}}(λ) )</td>
<td>PMOS gate width, expressed as a multiple of λ.</td>
<td>Set by designer</td>
</tr>
<tr>
<td>( F )</td>
<td>Number of fingers as NMOS and PMOS transistors (equal)</td>
<td>( \text{cell} \left( \frac{w_{\text{cell}}}{w_{\text{cell}} + w_{\text{cell}}} \right) )</td>
</tr>
<tr>
<td>( w_{\text{diff, NMOS}}(λ) )</td>
<td>Width of individual NMOS fingers, set equal to width of NMOS diffusion region.</td>
<td>( w_{\text{PMOS}} / F )</td>
</tr>
<tr>
<td>( w_{\text{diff, PMOS}}(λ) )</td>
<td>Width of individual PMOS fingers, set equal to width of PMOS diffusion region.</td>
<td>( w_{\text{PMOS}} / F )</td>
</tr>
<tr>
<td>( L_{\text{MOSFET}}(λ) )</td>
<td>MOSFET channel length</td>
<td>2</td>
</tr>
<tr>
<td>( S(λ) )</td>
<td>Spacing between gate poly and metal</td>
<td>1</td>
</tr>
<tr>
<td>( L_{\text{metal}}(λ) )</td>
<td>Width of metal above diffusion region contact</td>
<td>4</td>
</tr>
<tr>
<td>( B(λ) )</td>
<td>Border at right and left edges of cell</td>
<td>6</td>
</tr>
<tr>
<td>( L_{\text{diff, MOSFET}}(λ) )</td>
<td>Length of MOSFET diffusion region at the edges</td>
<td>( L_{\text{cell}} + S = 5 )</td>
</tr>
<tr>
<td>( L_{\text{diff, PMOS}}(λ) )</td>
<td>Per MOSFET length of diffusion region if between fingers</td>
<td>0.5 ( L_{\text{cell}} + 2S = 3 )</td>
</tr>
<tr>
<td>( L_{\text{pp, diff}}(λ) )</td>
<td>Total length of diffusion region (for active area calculation)</td>
<td>( 2L_{\text{diff, PMOS}} + 2 \left( \frac{L_{\text{cell}}}{2} - 1 \right) )</td>
</tr>
<tr>
<td>( L_{\text{active}}(λ) )</td>
<td>Total active area length</td>
<td>( L_{\text{cell}} + F_{\text{L}} + 8F_{\text{L}} + 11 )</td>
</tr>
<tr>
<td>( A_{\text{cell}}(λ^2) )</td>
<td>Standard cell width</td>
<td>( F_{\text{L}} \left( w_{\text{diff, PMOS}} + w_{\text{active}} \right) )</td>
</tr>
<tr>
<td>( A_{\text{active}}(λ^2) )</td>
<td>Total active area</td>
<td>( L_{\text{cell}} \left( w_{\text{diff, PMOS}} + w_{\text{active}} \right) )</td>
</tr>
<tr>
<td>( A_{\text{cell}}(2λ^2) )</td>
<td>Total cell area</td>
<td>( \text{cell} \left( \frac{w_{\text{cell}}}{w_{\text{cell}} + w_{\text{cell}}} \right) )</td>
</tr>
</tbody>
</table>
As an example of how a scalable $\lambda$-rule layout template can be used to estimate layout area, we give a complete illustration for the case of a sample inverter (see Fig. 23). Table IV summarizes the design rules and equations that can be used to derive CMOS area estimates.

REFERENCES


Ian O’Connor (S’95–M’98) was born in Cambridge, U.K., in 1968. He received the European M.Sc. degree in electronics engineering from the University of Essex, Essex, U.K., the Ph.D. degree in electronics from the University of Lille, Lille, France, in 1997, and the professorial dissertation (Habilitation à Diriger des Recherches) from Ecole Centrale de Lyon, Ecully, France, in 2005. He joined Ecole Centrale de Lyon in 1998 as an Associate Professor, and is currently head of the Design Group at the Lyon Institute of NanoTechnology, of which he is also the head of the vice-directors. From 1993 to 1997, he was a Senior Research and Teaching Assistant with the Institut Supérieur d’Électronique du Nord, Nord, France, and from 1997 to 2005, he was a Senior CAD Engineer with Philips Semiconductors, Southampton, U.K. His research interests include design methods and tools for physically heterogeneous systems on chip, and electronic application to novel architectures based on nonconventional devices. He has authored or coauthored well over 50 book chapters, journal publications, and conference papers and has presented invited papers at several major conferences.

Faress Tissafi-Drissi was born in Fes, Morocco, in 1977. He received the Ph.D. degree in microelectronics from Ecole Centrale de Lyon, Ecully, France, in 2004.

He is currently with Dolfini Integration, Meylan, France, as a Mixed-Signal Circuit Designer for systems on chip. He works particularly on the development of embedded memories for portable applications and scientific research in nanophotonics—optoelectronics areas.

Frederic Gaffiot was born in 1959. He received the M.S. degree in applied physics from Ecole Normale Superieure de Cachan, Cachan, France, and the Ph.D. degree in microelectronics from Ecole Centrale de Lyon, Ecully, France, in 1987.

He is currently with Ecole Centrale de Lyon. He was primarily an Associate Professor (1985–2000) and has been a full-time Professor since 2001. His research interests involve analog and mixed-signal circuit design, behavioral modeling, multiphysics design, optical interconnects, system on chip design, and wireless sensor networks.

Joni Dambre (S’99–M’04) was born in Gent, Belgium, in 1973. She received the M.S. degree in electronic engineering and the Ph.D. degree from Ghent University, Gent, Belgium, in 1996 and 2003, respectively. She is currently a Postdoctoral Researcher with the Department of Electronics and Information Systems (ELIS), Ghent University. Her research interests include platform and technology evaluation, system level design and prediction techniques, and novel interconnection technologies.
Michiel De Wilde was born in Leuven, Belgium in 1977. He received the M.S. degree in computer engineering and the Ph.D. degree in electronics and information systems from Ghent University, Ghent, Belgium, in 2000 and 2007, respectively. His research interests include design, modelling and system-level evaluation of short-range parallel optical interconnects, quantification of noise phenomena in integrated circuits, evaluation of FPGA architectures, and design of spatial indexing methods.

Joris Van Campenhout (S'01) received the M.S. degree in physics engineering from Ghent University, Gent, Belgium, in 2002, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include the design, fabrication, and characterization of silicon-based photonic integrated circuits. Over the last years, he has been working on the heterogeneous integration of InP-based microcavity lasers—based on photonic-crystal and microdisk structures—on a silicon platform.

Dries Van Thourhout (S'99–M'09) received the M.S. degree in physical engineering and the Ph.D. degree from Ghent University, Gent, Belgium, in 1995 and 2000, respectively. From October 2000 to September 2002 he was with Lucent Technologies, Bell Laboratories, Crawford Hill, NJ, working on the design, processing, and characterization of InP/InGaAsP monolithically integrated devices. In October 2002, he joined the Department of Information Technology (INTEC), Ghent University, where he is currently a member of the permanent staff of the Photonics Group. His research interests focus on the design, fabrication, and characterization of integrated photonic devices. Main topics involve silicon nanophotonic devices, heterogeneous integration of InP-on-silicon, and integrated InP-based optical isolators. He is also working on the development of new fabrication processes for photonic devices, e.g., based on focused ion beam etching and die-to-wafer bonding. He holds three patents, has authored and coauthored over 30 journal papers, and has presented invited papers at several major conferences.

Jan Van Campenhout (M'95) was born in Vleorzove, Belgium, in 1949. He received the M.S. degree in electromechanical engineering from Ghent University, Gent, Belgium, in 1972, and the M.S.E.E. and Ph.D. degrees from Stanford University, Stanford, CA, in 1975 and 1978, respectively. He is currently a Professor teaching courses in computer architecture, electronics, and digital design with the Faculty of Applied Sciences, Ghent University. His current research interests include the study and implementation of various forms of parallelism in computer systems, and their application in programming language support, computer graphics, and robotics. He is currently the head of the ELIS Department in the Faculty of Engineering.

Dirk Stroobandt (S'91–M'98) received the M.S. and Ph.D. degrees, both in electrotechnical engineering, from Ghent University, Gent, Belgium, in 1994 and 1998, respectively. Since October 2002, he has been a Professor with the Electronics and Information Systems Department, Ghent University. His research interests include a priori estimations of interconnection lengths in electronic systems and its applications to CAD, computer architecture evaluation, and design optimization.
SPECIAL SECTION ON SYSTEM-LEVEL INTERCONNECT PREDICTION

Guest Editorial ................................................................. J. Dambre and M. Hutton 853

SPECIAL SECTION PAPERS

Polaris: A System-Level Roadmapping Toolchain for On-Chip Interconnection Networks ............................................ V. Soteriou, N. Eisley, H. Wang, B. Li, and L.-S. Peh 855
Synthesis of Predictable Networks-on-Chip-Based Interconnect Architectures for Chip Multiprocessors .................................. S. Murali, D. Atienza, P. Meloni, S. Carta, L. Benini, G. De Micheli, and L. Ruffo 869
Design of an Interconnect Architecture and Signaling Technology for Parallelism in Communication ..................................... J. Kim, I. Verbaawhehe, and M.-C. F. Chang 881
Predicting Interconnect Delay for Physical Synthesis in a FPGA CAD Flow ................................................................. Y. Manohararajah, G. R. Chiu, D. P. Singh, and S. D. Brown 895
Stochastic Power/Ground Supply Voltage Prediction and Optimization via Analytical Placement ........................................... A. B. Kuhng, B. Liu, and Q. Wang 904
Fast Variational Interconnect Delay and Slew Computation Using Quadratic Models .................................................. X. Ye, F. Y. Liu, and P. Li 913

SPECIAL SECTION TRANSACTIONS BRIEFS

Post-Placement Interconnect Entropy ........................................ W. Feng and J. W. Greene 945
Routability of Network Topologies in FPGAs .............................. M. Saldana, L. Shannon, J. S. Yue, S. Sian, J. Craig, and P. Chow 948

REGULAR ISSUE PAPERS

Thermal-Aware Methodology for Repeater Insertion in Low-Power VLSI Circuits .................................................... J. C. Ku and Y. Ismail 963

TRANSACTIONS BRIEFS

Functional Processor-Based Testing of Communication Peripherals in Systems-on-Chip .................................................. A. Apostolakis, M. Psarakis, D. Gizopoulos, and A. Paschalidis 971
Code Decompression Unit Design for VLIW Embedded Processors ................................................................. Y. Xie, W. Wolf, and H. Lekatsas 975