
Wenjia Zhang, Bing Wang, Zhaomin Zhu, and Kenneth Eng Kian Lee
Singapore-MIT Alliance for Research and Technology (SMART)

Li Zhang
National University of Singapore

Jurgen Michel and Li-Shiuan Peh
Singapore-MIT Alliance for Research and Technology (SMART) and Massachusetts Institute of Technology

Chua Soo-Jin
Singapore-MIT Alliance for Research and Technology (SMART) and National University of Singapore

Editors’ notes:
This paper explores the integration of III-V-based LEDs and detectors in a traditional Si–CMOS process to power an on-die nanophotonic link. The work also shows that the LED-enabled on-chip optical link results in more energy-efficient communication than using an off-chip laser source to power the optical link.

—Sudeep Pasricha, Colorado State University and Yi Xu, Macau University of Science and Technology

Over the Past years, with Moore’s law supplying increasing number of transistors, industry has been boosting processor performance by increasing the number of on-chip cores. This brings about the challenge of designing a power-efficient on-die communication backbone, e.g., network-on-chip (NoC), for delivery of bits between processor cores and memories. Electrical copper interconnects have dominated on-chip communications in commercial processors and thus far satisfied the communication requirements of current multicore processors. However, as the number of cores grows, the chip power budget becomes increasingly constrained and electrical interconnects will be severely limited by their fundamental tradeoff of interconnect bandwidth and distance with power.

The power consumed by electrical interconnects relates to the capacitance of the link, supply voltage, and clock frequency. Since the link capacitance increases with link length and clock frequency (which impacts the link bandwidth), the power consumption of an electrical link grows with distance and bandwidth. Moreover, in order to improve the interconnect latency, long wires are routinely...
segmented into smaller sections, with repeaters inserted in between, increasing total wire coupling capacitance and thus the link power consumption. The scaling of CMOS process technology will not help to mitigate on-chip interconnect power consumption since the long wire capacitance, instead of the gate capacitance, dominates the load capacitance of on-chip interconnects, and the high wiring density on a processor chip results in higher parasitic interconnect capacitance, leading to ever-increasing dynamic switching power for electrical interconnects [1].

New interconnects are needed to ensure scalability to future many-core processors. Among recent disruptive technologies, optical interconnects have the potential to break the aforementioned bandwidth–distance–power tradeoff of electrical interconnects. In general, a complete optical interconnect link is composed of a light source for generating the information carrier, a modulator for electrical/optical (E/O) data transformation, a photodiode for detection, passive components for light guiding, and peripheral electronic devices for driving and biasing the photonic devices. In a photonic link, the light source is the most critical device as it consumes a substantial fraction of total link power. Previously proposed optical network solutions rely predominantly on off-chip lasers as the light source. However, laser sources consume significant power due to their high threshold current; even when the links are used sporadically, the power consumption of lasers remains constant as the communication data are typically modulated externally atop of the continuous wavelengths, resulting in high laser power consumption regardless of actual data transmission through the optical path. Prior studies hence conclude that laser-enabled optical interconnects have challenges replacing electrical interconnects as the on-chip communication fabric for the highly dynamic traffic patterns of many-core processors [1].

In this paper, we see promise in an alternative light source for optical on-chip interconnects: the directly modulated LED. We see two significant strengths in LED-enabled optical interconnects. First, the LED is a reliable light source that turns on without threshold current; therefore, negligible current conducts through the device when the LED is off, thus dissipating little power. Second, substantial power can be saved as an external modulator is no longer needed. External modulators require drivers with several amplification stages that consume a great deal of driving power, especially for high data rate modulation with stringent driving requirements. Furthermore, the insertion loss of a modulator, typically greater than 5 dB, worsens the optical power budget, requiring greater output power from the light source.

However, despite several efforts in LED device design for communications [2], [3], there still exist key challenges in realizing LEDs that can provide practical, low-power, and on-chip sources.

First, LEDs typically have low modulation bandwidth, typically lower than 1 Gb/s. This bandwidth limitation is fundamentally determined by the spontaneous radiative recombination lifetime of injected electrons or holes, presumably in the nanosecond range. However, recent successes of driving the LED to high frequencies have been achieved either by increasing the active layer concentration of electrons and holes [3], [4] or improving the bimolecular recombination [2]. For instance, a directly modulated photonic crystal nanocavity LED was demonstrated operating with a modulation speed of 10 GHz consuming less than 1 fJ/bit of energy [2]. Besides, multicore processor chip frequencies have been plateauing at low gigahertz to alleviate the power consumption of electronics.

Second, it is a challenge to realize a wide variety of photonic devices, including high-speed LEDs and detectors, and visible light transparent waveguides within a converged process platform. Silicon (Si) is widely considered as a future platform of choice for building optoelectronic devices, accommodating both Si–CMOS electronics as well as integrated photonics. For instance, IBM projected ten tera-floating-point operations per second (TFLOPS) on a 22-nm CMOS chip, with separate photonic, memory, and processing planes where the CMOS-compatible Si-based photonic plane is used for connecting various cores and routing over 70 Tbps of on-chip traffic and over 70 Tbps of off-chip traffic [5]. However, since silicon has an indirect bandgap that offers weak interaction between mobile charge carriers and photons, there still remains a hurdle for active photonic devices such as light sources to be fabricated using Si. On the other hand, III–V materials are particularly suitable for photonic devices. It will thus be desirable to have an on-wafer integration technology for on-chip optical interconnects where the
electrical devices are made from CMOS processing while the photonic devices are made of III–V semiconductors.

In this paper, we propose a LED-enabled on-chip optical network that leverages a novel CMOS-compatible III-nitride and Si-integrated process, bringing about ultralow power, small area footprint, and competitive performance versus electrical interconnects on a chip scale. The rest of the paper is organized as follows: Section III-Nitride and Si–CMOS Optical Link Design presents an entire optical link designed in a III-nitride (e.g., GaN) and Si–CMOS integration process where design considerations of III-nitride LED, photodetector, and visible light transparent SiN waveguide are discussed. Section Modeling of LED-Enabled Optoelectronic Link follows with detailed modeling of both its electrical and optical characteristics. Section Case Study: Optical Interconnected SMART Architecture simulates the optical single-cycle multihop asynchronous repeated travers-al (SMART) as a case study by running actual software applications, and evaluates our proposed network against laser-driven NoCs and state-of-the-art electrical NoCs.

III-nitride and Si–CMOS optical link design

As GaN materials need to be grown on a Si(111) substrate with matched lattice orientation whereas Si–CMOS circuits are typically fabricated on a Si(100) substrate, bonding technology is required to integrate Si–CMOS devices and GaN devices on a single wafer [6]. Considering that the growth temperature for GaN materials is too high (around 1000 °C) for the Si–CMOS devices to survive, the Si–CMOS devices need to be fabricated well after GaN device processing. The III-nitride and Si–CMOS integration process flow is similar to that of [6].

Figure 1a shows an optical link fabricated on the III-nitride and Si–CMOS integrated process where CMOS devices and circuits such as processors (with associated caches, routers, and link drivers) are fabricated on a silicon-on-insulator (SOI) wafer and the III-nitride photonic devices such as LEDs, SiN waveguides and photodetectors are processed on the GaN-on-Si wafer. The add–drop optical link (Figure 1b) is similar to a conventional electrical bus. The same III-nitride epitaxial growth layer structure (Figure 1c) can be employed for achieving triple-function optoelectronics devices: exhibiting light-emitting device, photodetector, and transparent pass-through [8]. This versatility of the device is attributed to the strong quantum confined Stark effect (QCSE) in polar (0001) InGaN/GaN multi-quantum well (MQW), which is as follows: the bandgap of the material is shifted with applied electric field and thereby the absorption coefficient versus wavelength is controlled with the applied bias [9].

**LED.** InGaN/GaN MQW LED structures have been widely used as a solid-state light source. When the device with the layer structure in Figure 1c is forward biased (Figure 1d-1), light will be generated and coupled into the waveguide. This integrated MQW LED is specially designed for high-speed on-chip communication due to microsize effects as well as a more efficient usage of injected current [3].

**Photodetector.** After propagating a distance along this waveguide, the optical signal is coupled into the same MQW structure as LED, now functioning as a photodetector absorbing at the LED’s emitting wavelength (Figure 1d-2). The responsivity of such a detector is about 0.1 A/W for the peak emission wavelength [8]. The excess unrecombined electron–hole pairs can be swept out with fast reverse biasing for enhanced high-speed response if necessary.

**Transparent mode.** This device is also capable of guiding the light through with minimal insertion loss by reverse biasing. Due to the increased QCSE as a result of reverse biasing, the overlap between electrons and holes in the MQWs is greatly reduced, leading to a correspondingly decreased absorption coefficient and small absorption loss (Figure 1d-3).

**Waveguide.** SiNx, with refractive index of ~2.1 and bandgap of ~5 eV, is chosen as the candidate of waveguide material that is transparent and low loss (around 1 dB/cm) in the working wavelength of 450 nm. Due to its high refractive index in contrast to SiO2, good light confinement is easily achieved. Besides, SiON, which is a mixture of SiN and SiO2, can supply a tunable refractive index and transparent wavelength, providing a better freedom of design. The waveguide, with core size of 360 nm × 180 nm, is evaluated using PhotonDesign Firmwave, showing that confinement for the TE mode is around 96% and the TM mode around 74% for the working wavelength.
Modeling of LED-enabled optoelectronic link

In this section, we model an entire optoelectronic link driven by a directly modulated III-nitride LED. We start with detailed modeling of the LED device growth structure and material composition to obtain its optical characteristics. Next, the electrical characteristics of the LED are modeled to determine its capacitance and, in turn, the required Si driver. The photonic component layout dimensions and rules are then modeled and captured in a process design kit (PDK) to allow detailed simulation of electrical circuits and the precise interconnection between the photonic and electrical devices.

LED photonic model

The LED model for optical interconnect implemented in PhotonDesign Picwave is used to compare the communication characteristics of MQW GaN LEDs. The epitaxial layers are represented in Figure 1c, and other critical device parameters such as waveguide length/width, top contact size, and resistivity of current flow model are varied through a sweep across the target design parameters. An output power monitor is configured to detect the fraction of spontaneously emitted light coupled out of the emitting section facet. The monitor wavelength is set to 450 nm with a response time of 0.01 ns.

The output power increases exponentially with increasing injected current as we can see from the inset of Figure 2a. Therefore, lower current is required for direct modulation due to this threshold-less characteristic, leading to lower power consumption in contrast to the case where external modulation is applied to a continuous lasing source.

Figure 1. (a) Optical link comprising LED, SiN$_x$ waveguide, and photodetector. (b) High-level representation of a multiple add-drop optical link with electrical routers controlling the triple-function devices. (c) Epitaxial growth layer structure for LED and PD. (d) Modes of the proposed triple-function device: (d-1) lighting mode; (d-2) detector mode; and (d-3) transparent waveguide mode.
It also shows that the smaller LEDs have better enhanced quantum efficiency, most likely due to high carrier concentration within reduced active layers. Heating effects, which were not considered in the simulation, would be more prominent since the heat dissipation is more difficult in smaller LEDs. However, this challenge can be resolved with improved packaging processes [3].

The frequency response of the light source is measured from the impulse response in the simulations. Figure 2c shows the frequency response of various device sizes recorded with different impulse amplitude. The LED's geometry is modeled as a square with the same length and width. The higher 3-dB bandwidth for LED with reduced size can be explained by the enhanced radiative recombination rate and small capacitance in smaller LEDs. This is in good agreement with the experimental measurements for an individual microdisk blue LED with diameter of 5 μm [3].

We also evaluate the effect of driving current on 1-Gb/s modulation quality. When the LED is directly driven by nonreturn zero (NRZ) data, the extinction ratio of the modulated data is varied with the value of current because the radiative recombination that generates the spontaneous emission is enhanced with higher injected current. The electrical eye diagrams at high current produce clearer eye openings than at low current. The Q-factor, the vertical opening of the eye relative to noise, is calculated to provide a single figure of merit to quantify the modulation quality of the LED device. Figure 2d shows the Q-factor of the eye diagram from the photodetector with various input driving currents. Assuming

![Figure 2. (a) Output power versus input current (L-I). (b) Voltage versus input current (V-I) of different sizes (10 and 20 μm). (c) Frequency response for different values of device size. (d) One-gigabit-per-second modulation quality (Q-factor) versus injected current for 10- and 20-μm LEDs.](image-url)
a target bit error rate of $10^{-9}$, similar to that of existing copper interconnect, a $Q = 6$ is desired, which translates to driving currents for 10- and 20-$\mu$m LED of about 0.06 and 0.2 mA with output power of 30 and 100 $\mu$W, respectively.

LED circuit model

A basic LED-equivalent circuit can consist of a combined depletion and diffusion capacitance, parallel/serial resistor, and build-in voltage source. The capacitance of the LED can be explained by charge storage during carrier flow. When the LED is switched on, the depletion capacitance dramatically reduces and the charges injected into the neutral region cause diffusion capacitance, which becomes the dominant factor in the switching speed of the device. The capacitance can be modeled as

$$C = A \frac{q \varepsilon \varepsilon_0 (N_D - N_A)}{2 (V_{bi} - V - \frac{KT}{q})}$$

where $A$ is the LED cross-sectional area, $N_D$ is the donor concentration, $N_A$ is the acceptor concentration, $\varepsilon_0$ is the permittivity (for GaN, $\varepsilon_0 = 9.5 \varepsilon_0$), $V_{bi}$ is the built-in potential, and $KT/q$ is the thermal voltage. For example, with this model, a $10 \times 10^{-3} \mu$m$^2$ GaN LED has estimated capacitance of around 6.3 fF.

The device capacitance is critical for determining the Si-driver size needed to meet the data-rate requirement [1], [10]. The Si driver is modeled as an inverter chain predriver followed by a final driver stage, and the energy per bit of the driver can be calculated as in [10].

PDK design

A PDK for a GaN LED is designed with an entire set of device information based on the proposed process. The mask layout of the LEDs, the electrical layout of the GaN-LED, and the Si driver are presented in Figure 3a and b. For the 0.25-µm technology node that we are targeting, layout design rules are also determined to permit design rule check (DRC) and layout versus schematic (LVS) check with conventional Cadence electrical VLSI design tools.

The Si drivers control the current and bias voltage of the LED to generate light, detect light, or transparently pass through light. The model of the Si driver is from Global Foundries’ PDK, while the device model of the GaN LED is implemented in verilog-A. The specific metal thickness of the via is 0.49 µm, sheet resistance is 0.077 Ω/µm$^2$, the plug thickness is 2 µ, and resistivity is $8.4 \times 10^{-2}$ Ω/µm$^2$, the plug thickness is 2 µ. Supposing the metal length is 20 µm and the width is 5 µm, this metal interconnect has a resistance of 0.32 Ω and a capacitance of 1.7 fF. The PDK development enables the same, conventional circuit simulation, place-and-route, and layout design process to be used for the CMOS electronic portion of the die with the hybrid Si and GaN process.

Case study:
Optical interconnected SMART architecture

Most previously proposed designs of photonic NoCs use lasers as sources and microring resonators as modulators, detectors, and routers [1], [11], [12]. Thus, they leverage multiple wavelengths with associated filters and design NoC architectures such as buses and token rings which suit 1-to-many connections. Unlike these designs, LED is an incoherent light source, so LED-based NoCs cannot use resonant devices for filtering, modulating and switching. Instead, LED-based links can only function with NoC architectures that can multiplex traffic flows atop 1-to-1 connections, i.e., where control and switching needs to be done with electrical routers. Yet, simply replacing electrical 1-to-1 links with our LED-based links can lead to exorbitant energy

Figure 3. (a) LED’s mask layout. (b) PDK design of GaN LED and Si driver.
consumed in the optical–electrical–optical (O/E/O) crossings that overwhelm the energy benefits.

Here, we present a case study exploring the suitability of a recently proposed NoC architecture SMART [7], which was originally architected for electrical clockless repeated links to realize a single-cycle data path across the entire die. SMART's ability to bypass intermediate routers within the same cycle allows our LED links to traverse long distances without the unnecessary O/E/O crossings in between, making the mesh topology feasible for our LED links. Any other NoC topologies that rely on 1-to-1 connections will also function with our LED links, but the energy benefits depend on the number of O/E/O crossings into electrical routers. In LED-enabled mesh (with SMART) and clos, we compare the NoC power performance of our LED links within the SMART mesh as well as a conventional clos topology.

SMART was proposed as a solution to break the latency barrier for NoCs, but it still consumes 28–32 fJ/b/mm, leading to a worst case transmission energy of 600 fJ to send a bit from one chip edge to the other on a typical 20 mm² chip. We see the potential of embedding LED-enabled photonic links into the SMART architecture to further break the power barrier of on-chip communications. SMART allows messages to dynamically arbitrate and create multihop bypass paths across the chip on-demand over a shared network fabric. Messages are only buffered at intermediate routers upon contention. By bypassing intermediate electrical routers, a message is allowed to transverse from source to destination electrical routers, avoiding the high energy overhead of intermediate electrical routers in most cases. Within the LED-enabled SMART, the electrical links are replaced with the proposed optical links containing LEDs, waveguides, and photodetector (PD), leveraging the distance-independent low-power transmission of photonics. SMART is set to prioritize bypassing messages over incoming ones, i.e., new messages are not injected into the NoC until existing messages already in the NoC are drained out. So, a message usually stays within the photonic plane from source to turning router, goes through the electrical turning router, then again zooms through the photonic plane from the turning to destination router, limiting the E/O/E. We can take an 8 × 8 optical interconnected mesh SMART NoC as an example: noting that two optical layers are generated to eliminate in-plane waveguide crossing loss, each layer containing eight 1 × 8 LED add–drop nodes as shown in Figure 1b; and optical data are converted to electrical signals in the turning router from one plane to another. Our III-nitride and Si–CMOS integration process enables the optical links to be closely integrated with the CMOS routers and processors.

Link energy evaluation

We evaluate the energy efficiency of our LED-enabled link against a baseline electrical clockless repeated link in the 45-nm node and a laser-enabled link using DSENT, a timing-driven NoC power modeling tool [1]. The laser-enabled optical link is composed of off-chip laser, microring modulator, receivers, and peripheral electrical devices. We use the electrical LED model described in the LED Circuit Model section to estimate the Si-driver size in DSENT. In particular, the effective capacitance of the LED (6.3 fF) and parasitic capacitance of the vias (1.7 fF) are used to size the driver and its power consumption. Waveguide loss is set to be 1 dB/cm, as discussed in the III-nitride and Si–CMOS Optical Link Design section, and responsivity for the detector is set to 0.1 A/W. For a short distances (< 8 mm), most power is consumed by the electrical circuits in the optoelectronic link. Therefore, it is shown in Figure 4a that the energy for electronic links increases linearly while the optical links remain almost the same regardless of the transmitting distance. The LED-enabled links outperform other designs with power efficiency of around 38 fJ/b.

64-core processor simulations

We run the parallel sections of all 64-threaded SPLASH-2 applications on an 8 × 8 multicore chip with shared L2 cache, averaging the multiple applications’ results. Two electrical NoC baselines are used: a state-of-the-art NoC with single-cycle-pipeline routers and a SMART NoC. All results are normalized against the single-cycle router. Note that both electrical NoC baselines are highly optimized, beating recent industry chip prototypes such as Intel 48-core SCC with a three-cycle router in latency and energy. Performance-wise, the electrical SMART delivers 5–8× lower latency than the single-cycle router electrical baseline, and our photonic SMART NoC maintains this performance benefit. Figure 4b graphs the energy consumption averaged across all SPLASH applications. The electrical SMART has a slight energy
advantage over the baseline single-cycle router due to savings in buffering at intermediate routers, while the photonic SMART substantially reduces the link and crossbar dynamic energy by 68% and 37% respectively, leading to overall energy savings of about 28% over the electrical SMART, across all applications.

LED-enabled mesh (with SMART) and clos

Next, we compare two alternative 64-tile network topologies with LED links in a 20 × 20-mm² chip. Besides 8 × 8 mesh (with SMART), we simulated an 8-ary, 3-stage 8 × 8 × 8 clos (without SMART) topology with LED links and an electrical equivalent, where router-to-router links are replaced by photonic links of equal latency and throughput (128 wires, each at 1 GHz) and core-to-ingress and egress-to-core links are electrical [7], [12]. The other network configurations and default technology parameters remain the same as in [1] and [7]. Clos is chosen as it works with 1-to-1 connections, and has few intermediate routers, minimizing O/E/O crossings.

All-optical switches are used for the laser-enabled network, assuming negligible switching power at the turning routers. Synthetic traffic to uniformly distributed destinations at varying injection rates is used to explore power performance tradeoffs in Figure 4c and d. (The SPLASH-2 applications earlier have low injection rates that will not highlight the benefits of the clos topology.) Even with the O/E/O power penalty when switching through electrical routers, the LED-enabled networks have energy advantage over the electrically based mesh/clos network by more than 50% while laser-enabled network only performs better than the electrical counterparts at high load due to its high non-data-dependent power. This advantage of LED-enabled links is attributed to excellent property of LED-enabled link: negligible power consumption when it is off and ultralow-power consumption when it works. We note though that, furthermore, comprehensive evaluations will have to be carried out for a clearer comparison between laser and LED-enabled NoCs.

Figure 4. (a) Energy efficiency for electronic repeated links, optical laser link, and optical LED link working at 1 GHz. (b) Normalized dynamic network energy for electrical SMART and optical SMART for SPLASH applications on a 64-core processor; normalized network power consumption varied with the achievable throughput for Mesh (c) and clos (d).
We proposed an LED-enabled on-chip optical network that leverages a novel III-nitride and Si–CMOS integrated process. This process allows transistor-level integration of CMOS circuits and III-nitride photonic devices, enabling fabrication of a complete optical link on-chip with light sources, waveguides, photodetectors, and driving circuits. From detailed modeling of the LED device growth structure and material composition, the simulation results, including output power, frequency response, and communication quality, are shown to meet the requirements of on-chip communications. The photonic component layout dimensions and rules are also modeled and captured in a PDK, enabling existing circuit simulation and place-and-route design flow to continue to be used for the CMOS electronics portion of the die. We embed our LED-enabled optical link design into the SMART NoC architecture, an architecture originally proposed for electrical interconnects but which is well matched for the characteristics of our LED-enabled links. Results from system level simulations of a 64-core processor show 28% more energy efficiency compared to fully electrical SMART NoC, which already beats prior electrical designs in energy and delay. In addition, more than 50% power saving is observed across the achievable throughput for mesh and close topologies. We hope these results based on a collaboration with our team members of material, process, device, circuit, and architecture researchers will pave the way for energy-efficient designs of future on-chip optical communication. We have since fabricated samples of the proposed LED MQW structure and are in the midst of testing and measurements.

Acknowledgment

We thank Dr. Z. Liu for detailed discussions regarding the material and process, and Dr. T. Krishna and O. Chen for their assistance with system-level simulations.

References


Zhaomin Zhu, biography not available at the time of publication.

Kenneth Eng Kian Lee, biography not available at the time of publication.

Li Zhang, biography not available at the time of publication.

Jurgen Michel, biography not available at the time of publication.

Li-Shiuan Peh, biography not available at the time of publication.

Chua Soo-Jin, biography not available at the time of publication.

Direct questions and comments about this article to Wenjia Zhang, Low Energy Electronic Systems (LEES), Singapore-MIT Alliance for Research and Technology (Smart), Singapore; wenjia@smart.mit.edu.