

# A Comparative Analysis of Front-End and Back-End Compatible Silicon Photonic On-Chip Interconnects

Ishan G Thakkar, Sai Vineel Reddy Chittamuru, Sudeep Pasricha  
Department of Electrical and Computer Engineering  
Colorado State University, Fort Collins, CO, U.S.A.  
{ishan.thakkar, sai.chittamuru, sudeep}@colostate.edu

## ABSTRACT

Photonic devices fabricated with back-end compatible silicon photonic (BCSP) materials can provide independence from the complex CMOS front-end compatible silicon photonic (FCSP) process, to significantly enhance photonic network-on-chip (PNoC) architecture performance. In this paper, we present a detailed comparative analysis of a number of design tradeoffs for CMOS front-end and back-end compatible devices for silicon photonic interconnects. A cross-layer optimization of multiple device-level and link-level design parameters is performed to enable the design of energy-efficient on-chip photonic interconnects using BCSP devices. The optimized design of BCSP on-chip links renders more energy-efficiency and aggregate bandwidth than FCSP on-chip links, in spite of the inferior opto-electronic properties of BCSP devices. Our experimental analysis compares the use of BCSP and FCSP links at the architecture level, and shows that the optimized design of the BCSP-based Firefly PNoC achieves  $1.15\times$  greater throughput and 12.4% less energy-per-bit on average than the optimized design of FCSP-based Firefly PNoC. Similarly, the optimized design of the BCSP-based Corona PNoC achieves  $3.5\times$  greater throughput and 39.5% less energy-per-bit on average than the optimized design of FCSP-based Corona PNoC.

**Categories and Subject Descriptors:** [Networks] Network on chip; [Hardware] Integrated Circuits/Interconnect: Photonic and optical interconnect

**Keywords:** Photonic network on chip; design tradeoffs; optimization; aggregate bandwidth; energy efficiency

## 1. INTRODUCTION

Recent advances in silicon photonics (SiP) based on the silicon-on-insulator (SOI) process have produced high performance building blocks such as modulators, detectors, filters, and switches that are highly desirable for high-bandwidth and energy-efficient on-chip photonic interconnects [1]-[4]. However, the SOI platform restricts SiP circuits to a single layer, which limits the number of devices that can fit on a chip. Also, the modern SOI process offers a very thin layer of buried oxide (BOX) (200nm thick BOX at 45nm and thinner for advanced technology nodes), which does not provide the necessary optical isolation required to guide light into SiP devices, resulting in large optical losses due to scattering [5]. To address these issues, recent efforts have proposed back-end integration of SiP devices with CMOS logic. In [6], electro-optic polymer and germanium, and in [7] III-V compounds are used as the active materials. However, fabrication of SiP devices using polymer based or III-V

compound based materials requires heterogeneous integration with CMOS logic, which is very costly, requiring specialized foundries.

As a solution to these limitations, Lee et al. in [5] discussed the use of back-end compatible silicon nitride (SiN) material to produce low-loss passive optical waveguides and the use of excimer laser annealed (ELA) quasi-single-crystalline polysilicon (pSi) and polycrystalline germanium (Ge) to produce active microring modulators and detectors. Traditionally, the photonics community has largely ignored pSi due to the challenges introduced by its high optical losses and inferior electrical properties. Similarly, the stress issues complicating the deposition of SiN films thick enough for guiding in the telecom wavelength range have limited the use of low-loss SiN waveguides only for visible wavelengths [5]. However, recent advances in back-end integration technology have led to several pSi and SiN devices being demonstrated with performance and loss values comparable to front-end integrated crystalline silicon (cSi) devices [8]-[11].

In this paper, we refer to SiP devices made of pSi and SiN materials as back-end compatible SiP (BCSP) devices, whereas we refer to SiP devices made of front-end integrated cSi material as front-end compatible SiP (FCSP) devices. BCSP devices provide independence from complex CMOS front-end processes. Moreover, the possibility of low-temperature multi-layer deposition of pSi and SiN materials on top of CMOS metallization layers, as demonstrated in [9], enables multi-level integration for 3D photonic networks-on-chip (PNoCs) on a logic chip. Thus, BCSP has a multitude of benefits over FCSP, which favors the use of such devices in the PNoCs of the future.

The design and characteristics of active and passive SiP devices control the feasibility, reliability, and performance of the entire SiP PNoC. Therefore, the designers of PNoCs should follow a strict set of device-level design guidelines to ensure good system performance. Existing device-level design guidelines, as presented in [12] and [13], are prepared for FCSP-based devices and systems. But the optical and electrical properties of BCSP devices are different from those of FCSP devices [5], which implies that a distinct set of design guidelines are required for BCSP systems. *For the first time, in this work we analyze a number of device-level tradeoffs for BCSP devices to derive design guidelines for BCSP-based PNoC architectures.*

From our analysis of device-level tradeoffs, we observed that the design of energy-efficient, low-noise, and high-aggregate-bandwidth BCSP interconnects requires cross-layer optimization of a number of interdependent device-level and link-level parameters. In recent years, several works have discussed such cross-layer optimization of parameters for FCSP interconnects [14]-[17][39]-[41]. In [39] and [14], the impact of fabrication-induced process variations and power-induced thermal variation on FCSP devices and its impact on the reliability, power dissipation, and performance of FCSP PNoCs was studied. Mohamed et al. in [15] presented analytical models of FCSP devices and analyzed the design tradeoffs for their applications at the network level. In [16], a high-aggregate-bandwidth microring link is analyzed to determine energy-efficiency and bandwidth-density for the link using best-of-class FCSP devices. Hendry et al. in [17] present physical layer analysis and modeling of FCSP-based dense wavelength division multiplexed (DWDM) bus architectures. In [40] and [41], optimized photonic link architectures comprised of FCSP devices are used to achieve high-bandwidth and energy-efficient data

transfers between core and off-chip memory. *Unlike any of these prior works, we perform a cross-layer analysis of design tradeoffs for BCSP interconnects and compare the results of this analysis with the results of a similar analysis for FCSP interconnects.* Our results provide a better understanding of available design choices for realizing energy-efficient and terabyte-per-second scale PNoCs.

We summarize the key contributions in this paper as follows:

- We present and analyze a number of device-level design tradeoffs for BCSP devices involving Q-factor, optical power loss in microring cavity, and modulator bit-rate as a function of radius;
- We characterize interdependence between various device-level and link-level design parameters of BCSP devices, and perform cross-layer optimization of these parameters, to realize energy-efficient and high-aggregate-bandwidth BCSP on-chip links;
- We perform a similar cross-layer analysis and optimization for FCSP devices and compare results with those for BCSP devices;
- We evaluate the impact of optimized designs of FCSP and BCSP links on the performance and energy-efficiency of two well-known PNoC architectures: Corona [18] and Firefly [19].

## 2. ANALYSIS OF DESIGN TRADEOFFS

A typical PNoC consists of microring resonators (MRs) that are coupled to one or more photonic DWDM bus waveguides (WGs) [42]-[44]. These MRs serve as modulators, filters, and switches. We direct the reader to [15] for more details on MR design and operation.

The feasibility, reliability, energy-efficiency, and performance of PNoCs depend on various device-level and link-level design parameters. Our goal in this section is two-fold: (1) to understand and analyze the tradeoffs present among various device-level and link-level design parameters of PNoCs; (2) to understand how these tradeoffs differ between BCSP and FCSP based PNoCs. As a first step towards achieving these goals, we present analytical models of BCSP and FCSP devices (Section 2.1). Then, using these models, we analyze the tradeoffs among various device-level (Section 2.2) and link-level (Section 2.3) design parameters for BCSP and FCSP devices.

### 2.1 BCSP and FCSP Device Modeling

In a PNoC, MRs are coupled to one or more DWDM bus WGs, and serve as modulators, filters, and switches [15]. For a passive component such as a filter, the MR can be considered as a looped photonic WG with a small diameter. For an active component such as a modulator, the MR's looped WG is doped such that it may be addressed as a PN junction device. The tradeoffs among the design parameters of a DWDM bus WG are mostly straightforward. In contrast, the resonant nature of an MR creates several complex tradeoffs among its design parameters. For this reason, in this subsection we present analytical device models for passive and active MRs. These models are equally relevant for BCSP and FCSP types of MRs, as they both have similar geometry, and work on the same principle.

#### Models for Passive Microring Resonators:

A passive MR acts as a bandpass filter, the characteristics of which are defined by the resonant wavelength ( $\lambda_r$ ), round-trip optical loss ( $a^2$ ), and Q-factor. The Q-factor of a passive MR that is coupled to a WG is known as loaded Q-factor  $Q_L$  [21], which is inversely proportional to the full width of its passband at half the maximum (FWHM) transmission. The  $Q_L$ ,  $a^2$ , and  $\lambda_r$  parameters, assuming a critical coupling of the MR to a WG, can be expressed as [21][24]:

$$Q_L = \frac{2\pi^2 n_g R a}{\lambda_r (1 - a^2)}, \quad (1)$$

$$\lambda_r = (2\pi R n_{eff})/m, \quad (2)$$

$$a^2 = \exp(-2\pi R(\alpha_i + \alpha_b + \alpha_d)), \quad (3)$$

$$\alpha_b = C1 * \exp(-C2 * R), \quad (4)$$

where,  $R$  is MR radius;  $m$  is the resonant mode number;  $n_{eff}$ ,  $n_g$ ,  $C1$ , and  $C2$  are constants; and  $\alpha_i$ ,  $\alpha_b$ , and  $\alpha_d$  are loss coefficients. The definitions and typical values of these constants are given in Table 1.

From Eq. (1)-(4), *the device-level parameters of a passive MR device such as round-trip optical loss ( $a^2$ ), resonant wavelength ( $\lambda_r$ ), and loaded Q-factor ( $Q_L$ ) ultimately depend on the MR radius ( $R$ ).*

#### Models for Active (Doped) Microring Resonators:

A doped MR acts as a modulator, a filter, or a switch, the characteristics of which are defined by the values of  $\lambda_r$ ,  $Q_L$ ,  $a^2$ , bit-rate, free-spectral range ( $FSR$ ), and modulation shift ( $\Delta\lambda_r$ ). Similar to passive MRs, Eq. (1)-(4) hold for doped MRs too. So, the values of  $Q_L$  and  $a^2$  depend on  $R$  for doped MRs as well.

Doped MRs are doped in a similar manner as PN junctions. The free carrier concentration in a PN junction based MR can be controlled by applying forward or reverse biased voltage across the junction. The change in free carrier concentration alters the optical properties of the MR owing to the free carrier dispersion (FCD) and the free carrier absorption (FCA) effects [25]. The FCD effect alters the refractive index  $n$  and the FCA effect alters the absorption related loss coefficient  $\alpha_d$ . The change in  $n$  in turn leads to a shift in the passband of the MR. The passband shift affects the light transmission from the source to the MR output, thereby achieving modulation, filtration, or switching of the input light signal. We assume the PN-junctions of doped MRs to be reverse-biased, as the doped MRs with reverse-biased PN-junctions render faster electrical response for high bandwidth modulation [27]. We also assume the doping concentrations of  $N_a = N_d = 3 \times 10^{18} \text{ cm}^{-3}$  ( $N_d$  for electrons in N-region and  $N_a$  for holes in P-region), as assumed in prior work [8].

We also study the effect of *MR radius* on bit-rate of a doped MR. As discussed in [27], the bit-period (and hence bit-rate) of a reverse-biased PN-junction based MR is limited either by the  $R_s C_j$  time constant (where  $C_j$  is junction capacitance and  $R_s$  is series resistance) or by the photon lifetime of the MR, depending on which of the two is greater.  $C_j$  depends on the junction area, which in turn depends on the *MR radius*. The photon lifetime for an MR device is given by  $\tau_p = (Q_L \lambda_r / 2\pi c)$ , which is a function of  $Q_L$  [27]. As explained earlier,  $Q_L$  of the MR depends on the radius ( $R$ ), which implies that the photon lifetime of an MR also depends on  $R$ . Moreover, the resonance of an MR cavity is cyclic in nature, and the free spectral range ( $FSR$ ; wavelength range between two successive resonances of an MR), is defined as [21]:  $FSR = \lambda^2 / 2\pi R n_g$ .

*In summary, the device-level parameters of a doped MR such as round-trip optical loss ( $a^2$ ), loaded Q-factor ( $Q_L$ ), bit-rate ( $R_s C_j$  time or photon lifetime), and FSR ultimately depend on the MR radius ( $R$ ).*

**Table 1: Definitions and typical values of some constants for MRs**

	Definition	Value	
		BCSP	FCSP
$n_{eff}$	Effective refractive index of MR [22]	2.49	2.45
$n_g$	Group refractive index of MR [22]	4.26	4.21
$n_{SiO2}$	Refractive index of SiO <sub>2</sub> cladding [8]	1.48	
$n$	Refractive index of an MR's looped WG core [37][38]	pSi	cSi
		3.48	3.47
$C1$	Coefficients based on the material and geometry of MR [22]	132	126
$C2$		10	10.1
$R_s$	Series resistance of MR [8][28] (in $\Omega$ )	750	250
$\alpha_i$	Intrinsic optical loss due to bulk defects and surface roughness in MR [8][23] (in $\text{cm}^{-1}$ )	3.87	2
$\alpha_d$	Optical absorption loss in MR (in $\text{cm}^{-1}$ )	0.23	0.23
$\alpha_b$	Bending loss due to MR curvature	Eq. (4)	
-	Cross-section dimensions of MR's looped WG	450nm×250nm	

### 2.2 Device-Level Design Tradeoffs

In this subsection, first, we present design tradeoffs for doped and passive MRs and then we present tradeoffs for passive WGs, for both FCSP and BCSP types of implementations.

#### Active/Passive Microring Resonators:

As concluded in Section 2.1, various device-level design parameters of passive and active (doped) MRs ultimately depend on *MR radius* ( $R$ ). This dependence of design parameters on  $R$  exists for both

BCSP and FCSP MRs, because MRs in both cases operate on the same principle. The values of coefficients  $C1$ ,  $C2$ ,  $\alpha_i$ ,  $n_g$ ,  $R_s$  and  $n_{eff}$  decide the degree by which various design parameters depend on  $R$ . The values of  $C1$ ,  $C2$ ,  $n_g$  and  $n_{eff}$  depend on the refractive index of MR materials and the device geometry.

BCSP MRs are made of pSi (core)-SiO<sub>2</sub> (cladding), whereas FCSP MRs are made of cSi (core)-SiO<sub>2</sub> (cladding), with both types of MRs having the same device geometry. The optical properties of pSi and cSi are marginally different, as pSi exhibits high intrinsic optical loss due to surface roughness, grain boundaries, and dangling bonds [5]. As a result, values of  $C1$ ,  $C2$ ,  $\alpha_i$ ,  $n_g$ ,  $R_s$  and  $n_{eff}$  differ between FCSP and BCSP MRs, causing the degree by which various device-level design parameters depend on  $R$  to differ for BCSP and FCSP MRs.

For this study, we modeled BCSP and FCSP MRs (both active/doped and passive) with the cross-sectional dimensions of 450nm×250nm, using the finite difference method [22]. For these models, we used the refractive index values  $n$  and  $n_{SiO_2}$  from Table 1 and calculated the values of  $C1$ ,  $C2$ ,  $n_g$  and  $n_{eff}$  for  $\lambda = 1600$ nm, which are also given in Table 1. We explain the reason behind using  $\lambda = 1600$ nm later when we explain the design tradeoffs for WGs. Using these values of the coefficients, we calculated the values of various design parameters using the equations presented in Section 2.1.

Figure 1 shows the various device-level design parameters such as  $R_sC_j$  time delay, photon lifetime, round-trip optical loss ( $a^2$ ), loaded Q-factor ( $Q_L$ ), and FSR versus the MR radius ( $R$ ) for BCSP and FCSP MRs. We use the equations given in [26] to model  $C_j$  for BCSP and FCSP MRs. From the figure, it can be observed that the degree by which the values of  $Q_L$ ,  $a^2$ ,  $R_sC_j$ , FSR and photon lifetime depend on the MR radius ( $R$ ) differs between BCSP and FCSP MRs. The round-trip cavity loss ( $a^2$ , shown with red lines in Figure 1(a)) of a BCSP MR is greater than that for an FCSP MR for all values of  $R$ . This is due to the higher loss coefficients for BCSP MRs (Table 1). The larger value of round-trip loss in case of a BCSP MR results in a smaller value of  $Q_L$  (green lines in Figure 1(a)). The smaller  $Q_L$  of a BCSP MR results in a broader passband compared to an FCSP MR, which leads to higher insertion loss for a BCSP MR. Nevertheless, our analysis in Section 3.4 finds that the optimal design of BCSP links made of BCSP MRs renders more energy-efficiency than the optimal design of FCSP links made of FCSP MRs.

As described in [27], the rise-time and fall-time, and hence the bit-period of an MR is controlled by either the  $R_sC_j$  time delay or the photon lifetime, depending on which one of the two is greater. From Figure 1(b), the photon lifetime (blue lines) of FCSP and BCSP MRs is greater than their  $R_sC_j$  time delay (green lines), which implies that the bit-rate (inverse of bit-period) of BCSP and FCSP MRs is limited by the photon lifetime. In addition, the photon lifetime of FCSP MRs is greater than BCSP MRs, which leads us to the important conclusion that the bit-rate of BCSP MRs is greater than bit-rate of FCSP MRs for all values of MR radius.

### Passive Waveguides (WGs):

Next, we discuss the design tradeoffs of FCSP and BCSP passive WGs. Typically, FCSP WGs are fabricated using cSi core and SiO<sub>2</sub> cladding, whereas BCSP WGs are made of SiN core and SiO<sub>2</sub> cladding. The SiN-SiO<sub>2</sub> WGs have very high propagation loss (about 6dB/cm) in the C-band due to N-H and Si-H bond absorption harmonics, therefore, SiN-SiO<sub>2</sub> WG systems are typically operated in the L-band (near 1600nm) where they exhibit lower propagation loss (about 1dB/cm) [5]. Because of this reason, we analyze all the device-level parameters discussed in the preceding subsection for the 1600nm operating wavelength. As discussed in [5], due to the ability of multilayer integration, superior coupling characteristics, and comparable propagation loss, the BCSP SiN-SiO<sub>2</sub> WGs outperform the FCSP cSi-SiO<sub>2</sub> WGs despite having higher scattering losses.

Furthermore, the maximum allowable optical power (MAOP) in

SiN-SiO<sub>2</sub> and cSi-SiO<sub>2</sub> WGs is limited due to the emergence of non-linearity effects at higher optical power, which incurs additional signal loss and degrades the performance of these WGs. The BCSP SiN-SiO<sub>2</sub> and FCSP cSi-SiO<sub>2</sub> WGs exhibit different types of nonlinear optical effects. The dominant nonlinear optical effects in the FCSP cSi-SiO<sub>2</sub> WGs are the two-photon absorption (TPA) effect and the resulting FCD and FCA effects [13]. The TPA induced FCA effect limits the MAOP in an FCSP cSi-SiO<sub>2</sub> bus WG to 100mW [13][17]. In contrast, due to the absence of free carriers in SiN material, the TPA effect and the resulting FCA effect are not present in BCSP SiN-SiO<sub>2</sub> WGs [29]. However, the dominant nonlinear optical effects in the FCSP SiN-SiO<sub>2</sub> WGs are the second and third harmonic generation, which limits the MAOP in a BCSP SiN-SiO<sub>2</sub> bus WG to 350mW [29]. It will be evident from the discussion in Section 2.3 that a higher value of MAOP ultimately results in a larger number of DWDM channels in a SiN-SiO<sub>2</sub> BCSP bus WG than in FCSP cSi-SiO<sub>2</sub> WGs.

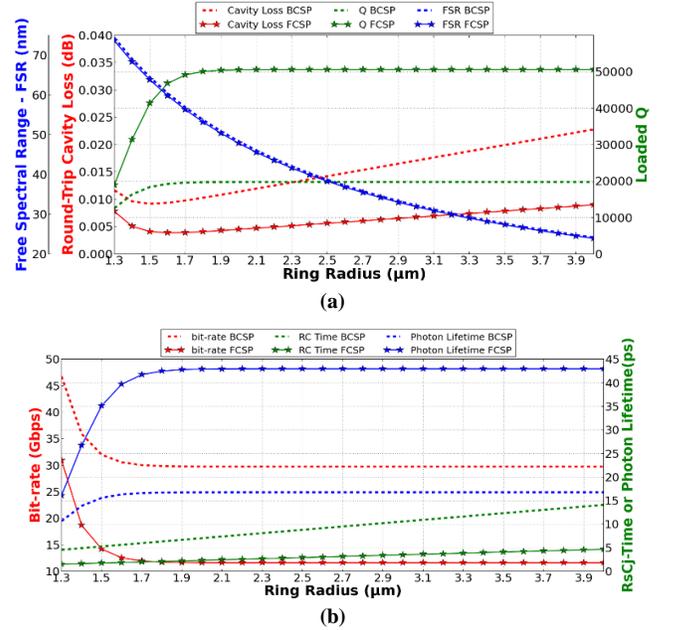


Figure 1. (a) Loaded Q factor, round-trip cavity loss, FSR, (b)  $R_sC_j$  time delay, photon lifetime, and bit-rate vs. MR radius for BCSP and cSi FCSP MRs. The curves of BCSP FSR and FCSP FSR are overlapped.

### 2.3 Link-Level Design Tradeoffs

In section 2.2, we presented the design tradeoffs among various device-level parameters such as MR radius,  $Q_L$ , bit-rate, nonlinear power limit, and FSR. In this subsection, we analyze how these parameters would affect design decisions at the higher link-level.

An on-chip SiP link typically comprises of a group of modulator MRs, a group of detector MRs with photodetectors, and a DWDM bus WG. The photonic signal transmission in on-chip SiP links is inherently lossy, i.e., the light signal is subject to losses such as insertion loss and modulation crosstalk related loss in modulator MRs, insertion loss and sideband truncation related loss in detector MRs, and propagation and bending loss in WGs. All wavelength channels of a DWDM WG are subject to these losses. To ensure that signals of all channels propagating through the SiP link reach their destination before attenuating below the sensitivity threshold of the detector (minimum detectable power), the aggregate loss of all the channels along that link must fall within an acceptable range. This constraint is called the optical power budget and can be calculated in dB as the difference between the MAOP and the detector sensitivity. The optical power budget in dB ( $P_{Budget}^{dB}$ ) determines how much loss can be present in the SiP link [13], which can be summarized as [17]:

$$P_{Budget}^{dB} \geq P_{Loss}^{dB} + 10 \log_{10}(N_\lambda), \quad (5)$$

where  $N_\lambda$  is the number of wavelength channels used in the link, and  $P_{Loss}^{dB}$  represents the sum of the loss contributions (in dB) incurred on a single channel by all the components (WG, detector and modulator MRs) present along the SiP link.

In this study, we assume the shot-noise limited sensitivity threshold of -22dBm for the FCSP photodetectors, as used in [17]. Due to the adverse effects of grain boundaries and dangling bonds, BCSP photodetectors are inherently more susceptible to noise than FCSP ones. Therefore, we assume a greater value of sensitivity threshold (-20dBm) for the BCSP photodetectors. From Section 2.2, the TPA-effect limited MAOP for an FCSP WG is 20dBm (100mW), whereas the harmonic generation effect limited MAOP for a BCSP WG is 25.4dBm (350mW). As a result, an FCSP link has  $P_{Budget}^{dB} = 42$ dB, whereas a BCSP link has  $P_{Budget}^{dB} = 45.4$ dB. The higher value of  $P_{Budget}^{dB}$  for the BCSP link allows a larger amount of aggregate loss ( $P_{Loss}^{dB} + 10\log_{10}(N_\lambda^{PB})$ ) to be present in the BCSP link than in FCSP links.

For a given value of single channel loss  $P_{Loss}^{dB}$ , the  $N_\lambda$  in Eq. (5) should be less than a threshold value to limit the aggregate loss of the link within the power budget ( $P_{Budget}^{dB}$ ). This threshold value (denoted as  $N_\lambda^{PB}$ ) gives a  $P_{Budget}^{dB}$ -limited number of channels per WG. Along with the  $P_{Budget}^{dB}$ , the FSR of the largest MR along the WG also limits the number of channels per WG. The FSR-limited number of channels is given as  $N_\lambda^{FSR} = FSR/CS$ . Here,  $CS$  represents channel spacing, which is the distance between two adjacent wavelength channels of the SiP link. The actual feasible number of channels ( $N_\lambda^{Act}$ ) per WG should be less than or equal to both  $N_\lambda^{PB}$  and  $N_\lambda^{FSR}$ . For a small enough value of  $P_{Loss}^{dB}$ , a given SiP link can have  $N_\lambda^{PB} > N_\lambda^{FSR}$ . In this case,  $N_\lambda^{Act}$  is the FSR-limited value  $N_\lambda^{FSR}$ . But, if the value of  $P_{Loss}^{dB}$  is greater than some threshold, then  $N_\lambda^{PB}$  becomes less than  $N_\lambda^{FSR}$ , and  $N_\lambda^{Act} = N_\lambda^{PB}$ . Thus, the actual number of channels ( $N_\lambda^{Act}$ ) that are available for use per WG is  $N_\lambda^{Act} = \min_{N_\lambda > 0}(N_\lambda^{FSR}, N_\lambda^{PB})$ .

In this study, we assume the cross-sectional dimensions of 450nm×250nm and WG propagation loss of 1dB/cm for both BCSP and FCSP WGs. We calculate the insertion loss and crosstalk related power penalty for the modulator MRs using the method described in [12], [13]. Moreover, to calculate the insertion loss and sideband truncation related power penalty for detector MRs, we use the experimentally validated analytical method described in [30]. From [12], the insertion loss and the crosstalk power penalty of modulator MRs depend on the  $Q_L$ , channel spacing ( $CS$ ), and modulation shift ( $MS$ ).  $MS$  is the amount by which the passband of a modulator shifts while modulating a signal. From [30], the insertion loss and the power penalty due to sideband truncation of MR detectors depend on the  $Q_L$ ,  $CS$ , and bit-rate ( $BR$ ). Thus, the link-level design parameters  $P_{Loss}^{dB}$  and  $N_\lambda$  depend on some link-level design parameters such as  $CS$ ,  $MS$ , and  $P_{Budget}^{dB}$ , as well as on some device-level design parameters such as  $Q_L$  and  $BR$  of MRs.

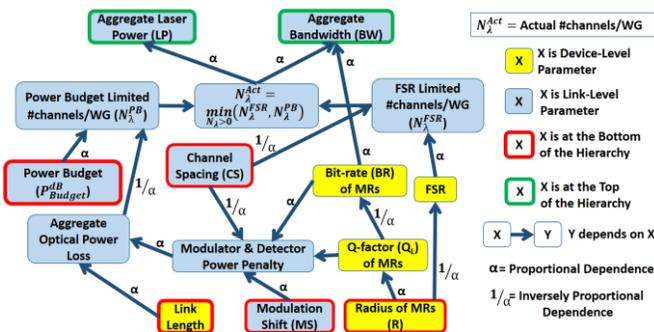


Figure 2. Interdependence among various link-level and device-level design parameters of on-chip SiP interconnects.

The observation above implies that the various device-level and link-level design parameters are interdependent. Figure 2 shows this

interdependence among various design parameters of SiP links. The figure shows how the channel spacing ( $CS$ ), modulation shift ( $MS$ ), link-length,  $P_{Budget}^{dB}$ , and MR radius ( $R$ ) do not depend on any other parameter in the dependence hierarchy. The combination of these five parameters in turn controls all the other parameters, which ultimately affects the aggregate bandwidth and power of the SiP link.

Consider Figure 3 to understand how the link-level design parameters such as channel loss ( $P_{Loss}^{dB}$ ) and aggregate bandwidth depend on the power budget ( $P_{Budget}^{dB}$ ), MR radius ( $R$ ), link-length,  $MS$  and  $CS$ . Figures 3(a), 3(c) show aggregate bandwidth versus  $R$  and  $CS$ , whereas Figures 3(b), 3(d) show  $P_{Budget}^{dB}$  and  $P_{Loss}^{dB}$  values versus  $R$  and  $CS$ , for 5cm long BCSP and FCSP links, with  $MS=6$ pm. From Figures 3(b), 3(d), the FCSP link has  $P_{Budget}^{dB} = 42$ dB, whereas the BCSP link has  $P_{Budget}^{dB} = 45.4$ dB for all the values of  $R$  and  $CS$ .

The maximum aggregate bandwidth of 1.47Tbps for the BCSP link occurs at  $R=1.9\mu\text{m}$  and  $CS=150$ pm (Figure 3(a)), which corresponds to  $P_{Loss}^{dB}$  of 28.3dB (Figure 3(b)) and Q-factor of 20,000. The maximum aggregate bandwidth of 1.93Tbps for the FCSP link occurs at  $R=2.1\mu\text{m}$  and  $CS=150$ pm (Figure 3(c)), which corresponds to  $P_{Loss}^{dB}$  of 19.7dB (Figure 3(d)) and Q-factor of 52,000. The smaller Q-factor renders higher power penalty due to MR sideband truncation for the BCSP link [30], which results in greater  $P_{Loss}^{dB}$  for the BCSP link than the FCSP link. Based on  $P_{Loss}^{dB}$  and  $P_{Budget}^{dB}$  values, the BCSP link and the FCSP link result in  $P_{Budget}^{dB}$  limited  $N_\lambda^{Act}$  values of 51 and 169 respectively. Thus, the BCSP link has less number of channels per waveguide. Moreover, the values of Q-factor translate into channel bit-rate values of 28.7Gbps and 11.4Gbps for the BCSP link and the FCSP link respectively, which results in less maximum aggregate bandwidth of 1.47Tbps for the BCSP link than the maximum aggregate bandwidth of 1.93Tbps for the FCSP link.

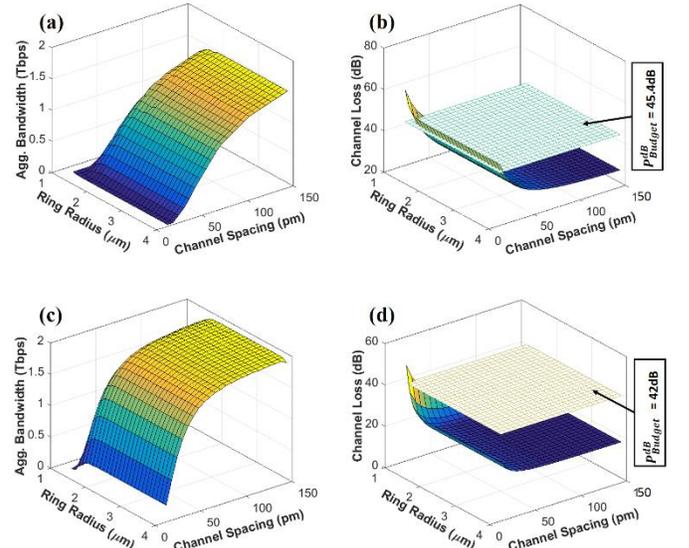


Figure 3. Aggregate bandwidth versus MR radius ( $R$ ) and channel spacing ( $CS$ ) for (a) a BCSP link and (c) an FCSP link. Power budget ( $P_{Budget}^{dB}$ ) and channel loss ( $P_{Loss}^{dB}$ ) versus  $R$  and  $CS$  for (b) a BCSP link and (d) an FCSP link. All plots are for 5cm link-length and  $MS$  of 6pm.

Thus, for the given values of link-length=5cm,  $MS=6$ pm, and  $P_{Budget}^{dB}$ , the values of  $R$  and  $CS$  ultimately control  $P_{Loss}^{dB}$  and aggregate bandwidth of the BCSP and FCSP links. Similarly, for given values of  $R$ ,  $P_{Budget}^{dB}$ , and  $CS$ , the link-length and  $MS$  can be shown to affect the ultimate values of  $P_{Loss}^{dB}$  and aggregate bandwidth. Thus, it can be concluded that the combination of the parameters  $R$ ,  $CS$ ,  $MS$ , link-length, and  $P_{Budget}^{dB}$  controls all the other parameters in the dependence hierarchy in Figure 2, which ultimately affects the aggregate bandwidth and  $P_{Loss}^{dB}$  of the SiP link. However, note that the values of link-length and  $P_{Budget}^{dB}$  cannot be varied for link optimization, as  $P_{Budget}^{dB}$  has a fixed value based on the underline device technology

(FCSP or BCSP), and the link-length has a fixed value based on the layouts of and the distance between the source and destination. For this reason, the parameters  $R$ ,  $CS$ , and  $MS$  are the only independently optimizable parameters in the dependence hierarchy in Figure 2.

Lastly, as evident from Figure 3, the decrease in  $CS$  results in the decrease of aggregate bandwidth but the increase of  $P_{Loss}^{dB}$ . Similarly, the increase in  $R$  results in the increase of aggregate bandwidth but the decrease of  $P_{Loss}^{dB}$ . Along the same lines, the increase in  $MS$  also affects the aggregate bandwidth and  $P_{Loss}^{dB}$  in opposite manners. Thus, it can be inferred that *the parameters  $R$ ,  $CS$  and  $MS$  affect different parameters of the dependence hierarchy in different ways. Therefore, it is imperative to optimize all three of them simultaneously, to achieve energy-efficient and high-aggregate-bandwidth on-chip SiP links.* The next section discusses such an optimization step.

### 3. CROSS-LAYER OPTIMIZATION

In this section, we present a cross-layer optimization of various device-level and link-level parameters for BCSP and FCSP interconnects. These parameters depend on one another as shown in Figure 2.

#### 3.1 Problem Formulation

As the  $MR$  radius ( $R$ ),  $CS$ , and  $MS$  are the only independently optimizable parameters in the dependence hierarchy given in Figure 2, we use all possible values of these three variables as an input to our problem of parameter optimization. In Figure 1, the  $Q_L$  of the MRs saturates for a radius of about 3-4 $\mu$ m. Moreover, researchers have demonstrated in [23] that the minimal radius to obtain an intrinsic  $Q$  of 20,000, which corresponds to an optical bandwidth of 20GHz around the wavelength of 1.55 $\mu$ m, is 1.37 $\mu$ m. Furthermore, for any MR radius of greater than 4 $\mu$ m, the FSR becomes very small leading to an undesirably small value of  $N_{\lambda}^{FSR}$ , which results in poor aggregate bandwidth. Due to these reasons, we define the set of all possible viable values of  $MR$  radius  $R = \{r | r \in Q^+; r \text{ is in } \mu\text{m}; 1.3\mu\text{m} \leq r \leq 4.0\mu\text{m}; (r/0.1) \in N\}$ , which has 28 elements. We aim to design SiP interconnects in ultra-dense WDM (UDWDM) regime, for which the  $CS$  is usually kept smaller than 25GHz or 200pm [31]. Therefore, we define the set of all possible values of  $CS$  as  $\Delta = \{\delta | \delta \in N; \delta \text{ is in } pm; 12pm \leq \delta \leq 150pm; (\delta \bmod 6) = 0\}$ , which has 23 elements. Finally, as discussed in [12], the value of  $MS$  should be less than half the value of  $CS$  to limit worst-case insertion loss for modulator MRs. Therefore, to limit  $MS$  up to half of the  $CS$ , we define the set of all  $MS$  values  $X = \{x | x \in N; x \text{ is in } pm; 6pm \leq x < 75pm; (x \bmod 6) = 0\}$ , which has 10 elements. The individual values for  $R$ ,  $\Delta$  and  $X$  combine to make a triplet in  $28 \times 23 \times 10 = 6440$  different ways. We create a set  $Y$  of these triplets,  $Y = \{(r_1, \delta_1, x_1), (r_1, \delta_1, x_2), \dots, (r_{28}, \delta_{23}, x_{10})\}$  and give it as an input to our cross-layer optimization problem.

#### 3.2 Problem Objective and Constraints

The main objective of our optimization problem is to design a single-WG SiP link of a given length with minimized *aggregate energy-per-bit (EPB)*. The *aggregate EPB* is the sum of *static EPB (SEPB)* and *dynamic EPB (DEPB)*. We obtain *SEPB* by dividing the aggregate laser power by aggregate bandwidth. The *DEPB* here represents *DEPB* of MRs. We calculate the *DEPB* of an MR from the required amount of charge depletion  $\Delta q$  to achieve corresponding  $MS$  using the equations given in [34]. As implied from the discussion in [12], the value of  $MS$  should be less than half the value of  $CS$  to limit the worst-case insertion loss for modulator MRs below an acceptable level, which is the constraint of the optimization problem. Out of 6440 total triplets of  $Y$ , 2268 triplets have  $MS > (CS/2)$ , so they violate this constraint. Therefore, we remove these 2268 triplets from  $Y$  and define a new input set  $Y'$  with the remaining 4172 triplets.

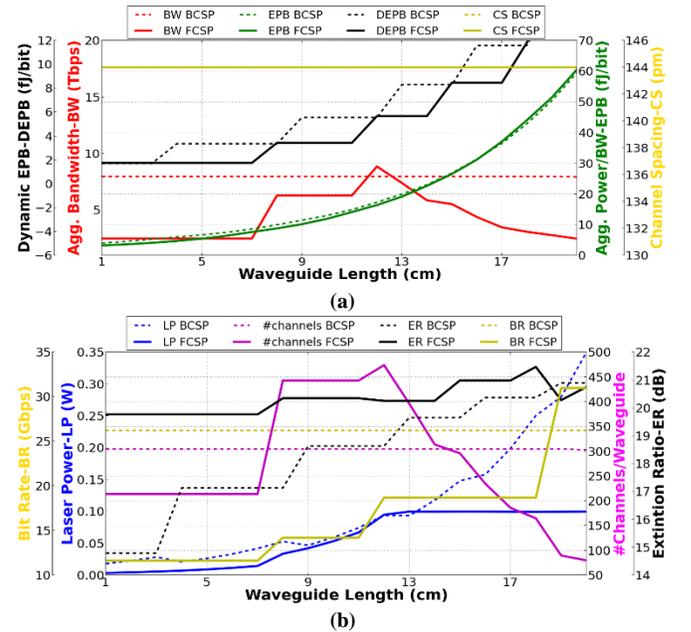
#### 3.3 Optimization Approach

For each triplet of the constrained input set  $Y'$ , first we calculate the  $Q_L$ ,  $FSR$ , and *bit period* ( $2 \times R_s C_j$  or  $2 \times \text{photon lifetime}$ ) using the

methods and equations presented in Section 2.1. Using the values of  $Q_L$  and *bit period*, we then calculate the total channel loss  $P_{Loss}^{dB}$  (in dB) using the methods described in Section 2.2. Then, based on the optical power budget, we calculate  $N_{\lambda}^{Act}$  as described in Section 2.2. Next, for each triplet, we calculate the *MR bit-rate (BR)* by inverting *bit-period*. The actual feasible number of channels  $N_{\lambda}^{Act}$  is multiplied by the *BR* to obtain the *aggregate bandwidth (BW)* per WG. Using the calculated value of  $P_{Loss}^{dB}$  and  $P_{Budget}^{dB}$ , we calculate the total optical/laser power required to achieve the *BW*. We divide total laser power by the achieved *BW* to obtain *SEPB*. We add *SEPB* and *DEPB* to obtain *aggregate EPB*. Lastly, we find an optimal triplet with minimum *aggregate EPB* out of all triplets of  $Y'$ . We use an exhaustive search approach, because it guarantees to find the optimal solution for the marginally small size of the constrained input set  $Y'$ .

#### 3.4 Comparison of Optimized BCSP and FCSP Links

To understand the available design choices for realizing energy-efficient and terabyte-per-second scale SiP interconnects with BCSP and FCSP, we optimize BCSP and FCSP links of 20 different lengths in the range from 1cm to 20cm using our cross-layer optimization framework. The results of this optimization are shown in Figure 4, which plots the values of various parameters obtained for the optimized BCSP and FCSP links of 20 different lengths (x-axes).



**Figure 4. (a) Aggregate bandwidth (BW), aggregate energy-per-bit (EPB), dynamic EPB (DEPB), and channel spacing (CS); (b) laser power (LP), number of channels per WG ( $N_{\lambda}^{Act}$ ), extinction ratio (ER) and bit-rate (BR) values obtained for the optimized BCSP and FCSP links of 20 different lengths. The traces of CS BCSP and CS FCSP are overlapped.**

From the figure, it can be observed that the laser power ( $LP$ ) for the BCSP link increases with increase in link-length. This is because the WG propagation loss (in dB) increases with increase in link-length, which in turn increases aggregate loss in the link, thus requiring higher  $LP$ . However, the  $BW$  of the BCSP link remains constant at 7.9Tbps for all link-lengths. As shown in Figure 2, the  $BW$  depends on only two parameters:  $BR$  and  $N_{\lambda}^{Act}$ . This implies that both  $BR$  and  $N_{\lambda}^{Act}$  should be constant for all link-lengths. As evident from Figure 4,  $BR$  and  $N_{\lambda}^{Act}$  actually remain constant at 26Gbps and 305 respectively for all link-lengths. Now,  $N_{\lambda}^{Act}$  is equal to the minimum of  $N_{\lambda}^{FSR}$  and  $N_{\lambda}^{PB}$  (Figure 2), which implies that either  $N_{\lambda}^{FSR}$  or  $N_{\lambda}^{PB}$  should be constant for all link-lengths. But, as the aggregate loss in the link increases,  $N_{\lambda}^{PB}$  should decrease to meet the power budget constraint in Eq. (5). This implies that  $N_{\lambda}^{FSR}$  remains constant for all link-lengths, which in turn keeps  $N_{\lambda}^{Act}$  constant. As a result,  $N_{\lambda}^{FSR} <$

$N_{\lambda}^{PB}$ , and  $N_{\lambda}^{Act}$  is the FSR-limited  $N_{\lambda}^{FSR}$ . Similarly, for all the FCSP link-lengths below 8cm, the  $N_{\lambda}^{FSR}$  is less than  $N_{\lambda}^{PB}$ , as the  $BW$ ,  $BR$  and  $N_{\lambda}^{Act}$  (FSR-limited) are constant at 2.5Tbps, 11.5Gbps and 214 respectively. Thus, it can be concluded that the FSR-limited value of  $N_{\lambda}^{Act}$  achieves constant  $BW$  for BCSP links irrespective of the link-length and link losses.

For FCSP, at link-length of 8cm, the  $BW$  of the FCSP link shoots up to 6Tbps from 2.5Tbps. So, as evident from Figure 2, the increase in either  $BR$  or  $N_{\lambda}^{Act}$  should be the cause of it. From Figure 4, at link-length of 8cm both  $BR$  and  $N_{\lambda}^{Act}$  increase to 14Gbps and 440 respectively, the combined effect of which increases the  $BW$ . For FCSP link-lengths between 8cm and 12cm, as shown in Figure 4, both  $BR$  and  $N_{\lambda}^{Act}$  keep increasing with increase in link-length, which results in the increase of  $BW$  with increase in link-length. In addition, the  $LP$  also keeps increasing with link-length. However, for FCSP link-lengths beyond 12cm,  $BW$  decreases with increase in link-length, in spite of the increase in  $BR$ . This is due to decreasing  $N_{\lambda}^{Act}$ . As  $N_{\lambda}^{PB}$  becomes less than  $N_{\lambda}^{FSR}$ ,  $N_{\lambda}^{Act}$  becomes power budget-limited. As the  $LP$  is saturated at  $MAOP$  of 100mW for FCSP link-lengths beyond 12cm,  $N_{\lambda}^{Act}$  keeps decreasing with increase in link-length, because of the increase in aggregate link loss with increase in link-length. This observation implies that for an FCSP link whose  $N_{\lambda}^{Act}$  is limited by the power budget, the  $BW$  decreases with increase in link-length and the  $LP$  remains constant at the  $MAOP$ . Decreasing  $BW$  at constant  $LP$  causes a deterioration in SEP. These observations can be generalized to hold true for BCSP links as well, because BCSP and FCSP links operate on the same principle. Thus, it can be concluded that to design an FCSP or BCSP link to achieve high  $BW$  irrespective of the link-length and link losses, all link-level and device-level design parameters in Figure 2 should be optimized to achieve an FSR-limited value of  $N_{\lambda}^{Act}$ . For that, the channel loss  $P_{Loss}^{dB}$  (Eq. 5) should be smaller than a certain threshold value to allow  $N_{\lambda}^{PB}$  to be greater than  $N_{\lambda}^{FSR}$ .

From Figure 4, the optimal value of  $CS$  for all link-lengths for both the BCSP link and the FCSP link is 144pm. For all link-lengths, the BCSP link has greater *dynamic EPB* than the FCSP link. This is because the BCSP link has greater values of optimized modulation shift (not shown in the figure) than the FCSP link. Moreover, Figure 4 also plots *aggregate EPB* and extinction ratio ( $ER$ ). Extinction ratio is defined as the ratio of the optical power in the bus WG during logic “1” state to the optical power during logic “0” state. As evident from the figure, BCSP links have inferior  $ER$  compared to the FCSP links. This is because, as shown in Figure 1, the BCSP MRs have smaller  $Q_L$  and greater cavity loss than the FCSP MRs, which results in lower optical power in the bus WG for logic “1”, thereby decreasing the  $ER$ . The inferior  $ER$  for the BCSP link decreases the signal power and increases its susceptibility to noise. Furthermore, the aggregate EPB values obtained for the BCSP links is quite comparable to those obtained for the FCSP links. Therefore, it can be concluded from these observations that the optimized design of a BCSP link yields more aggregate bandwidth with comparable aggregate EPB, but an inferior extinction ratio than the optimized design of an FCSP link.

## 4. EVALUATION

### 4.1 Evaluation Setup

We performed benchmark-driven simulation-based analysis to evaluate the impact of FCSP and BCSP devices on the performance and efficiency of two well-known crossbar PNoC architectures: Corona [18] and Firefly [19]. We modeled and simulated the Corona and Firefly PNoCs with FCSP and BCSP devices using an in-house cycle-accurate NoC simulator. We evaluated performance for a 256 core single-chip architecture at a 22nm CMOS node. We used real-world traffic from applications in the PARSEC benchmark suite [32] in our analysis. GEM5 full-system simulation [33] of parallelized PARSEC applications was used to generate traces that were fed into our cycle-accurate NoC simulator. We set a “warm-up” period of 100M instructions and then captured traces for the subsequent 1B instructions.

First, based on geometric analysis, we estimated the maximum length of the crossbar WG in both Firefly and Corona PNoCs. The maximum length of the single-write-multiple-read (SWMR) WG in Firefly PNoC is 8cm. This 8cm long SWMR WG between a source and destination node passes through 6 intermediate inactive nodes. Similarly, the maximum length of the multiple-write-single-read (MWSR) WG in Corona PNoC is 12cm. This 12cm long WG between a source and a destination node passes through 62 intermediate inactive nodes. Each node along the crossbar WGs of both the Corona and Firefly PNoCs has arrays of modulator and detector MRs.

We model two different variants of Corona and Firefly PNoCs along with the baseline variants. One variant of Corona and Firefly each uses BCSP devices (referred to as Corona-BCSP and Firefly-BCSP), whereas the other variant uses FCSP devices (referred to as Corona-FCSP and Firefly-FCSP). The baseline variants also use the same type of front-end compatible MRs and WGs as used in the FCSP variants of the PNoCs. However, we optimize the design parameters of the FCSP variants (Firefly-FCSP and Corona-FCSP) using our cross-layer optimization framework, whereas the design parameters of the baseline variants are taken from [18] and [19] and are not optimized. We keep the number of WGs and basic floorplan of the architectures constant across all three variants. We optimized the crossbar data WG designs of all the variants of both PNoCs using the cross-layer optimization described in Section 3, and obtained the maximum allowed number of channels  $N_{\lambda}^{Act}$  for all of them. Here,  $N_{\lambda}^{Act}$  represents the maximum allowed DWDM degree for a given power budget. We also obtain the optical loss values and dynamic EPB values from our optimization framework. Further, we considered a fixed packet size of 512 bits across all the variants of Corona and Firefly architectures.

Table 2 summarizes the DWDM degree, optical loss, and dynamic EPB values for the different variants of the Firefly and Corona PNoCs that we consider. Our optimization framework obtains the optimal modulation shift (MS) of 18pm, 24pm, 54pm, and 72pm for the Firefly-FCSP, Firefly-BCSP, Corona-FCSP, and Corona-BCSP respectively, which results in the dynamic energy values of 3.5pJ/bit, 5.5pJ/bit, 15pJ/bit, and 20pJ/bit for the Firefly-FCSP, Firefly-BCSP, Corona-FCSP, and Corona-BCSP respectively.

**Table 2: Packet size, DWDM degree, optical loss and per bit dynamic energy for different variants of Firefly and Corona PNoC architectures.**

Configuration	Maximum waveguide DWDM	Selected waveguide DWDM	Optical loss data WGs (in dB)	Dynamic energy (in fJ/bit)
Firefly Baseline	64	64	-41.64	1.1
Firefly FCSP	215	128	-39	3.5
Firefly BCSP	260	256	-43	5.5
Corona Baseline	64	64	-51.4	1.1
Corona FCSP	5	4	-42	15
Corona BCSP	20	16	-44.4	20

### 4.2 Evaluation Results for Firefly PNoC

We used the reservation-assisted Firefly PNoC architecture with 64 DWDM as the baseline and compared it with two variants: Firefly-BCSP and Firefly-FCSP. As shown in Table 2, the Firefly-BCSP and Firefly-FCSP have maximum DWDM degree of 260 and 215 respectively. These values of DWDM degree are FSR-limited and we have obtained them for  $CS=0.15$ nm from our optimization framework. Prior works [35] and [36] have demonstrated 20 GHz-spaced (0.2nm-spaced), 200nm-wide comb sources, which are capable of sourcing a WG with DWDM degree of 1000 (total 1000 channels per WG). This implies that it is feasible for the Firefly-BCSP and the Firefly-FCSP to have DWDM degree of 260 and 215 respectively. However, we choose the DWDM degrees of the PNoCs to be factors of the packet-size of 512 bits. Therefore, we select the DWDM degree of the Firefly-BCSP and the Firefly-FCSP to be 256 and 128 respectively (Table 2). Moreover, to facilitate simultaneous traversal of 512 bits (entire

packet) from source node to destination node in Firefly-BCSP, we have considered two SWMR WGs as a group with each WG having 256 DWDM. Further, for reasonable comparison of Firefly-BCSP with Firefly-FCSP and Firefly (baseline), we also considered two SWMR WGs as a group in these architectures as well.

The average throughput and aggregate energy-per-bit (EPB) for all three variants of the Firefly PNoC architecture across 12 multi-threaded PARSEC benchmarks are presented in Figure 5 and Figure 6 respectively. As evident from Figure 5, Firefly-BCSP and Firefly-FCSP yield 36.4% and 19.1% higher throughputs respectively on average over the baseline Firefly. The larger value of DWDM degree for Firefly-BCSP results in greater throughput compared to Firefly-FCSP and baseline Firefly. We calculate aggregate EPB values using the same method as used in our optimization framework described in Section 3. From Figure 6, Firefly-BCSP and Firefly-FCSP yield 26.4% and 15.9% less aggregate EPB respectively on average over the baseline Firefly. Firefly-BCSP achieves 1.15 $\times$  greater throughput and 12.4% less EPB than Firefly-FCSP. The greater throughput for Firefly-BCSP results in a lower value of aggregate EPB compared to Firefly-FCSP and baseline Firefly. The smaller value of aggregate EPB obtained for Firefly-BCSP implies that Firefly-BCSP is more energy-efficient than Firefly-FCSP.

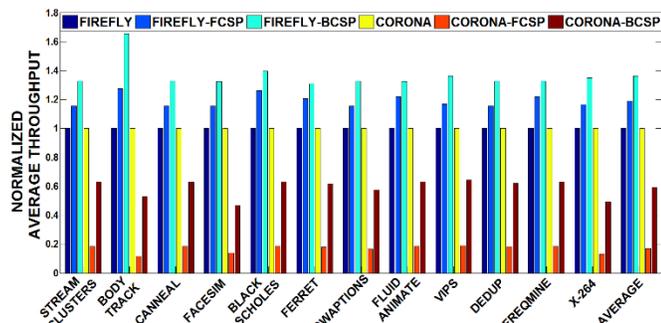


Figure 5. Throughput comparison for different variants of Firefly and Corona PNoCs. Results are shown for PARSEC application workloads and normalized wrt baseline architectures.

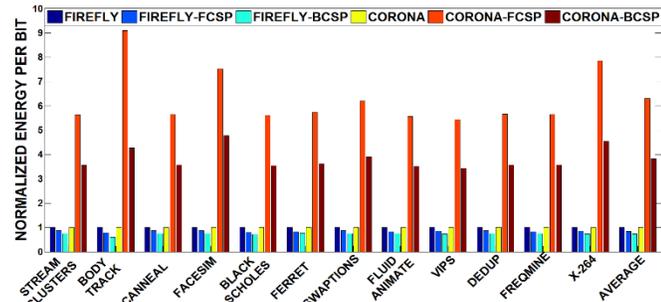


Figure 6. Energy-per-bit (EPB) comparison for different variants of Firefly and Corona architectures. Results are shown for PARSEC application workloads and normalized wrt baseline architectures.

### 4.3 Evaluation Results for Corona PNoC

We performed a similar analysis for the Corona PNoC architecture with token-slot arbitration and 64 DWDM as the baseline and compared it with two variants Corona-BCSP and Corona-FCSP. As shown in Table 2, Corona-BCSP and Corona-FCSP have a power-budget limited DWDM degree of 20 and 5 respectively. As mentioned earlier, the crossbar WG of Corona is 12cm long and it passes through 62 intermediate nodes, which in turn increases the optical loss resulting in smaller values of DWDM degree compared to Firefly. Moreover, the baseline Corona has optical loss of 51.4dB (Table 1), which is significantly larger than the optical power budget of 42dB for FCSP WGs. *This implies that the DWDM degree of 64 used in the baseline Corona architecture is not feasible from a practical implementation perspective.*

The average throughput and aggregate EPB for all three variants

of the Corona architecture across 12 multi-threaded PARSEC benchmarks are presented in Figure 5 and Figure 6, respectively. As the baseline Corona PNoC is not feasible, the results shown in Figure 4 and Figure 5 for the baseline Corona configuration are not practically achievable. As evident from Figure 5, Corona-BCSP and Corona-FCSP yield 40.8% and 83.1% less throughput respectively on average over the baseline Corona configuration. The baseline has a larger (but impractical to achieve) DWDM degree, which results in larger values of throughput for it compared to Corona-BCSP and Corona-FCSP. As evident from Figure 6, Corona-BCSP and Corona-FCSP yield 3.82 $\times$  and 6.31 $\times$  greater aggregate EPB respectively on average over the baseline. The greater DWDM degree of 64 (although impractical) results in greater throughput for the baseline, and consequently a lower value of aggregate EPB compared to Corona-FCSP and Corona-BCSP. Similarly, greater DWDM degree for Corona-BCSP yields 3.5 $\times$  greater throughput for it compared to Corona-FCSP. The greater throughput results in 39.5% less EPB for Corona-BCSP compared to Corona-FCSP.

To summarize the major findings from our experiments, we showed that Firefly-BCSP and Corona-BCSP yields greater throughput and less aggregate EPB than Firefly-FCSP and Corona-FCSP respectively, implying that BCSP links perform better and are more energy-efficient than FCSP links. The smaller values of DWDM degree obtained for Corona-FCSP and Corona-BCSP corroborate our previous observation (Section 2.3) that the power budget and optical loss of BCSP and FCSP links limit the maximum allowable DWDM degree, which in turn constrains the practically achievable aggregate bandwidth and energy-efficiency in PNoCs such as Corona.

## 5. CONCLUSIONS

This paper presented a detailed comparative analysis of a number of design tradeoffs for CMOS front-end (FCSP) and back-end (BCSP) compatible silicon photonic devices. The results of the cross-layer optimization of multiple device-level and link-level design parameters indicate that BCSP interconnects yield more throughput with comparable energy-efficiency compared to FCSP interconnects. The optimized design of BCSP-based Firefly and Corona photonic network-on-chips (PNoCs) yield 1.15 $\times$  and 3.5 $\times$  greater throughput with 12.4% and 39.5% more energy-efficiency than the optimized design of FCSP-based Firefly and Corona PNoCs respectively. The greater throughput and comparable energy-efficiency obtained for BCSP links favor their use in the terabyte-per-second scale silicon photonic interconnects in future PNoCs. However, the inferior extinction ratio for BCSP links necessitates a reduction of intrinsic optical losses present in BCSP devices. Moreover, the sources of crosstalk and noise in BCSP interconnects that threaten the reliability of communication need to be thoroughly investigated and mitigated.

## ACKNOWLEDGMENTS

This research is supported by grants from SRC, NSF (CCF-1252500, CCF-1302693), and AFOSR (FA9550-13-1-0110).

## REFERENCES

- [1] S. Feng et al., "Silicon Photonics: from a Microresonator Perspective," in *Laser Photonics Rev.*, vol. 6, no. 2, pp. 145–177, 2012.
- [2] D. J. Thomson et al., "High contrast 40Gbit/s optical modulation in silicon," in *Opt. Express*, vol. 19, no. 12, p. 11507, 2011.
- [3] S. Liao et al., "36 GHz submicron silicon waveguide germanium photodetector," in *Opt. Express*, vol. 19, no. 11, p. 10967, 2011.
- [4] L. Vivien et al., "40Gbit/s germanium waveguide photodiode," in *OFC/NFOEC*, 2013.
- [5] Y. Lee and M. Lipson, "Back-End Deposited Silicon Photonics for Monolithic Integration on CMOS," in *IEEE JSTQE*, vol. 19, no. 2, 2013.
- [6] I. A. Young et al., "Optical I/O Technology for Tera-Scale Computing," *IEEE JSSC*, vol. 45, no. 1, pp. 235–248, 2010.
- [7] J. M. Fedeli et al., "Integration issues of a photonic layer on top of a CMOS circuit," in *Proc. SPIE* [Online], 2006.
- [8] K. Preston et al., "Polysilicon photonic resonators for large-scale 3D integration of optical networks," in *Opt. Express*, vol. 15, no. 25, 2007.

- [9] N. Sherwood-Droz et al., "Scalable 3D dense integration of photonics on bulk silicon," in *Opt. Express*, vol. 19, no. 18, p. 17758, 2011.
- [10] A. Gondarenko et al., "High confinement micron-scale silicon nitride high Q ring resonator," in *Opt. Express*, vol. 17, no. 14, p. 11366, 2009.
- [11] K. Preston et al., "Deposited silicon high-speed integrated electro-optic modulator," in *Opt. Express*, vol. 17, no. 7, p. 5118, 2009.
- [12] N. Sherwood-Droz et al., "Device Guidelines for WDM Interconnects Using Silicon Microring Resonators", in *WINDS*, 2010.
- [13] K. Preston et al., "Performance guidelines for WDM interconnects based on silicon microring resonators," in *CLEO*, 2011.
- [14] Z. Li et al., "Device modeling and system simulation of nanophotonic on-chip networks for reliability, power and performance," *DAC*, 2011.
- [15] M. Mohamed et al., "Modeling and analysis of micro-ring based silicon photonic interconnect for embedded systems," in *CODES+ISSS*, 2011.
- [16] N. Ophir et al., "Silicon Photonic Microring Links for High-Bandwidth-Density, Low-Power Chip I/O," in *IEEE Micro*, 2013.
- [17] R. Hendry et al., "Physical layer analysis and modeling of silicon photonic WDM bus architectures," in *Proc. HiPEAC Workshop*, 2014.
- [18] D. Vantrease et al., "Light speed arbitration and flow control for nanophotonic interconnects," in *Micro*, 2009.
- [19] Y. Pan et al., "Firefly: Illuminating Future Network-on-chip with Nanophotonics," in *ISCA*, 2009.
- [20] K. Padmaraju et al., "Intermodulation crosstalk characteristics of WDM silicon microring modulators," in *IEEE PTL*, vol. 26, no. 14, 2014.
- [21] W. Bogaerts et al., "Silicon microring resonators," *Laser Photonics Rev.*, vol. 6, no. 1, pp. 47–73, 2012.
- [22] A. Fallahkhair et al., "Vector Finite Difference Modesolver for Anisotropic Dielectric Waveguides," *J. Light. Technol.*, vol. 26, no. 11, 2008.
- [23] Q. Xu et al., "Silicon microring resonators with 1.5- $\mu\text{m}$  radius," in *Opt. Express*, vol. 16, no. 6, p. 4309, 2008.
- [24] G. T. Reed and A. P. Knights, *Silicon Photonics: An Introduction*. John Wiley & Sons, 2004.
- [25] R. Soref et al., "Electrooptical effects in silicon," *IEEE JQE*, vol. 23, no. 1, pp. 123–129, 1987.
- [26] D.A. Neamen, *Semiconductor Physics And Devices: Basic Principles*, 4th edition. New York, NY: McGraw-Hill, 2011.
- [27] Q. Xu et al., "Micrometre-scale silicon electro-optic modulator," *nature*, vol. 435, no. 7040, pp. 325–327, 2005.
- [28] G. Li et al., "25Gb/s 1V-driving CMOS ring modulator with integrated thermal tuning," *Opt. Express*, vol. 19, no. 21, p. 20435, 2011.
- [29] J. Levy, "Integrated nonlinear optics in silicon nitride waveguides and resonators," *PhD Thesis*, 2011.
- [30] M. Bahadori et al., "Optimization of Microring-based Filters for Dense WDM Silicon Photonic Interconnects," *IEEE OI*, 2015.
- [31] J. Reis et al., "Architectural optimization of coherent ultra-dense WDM based optical access networks," in *OFCC*, 2011.
- [32] C. Bienia et al., "The PARSEC Benchmark Suite: Characterization and Architectural Implications," in *PACT*, 2008.
- [33] N. Binkert et al., "The Gem5 Simulator," in *Comp. Arch. News*, 2011.
- [34] R. Wu et al., "Compact Modeling and System Implications of Microring Modulators in Nanophotonic Interconnects," in *SLIP*, 2015.
- [35] A. Johnson et al., "Chip-based frequency combs with sub-100 GHz repetition rates", in *Optics Letters*, OSA, 2012.
- [36] D. Moss et al., "New CMOS-compatible platforms based on silicon nitride and Hydex for nonlinear optics", in *NPHOTON*, 2013.
- [37] D. Pierce et al., "Electronic Structure of Amorphous Si from Photoemission and Optical Studies", in *Phys. Rev. B*, 1972.
- [38] C. Salzberg et al., "Infrared Refractive Indexes of Silicon Germanium and Modified Selenium Glass", in *J. Opt. Soc. Am.*, 1957.
- [39] S. Chittamuru, I. Thakkar, S. Pasricha, "Process Variation Aware Crosstalk Mitigation for DWDM based Photonic NoC Architectures", in *ISQED*, 2016.
- [40] A. Udipi et al., "Combining memory and a controller with photonics through 3D-stacking to enable scalable and energy-efficient systems", in *ISCA*, 2011.
- [41] I. Thakkar, S. Pasricha, "3D-ProWiz: An Energy-Efficient and Optically-Interfaced 3D DRAM Architecture with Reduced Data Access Overhead", in *TMSCS*, 2015.
- [42] S. Bahirat, S. Pasricha, "Exploring Hybrid Photonic Networks-on-Chip for Emerging Chip Multiprocessors", in *CODES+ISSS*, 2009.
- [43] S. Pasricha, S. Bahirat, "OPAL: A Multi-Layer Hybrid Photonic NoC for 3D ICs", in *ASPAC*, 2011.
- [44] S. Bahirat, S. Pasricha, "A Particle Swarm Optimization Approach for Synthesizing Application-specific Hybrid Photonic Networks-on-Chip", in *ISQED*, 2012.