A Process Variation Tolerant Method for Nanophotonic On-chip Network

YI XU, Macau University of Science and Technology
JUN YANG, University of Pittsburgh
RAMI MELHEM, University of Pittsburgh

Nanophotonic networks, a potential candidate for future networks on-chip, have been challenged for their reliability due to several device-level limitations. One of the main issues is that fabrication errors (a.k.a. process variations) can cause devices to malfunction, rendering communication unreliable. For example, microring resonator, a preferred optical modulator device, may not resonate at the designated wavelength under process variations (PV), leading to communication errors and bandwidth loss.

This paper proposes a series of solutions to the wavelength drifting problem of microrings and subsequent bandwidth loss problem of an optical network, due to PV. The objective is to maximize network bandwidth through proper arrangement among microrings and wavelengths with minimum power requirement. Our arrangement, called "MinTrim", solves this problem using simple integer linear programming, adding supplementary microrings and allowing flexible assignment of wavelengths to network nodes as long as the resulting network presents maximal bandwidth. Each step is shown to improve bandwidth provisioning with lower power requirement. Evaluations on a sample network show that a baseline network could lose more than 40% bandwidth due to PV. Such loss can be recovered by MinTrim to produce a network with 98.4% working bandwidth. In addition, the power required in arranging microrings is 39% lower than the baseline. Therefore, MinTrim provides an efficient PV-tolerant solution to improving the reliability of on-chip photonics.

CCS Concepts: • **Networks** → **Network on chip**; *Network reliability*; • **Computer systems organization** → *Reliability*;

Additional Key Words and Phrases: Nanophotonics, reliability, process variation (PV), chip multiprocessor (CMP), network-on-chip (NoC)

ACM Reference format:
Yi Xu, Jun Yang, and Rami Melhem. 2018. A Process Variation Tolerant Method for Nanophotonic On-chip Network. *ACM J. Emerg. Technol. Comput. Syst.* 9, 4, Article 39 (September 2018), 24 pages.
https://doi.org/0000001.0000001

The previously published paper appeared in the 39th International symposium on computer architecture, Portland, OR, 2012. This work is supported by the National Science Foundation under grant 0747242, 1012070 and the Science and Technology Development Fund of Macau under Grant No.039/2013/A2.
Author's addresses: Y. Xu, Space Science Institute, Macau University of Science and Technology, Macau; J. Yang, Department of Electrical and Computer Engineering, University of Pittsburgh; R. Melhem, Department of Computer Science, University of Pittsburgh.

1550-4832/2018/9-ART39 $15.00
https://doi.org/0000001.0000001

1 INTRODUCTION

Electrical on-chip networks are hitting great challenges in power, latency and bandwidth density with technology scaling. Such challenges are especially pronounced in the era of multi-core computing where high bandwidth, low power, and low-latency global transmission are required. For those reasons, and because of recent breakthroughs in nanophotonic devices, optical interconnection is again considered as a potential on-chip network for future many-core microprocessors. Many studies have been performed on network topology, optical routers, cache coherency [8, 13, 16, 21, 30, 31, 41–43, 50], as well as chip-to-chip and CPU-DRAM communication using optics [3, 14, 19].

While optical interconnect provides many promising features, there are also fundamental challenges in integration and fabrication of those devices to providing robust and reliable on-chip communication. Among many challenges, the *thermal sensitivity* and *process variations* (PV) of silicon photonic devices are the key difficulties. Thermal sensitivity refers to the changes in refractive index of optical components, e.g. photonic microring (μring) resonator, due to temperature fluctuations, such that those components fail to resonate designated wavelengths in the waveguide. Studies have reported that μring's resonance wavelength typically drifts by $\sim$0.1nm/°C [33, 34, 53], while chip temperature could fluctuate well beyond 30°C. PV refers to variations of critical physical dimensions, e.g. thickness of silicon, width of waveguide, caused by lithography imperfection and etch non-uniformity of devices [37]. Those variations will directly affect the resonant wavelengths of a μring [14, 29, 36, 46], a critical optical component used as a modulator, a filter or a switching element. Several recent laboratory measurements have reported that wavelength drifts of μrings due to PV (termed PV-drift for short) are quite significant. For example, as much as $\sim$4.79nm of PV-drift within a wafer has been observed in a demonstration of a photonic platform leveraging the state-of-the-art CMOS foundry infrastructure [29]. A recent work [37] has also reported a standard deviation of 0.55nm for two μrings that are only 1.7mm apart. The analytical method obtained the worst-case resonance wavelength drift for 60 identical μrings on a 2.1×4.5 mm^2 chip fabricated by the electron beam lithography system with a high resolution of 2 nm is 2.11 nm [26]. In addition, the wavelength deviation increases almost linearly with the physical distance between μrings, which has been observed from 371 identical resonators on a 16×9 mm chip fabricated by IMEs silicon photonics foundry [7]. To achieve high network bandwidth, the spacing between adjacent wavelengths, denoted as $\Delta\lambda$, is $\sim$0.8nm [38] or lower [9, 27] in a wavelength division multiplexing (WDM) enabled optical interconnect. A previous study shows that when PV-drift is over 1/3 of $\Delta\lambda$, the bit-error-rate of optical transmission would increase from 10^{-12} to 10^{-6} [22]. Larger PV-drifts and thermal variations would bring the μring to resonate at a completely different wavelength that is several channels away. As a result, drifted μrings cannot be used for communication since they will create erroneous signals. Hence, network nodes that do not have all working μrings would lose bandwidth in communication.

At present, there are two types of techniques that can restore the resonance frequency of μrings. The first type is post-fabrication physical trimming, where high-energy particles such as UV light or electron beam is used to adjust the refractive index of μrings [12, 20, 25, 39] or effective refractive index of the waveguide [36] to achieve resonance correction. However, such techniques require trimming to be carefully tuned for individual μring. Given that the number of μrings on-chip is on the order of thousands to millions [1, 14, 17, 31, 43], it is unclear if such physical trimming is practical for volume production. In addition, physical

trimming may create degradation of the quality factor, "Q", of a μring, bouncing of corrected wavelength, and faster aging of the trimmed devices [36]. The second type of techniques for restoring the resonance frequency is power trimming, in which heating or current injection into a μring is used to correct its resonance wavelength. The former causes the wavelength to shift towards the red end and the latter towards the blue end of the resonance spectrum. Although power trimming could address the drifts introduced by both PV and thermal variations, it can result in significant power consumption so as to nullify the power advantage that ideal on-chip optical interconnects are projected to have [9, 22, 27, 28, 54]. In addition, current injection has very limited correction range, as it would generate thermal runaway beyond the trimming range [9, 22, 27]. Nevertheless, power trimming has been considered necessary for tackling thermal sensitivity, as demonstrated in the "Sliding Ring Window" technique [27]. Hence, we will assume that power trimming is already in place for thermal sensitivity, and propose techniques to minimize the total tuning power required for correcting PV-drifts in this paper.

We proposed an architectural methodology in [51] to salvage network bandwidth loss due to PV-drifts. Our goal in [51] is to maximize the number of usable wavelengths for all nodes, each wavelength being resonant with one μring while minimizing the power required in trimming, hence the name "MinTrim". The first step of "MinTrim" tackles the limitation of current injection, and trims a μring to a nearby wavelength rather than the nominal one. Integer linear programming (ILP) is used to maximize the likelihood of successful trimming with minimum trimming power. This step can add 20$\sim$30% of total bandwidth over the baseline. The next step further mitigates PV-drifts by provisioning additional μrings in the ILP framework, which brings more opportunities to finding a nearby μring that can be trimmed to a wavelength. This step can bring up the network bandwidth to $\sim$90% if the number of μrings is doubled, while further reducing power requirement. Finally, we allow flexible wavelength assignment for each network node, as long as each one can be allocated with enough wavelengths, to give more freedom to trimming. This step can produce a network of 98.4% working bandwidth, salvaging most of the lost bandwidth in the baseline. The total trimming power required to achieve such bandwidth is 39% less than the baseline.

In this paper, we have made following improvements over the original design in [51]: (1) Not only the cumulative impacts of a serial of proposed approaches but also the effectiveness of each individual scheme has been evaluated and analyzed to show the bandwidth achieved by each scheme. (2) The discussions on the network crossbar with different configurations including the extreme cases such that no node shares the same waveguide has been added in evaluation section. (3) Different types of power trimming technique is considered. We evaluated the effectiveness of MinTrim with heating-only trimming and proposed a new supplementary ring mapping strategy, which leads to better performance than the approaches of [51]. (4) The strategy to differentiate the impacts of modulators and receivers on network bandwidth has been added in section 3.2 and the corresponding simulation result is discussed in section 5.2. (5) The performance of wrap-around scheme [5, 11] have been evaluated and compared with MinTrim scheme.

2 BACKGROUND AND PRIOR WORK

The key elements in an optical network includes a laser source, which generates laser of different wavelengths; waveguides, which propagate laser signals across the chip; modulators, which imprint binary signals on laser of certain wavelengths, and detectors, which receive optical signals and convert them to electrical signals. The laser source is responsible for

generating phase-coherence and equally spaced wavelengths. It is expected that such laser source could produce 64 or even more wavelengths per waveguide for a dense wavelength division multiplexing (DWDM) network [18, 47]. For modulation, μring resonators are typically preferred over other modulators due to their high modulation speed, low power and small footprint [23, 48, 49]. The same ring structure can be used as a wavelength selective detector to extract light out of the waveguide, if the μring is doped with a photo-detecting material such as germanium. The resonant light will be absorbed by the germanium and converted into electrical signal.

2.1 A Motivating Example

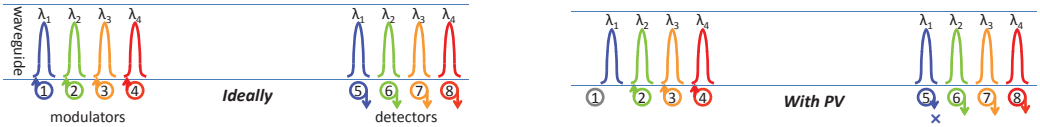


Fig. 1. Bandwidth loss due to PV-drift.

If μrings are fabricated perfectly, a sender and a receiver can modulate and extract optical signals correctly without any loss. The upper part of Fig. 1 illustrates such an ideal scenario where the sender uses microrings #1 $\sim$ #4 to modulate their nominal four wavelengths $\lambda_1 \sim \lambda_4$, and the receiver uses microrings #5 $\sim$ #8 to detect and extract the same wavelengths respectively. Note that ring #5 and #1 have the same resonance, so do #6 and #2 etc. Under ideal situation, both sender and receiver can utilize 100% of their bandwidth for transmission. When PV is present, some μrings are off from their resonance due to imprecise dimension, e.g. waveguide width. Fig. 1 shows the same example with μring #1 being off from λ_1. As a result, it cannot resonate at λ_1, downgrading the sender's bandwidth to 75%. Consequently, ring #5 at the receiver cannot receive any signal. Such a bandwidth loss is a static loss meaning that this sender loses 25% bandwidth permanently.

2.2 Current Approaches and Challenges

There are mainly two types of approaches to trimming the drifted resonant wavelength of μrings. The first one is power trimming. Heating and carrier injection can shift the resonant wavelength of a μring up and down respectively [1]. In Figure 1, μring #1 can be corrected towards red using heating. This type of method can fine tune the resonance of μrings. However, there are fundamental limitations to power trimming:

Challenge 1: Power trimming incurs high static power consumption. Many existing work have shown that the static power for trimming the μrings is a significant portion, or even dominant portion, of the total optical network power. For example, the Corona network in 17-nm technology from HP [1, 43] is estimated to consume $\sim$26W in power trimming, out of $\sim$48W of total network power. Even with the most optimistic μring heating efficiency, e.g. using in-plane heaters and air-undercut [13, 14], it is estimated that μring heating still consumes 38% of total network power [30]. For this reason, many work also focused on reducing the amount of μrings on-chip to reduce the power needed for trimming [13, 30].

Challenge 2: Power trimming can only correct limited resonance drifts. Even though the resonance wavelength can be corrected towards red or blue, blue shifts is still limited no matter how much power we are willing to pay. This is because blue shifts is achieved through carrier injection, which heats up the μrings and causes red shifts that

need further carrier injection for correction, forming a positive feedback loop and thermal runaway [27]. In addition, more carrier injection degrades the extinction ratio and creates more power loss of the signal, e.g. ∼0.4nm tuning in wavelength results in 1dB signal loss and significant degradation of quality factor (Q) and extinction ratio (ER) of the microring resonator [9, 22, 45]. Hence, the achievable amount of blue shift is far less than of red shift [27]. For this reason, many work just use heating to keep all μrings at a constant temperature [13, 14, 29], which should be close to the peak temperature of the chip to avoid blue shifts.

The second class of trimming is done post-fabrication by changing its refractive index of the μring directly, or adjusting the stress level of the cladding material. The advantage of such physical trimming is that, if successful, no additional power is required for correcting PV-drifts. However, the challenge is:

Challenge 3: Physical trimming is immature and less commercially practical. All physical trimmings require precise control of irradiation dose and energy, which is different from μring to μring. Given that there are thousands to millions of μrings on-chip, it is currently difficult to do physical trimming in mass fabrication which is critical for commercial purposes. Whereas, the power trimming saves *tuning effort* from that required for physical trimming with the receive-data driven control circuit [11], which can tune the μrings without external intervention. Second, in SOI technology, trimming the cladding material (SiO_2) is unstable as a subsequent red shift of 0.15nm was observed 5 days after the irradiation. Moreover, the quality factor Q of the μring decreased by 21∼41.2% with a 1∼2nm correction [36], which would increase the BER of the optical signal or require higher laser source power to overcome signal attenuation.

There are also proposals that do not rely on physical or power trimming to overcome PV. A dynamic regulation method was proposed [22] in which adjusting chip temperature is used to compensate chip-wise PV-drifts (i.e. systematic variations). For example, if the PV-drift of μrings in a chip region are toward blue, then the regulator would heat up, i.e., red shift, the region via e.g., dynamic voltage/frequency scaling (DVFS). Such coarse-grained regulation cannot overcome random PV-drifts, e.g., both red and blue drifts, among different μrings within the region. Also, DVFS comes at non-trivial performance cost, especially when cooling the chip region is required. Nitta *et al.* proposed to use error detection/correction code to tackle faulty μrings that are due to either PV-drifts, or temperature induced resonant wavelength drifts, or insufficient trimming [28]. However, such schemes can only handle small number of faulty μrings since the overhead of error correction coding, in both performance and extra optical bandwidth requirement, would be daunting otherwise. As we will show in our experiments, even conservative estimation of PV-drifts indicates that more than half the μrings could become faulty, which would require complex ECC codes. Given the constraints of on-chip interconnects, the decoding and correction process is too slow. A tuning control circuit that allows μring to resonate at its closet wavelength instead of the original assigned one through bit re-shuffling was developed [11]. We adopt the same circuit design in this paper and use their tuning strategy as one of the baselines to compare against ours.

Next we describe our proposed suite of solutions starting from improving the success rate of trimming while minimizing the static power, to ultimately provisioning near-full bandwidth for an optical network under PV.

3 MINTRIM: MAXIMIZE NETWORK BANDWIDTH WITH MINIMUM TRIMMING POWER

The first drawback of power trimming is high static power, since all μrings need to be kept at a constant temperature to be functional, which would require continuous heating power or current injection power (effective "cooling" through power) to cancel the effect of on-chip temperature fluctuation. With PV, a μring may be off its nominal resonant wavelength, so *additional power trimming* is required to correct it back, on top of the power to keep it thermally stable, exacerbating the already high static power of the optical network. A μring's resonance wavelength typically drifts by $\sim$0.1nm/°C [33, 34, 53]. Hence, an average of 1nm of PV-drift [9, 36, 48] would require equal amount of power for regulating the μring temperature within 10°C fluctuation range. Hence, PV-drifts add significant power overhead to the network, which is what we will minimize in MinTrim.

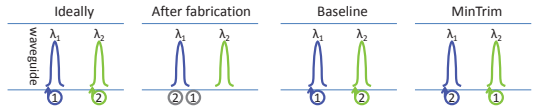
Second, even with unlimited power supply, current injection can shift the resonant wavelength towards the blue end of the spectrum, but can also degrade trimming efficiency and even trigger thermal runaway [9, 22, 27]. Hence, it can only correct small PV-drifts, e.g. 0.4nm which also results in 1dB signal loss [22]. With PV, a μring's resonant wavelength may be shifted towards red beyond the correctable range. This is the main reason for the network to lose bandwidth since such μrings and the corresponding nominal wavelengths cannot be used. As we will show later, our sample network architecture loses more than 40% bandwidth because 32% of the μrings are uncorrectable due to PV. MinTrim strives to turn uncorrrectable into correctable scenarios to achieve maximum bandwidth.

We discuss MinTrim using three types of wavelength-μring organization of optical buses and crossbars, namely single-writer-multiple-reader (SWMR), multiple-writer-single-reader (MWSR), and multiple-writer-multiple-reader (MWMR) [2, 16, 21, 31, 43, 50]. In SWMR/MWSR, network nodes have exclusive sets of wavelengths for transmitting/receiving data. In these two architectures, modulators and detectors of each node use complementary sets of wavelengths. In MWMR, all modulators and detectors of a node use all wavelengths, increasing the network bandwidth over the other two. Both MWSR and MWMR require arbitration before sending data while SWMR does not. MinTrim is applicable to all these three architectures.

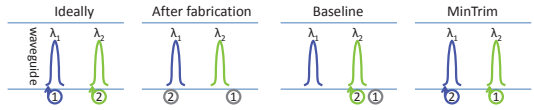
3.1 An Optimization Problem

The first step in MinTrim is developed based on the observation that a μring does not have to be trimmed to its nominal wavelength as it may be far from the μring's resonant wavelength. With PV, the distribution of the resonant wavelengths of μrings are somewhat random. Hence, as long as we can generate an association between μrings and wavelengths, such that the number of usable wavelengths for each node is maximized, then we can achieve the highest bandwidth. In order to keep the trimming power low, the most intuitive way is to trim a μring to a *nearby* wavelength, rather than its nominal wavelength, to reduce the trimming distance which linearly affects the trimming power. More importantly, such nearby-mapping can reduce the number of uncorrectable μrings as their trimming distances are now smaller.

Fig. 2 illustrates these two advantages with a simple example. Here the nominal wavelengths of μring#1 and #2 are λ_1 and λ_2 respectively. In 2(a), suppose PV causes μring#1 and #2 to be closer to λ_2 and λ_1 respectively. The baseline design trims the two μrings back to their nominal wavelengths. In MinTrim, μring#1 will be trimmed to λ_2, and μring#2 to λ_1, which clearly consumes less trimming power than in the baseline. In 2(b), suppose



(a) Less trimming power for PV.



(b) Higher bandwidth.

Fig. 2. Two advantages of trimming μrings to a nearby wavelength.

μring#1's resonant wavelength is too far from λ_1 to be correctable using current injection. The baseline would lose λ_1 as no μring can resonate at it, but MinTrim would actually make μring#1 correctable by trimming it to λ_2 since it is closer, and μring#2 to λ_1, salvaging all available bandwidth.

However, if the resonant wavelength of a μring is roughly in the middle of two channels, say λ_i and λ_{i+1}, MinTrim needs to determine which wavelength should the μring be trimmed to. The decision is based on which map would generate higher bandwidth and require lower trimming power. The network bandwidth is defined as the number of working channels (pair-wise tuned senders and receivers), summed over all possible sender-receiver pairs of the network. This is important because under PV, a sender and a receiver must have the same λ's to communicate. Hence, only the common λ's between the two nodes are counted towards effective bandwidth. Total network bandwidth of a perfect network without PV is 100%, and MinTrim strives to approach that.

Since the decision for one μring affects other μrings, MinTrim needs to generate a globally optimal solution, which can be solved by an optimization tool such as integer linear programming (ILP). ILP is a powerful method for optimizing a certain objective function through determining a set of decision variables, subject to some constraints. Note that MinTrim is a post-fabrication procedure to alleviate the PV-induced damage. No further reconfigurations are needed to tackle PV at runtime. However, additional adjustment [52] is required for thermal drifting. Hence, running an optimization algorithm incurs only a one-time cost, and is worthwhile since it improves the yield of the chip effectively. We will now describe how to formulate MinTrim into an ILP problem by defining the decision variables, objective functions and constraints.

Decision Variables. Since we are trying to decide which wavelength should a μring be trimmed to, the decision variables of our problem are simply boolean variables, map(r_n, w_m, *node*), representing whether μring r_n of a *node* should be trimmed to wavelength w_m, 1 being yes and 0 being no.

Objective Function. MinTrim tries to achieve two objectives: maximal bandwidth and minimal trimming power. Given that ILP can only maximize (or minimize) one goal, we let maximal bandwidth take higher priority over minimal power, but the reverse can also be formulated under a different chip design goal. That is, if there are two solutions, one with higher bandwidth and the other with lower trimming power, MinTrim will return the former as the solution. To achieve this, we iteratively run ILP with a bandwidth in descending order starting from 100%. The granularity of decreasing bandwidth is losing one wavelength

for a node at a time. The algorithm terminates when a solution is found, i.e. the requested bandwidth is satisfied and the trimming power is the lowest within the available solution pool. Trimming power is calculated as follows:

$$\sum_{\substack{\forall n,\forall m\\ \forall node}} map(.) \times \begin{cases} TP_c \times (\lambda_{act}[r_n] - w_m) & \text{if} \lambda_{act}[r_n] \geq w_m, \\ TP_h \times (w_m - \lambda_{act}[r_n]) & \text{if} \lambda_{act}[r_n] < w_m. \end{cases} \tag{1}$$

where $\lambda_{act}[r_n]$ is a parameter of r_n to represent the actual wavelength of r_n post fabrication. The difference between actual and target wavelength, w_m, determines how much trimming power is required. The coefficients, $TP_c = 0.13$mW/nm and $TP_h = 0.24$mW/nm, are unit power required for current injection and heating respectively [27]. We will use $map(.)$ rather than the full length of the *map* function for brevity, since they are all in one form.

Constraints. There are two constraints on trimming μrings to wavelengths. For every node using a waveguide, (1) every μring of the node should resonate with at most one wavelength in the waveguide; and (2) every wavelength in the waveguide should be resonant with at most one μring of the node:

$$\forall r_n,\ \forall node, \quad \sum_{w_m \in \{\text{all } \lambda\text{'s}\}} map(.) \leq 1 \tag{2}$$

$$\forall w_m,\ \forall node, \quad \sum_{r_n \in \{\text{modulators in } node\}} map(.) \leq 1, \quad \sum_{r_n \in \{\text{detectors in } node\ \}} map(.) \leq 1 \tag{3}$$

To enforce that modulators and detectors of each node use complementary set of wavelengths in SWMR and MWSR, we have:

$$\forall node,\ \text{Let } S = \{\lambda\text{'s assigned to } node \text{ for modulation}\},$$

$$\forall w_m \notin S,\ \sum_{r_n \in \{\text{modulators in } node\}} map(.) = 0 \tag{4}$$

$$\forall w_m \in S,\ \sum_{r_n \in \{\text{detectors in } node\}} map(.) = 0 \tag{5}$$

Those are not needed for MWMR since it does not have this constraint.

Another set of important constraint is on the trimming distance. In this paper, we assume 0.4nm as the constraint for current injection [22]. For trimming through heating, the constraint depends on the chip power budget since heating power increases linearly with trimming distance. A 2nm of wavelength shift requires the temperature of the μring to be 20°C above the ambient temperature [27]. In addition, allowing a wide range of heating brings challenges to thermal insulation among the μrings. Therefore, in this paper, we will put constraints on trimming distance through current injection and heating, termed "Llimit" and "Rlimit", respectively and show in the results the trend of trimming power and network bandwidth with varying RLimit and the fixed Llimit of 0.4 nm. Hence, the constraints for trimming distance are:

$$\forall n,\ \forall m,\ \forall node,$$

$$map(.) \times (\lambda_{act}[r_n] - w_m) \leq \text{Llimit},\ \text{if } \lambda_{act}[r_n] \geq w_m, \tag{6}$$

$$map(.) \times (w_m - \lambda_{act}[r_n]) \leq \text{Rlimit},\ \text{otherwise}. \tag{7}$$

In addition, the constraint for bandwidth is:

$$\forall node,$$

$$\sum_{r_n \in \{\forall \mu\text{rings in } node\}, w_m \in \{\text{all } \lambda\text{'s}\}} map(.) \geq Bandwidth_{min} \tag{8}$$

where Bandwidth_{min} is reduced incrementally, starting from 100%, during the interactive search procedure.

This first ILP step is able to dramatically improve the success rate of trimming μrings and the number of usable wavelengths. As will be shown later, the number of usable μrings

improved from 68% in the baseline to 97%, resulting in a bandwidth increase from 59% to 81%. To salvage the remaining bandwidth loss, we now introduce the next step in MinTrim.

3.2 Supplementing μrings with Spares

The next simple method is to supplement the existing μrings with spares. Having more μrings creates more opportunities for selecting correctable μrings, as illustrated in Fig. 3 where μring#1 is supplemented with #2 which is closer to μring#1's nominal wavelength λ_1, under PV. MinTrim will trim μring#2 to λ_1. The rationale behind this idea is that when fabricating two μrings of the same nominal wavelength instead of one, there is always a better one for MinTrim to pick. The advantages are again two fold: (1) reduced trimming power and (2) improved successful trimming; since the μring with less trimming distance will be selected. Incorporating spare μrings in ILP formulas is as simple as increasing the set of modulators and detectors in Equation (2)-(5), without any further changes.

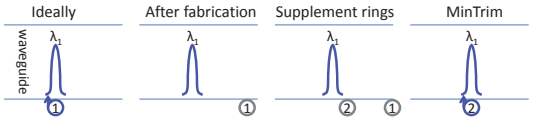


Fig. 3. Supplementing μrings with spares.

The first question to address is how many spares to provide for a node with N μrings and resonant wavelength $\lambda_1, ..., \lambda_N$. We do not have to backup every μring because many of them might already be good enough. Suppose we provide M supplemental μrings, M<N. Ideally, these M μrings should be the backups for those with large PVs. Unfortunately these are not known prior to fabrication and, hence, there are a number of alternatives for assigning nominal wavelengths to μrings. For instance, we can assign N μrings to $\lambda_1, ..., \lambda_N$ and assign the remaining M μrings to M wavelengths chosen uniformly among $\lambda_1, ..., \lambda_N$. However, this alternative is likely to benefit only two wavelengths closest to a spare's resonant wavelength. Hence, in our experiments, we will explore the following strategies:

(1) The nominal wavelengths of all N+M μrings are uniformly distributed across the entire wavelength spectrum $\lambda_1 \sim \lambda_N$, to hopefully generate the best coverage. We term this strategy `Even` as shown Fig. 4(a).
(2) Observing that it is more difficult for MinTrim to correct μrings on the two ends of the wavelength spectrum because they can only be trimmed in one direction while others can be trimmed towards either red or blue, it is also natural to supplement μrings on the two ends with more spares than in the middle. We term this strategy `Double_ends_even_middle`, or DEEM, meaning that we assign 2R spares with nominal wavelengths of $\lambda_1 \cdots \lambda_R$ and $\lambda_{N-R+1} \cdots \lambda_N$, and distribute the remaining M-2R μrings across the spectrum of $\lambda_{R+1} \sim \lambda_{N-R}$. Fig. 4(b) shows an example of DEEM.
(3) If M = N, then two μrings can be assigned to each of $\lambda_1 \cdots \lambda_N$. We term this strategy `Double` as illustrated by Fig. 4(c).

In SWMR optical crossbar architecture, modulators have larger impact on network bandwidth than receivers. Because losing one modulator results in the bandwidth loss of all links connected to the local node. Whereas the failure of one receiver only causes the bandwidth degradation for one link between two nodes. In addition, the modulators is much less than the receivers at each node. Due to these reasons, adding redundant modulator is

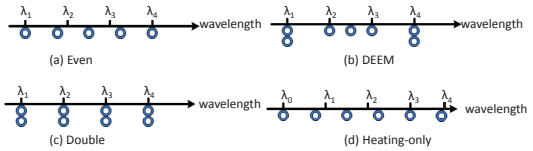


Fig. 4. Different strategies for spare μrings placement.

more cost-efficient than supplying spare receivers. Hence, when M > N, we implement a triple-sender-double-receiver strategy termed as `3S2R` that required 4 more modulators per node per waveguide over `Double`.

For the optical network adopting the heating-only trimming method, λs can only be trimmed towards red. To improve the possibility of successful wavelength mapping with heating, we proposed to supply the extra rings at the left side of the spectrum in addition to the ones inside the spectrum. Fig. 4(d) shows that the ring of resonant wavelength λ_0 that does not belong to the designated wavelength set are added to optical network. This strategy could handle the conditions when the red shifts caused by PV can not be corrected due to the limitation of trimming. Then the supplementary rings with smaller λs, e.g. λ_0 can be trimmed to the ones inside the spectrum, e.g. λ_1 with heating. We assumed that K supplemental μrings outside the spectrum range, K<N. The experiment results in section 5 indicate that a small value of K can result in 20% of bandwidth improvement.

The second question relates to the possibility that adding more μrings may increase the power consumption of the network. If N out of N+M μrings are selected by MinTrim, the remaining M μrings may cause ∼1dB light loss each in the waveguide [14], especially when a μring is close to a wavelength. For this reason, those M μrings should be tuned off, by bringing their resonance wavelengths to the closest mid-points between two channels. When a μring is tuned off, it generates ∼1e-3dB/1.7e-2dB light loss in the waveguide [40] at non-neighboring λ/neighboring λ, respectively, which results in a total of 0.023% or 0.39% laser power loss. Since off-tuning is done through trimming, this amount of trimming power overhead is more of a concern. We measured through our experiments that the average trimming distance for this part is 0.205nm, about $\Delta\lambda/4$ since the trimming distance is within $[0,\Delta\lambda/2]$. In fact, our experimental results will show that more spare μrings lead to total trimming power (trimming N μrings + tuning off M μrings) reduction because the power required to trim one μring by 2nm through heating is equivalent to the power for tuning off 9 unselected μrings. As a result, the `DEEM` strategy of sparing results in the best bandwidth, ∼90%, with the lowest power requirement, as will be shown in Section 5. Last, the spare μrings do not increase the die area since the waveguides extend across the entire die and there is plenty of space between μrings to accommodate spares.

3.3 Flexible Wavelength Assignment for Network Nodes

To recover the remaining bandwidth, we develop the third step of MinTrim. Observe that in both SWMR and MWSR, a node (either a modulator in SWMR or a detector in MWSR) does not use all wavelengths in a waveguide to transmit or receive data. Each node is assigned N/X wavelengths for transmission (SWMR) or receiving (MWSR), where N is the total number of wavelengths in a waveguide shared by X network nodes. With perfect fabrication process, i.e. no PV, it does not matter which N/X wavelengths are assigned to each node. With PV, however, determining which N/X wavelengths are assigned to a node is crucial

since a wavelength may not be usable by one node but usable by another. Hence, a node should be assigned with those N/X wavelengths that are usable by this node. We term this technique flexible_wavelength_to_node_assignment, or `Flexible_assignment`.

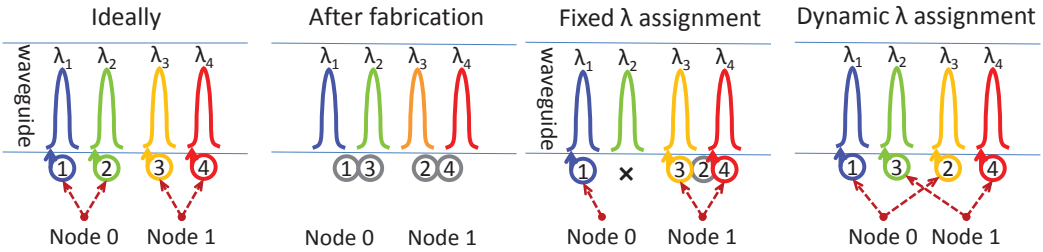


Fig. 5. A case for flexible assignment between wavelengths and nodes.

Fig. 5 explains why having a flexible assignment is effective in bandwidth recovery. Node_0 has two μrings: #1 and #2, and Node_1 has #3 and #4. Due to PV, the resonance wavelengths of the four μrings are drifted as shown in "After fabrication". When original wavelength-node assignment is used ("Fixed λ assignment"), ILP will search *within the local pool of μrings* to find ones that can resonate at λ_1 and λ_2. Since μring #2 is drifted beyond correctable range of current injection, λ_2 becomes unusable. However, note that the optimum assignment between these wavelengths and nodes is: (μring) #1$\rightarrow \lambda_1$, #3$\rightarrow \lambda_2$, #2$\rightarrow \lambda_3$, #4$\rightarrow \lambda_4$. With a fixed wavelength-node assignment, Node_0 cannot use λ_2 because μring #3 is physically local to Node_1. However, μring #2 is physically local to Node_0, and can resonate at λ_3, Node_0 can hence use λ_1 and λ_3, Node_1 can use λ_2 and λ_4, as shown in the figure.

To achieve flexible assignment between wavelengths and nodes, we extend the ILP formulation with new constraints. First of all, while Equation (2) and (3) still hold, Equation (4) and (5) cannot be used since the set of modulating wavelengths of each node is no longer pre-defined. MinTrim needs to search for such set for each node. A new constraint we establish is that a wavelength can be assigned by at most one node:

$$\forall w_m, \sum_{\forall node} \sum_{\substack{r_n \in \{\text{modulators (SWMR) } in \text{ } node\} \text{ or} \\ r_n \in \{\text{detectors (MWSR) } in \text{ } node\}}} map(.) \leq 1 \tag{9}$$

For detectors in SWMR, their resonant wavelengths should be the union of all modulating wavelengths of all other nodes. The same principle applies to modulators in MWSR.

$$\text{Let} R = \{\forall \text{dectecotrs (SWMR)}\} \text{ or } \{\forall \text{ modulators (MWSR)}\}$$
$$\forall w_m, \ \forall node,$$
$$\sum_{r_n \in R} map(.) \leq \sum_{r_n \neq node} \sum_{r_n \in \bar{R}} map(.) \tag{10}$$

Finally, since we have spare μrings which can also be applied with flexible assignment, we define the following constraint to avoid having too many modulators or detectors per node.

$$\forall node, \sum_{r_n \in \bar{R}} \sum_{\forall w_m} map(.) \leq N/X \tag{11}$$

As we will show in our results, flexible assignment can recover almost all the remaining lost bandwidth. Lastly, MWMR does not need this step, so only the first two steps (ILP with spares) will be sufficient. This is because both modulators and detectors already have the full bandwidth spectrum to resonate. There is no need to reassign wavelengths among nodes since every node already has all available wavelengths.

4 MODELING PV OF MICRORINGS

To evaluate the effectiveness of the proposed MinTrim, we first need to model an optical network subject to PV. The resonant wavelength of a μring is determined by several factors including material used for waveguide and cladding, waveguide cross-section dimensions, circumference of the μring, temperature etc. [36, 37]. Especially, for a fixed material and constant temperature, the wavelength is sensitive to the width and height of the waveguide. The variations of the wavelength is approximately linear to the width variation and height variation in waveguide. For example, 1nm of variation in width and height leads to 0.58∼1nm [22, 37, 44] and ≈2nm [37] shift in resonance wavelength of the μring respectively. Due to fabrication imperfection, the variations of critical physical dimensions, such as width, or height of silicon are inevitable. Hence, to characterize the PV of resonance wavelength of μrings, we will develop a variation model for the physical dimensions of the optical waveguide. Recent laboratory fabrications of optical devices show that physical dimensions variations can be classified into die-to-die, or D2D, (a.k.a. inter-die) and within die, or WID (a.k.a. intra-die) variations [29, 37]. The D2D variation refers to non-uniformity of devices between dies that are on the same wafers. This is generally caused by the fabrication tool and process design. The WID variation refers to such non-uniformity between identical devices within a single die. This is generally caused by die-level processes such as lithography and dry etch [15, 32]. Within a die, each step of the process may create spatial (systematic) and random variations in physical dimensions of the waveguide. Since the characteristics of the variations in optical devices are close to process variations in CMOS devices [6, 24] which also present D2D, WID including systematic and random variations among transistors, we adopt VARIUS [35], a PV modeling infrastructure for CMOS technology, based on the statistic tool `R` and its package `geoR` to model both WID and D2D variations.

VARIUS uses Normal (Gaussian) distribution to characterize on-chip process variations. The key parameters are mean (μ), variance (σ^2), and density (ϕ) of a variable that follows Normal distribution. Since wavelength variations are approximately linear to dimension variations of waveguide, we assume they follow the same distribution. The mean (μ) of wavelength variation of a μring is its nominal wavelength. We use a spectrum of 64 wavelengths in a WDM network starting at 1550nm [29] and a channel spacing of 0.8nm. Hence, those wavelengths are the means for each μring modeled.

The variance (σ^2) of wavelength variation is determined based on laboratory fabrication data [29, 37] and our target die size. Since optics are more cost-effective for many-core CMPs, we choose to model a 64-core chip with die size $400mm^2$ [16, 43]. There are no readily available variation data for such a die size, and measurements for small die sizes cannot be directly used because variations in small region is different from those in a large region. In [37], the standard deviation, σ, is 0.15nm for two μrings that are only 25μm apart, and 0.55nm if they are 1.7mm apart. The former characterizes the random variations within a die, and the latter describes systematic variations for a small die, e.g. $2{\times}2mm^2$. The D2D die variation in a $200mm$ wafer is also reported to be 1.08nm. To derive corresponding parameters for a $400mm^2$ die, we first generated 3K dies of $2{\times}2mm^2$ using the above variation parameters: $\sigma_{D2D} = 1.08nm$, $\sigma_{WID-systematic} = 0.55nm$, $\sigma_{WID-random} = 0.15nm$. Then we sort the dies according to their resulted mean values, and selected 100 (400/4) dies with close mean values to assemble a large die, $400mm^2$. This is because previous experiments demonstrated strong within-die spatial correlations of dimension variations [29, 37]. Hence, the 100 small dies that are next to each other should be strongly correlated as well. From the assembled large die, we then derive the WID and D2D variations that are used in our experiments.

Finally, the density ϕ is a parameter that determines the range of WID spatial correlation. It is expressed as a fraction of chip's length in one dimension in VARIUS. As the spatial correlation of two devices decreases as their distance grows, ϕ is the distance at which the correlation drops to zero. Typical value for ϕ is 0.5/1.0 and for a large/small die.

	WID Variation (nm)		D2D Variation (nm)	
	small die	large die	small die	large die
PV_1	0.57 [37]	0.61	1.08 [37]	1.01
PV_2	0.37 [29]	0.39	1.6 [29]	1.40

Table 1. Two sets of PV parameters. WID variation=$\sqrt{\text{systematic var.}^2 + \text{random var.}^2}$ [35].

We generated two sets of variation parameters based on two different fabrications results [29, 37], using the same methodology since both of them use small dies ($2\times2.2mm^2$) in [29]. Table 1 compares the published results and our derivation for larger die sizes. As we can see, when dies size is larger, WID variation increases since some portion of D2D variation is now WID. Consequently, D2D variation decreases a little since it loses a portion to WID. We input these two sets of parameters into VARIUS and generated 100 sample dies of $400mm^2$ each. Each sample contains over one million points indicating the wavelengths of μrings. We then extracted those along the optical waveguide according to the physical layout of an optical crossbar [2]. The total number of points picked from the samples is equal to the number of μrings. Fig. 6 shows the distribution of resonant wavelength shifts under PV_1 and PV_2. As we can see, the total effective variance, including both WID and D2D, of PV_2 is larger than of PV_1, so the bell shaped distribution is wider than for PV_1, meaning that more shift is present on-die which creates more bandwidth loss.

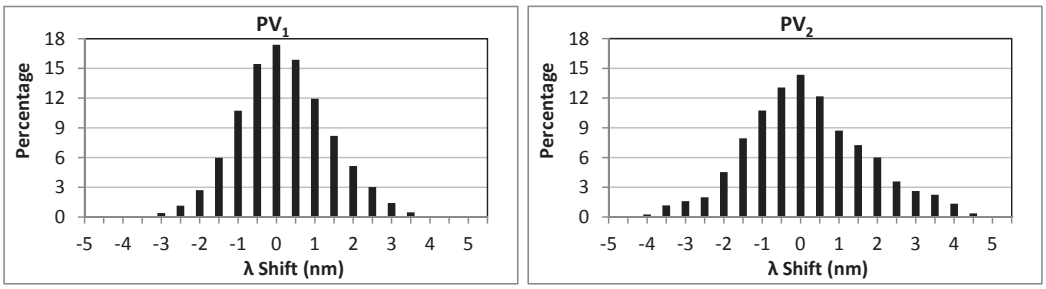


Fig. 6. Distribution of resonant wavelength shift for two sets of PV parameters in Table 1.

5 EVALUATIONS AND RESULTS

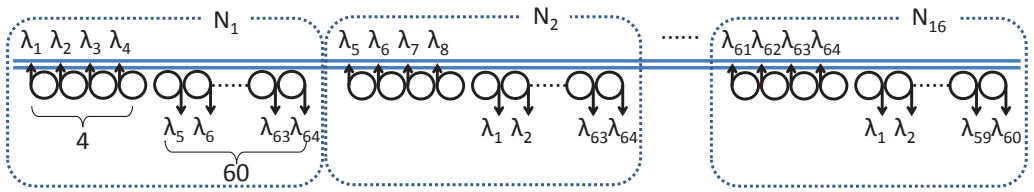


Fig. 7. An SWMR network architecture used for evaluating MinTrim.

We use an SWMR crossbar, shown in Fig. 7, as an example to demonstrate the effectiveness of MinTrim, although it is applicable to MWSR and MWMR as elaborated in Section 3. Our optical network is composed of 4 identical waveguides, each supporting 64 wavelengths denoted by $\lambda_1, ..., \lambda_{64}$. Each waveguide is shared by 16 network nodes. Since this is a single writer architecture, each node is exclusively assigned 4 λ's for transmission. Hence no contention can occur during a write. Four μrings are used as modulators to resonate with these 4 λ's. Every node can simultaneously read from all other 15 nodes, hence "multiple readers", requiring a total of 60 λ's for reception, and 60 μrings detectors as shown in the figure. Therefore, 4K μrings in total are used for the SWMR or MWSR crossbar, while MWMR crossbar require 8K μrings if each waveguide also supports 64 wavelengths. However, the number of wavelengths or waveguide can be reduced considering that wavelength sharing between different nodes improves its utilization.

The physical layout of the crossbar employed here is a symmetric design as each waveguide has exactly the same placement of μrings next to it. There are asymmetric designs such letting a subset of network nodes share one waveguide. For example, each node sends data via 16 wavelengths traversed in one specific waveguide instead of 4 wavelengths per waveguide and the light transmitted in each waveguide can be modulated by μrings connected to different set of nodes. However in such case, the bandwidth loss might be a little less as it is less likely to have 16 failed rings than 4 ones. Besides simplicity, the reason that we select the symmetric layout is to show that the proposed solutions are able to recover most of network bandwidth even the design is vulnerable to PV. In addition, ILP and sparing still apply to other configurations. Flexible wavelength mapping requires some minor modification, such as letting four nodes share 64 wavelengths in the waveguide instead of 16 nodes. However, if the connection of each node is separated by using different bundles of waveguides to obtain high transmission bandwidth, only ILP and spare can be applied to the optical network. Later we will show that previous two schemes dominate the contribution of bandwidth improvement, so we expect that the final bandwidth to still approach 98% because our baseline configuration has more bandwidth loss than the original settings.

The variations of all μrings are generated as described in the previous section. Results are averaged over 100 sample dies. MinTrim computes solutions using the state-of-the-art ILP solver `lpsolve` [4]. The constraints and objective functions in the ILP problem are formulated using the front-end AMPL language [10].

We use *total network bandwidth* as a metric to evaluate MinTrim under different settings. The total network bandwidth is defined as the number of working channels (pair-wise tuned senders and receivers), summed over all possible sender-receiver pairs of the network. This is important because under PV, a sender and a receiver must have the same λ's to communicate. Hence, only the common λ's between the two nodes are counted towards effective bandwidth. As we can see, to have high total bandwidth, each node must be able to use as many λ's as possible. Total network bandwidth of a perfect network without PV is 100%, and MinTrim strives to approach that.

In addition, we measure the power consumption of the network since another major advantage of MinTrim is power reduction. The power trimming techniques we employ requires 0.13mw/nm for current injection [1] and 0.24mw/nm for heating [27]. We assume current injection can correct up to $0.5\Delta\lambda$ towards blue [22] for power trimming. For the design just use heating to keep all μrings at a constant temperature [13, 14, 29], no blue shifts are allowed. For Rlimit (or Rlimit for short), we assume that the chip has certain power budget that limits this amount and we gradually relax such constraint to see if, using

MinTrim, a large power budget is necessary to achieve high network bandwidth. Power measurement includes both the trimming power used to correct μring's λ's and the power required to tune-off unused μrings, i.e., power overhead.

5.1 Baseline Bandwidth Results

With PV, large amount of μrings are off from their nominal resonance λ, leading to a significant bandwidth loss. Assume an optimistic error tolerance of 10% $\Delta\lambda$, i.e., if the actual λ of a μring is within 10% of the nominal λ, the μring can still work. If no trimming is applied, the average total bandwidth is only 0.6% for both PV_1 and PV_2. In other words, the network does not work at all. Hence, we adopt power trimming in our baseline, and first compare two different ways of such trimming: (1) trim the μring to the *closest* λ; and (2) trim the μring to its nominal λ, both under trimming distance constraints. Note that (1) is different from trimming a μring to a nearby λ, as is done in MinTrim, because a nearby μring may not be the closest one, and searching for a good nearby λ requires global optimization. Trimming to the closest λ minimizes trimming power, but does not optimize bandwidth. The bandwidths after these trimmings are shown in Fig. 8. As we can see, the

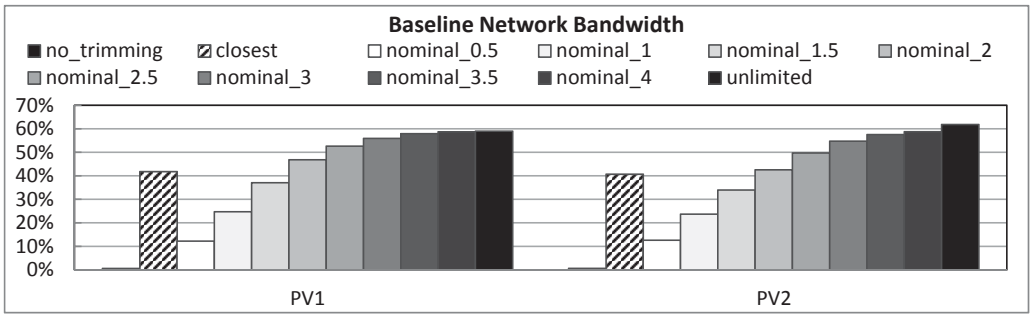


Fig. 8. Average baseline network bandwidth comparison. Numbers following `nominal_` are Rlimit in unit of $\Delta\lambda$.

"closest" bars can recover bandwidth from 0.6% in "no_trimming" to ~42%. The advantages of "closest" is that it does not require large trimming distance, and has the lowest trimming power as will be shown in Fig. 13(a). It loses bandwidth when (1) more than one μrings are trimmed to the same λ, so one has to be removed and no spare is available for making this up; and (2) a sender and a receiver's μrings are trimmed to different λ's, and only the common λ's can be used for communication. Those two cases can be avoided by trimming the μrings to their nominal λ's, labeled as "nominal_Rlimit". However, the "nominal"s also have limited capability in bandwidth recovery under tight heating power budget, e.g. below $2\Delta\lambda$. Progressively better bandwidth can be achieved when we relax Rlimit: 59%~62% for PV_1 and PV_2 respectively with unlimited Rlimit. We will use both "closest" and "nominal_"s in our later results. Also, although PV_1 and PV_2 show noticable λ shift distribution (Fig. 6), the resulting baseline bandwidths differ only slightly. We will show results for PV_1 in the following discussion for clarity.

5.2 MinTrim Bandwidth Results

First step: ILP. When ILP is applied, great bandwidth improvement can be achieved immediately, as shown in Fig. 9. The error bars show the minimum and maximum results

from the 100 samples we experimented with. ILP achieves a bandwidth of 74% and 81% when Rlimit is $2\Delta\lambda$ and unlimited respectively. The reason of this improvement was illustrated in Fig. 2: ILP can reduce the uncorrectable μrings by finding a good nearby λ. However, the improvement diminishes with a larger power budget. This problem can be addressed by having spare μrings as shown below.

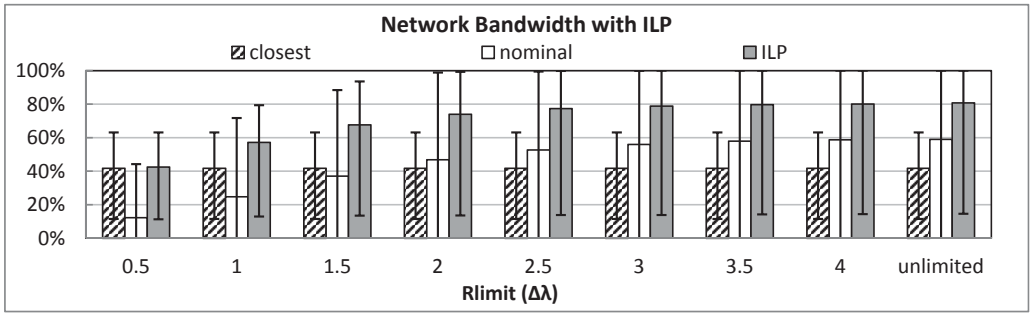


Fig. 9. Bandwidth comparison among "closest", "nominal", and ILP-only.

Second step: Using spare μrings. Each node in baseline and ILP only has 64 μrings. We now show the bandwidth results with different number of spare μrings, 16, 32, 48, and 64, on top of the original 64. When the number of spares is less than 64, we use the `Even` distribution as introduced in Section 3.2. When there are 64 spares, we applied all three distribution methods: `Even`, `Double` and `DEEM`. Recall that each node originally has 4 modulators and 60 detectors. With `DEEM`, we `Double` the 4 modulators, and 8 detectors (4 on each end of spectrum), and use `Even` for the remaining 104 filters. We treat modulators and detectors separately because they are built differently. As we can see from Fig. 10, having spares

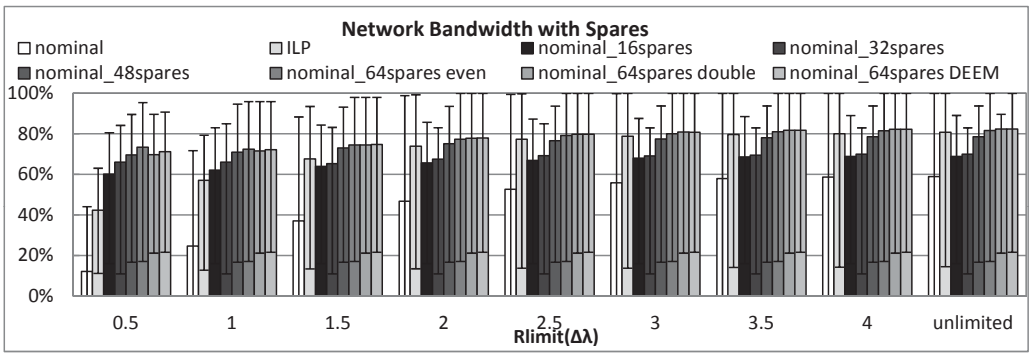


Fig. 10. Bandwidth comparison among "nominal", ILP and varying amount of sparing with nominal mapping.

effectively recovers more bandwidth than using "nominal" alone. More spares result in more improvement. From nominal to having 64 spares using `Even`, the bandwidth improvements are 500%∼38% when Rlimit increases from $0.5\Delta\lambda$ to unlimited. The `Double` method is more effective than `Even` because doubling μrings at their nominal λ's have higher chances of getting a working μring, as indicated by the λ shift distribution in Fig. 6. Whereas, in `Even`, the nominal λ's are not the 64 channels in the waveguide. Finally, the `DEEM` method stands

out as the best one because the μrings on the ends of a spectrum are more difficult than in the middle. So doubling those μrings while using `Even` for middle μrings, given the same number of spares as in `Double`, achieves the best tradeoff. The bandwidth of `DEEM` reaches 73%∼82% when Rlimit increases from $0.5\Delta\lambda$ to unlimited. Fig. 10 also shows that without additional μrings, ILP is able to recover similar amount of bandwidth as the approach of spare rings under loose trimming constraint(Rlimit $\geq 3\delta\lambda$).

Third step: Flexible λ assignment to nodes. To evaluate the effectiveness of the proposed third scheme, we applied the flexible assignment between λ's and network nodes to "nominal" and compare it with the other two schemes: ILP and "spare"—64 μspare rings with `DEEM` in Fig. 11. Not surprisingly, providing flexibility in λ-node assignment generates more bandwidth than the baseline. As shown in Fig. 11, all the three schemes could improve the network bandwidth significantly, while having spares performs slightly better than ILP and flexible mapping. In addition, applying 64 spare μrings to ILP can achieve 18%∼35% more bandwidth than using ILP alone. Adding flexible mapping on top of ILP with spares increases bandwidth by 8%∼12%. When Rlimit is $2.5\Delta\lambda$, the bandwidth is 98.2%, close to 98.4% at unlimited Rlimit. Hence, with flexible assignment, having a power budget corresponding to Rlimit=$2.5\Delta\lambda$ is sufficiently good. The bandwidth achieved by the strategy `3S2R` does not contribute more bandwidth (only 0.1%) improvement since the flexible λ assignment has already been able to improve the wavelength matching rate of modulators and no extra spares is necessary. For clarity reason, we didn't include the result in the figure. More interestingly, the flexible assignment scheme results in much smaller error range (from 100 samples we generated), meaning that by allowing the nodes to select most suitable λ's, the success rate of finding an assignment with high bandwidth is increased. This indicates that MinTrim provides a robust method to salvage network bandwidth under PV.

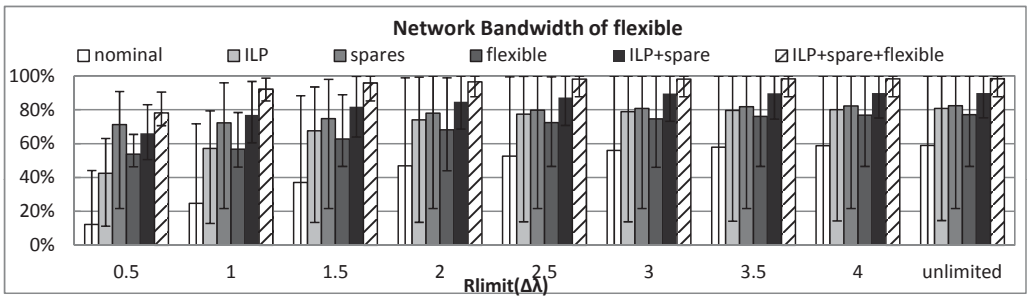


Fig. 11. Bandwidth comparison between fix and flexible wavelength assignment.

5.3 Compared to Wrap Around Scheme

Given that the resonance of μring repeats in each free spectral ranges (FSR), the separations between peaks of wavelength transmissivity, prior works exploited the ring resonance repetition by wrapping around the next resonance for rings [5, 11]. Fig. 12 shows an example that applying wavelength warp around scheme to improve the successful rate of trimming. After fabrication, the resonances of ring #1 $\sim$ #4 all shift toward red due to systematic variation. The resonance of ring #4 in next FSR is drawn with the dotted line, which is close to λ_1—the first wavelength inside the spectrum. Ring #4 is trimmed to λ_1 instead of λ_4 to meeting trimming constraint. Then Ring #1∼ #3 are shifted by one channel. Through

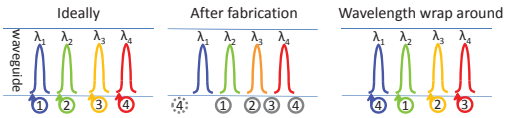


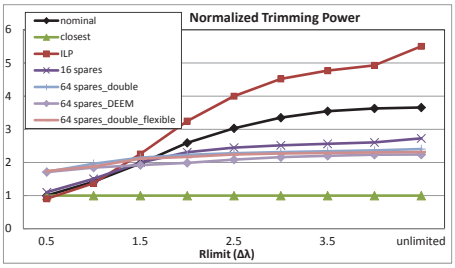
Fig. 12. A case for wavelength wrap around.

this way, there are more opportunity to find closer rings to improve bandwidth and reduce trimming distance.

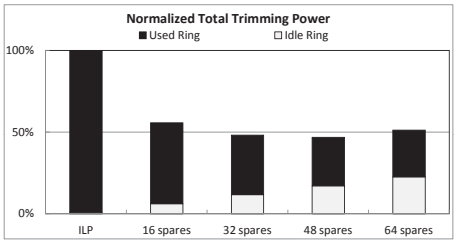
To evaluate the wrap around scheme, we applied it in addition to our MinTrim design and observed that the bandwidth improvement is only 0.2% over MinTrim on average. The reason is that wrap around mainly addresses the systematic variations, whereas ILP and flexible wavelength mapping in MinTrim can also mitigate this type of PV. Furthermore, wrap around approach requires delicate design such that the range of wavelength spectrum should be close to and no larger than the size of FSR. Plus FSR of each ring is different from each other which makes the design complicated. While MinTrim is able to overcome PV without the requirement on optical network settings.

5.4 MinTrim Power Consumption Results

The other major advantage of MinTrim is the trimming power reduction due to decreased total trimming distance. Fig. 13(a) shows the power comparison among different schemes normalized to baseline schemes "nominal" at Rlimit=$0.5\Delta\lambda$. For clarity, we do not show all sparing settings because their results overlap heavily in the figure. As we can see, baseline-closest requires lowest power among all schemes, but it can only achieve 41.8% of total bandwidth. MinTrim-ILP at $0.5\Delta\lambda$ consumes even lower power (10% lower) than in baseline while achieving similar bandwidth (42.4%). However, ILP consumes the highest power, and baseline-nominal is the 2^{nd} highest among all when Rlimit increases because they both can trim more μrings at further distances but ILP has higher priority in bandwidth so it trims more μrings at further distances than in baseline. Once we add spare μrings, e.g. starting at 16 spares, the power consumption immediately drops at all Rlimits beyond $1.5\Delta\lambda$. This is because solutions can be found with closer μrings that help to decrease trimming distance. However, before $1.5\Delta\lambda$, higher power is consumed because again, higher bandwidth is more important. So MinTrim halts when there is a solution for high bandwidth, even when the power is higher. Overall, having more than 32 spares consumes about the same power, with 48-spares being the lowest. For example, A 37%/39% power reduction is observed for using 48 spares, compared with "nominal" when Rlimit is $3\Delta\lambda$/unlimited. `Double`, `DEEM` and flexible assignment do not differ significantly. The conclusion from these results is that having spares is effective in lowering power and improving bandwidth. Fig. 13(b) shows the power breakdown for MinTrim, between trimming useful μrings (used ring) and tuning-off unused μrings (idle ring), with different number of spares from 0 to 64. The results are normalized to total trimming power of ILP, i.e. with 0 spares. The trend clearly shows that although adding spares increases the power for off-tuning unused μrings, the amount of active power for trimming useful μrings is greatly reduced, resulting in a large total reduction. Also, having 64 spares is sufficient because having more spares would slowly increase the total power because the useful power is stabilizing while the off-tuning power increases steadily.



(a) Trimming power normalized to baseline-nominal.



(b) Power breakdown with Rlimit=3Δλ.

Fig. 13. Power analysis of different MinTrim schemes.

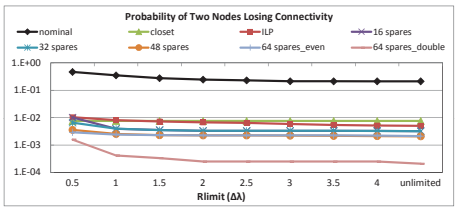


Fig. 14. Probability of losing connectivity between two nodes.

5.5 MinTrim Quality Assessment through Network Connectivity Evaluation

As discussed earlier, MinTrim with flexible λ-node assignment is a robust method for improving network bandwidth because its worst cases (worst solutions due to severe PV in the 100 generated samples) are much better than using fixed assignment. Since the achieved bandwidth is still not 100%, another important metric is the probability of completely losing connectivity between two nodes. That is, no single λ is common between the two nodes. Fig. 14 shows such probability on logarithmic scale. The data are collected through counting how many such pairs occur in the entire 100 samples, each having 16×15 node pairs. We did not find any disconnected node pairs in "64-spares-DEEM" and "64-spares-double-flexible". Although MinTrim does not guarantee connectivity, our experiments do show that the probability for the two schemes are very low. The next best scheme is "64-spares-double" using fixed λ-node assignment. The probability of losing one pair is $10^{-4} \sim 10^{-3}$. The next batch of schemes have similar probabilities: $10^{-3} \sim 10^{-2}$. These schemes include those with spares, ILP, and baseline-closest. Baseline-nominal has the highest disconnection rate, nearly 2 orders of magnitude worse than other schemes. This is because pair-wise disconnection often occurs in worst PV scenarios. The worst cases for "nominal" is worse than for "closest" and ILP, as shown in Fig. 9. For example, if all μrings of a node drifted too far to be correctable, "nominal" bails out but "closest" and ILP may still find a solution. In summary, MinTrim with enough spares and flexible assignment are among the best schemes in terms of network connectivity.

We illustrate the bandwidth improvement of MinTrim with an randomly selected sample in Fig. 15 with the X axle being the index of the network node. For a specific node, it may have different number of wavelengths to communicate with other nodes because the wavelengths used for transmission might not be available at each receiver node. Hence, Fig. 15 shows the minimum and maximum connection bandwidth of each network node under baseline design

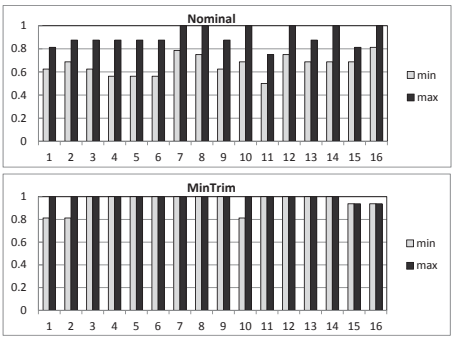


Fig. 15. An example to show the maximum and minimum connection bandwidths of each node for "nominal" and MinTrim, respectively.

and MinTrim, respectively. We can observe that after applying MinTrim, both worst and best case connection bandwidth is improved and become uniform.

5.6 Heating-only Trimming

Normalized Bandwidth

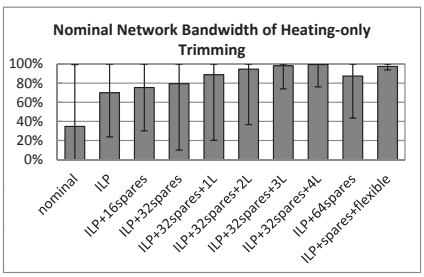


Fig. 16. Normalized bandwidth achieved by heating-only trimming.

Since correction ability on blue shift is far less than of red shift [27], many work only use heating to keep all μrings at a constant temperature [13, 14, 29], which should be close to the peak temperature of the chip to avoid blue shifts. However, it could not alleviate the red shifts introduced by PV. We measured the normalized network bandwidth achieved by baseline design, illustrated in Fig. 16 with RLimit being $2\Delta\lambda$. Compared to the one allowing $0.5\Delta\lambda$ of blue shifts shown in Fig. 8, bandwidth is degraded by 12%. After applying ILP, the normalized bandwidth reaches 70%. Adding spare rings also helps improve the successful rate of associating μrings and wavelengths, which leads to higher network bandwidth. In Fig. 16, "ILP+32spares+KL" indicates K extra rings at the left side of the spectrum. We can observe from the figure, the network bandwidth is close to 100% with only 4 μrings, which means that the asymmetric placement of μrings corresponding to imbalanced trimming ability can effectively mitigate the bandwidth loss. Whereas supplying more rings inside the spectrum indicated by "ILP+64spares" can only produce 87% of network bandwidth on average and flexible wavelength mapping leads 10% more of bandwidth improvement, which is still less effective than asymmetrical spare ring strategy.

Trimming Power

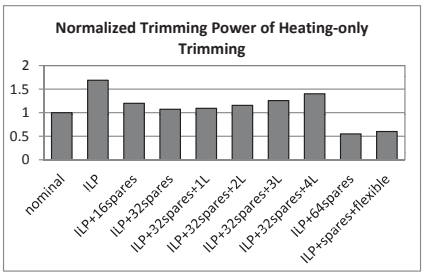


Fig. 17. Normalized trimming power required by heating-only trimming.

Fig. 17 shows the comparisons on the trimming power generated by baseline and MinTrim under heating-only trimming method at Rlimit=$2\Delta\lambda$. MinTrim-ILP consumes higher power than "nominal" since it is able to correct much more μrings, which results in larger cumulative trimming distance. The power consumption immediately drops when spare μrings approach is applied, same as the normal power trimming method. However, power cost increases quickly with the number of the μrings placed at the left side of the spectrum because even adding one μring might have significant impact on the μring mapping solution generated by ILP. While supplying μrings inside the wavelength spectrum could help reduce both bandwidth loss and power consumption by 45% and 63% compared to baseline design, respectively. Overall, all the spare ring strategies can effectively mitigate bandwidth loss. The tradeoff is that adding rings inside the spectrum can reduce trimming power significantly but requires doubling the μrings while supplying a couple of rings with resonate wavelengths outside the spectrum is enough to provide nearly full bandwidth but needs higher trimming power.

6 CONCLUSIONS AND FUTURE WORK

We have shown that PV in optical networks is a serious problem. A network can be paralyzed by PV due to variations of device dimensions and changes in resonance wavelength of μrings. Current power trimming techniques cannot solve this problem, as shown by our experiments. Our proposed technique, MinTrim, is shown to be effective in tolerating PV. The key ideas of MinTrim include using redundancy and allowing flexibility, which are natural approaches to handling variations. MinTrim improves bandwidth from 59% to 98.4% in the best cases. We also found that using redundancy is not only effective in improving bandwidth, but also in reducing power consumption which is a critical factor in optical network. A 39% trimming power reduction is observed through MinTrim.

Further improvement of MinTrim is possible with advanced optical technologies. For example, although MinTrim has flexibility in assigning λ's to nodes, further improvement is possible if the laser source frequency can also be tuned to increase the chance of successful mapping. For network architectures that do not belong to SWMR, MWSR, or MWMR, we emphasize that the first two steps of MinTrim, ILP and sparing, can always be applied. Hence, MinTrim is a general method that can be tailored to a network architecture. Finally, although MinTrim was proposed to target only the PV problem in this paper, new techniques based on MinTrim can be developed, with a focus on reducing the computation cost to be applicable at runtime so that temperature (dynamic) variations are addressed.

REFERENCES

[1] J. Ahn, M. Fiorentino, R. G. Beausoleil, N. Binkert, A. Davis, D. Fattal, N. P. Jouppi, M. McLaren, C. M. Santori, R. S. Schreiber, S. M. Spillane, D. Vantrease, and Q. Xu. 2009. Devices and architectures for photonic chip-scale integration. *Appl. Phy. A* (2009). Issue 95.

[2] C. Batten. 2010. Designing Nanophotonic Interconnection Networks. In *Workshop on the Interaction between Nanophotonic Devices and Systems*.

[3] Scott Beamer, Chen Sun, Yong-Jin Kwon, Ajay Joshi, Christopher Batten, Vladimir Stojanović, and Krste Asanović. 2010. Re-architecting DRAM memory systems with monolithically integrated silicon photonics. In *ISCA '10*. 129–140.

[4] M. Berkelaar, K. Eikland, and P. Notebaert. 2007. LP solve: Opern Source (Mixed-Integer) Linear Programming System. (2007). http://lpsolve.sourceforge.net/5.5/

[5] Nathan L. Binkert, Al Davis, Norman P. Jouppi, Moray McLaren, Naveen Muralimanohar, Robert Schreiber, and Jung Ho Ahn. 2011. The role of optics in future high radix switch design. In *ISCA*. 437–448.

[6] Y. Cao and L. T. Clark. 2005. Mapping Statistical Process Variations Toward Circuit Performance Variability: An Analytical Modeling Approach. In *Design Automation Conference*. 658–663.

[7] L. Chrostowski, X. Wang, J. Flueckiger, Y. Wu, Y. Wang, and S. T. Fard. 2014. Impact of fabrication non-uniformity on chip-scale silicon photonic integrated circuits. In *Optical Fiber Communications Conference and Exhibition (OFC)*. 1–3.

[8] Mark J. Cianchetti, Joseph C. Kerekes, and David H. Albonesi. 2009. Phastlane: a rapid transit optical routing network. In *Proceedings of the 36th annual international symposium on Computer architecture (ISCA '09)*. 441–450.

[9] R. K. Dokania and Alyssa B. Apsel. 2009. Analysis of challenges for on-chip optical interconnects. In *GLSVLSI '09*.

[10] Robert Fourer, David M. Gay, and Brian W. Kernighan. 2002. *AMPL: A Modeling Language for Mathematical Programming, 2nd ed.* Duxbury Press Publishing Company.

[11] M. Georgas, J. Leu, B. Moss, Chen Sun, and V. Stojanovic. 2011. Addressing link-level design tradeoffs for integrated photonic interconnects. In *Custom Integrated Circuits Conference (CICC), 2011 IEEE*. 1 –8.

[12] H. Haeiwa, T. Naganawa, and Y. Kokubun. 2004. Wide range center wavelength trimming of vertically coupled microring resonator filter by direct UV irradiation to SiN ring core. *IEEE Photonics Technology Letters* 16 (2004), 135–137. Issue 1.

[13] A. Joshi, C. Batten, Yong-Jin Kwon, S. Beamer, I. Shamim, K. Asanovic, and V. Stojanovic. 2009. Silicon-photonic clos networks for global on-chip communication. In *3rd ACM/IEEE International Symposium on Networks-on-Chip*. 124 –133.

[14] A. Joshi, J. Orcutt, A. Khilo, B. Moss, C. Holzwarth, M. Popovic, Hanqing Li, H. Smith, J. Hoyt, F. Kartner, R. Ram, V. Stojanovic, and K. Asanovic. 2008. Building manycore processor-to-dram networks with monolithic silicon photonics. In *Hot Interconnects*. 21–30.

[15] J. Karttunen, Jyrki Kiihamäki, and Sami Franssila. 2000. Loading effects in deep silicon etching. In *International Society of Optical Engineering*, Vol. 4174. 90–97.

[16] Nevin Kirman, Meyrem Kirman, Rajeev K. Dokania, Jose F. Martinez, Alyssa B. Apsel, Matthew A. Watkins, and David H. Albonesi. 2006. Leveraging Optical Technology in Future Bus-based Chip Multiprocessors. In *Proceedings of the 39th Annual IEEE/ACM International Symposium on Microarchitecture*. 492–503.

[17] N. Kirman and José F. Martínez. 2010. A power-efficient all-optical on-chip interconnect using wavelength-based oblivious routing. In *ASPLOS '10*. 15–28.

[18] Brian R. Koch, Alexander W. Fang, Oded Cohen, and John E. Bowers. 2007. Mode-locked silicon evanescent lasers. *Optics Express* 15, 18 (2007).

[19] Pranay Koka, Michael O. McCracken, Herb Schwetman, Xuezhe Zheng, Ron Ho, and Ashok V. Krishnamoorthy. 2010. Silicon-photonic network architectures for scalable, power-efficient multi-chip systems. *SIGARCH Comput. Archit. News* 38 (2010). Issue 3.

[20] Yasuo Kokubun, Naoki Kobayashi, and Tomoyuki Sato. 2010. UV trimming of polarization-independent microring resonator by internal stress and temperature control. *Optics Express* 18 (2010), 906–916. Issue 2.

[21] George Kurian, Jason E. Miller, James Psota, Jonathan Eastep, Jifeng Liu, Jurgen Michel, Lionel C. Kimerling, and Anant Agarwal. 2010. ATAC: a 1000-core cache-coherent processor with on-chip optical

network. In *Proceedings of the 19th international conference on Parallel architectures and compilation techniques*. 477–488.

[22] Zheng Li, Moustafa Mohamed, Xi Chen, Eric Dudley, Ke Meng, Li Shang, Alan R. Mickelson, Russ Joseph, Manish Vachharajani, Brian Schwartz, and Yihe Sun. 2010. Reliability modeling and management of nanophotonic on-chip networks. *IEEE Transactions on Very Large Scale Integration Systems* (2010), 98–111.

[23] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia. 2004. A High-Speed silicon optical modulator based on a metal-oxdie-semiconductor capacitor. *Nature* 427 (2004), 615–618.

[24] S. R. Nassif. 2001. Modeling and forecasting of manufacturing variations. In *Asia and South Pacific Design Automation Conference*. 145–149.

[25] Y. Nasu, M. Kohtoku, M. Abe, and Y. Hibino. 2005. Birefringence suppression of UV-induced refractive index with grooves in silica-based planar lightwave circuits. *Electronics Letters* 41 (2005), 1118–1119. Issue 20.

[26] M. Nikdast, G. Nicolescu, J. Trajkovic, and O. Liboiron-Ladouceur.

[27] C. Nitta, M. Farrens, and V. Akella. 2011. Addressing system-level trimming issues in on-chip nanophotonic networks. In *IEEE 17th International Symposium on High Performance Computer Architecture (HPCA)*. 122–131.

[28] C. Nitta, M. K. Farrens, and V. Akella. 2011. Reilient Microring Resonator Based Photonic Networks. In *Proceedings of the 44th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*.

[29] J. S. Orcutt, A. Khilo, C. W. Holzwarth, M. A. Popovic, H. Li, J. Sun, T. Bonifield, R. Hollingsworth, F. X. Kartner, H. I. Smith, V. Stojanovic, and R. J. Ram. 2011. Nanophotonic integration in state-of-the-art CMOS foundries. *Optics Express* 19 (2011). Issue 3.

[30] Yan Pan, J. Kim, and G. Memik. 2010. FlexiShare: Channel sharing for an energy-efficient nanophotonic crossbar. In *High-Performance Computer Architecture*. 1–12.

[31] Yan Pan, Prabhat Kumar, John Kim, Gokhan Memik, Yu Zhang, and Alok Choudhary. 2009. Firefly: Illuminating future network-on-chip with nanophotonics. In *Proceedings of the 36th annual international symposium on Computer architecture*.

[32] S. Postnikov, S. Hector, C. Garza, R. Peters, and V. Ivin. 2003. Critical dimension control in optical lithography. *Microelectronic Engineering* 69, 2-4 (2003), 452–458.

[33] C. Qiu and Q. Xu. 2011. Wavelength tracking with thermally controlled silicon resonators. *Optics Express* 19 (2011). Issue 6.

[34] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson. 2010. Silicon optical modulators. *Nature Photonics* 4 (2010), 518 – 526.

[35] Smruti R. Sarangi, Brian Greskamp, Radu Teodorescu, Jun Nakano, Abhishek Tiwari, and Josep Torrellas. 2008. VARIUS: A Model of Process Variation and Resulting Timing Errors for Microarchitects. *Semiconductor Manufacturing, IEEE Transactions on* 21, 1 (feb. 2008), 3 –13.

[36] J. Schrauwen, D. Van Thourhout, and R. Baets. 2008. Trimming of silicon ring resonator by electron beam induced compaction and strain. *Optics Express* 16 (2008), 3738–3743. Issue 6. https://doi.org/10.1364/OE.16.003738

[37] S. K. Selvaraja. 2011. *Wafer-Scale Fabrication Technology for Silicon Photonic Integrated Circuits*. Ph.D. Dissertation. Ghent University.

[38] N. Sherwood-Droz, K. Preston, J.S. Levy, and M. Lipson. 2010. Device guidelines for WDM interconnects using silicon microring resonators. In *Workshop on the Interaction between Nanophotonic Devices and Systems*.

[39] S. Ueno, Naganawa T., and Kokubun Y. 2005. High UV sensitivity of SiON film and its application to center wavelength trimming of microring resonator filter. *IEICE Transactions on Electron* E88-C, 5 (2005), 998–1004.

[40] D. Vantrease. 2010. *Optical Tokens in Many-Core Processors*. Ph.D. Dissertation. University of Wisconsin-Madison.

[41] D. Vantrease, N. Binkert, R. Schreiber, and M.H. Lipasti. 2009. Light Speed Arbitration and Flow Control for Nanophotonic Interconnects. In *Proceedings of the 42 Annual IEEE/ACM International Symposium on Microarchitecture*. 304–315.

[42] Dana Vantrease, Mikko H. Lipasti, and Nathan Binkert. 2011. Atomic Coherence: Leveraging nanophotonics to build race-free cache coherence protocols. In *Proceedings of the 2011 IEEE 17th International Symposium on High Performance Computer Architecture*. 132 –143.

[43] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N.P. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R.G. Beausoleil, and J.H. Ahn. 2008. Corona: System Implications of Emerging Nanophotonic Technology. In *35th International Symposium on Computer Architecture*. 153–164.
[44] K. Williams and P. Watts. 2011. Optical interconnects for NOCS and off-chip communications. Tutorial, Intl. Symp. on Networks-on-Chip. (2011).
[45] R. Wu, C. H. Chen, C. Li, T. C. Huang, F. Lan, C. Zhang, Y. Pan, J. E. Bowers, R. G. Beausoleil, and K. T. Cheng. 2015. Variation-aware adaptive tuning for nanophotonic interconnects. In *2015 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. 487–493.
[46] Danxia Xu. 2011. Polarization Control in Silicon Photonic. *Topics in Applied Physics* (2011), 31–70.
[47] Q. Xu, D. Fattal, and R. G. Beausoleil. 2008. Silicon microring resonators with 1.5-μm radius. *Optics Express* (2008), 4309–4315.
[48] Q. Xu, S. Manipatruni, B. Schmidt, J. Shaky, and M. Lipson. 2007. 12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators. *Optics Express* (2007), 430–436. https://doi.org/10.1364/OE.15.000430
[49] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson. 2005. Micrometre-scale silicon electro-optic modulator. *Nature* (2005), 325–327.
[50] Y. Xu, Y. Du, Y. Zhang, and J. Yang. 2011. A composite and scalable cache coherence protocol for large scale CMPs. In *International Conference on Supercomputing*. 285–294.
[51] Y. Xu, J. Yang, and R. Melhem. 2012. Tolerating Process Variations in Nanophotonic On-chip Networks. In *The 39th International symposium on computer architecture*.
[52] Y. Xu, J. Yang, and R. Melhem. 2015. BandArb: mitigating the effects of thermal and process variations in silicon-photonic network. In *Proceedings of the 12th ACM International Conference on Computing Frontiers*.
[53] S.J.B. Yoo. 2010. CMOS-compatible silicon photonic integrated systems in future computing and communication systems. In *OptoeElectronics and Communications Conference (OECC), 2010 15th*.
[54] Y. Zheng, P. Lisherness, M. Gao, J. Bovington, K. T. Cheng, H. Wang, and S. Yang. 2012. Power-efficient calibration and reconfiguration for optical network-on-chip. *IEEE/OSA Journal of Optical Communications and Networking* 4, 12 (Dec 2012), 955–966.

Received June 2017; revised Auguest 2017; accepted November 2017