MIN: A Power Efficient Mechanism to Mitigate the Impact of Process Variations on Nanophotonic Networks

Majed Valad Beigi and Gokhan Memik
Department of EECS, Northwestern University, Evanston, IL majed.beigi@northwestern.edu and memik@eecs.northwestern.edu

ABSTRACT
In this paper, we introduce MIN, a novel method for assigning wavelengths to nodes dynamically on a nanophotonic network to minimize the impact of process variations (PVs). Among the available wavelengths on a waveguide, a subset of them, called bubbles, are left unintentionally unused. These bubbles are then borrowed by nodes dynamically to improve the channel utilization. We present an express wavelength regulation approach to manage these assignments efficiently. Evaluation results reveal that MIN can recover 80% of bandwidth loss due to PVs and achieve 41% trimming power reduction compared to state-of-the-art alternative approaches.

Categories and Subject Descriptors
C.1.2 [Computer System Organization]: Process Architectures – Interconnect architectures

General Terms
Design, Reliability, Performance

Keywords
Nanophotonic, Process variations, Networks-on-chip

1. INTRODUCTION
The advances in deep sub-micron technology have enabled the integration of a variety of components onto a single chip. Furthermore, ITRS [1] has predicted that the number of processing cores in multiprocessor systems-on-chips (MPSoCs) will reach 1,000 by 2020. Therefore, with increasing on-chip networks size, the energy efficiency, latency, and high bandwidth of on-chip networks are becoming even more important [2]. Electrical networks have several limitations in such large scales and optical interconnect networks is considered one of the most important potential alternative for future many core microprocessors [3-5].

Optical interconnects possess many promising characteristics. However, they also face several key challenges. Arguably, the most important challenge is the reliability of nanophotonic on-chip network components. Typically, there are two crucial obstacles for reliable nanophotonic-on-chip communications: Process variations (PVs) and thermal sensitivity of silicon photonic devices [6-8, 22]. PVs refer to inaccuracies in fabrication and changes in physical dimensions such as thickness of silicon or, width of waveguide caused by lithography imperfection and etch non-uniformity of devices [7]. On the other hand, thermal sensitivity relates to spatial and temporal variation in the refractive index of silicon-photonic devices [9]. Those changes will directly affect the microring resonator (µring) [10-12], which is extremely compact (3-10 µm radius). Under process and thermal variations, µrings fail to resonate at the designated wavelengths. This leads to communication errors and bandwidth loss. Several studies have reported that a 1 nm variation in the µring width can shift a µring’s resonance by approximately 0.5 nm. Similarly, a single degree change in temperature can shift a µring’s resonance by roughly 0.1 nm [13, 14].

Moreover, the spacing between two adjacent wavelengths is less than 1 nm [9] in a wavelength division multiplexing (WDM) enabled optical interconnect; this means µring diameter variation of 2 nm or thermal variation of 10°C would bring the µring to resonate at the neighboring wavelength. If a µring resonates at an unassigned wavelength, the data sent on the new wavelength is corrupted. In addition, the original wavelength is not used causing a decrease in the achievable bandwidth.

There exist two common techniques to eliminate resonance frequency shifts of µrings. The first one is post-fabrication physical trimming, where the refractive index of a µring is adjusted by employing high energy particles such as UV light or electron beam to correct the resonance [7, 8, 15]. However, since this technique requires trimming for each µring individually, it is unclear that it is practical for high volume production [7, 10]. In addition, this technique does not eliminate resonance shifts due to temperature variations. The second technique is power trimming. Specifically, to correct the resonance wavelength, µrings are either heated (for red-shifting, i.e., increase their resonance wavelength) or carrier injection is applied (for blue-shifting, i.e., to reduce the resonance wavelength). The most important problem with this approach is the high power consumption [6, 10, 13]. For example, a total of ~26W is necessary for trimming in the Corona network, which constitutes 54% of the total power dedicated to the network [5]. In this paper, we present techniques to minimize the trimming power needed for correcting resonant wavelengths.

We propose a novel methodology, called MIN, to reduce the impact of PVs in optical networks while keeping bandwidth at maximum and power consumption at minimum. The objective of this approach is to maximize the number of usable wavelengths for all nodes through properly arranging µring resonant wavelengths while considering the power consumption. To reach this goal, the first step of our approach is based on the idea that a µring should be trimmed to a nearby wavelength rather than the one statically assigned to it. Hence, we divide the wavelengths in a waveguide to a number of regions where their wavelengths are determined based on the power trimming limitation of current injection and heating. We also define a region within which the resonant wavelength can be shifted. These shifts are made possible by our initial assignment of the wavelengths. Specifically, MIN divides the wavelengths into regions and leaves a few wavelengths unused within each region. These unused wavelengths are called bubbles. Depending on the traffic rate and the conditions of the µrings, these bubbles can be allocated to different nodes. This allocation is performed at fixed intervals. At each time epoch, we perform the allocation of the bubbles to requesting nodes while minimizing the power needed for trimming. We give higher priority to nodes under heavy PVs than others. It is worthy to mention that using bubbles creates more opportunities for selecting fine-tuned µrings, which result in less trimming power to correct their resonance wavelengths. Further, the wavelength assignments are monitored periodically by an express wavelength regulation approach that is triggered to manage the wavelength assignment and also arbitrate among multiple nodes requesting access to the same bubbles efficiently.

The idea of dynamically assigning wavelengths to µrings has been explored recently: Nitta et al. and Xu et al. have shown the benefits of allowing µrings to resonate at its closest wavelength instead of the originally assigned one [6, 7]. However, these approaches rely on additional hardware cost (e.g., using additional µrings) and/or are
power hungry [8]. Our results, described in Section 4, reveal that our approach achieves 41% lower consumption and 61% higher bandwidth in comparison against these state-of-art methods. We also show that this is achieved with minimal area overhead; the hardware necessary to implement MIN causes 3.5% area overhead compared to the baseline design.

The rest of this paper is organized as follows. Section 2 describes our methodology. The experimental results are presented in Section 3 and Section 4 concludes the paper with a summary.

2. MIN Methodology

In this section, we present the proposed MIN methodology. As we mentioned earlier, µrings are drifted from their resonance due to process and thermal variations. This drift can be corrected by heating or current injection. Both of these methods are power hungry. As a result, the total trimming power increases with the total number of µrings. Moreover, even using efficient µring heating approaches such as in-plane heaters and air-undercut [16, 17] does not result in significant reduction: it is estimated that µring heating power still consumes 38% of total network power with efficient trimming hardware [4]. Another important property of these techniques is that the power consumption varies with the amount of wavelength shift: on average, the power needed for blue- and red-shifting of a 3-µm µring are 130 µW/nm and 240 µW/nm, respectively [6, 10]. Hence, minimizing the amount of shift results in reduced trimming power consumption.

Our technique, described in the next section, tries to achieve that. In addition to the power limitation, another important limitation is the amount of shift possible with this technique. Specifically, current injection to shift the resonance wavelength towards the blue end of spectrum causes signal loss: 0.4nm shift results in 1 dB signal loss [7,13]. As a result, the current injection can be applied only for short shifts. By using bubbles and efficient wavelength assignment to use less trimming power under PUs, our technique can reduce the power and increase the bandwidth. Here, in order to describe the MIN approach, we first provide an overview of the proposed design. Second, the details of the MIN technique are described. Third, we explain the proposed wavelength assignment mechanism and the wavelength regulation.

2.1 Overview of MIN design

There are three alternatives to organize the optical channels to build nanophotonic crossbars, namely SWMR, MWSR, and MWMR [3-5]. In this paper we employ MWMR [4] as the baseline crossbar design. We have chosen this topology as it provides the most energy-efficient topology compared to SWMR and MWSR. We slightly modify the topology in order to support the MIN approach. The channel defined here is a set of wavelengths used to transfer packets. MWMR, which is a dynamic channel allocation scheme, can improve channel utilization and network throughput with channel sharing. Although MWMR requires additional router complexity and more µrings compared to SWMR and MWSR, the added flexibility and reduced overall power consumption makes it more attractive than other topologies. In addition, the large number of µrings in MWMR allows the design to be provisioned with any number of channels. As a result, PUs can be handled efficiently without the need of adding extra µrings. Finally, we must highlight that our approach can be easily applied to SWMR and MWSR.

In our design, a set of wavelengths is devoted for establishing bubbles to be employed in the presence of PUs. Bubbles operate similar to other wavelengths in the network but they can be borrowed by nodes for a determined period of time. However, if a node with PUs requests an idle bubble, it will have priority over other nodes. In MIN, we use a single-round data channel [4] because it reduces the waveguide length, which reduces the overall energy consumption.

2.2 MIN Mechanism

The first step in our approach is based on the idea that a µring does not have to be trimmed to a predefined wavelength that may be far from its current resonating wavelength. We first divide the waveguide into regions; each µring is shifted within its region. Specifically, for each interval, we find the region that a µring belongs to for that interval (µring resonates at a wavelength in that region). Since we focus on PUs, there is specifically one region for each µring throughout execution. If thermal variations are taken into consideration, the region should be periodically determined. The technique for computing regions is divided into two sequential phases. Figure 1 shows the different phases. Initially, the technique receives the waveguide information including the number of wavelengths and the distance between them. Besides, it also receives restrictions for trimming in both direction and the power for a given shift amount. The length of the region is denoted by d and is determined based on the number of wavelengths and the number of bubbles in the network.

In the next step, one wavelength is assigned as a bubble in each region to provide a flexible wavelength assignment. The bubble is regularly put in the center of each region so that it is easily reachable from each µring that might be trimmed to it. It should be noted that, in some extreme cases regions may have an overlap; for example, if the number of wavelengths are not divisible by the number of regions, some wavelengths may fall into two regions. Since MIN uses trimming algorithm only inside the region, the policy for trimming are accomplished based on the unit power for current injection and heating. In MIN, we try to trim a µring inside of the region. The reason we define (and remain within) a region is to limit the amount of shift necessary and to minimize the cost of arbitration. In other words, in each region a µring is trimmed to a nearby bubble rather than its actual wavelength, thus trimming distance is reduced and trimming power is decreased. Moreover, the mechanism of trimming in each region can reduce the number of uncorrectable µrings as their trimming distances are now smaller.

The pseudo-code of the proposed policy is shown in Figure 2. In this pseudo-code $\lambda_{idle}$ represents the available idle wavelength in the region, $b_i$ shows the bubble in region $r$, $\mu_{current}$ illustrates the current µring location and $\sigma$ denotes the current injection trimming constraint and is assumed to be 0.4 [13]. Additionally, function $\mu_a. B_i$ shows the absolute value of the difference between wavelength $a$ and $b$. $\text{Current injection}()$ and $\text{heating}()$ are the functions that operate for current injection and heating trimming techniques, respectively. As can be seen from pseudo-code, the policy first checks the difference between the current µring location and its nearby bubble and also the difference between the current µring location and its nearby available idle wavelength. If one of these distances is smaller than the other one the µring will be trimmed to that wavelength (actual one or bubble). In this condition if this difference is less than $\sigma$, the current injection...
trimming technique is triggered to tune it; otherwise the heating approach is performed to shift the µring.

Figure 3 illustrates the trimming approach with a sample example. As depicted in the figure, there are four wavelengths which are distributed into two regions. Each region includes two wavelengths, λ1 and λ2 are bubbles while λ3 and λ4 are regular wavelengths. Here the assigned wavelength of µring #2 and #4 are λ1 and λ3 respectively. In Fig. 3-b, PVs cause µrings #2 and #4 to be closer to λ1 and λ3 respectively. The baseline design trims the two µrings back to their actual wavelengths as illustrated in Fig. 3-c. However, with MIN, µring #2 will be trimmed to λ1 in region #1 and µring #4 to λ3 in region #2 which clearly consumes less trimming power than the baseline. As we mentioned before, in each region one bubble is dedicated for helping µrings with PVs. However, in high load traffic this bubble can be borrowed by any node for a period of time. So, if there are no PVs, bubbles can be used efficiently. In other words, our approach does not place a limit on the utilization of the network. In the case that PVs occur and there is no available wavelength in a region for a µring to be trimmed to, the bubble is released from its current node immediately after the time epoch and assigned to the µring affected by PVs. As we will describe in the next section, this process is completed within a few cycles, hence there is no possibility of long durations of starvation.

2.3 Wavelength Regulation

Figure 4 illustrates the wavelength regulation established by MIN which is triggered for each node. The wavelength regulator is placed at the end of the waveguide as the home node. Therefore, no additional physical connectivity is required. As seen in the figure, each node employees token arbitration similar to the baseline design to send its requests to use wavelengths. However, some modifications are needed to support assignment of bubbles. With no PVs, wavelengths are assigned to each node similar to the baseline design. Nevertheless, we record the region that each wavelength is located at. In some conditions, such as high traffic rate, nodes may send requests to borrow bubbles. As we mentioned previously, this happens at the boundary of time epochs of T cycles each. Each node can borrow a bubble for one epoch. After that period, if a node wants to keep the bubble it should send its request again to a wavelength regulator and the regulator decides whether to assign the bubble to the node or not.

MIN first checks the possibility that a µring can be trimmed to nearby idle wavelength in its region, if this wavelength is not a bubble, a µring will be trimmed to a wavelength in its region. However, if this wavelength is a bubble and the bubble is idle it will be shifted to it. If all the wavelengths (bubble and other wavelengths) in the region are busy, node sends its request to the regulator to trim the µring to the bubble. Once the regulator receives the node's request, it checks the bubble in the node region and if the bubble has been borrowed by another node for bandwidth purpose (node), the regulator releases the bubble immediately after finishing its epoch time and the µring of node, is easily trimmed to the bubble. The size of the epoch should be determined efficiently to prevent increasing the time that a µring has to wait before having access to the bubble. However, during this wait time, the regulator checks the region periodically and if a nearby wavelength is released before finishing the epoch time, this wavelength is used by the µring. Hence, the number of usable wavelengths for each node is maximized and we can achieve the highest bandwidth.

2.3.1 Express wavelength regulation

With perfect fabrication process, it does not matter which wavelengths are assigned to each node. With PVs, however, determining which wavelengths are assigned to a node is crucial since a wavelength may not be usable by one node but usable by another. Therefore, in order to achieve efficient wavelength assignment, we define Associate Control Logic (ACL) that is kept at the home node. ACL manages the wavelength assignment for multiple nodes. Since a request from the nodes is sent to the ACL in the home node, it first checks the wavelength availability and then according to the proposed approach, it assigns a suitable wavelength to the node. If multiple nodes send requests to use the wavelengths, it stores their information in its tables and assigns an appropriate wavelength dynamically to each node based on the nodes priority. ACL includes four modules: 1) wavelength control table, 2) the wavelength availability tracker, 3) node arriving/departing pointer logic, and 4) wavelength dispenser. All four modules operate independently and in parallel. The operation of each module is described as follows:

1) The wavelength control table is the core of the ACL. It is a compact table holding IDs of all nodes (Node-ID). The wavelength control table is organized by wave-ID, bubble-ID and region-ID which are established for assigning wavelengths efficiently. While the requests are sent from the nodes, their IDs are placed to the table. In each epoch, the table information is updated. If a wavelength is idle, a tag of the node relating to Node-ID is NULL. When a wavelength is allocated to the node, ACL updates the control table accordingly. In each region one wavelength is defined as a bubble and the bubble-ID of the rest of the wavelengths in each region are NULL.

2) Wavelength availability tracker keeps track of all the wavelengths in the wavelength control table. If a node makes a request, the tracker selects a wavelength based on the region-ID and the available wavelengths (which can be found by monitoring the node-ID and the wave-ID in the wavelength control table).

3) Arriving/Departing node pointer logic directly controls the node requests. It is also directly linked to wavelength control table. If multiple nodes want to access the same bubble, their requests are monitored in the node arriving/departing table and the information about them is stored in the table. For each incoming request from the node, we store the priority and the request order for the node.

The Node-ID is obtained according to ID of each node while the FIFO-ID is determined based on the time that each node has sent its request. We also rank each node considering both FIFO-ID and Priority tag. As for the regular wavelengths that are not bubbles, the values of FIFO-ID and Priority are same. However, if the wavelengths are bubbles, the Priority is calculated as follows:

\[
\text{Priority} = \text{FIFO-ID} \times \text{Max} (\text{FIFO-ID})
\]

where max (FIFO-ID) is the maximum number among all the FIFO-ID values. By employing the priority tag, the bubbles can be assigned to the nodes affected by PVs immediately. Clearly, the nodes with higher priority values will be given priority over nodes with lower priority value.

4) The wavelength dispenser is responsible for dispensing free wavelengths to requesting nodes. The wavelength dispenser dynamically assigns wavelengths to the requesting node; they are returned to the dispenser upon release. Based on the information provided by the wavelength availability tracker, the wavelength dispenser decides whether to grant a wavelength or not.

We have implemented all the MIN structures in VHDL and compiled it using Xilinx ISE tools. Based on the results, we estimate the overall area cost of the structures to be 3.5% of the baseline interconnection network.

3. EVALUATION

In order to evaluate the proposed MIN, we first need to provide a model of the impact of PVs on the optical network. To characterize the impact of PVs on unassigned wavelengths, we have utilized a variation model for the physical dimensions of the optical waveguide based on die-to-die (D2D), a.k.a. intra-die, and within die (WID), a.k.a. inter-die, variations [18]. We generate a set of variation parameters based on WID and D2D results [21] since both of them use small dies (2×2 mm²) with 0.537nm WID variation and 1.08nm D2D variation. As for large dies, a PV modeling infrastructure for CMOS technology, called VARIUS [20] is used to model both WID and D2D variations [7]. The results from the small die in [21] are placed into VARIUS. Then VARIUS generates 100 sample dies of 400 mm² each. On average, these dies have 0.61nm WID variation
and 1.01nm D2D variation. In the results we present in this section, we use both results from [7] and [21] for small and large dies to evaluate the effectiveness of the proposed approach. Moreover, we modify the cycle accurate network simulator booksim [19] to compare the proposed schemes with alternatives. We use a MMWR crossbar and simulate crossbar size of K=16 with single flit packets of 64 bytes. We assume Dense Wavelength Division Multiplexing where 64 wavelengths are transmitted in a single waveguide (in both directions). The power consumption of the network includes the trimming power, which is 0.13 mw/nm for current injection and 0.24 mw/nm for heating [6, 10]. We also take the distance between adjacent wavelengths, ∆λ, into consideration. We assume that current injection can correct up to 0.5∆λ towards blue shift [13].

Figure 5 shows the power consumption among different schemes for the two ∆λ values (0.5 and 1) and with different region sizes. The epoch size is 128 cycles. The results are normalized to Closest trimming scheme. In the Closest scheme, the trimming is done based on shifting a group of m rings to their closest λs (similar to sliding window scheme by Nitta et al. [6]) while the 'Actual' trims the m rings to their originally assigned wavelength. We test two different configurations of our scheme; d=2 and d=4, where d is the region size described in the previous section. Although the Closest approach eliminates large trimming distances, MIN with d=2 consumes less trimming power compared to it. Specifically, 41% power reduction is observed in MIN in comparison to the Closest approach. This achievement is obtained because we provide freedom for each m ring by employing a bubble and also triggering efficient wavelength regulation mechanism to manage the bubbles. However, as depicted in the figure, if the length of region increases, the trimming power in our approach rises too. The reason for such increases relies on boosting the trimming distance between the nodes in the region and the bubble.

![Figure 5. Normalized trimming power for two ∆λ values: (a) 0.5 and (b) 1.](image)

Figure 5. Normalized trimming power for two ∆λ values: (a) 0.5 and (b) 1.

Figure 6 illustrates the network bandwidth comparison between MIN and other schemes. The results are also normalized based on the 'Actual' trimming scheme. The 'Actual' scheme is assumed that it can achieve 100% bandwidth after trimming. The network bandwidth metric in our evaluation is obtained based on the number of working channels considering both tuned senders and receivers. In this figure in addition to comparing MIN with other schemes, we also contrast the proposed approach in different regions and epoch sizes to understand the effect of these parameters efficiently. As depicted in the figure, the proposed methodology achieves nearly 80% of network bandwidth in both ∆λ values. The Closest approach, on the other hand, can utilize only 24% of the bandwidth. The reason of this improvement was illustrated in Figure 3: MIN can reduce the uncorrectable m ring by finding a nearby wavelength without disrupting any other communication. It is worthy to note that the high network bandwidth achieved with our approach is accomplished without using spare m rings. This also helps in obtaining lower power consumption. Additionally, as illustrated in the figure, MIN acquires higher bandwidth in epoch size of 128 compared to epoch size of 256. It is intuitive that a smaller epoch size results in higher bandwidth: if more than one m ring request the same bubble and the bubble is grabbed by a node for a long period (epoch time), it will lead to underutilization of the network. On the other hand, if the epoch size is too small, it will cause frequent requests and increase the power consumption.

In summary, Figures 5 and 6 indicate that MIN provides a robust mechanism to achieve high network bandwidth and low trimming power under PVs compared to recently proposed alternatives.

4. CONCLUSION

This paper presents a novel methodology, called MIN, to alleviate the impact of PVs on optical networks. We improve the level of flexibility and accuracy of wavelength assignment by providing efficient wavelength arrangement and regulation. The proposed approach allocates bubbles and allows nodes to borrow wavelengths in order to reduce trimming power and achieve maximal bandwidth. Evaluation results show that the proposed methodology results in 41% less power consumption and over 3x improvement in bandwidth compared to state-of-the-art approaches. Hence, it provides a significant level of reliability with lower hardware overhead, latency and energy compared to alternative designs.

5. REFERENCES