Post-Fabrication Reconfiguration for Power-Optimized Tuning of Optically Connected Multi-Core Systems

Yan Zheng[1,2], Peter Lisherness[2], Saeed Shamshiri[2], Amirali Ghofrani[2], Shiyuan Yang[1], and Kwang-Ting Tim Cheng[2]

1. Dept. Automation, Tsinghua University, Beijing, 100084, China
2. Dept.Electrical and Computer Engineering University of California, Santa Barbara, Santa Barbara, CA 93106 U.S.A.
e-mail : zheng-y03@mails.thu.edu.cn, peter@ece.ucsb.edu, {saeed, ghofrani}@umail.ucsb.edu, ysy-dau@tsinghua.edu.cn, timcheng@ece.ucsb.edu

Abstract— Integrating optical interconnects into the next-generation multi-/many-core architecture has been considered a viable solution to addressing the limitations in throughput, latency, and power efficiency of electrical interconnects. Optical interconnects also allow the performance growth of inter-core connectivity to keep pace with the growth of the cores' processing ability. However, variations in the fabrication process significantly impair an optical network's communication quality. Existing post-fabrication tuning methods, which are based on adjusting the voltages and temperatures, have very limited tunability and require excessive power to fully compensate for the variation. In this paper, we study the sources and severity of process variation, and propose two methods to enhance the robustness of an on-chip optical network: 1) adding spare modulators and detectors for post-fabrication reconfiguration and low-power tuning, and 2) introducing a combined detector/modulator structure for a more robust network topology. Simulation results show that employing both methods can reduce the tuning power from hundreds of watts to 6W while maintaining a throughput of 99.7%. To maintain a throughput of 50%, the tuning power can be further reduced to only 12mW.

I. Introduction

Multi-core systems are widely accepted as the dominant architecture of processors and system-on-chips [1]. Optical interconnect, which can provide high bandwidth density and low power consumption, have been proposed to meet future multi-core systems' bandwidth and performance demands [2].

One optical device, the micro-ring resonator, is widely used in on-chip photonic networks due to its functional flexibility: it can function as a modulator [3], filter [4], switch [5], or laser 0. Because it is a wavelength-specific device, it also allows Dense Wavelength Division Multiplexing (DWDM) without the need for optical mux/demux. This increases the bandwidth density significantly, providing multiple independent and simultaneous data channels on a single waveguide. Ring resonators also do not require an optical gain material, and can therefore be manufactured in Si/SiO2 using existing silicon fabrication equipment.

As a case study in this paper, we focus on *Corona* [7], an optically connected 3D multi-core system targeting the 16nm process node in 2017. The system consists of 256 cores, organized in 64 clusters, providing peak floating-point performance of 20 teraflops, with all clusters connected by all-optical, high bandwidth DWDM waveguides with micro-ring modulators and detectors.

These ring resonators suffer from several kinds of variation, including both fabrication process variation (static variation) and run-time thermal variation (dynamic variation). DWDM requires all the ring resonators of the same data channel (which consists of 63 modulators and 1 detector in the *Corona* system) to have the same center resonant wavelength λ_R, which must further match the center wavelength of the data channel λ_C (determined by the light source). Variations resulting in resonant wavelength shifts can cause a mismatch in the resonant wavelength between the modulators and the detector, leading to a higher bit error rate (BER) and degrading the system performance, even to the point of complete failure [8].

In order to compensate for these variations and to obtain better communication quality, there are two major schemes for tuning the resonant wavelength: bias voltage tuning [9] and local heating [10]. Each of these methods has a limitation: 1) the tuning range of the bias voltage method is very limited (e.g. less than 1nm in [11]), and 2) local heating is very power hungry (in Section 3 we expound on this).

In this paper, we target only static process variation, for which the network needs to be configured only once after fabrication. To address dynamic thermal variations, Zheng et al. [8] use system-level thermal management with device tuning to regulate the temperatures of the processor cores on-line, which efficiently reduces the dynamic thermal variation.

To compensate for process variation, we study the sources and magnitude of variation in the resonant wavelength, and propose two post-fabrication reconfiguration schemes to enhance the yield and achieve a more reliable communication network without incurring high tuning power.

The contributions of this paper include:

1. We propose a spare-ring-based reconfiguration scheme to compensate for static variations. By adding several spare rings to each of the original modulators and detectors, we could choose the one whose resonant wavelength is closest to the channel wavelength and disable the others. The idea of adding spares to improve reliability has already been widely used in memories [12], multi-core chips [13], electrical network-on-chips [14], and through-silicon-vias [15].To our knowledge, this idea has never been studied for optical interconnects.

2. We propose the use of a combined modulator/detector in place of the regular modulators throughout the system. Such a

structure significantly strengthens the network connectivity and enhances the resilience of the communication fabric.

The rest of the paper is organized as follows: Section II describes the communication mechanism of an optically connected multi-core system. Section III discusses the variation in resonant wavelength induced by static process variation and the reliability challenges for optically connected systems like Corona. The proposed post-fabrication reconfiguration scheme based on spare rings is described in Section IV, followed by the combined modulator/detector structure, along with an analysis of its performance and power consumption, in Section V. Section VI shows the experimental results. Finally, Section VII concludes the paper.

II. Optically Connected Multi-Core Systems

A. *Ring Resonator Based Communication System*

Fig. 1 shows the inter-core communication mechanism in an optically-connected 3D chip [7]. This 3D structure is composed of three dies: a digital die for the cores, a photonic die for modulators and detectors, and a mixed-signal die to connect the cores to the optical devices and provide control over them. Digital data from cores, passed through DACs provide two-level bias voltage for each modulator. Switching the bias voltage brings a resonator in and out of resonance either to absorb a specific wavelength of the light carried on the waveguide (i.e. modulate logic-0), or to let the light pass (i.e. modulate logic-1). Receivers are SiGe-doped ring resonators. SiGe absorbs the light that couples into the ring and generates a photocurrent, which is further converted to a digital signal by an ADC before reaching the destination core.

DWDM enables high-bandwidth waveguides by sharing a waveguide among multiple data channels on different wavelengths. Thus, one core can simultaneously send data through multiple channels on a single waveguide using multiple modulator-detector pairs.

B. *Specific Case Study: Corona*

Shown in Fig. 2, *Corona* [7] has 64 four-core clusters connected by 64 bundles of waveguides. Each waveguide bundle is composed of four waveguides, and each of those carries 64 wavelength channels. An external light source injects these 64 wavelengths into an on-chip power waveguide. Each cluster has a splitter which splits a fraction of light from the power waveguide into its 4-waveguide bundle, providing a total of 256 data channels per cluster (64 channels for each waveguide). All waveguides travels through all clusters;

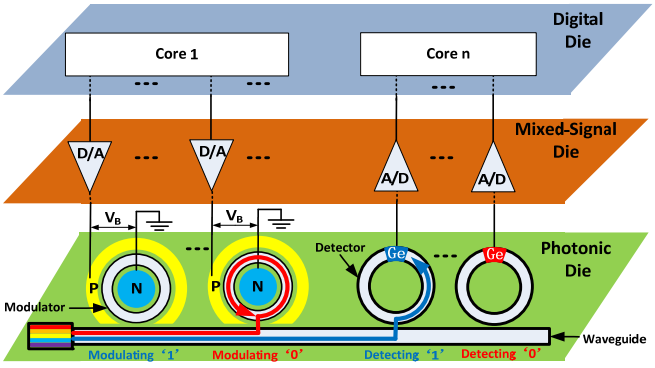

Fig. 1. Inter-core communication in an optically-connected 3D chip

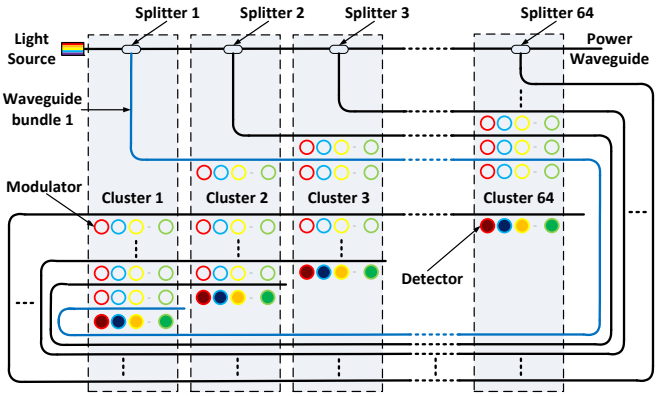

Fig. 2. Optical Communication in Corona

nevertheless, on each waveguide, only one cluster has detectors, while every other cluster has only modulators. Therefore, if a cluster wants to transmit data to cluster *m*, it can only use waveguide bundle *m*, the only waveguide on which cluster *m* has detectors. Using ring resonators with modulation rates of 10 GB/s [16], *Corona* could achieve a bandwidth of 320GB/s for each waveguide bundle, for a total system throughput of 20TB/s.

III. Reliability challenges

A. *Process Variation*

Reliable communication in an optical channel requires modulators and detectors to work on exactly the same wavelength. The resonant wavelength of a ring resonator, λ_R, is very sensitive to fabrication and thermal variations; 1nm change of the ring radius would lead to a shift of 0.583nm in the resonant wavelength [8], and 1°**C** change in temperature would lead to a shift of 0.11nm [17]. Deviation in λ_R will cause a mismatch among modulators, detectors, and the dedicated channel wavelength λ_C, which would in turn result in a higher bit error rate, and can cause the entire channel to fail.

Fig. 3 shows the ring yield in Corona as resonant wavelength variation increases, assuming no measures are taken to tune or correct it. In this experiment, the radius of ring resonators is 5um [16], and we assume that a deviation of 0.02 nm in λ_R makes a resonator inoperable (based on a 0.04 nm full-width spectral bandwidth [16]). Also, we assume that the variation follows a normal distribution. The horizontal axis of Fig. 3 indicates the standard deviation of the resonant wavelength. Based on our simulation results, *Corona* is very vulnerable to variation: a 0.03nm variation in resonant wavelength would already cause 50% performance degradation.

The effects of process and thermal variations depend on a number of factors including waveguide geometry, material, and fabrication process. In this paper, we only target process variation and consider the aggregate effect as variation in resonant wavelength.

Resonant wavelength deviation caused by process variation is a combination of random and systematic errors. For example, the thickness of the waveguide changes slowly

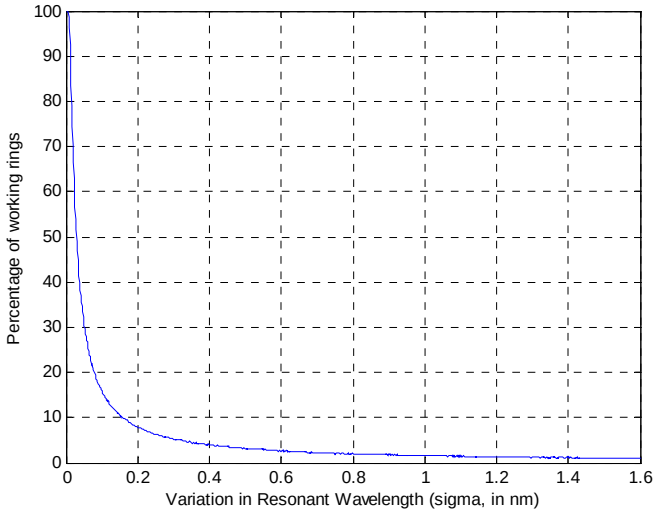

Fig. 3. Ring yield without tuning in Corona

across the die; therefore, a group of rings close to each other will experience strongly correlated variation, and their resonance wavelength will tend to shift together. However, the correlation becomes weaker for rings farther apart. That is, the resonant wavelengths of the rings belonging to the same core shift almost in the same fashion, while two rings belonging to distinct cores could be very different.

B. *Tuning Methods*

There are two major tuning methods to compensate for the aforementioned process variations: local heating and bias voltage tuning.

Local heating: heaters are built on top of or beside the rings to control their resonant wavelength. Recent studies show that a local heater can shift the resonant wavelength by 100 nm or more with a cost of 11.5 mW/nm [18]. With a titanium heater the tuning power could be improved to 2.4mW/nm [10].

Bias voltage tuning: adjusting the forward bias voltage of a ring resonator changes the density of electron-hole-pairs which in turn changes the refractive index of the waveguide and disturbs the resonant wavelength. Even though the power consumption of the bias voltage tuning method is one order of magnitude lower than that of local heating, a higher bias voltage leads to a smaller extinction ratio (i.e. the ratio of light coupled in the ring) and causes lower communication quality. As a result, the tuning range of the bias-voltage-tuning method is very limited and could be as little as 1nm [9].

These two tuning methods are complementary. Local heating offers a larger tuning range but is power hungry and relatively slow. On the other hand, the bias-voltage-tuning method, fast and power efficient, has a limited tuning range. To minimize the overall power consumption, we could employ a hybrid solution which uses local heating to coarse-tune and bias-voltage tuning to fine-tune the resonant wavelength to the center of the channel's wavelength. However, even with the hybrid solution, the power required for tuning the entire system is still too high.

Recent studies showed that the total tuning power for the entire multi-core system should be less than 10W; otherwise, there will be no power consumption advantage in using optical interconnects [19]. In *Corona*, there are over one million ring resonators. Thus, the average per-ring tuning power budget would be under 10uW. This power budget completely eliminates the possibility of using local heating for tuning. Using the bias voltage tuning method, this power budget can only adjust the resonant wavelength up to 0.025nm, which is definitely not enough.

Fig. 4 shows the tuning power required for *Corona* using the aforementioned hybrid solution. This simulation result confirms that the total tuning power will easily exceed the power budget (i.e. 10W).

IV. Post Fabrication Reconfiguration I: Spare Rings

A. *Overview and Cost Analysis*

To make the optical communication reliable without exceeding the power budget, we propose to include spare rings for post-fabrication reconfiguration. The idea is to design multiple spare rings for every modulator and detector. After fabrication, we test the resonant wavelength of each ring. Among those spare rings, we select the one with the closest wavelength to the channel's wavelength and disable the rest. Similar to memory repair [12], a laser beam could be used to disable the rings which are not selected after fabrication. Choosing the best ring out of multiple available spares significantly reduces the effective resonant wavelength variation and thus lowers the tuning power consumption. The extra spares need to be completely disabled to prevent their stealing of any light from the waveguide, as the loss of light would lower the communication quality. Each group of spare rings would share the same control logic (see Fig. 5). When we disable a ring, we would also cut its trace so that it doesn't consume any power in normal operation.

ITRS [20] shows that the size of a high performance processor could be as large as 260mm^2. The optical components do not share a plane with the logic (they are likely on separate stacked dies), so this entire area is available for optical devices. To implement a ring resonator of 1.5um radius [21], including the overhead of waveguides, control pads, and traces, the active footprint of a ring would be around 25 um^2. Therefore, on a 260mm^2 die, we can fabricate over 10 million resonators. This capacity allows up to 10 rings (9 spares) for each of the 1 million modulators and detectors in *Corona*.

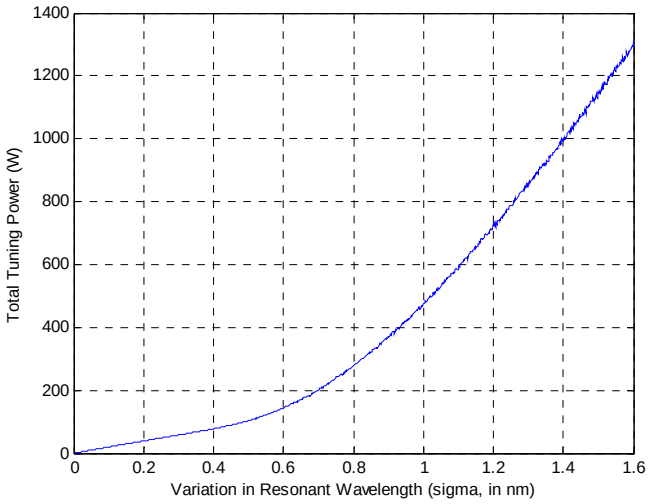

Fig4. Tuning power required for Corona w.r.t. the variation

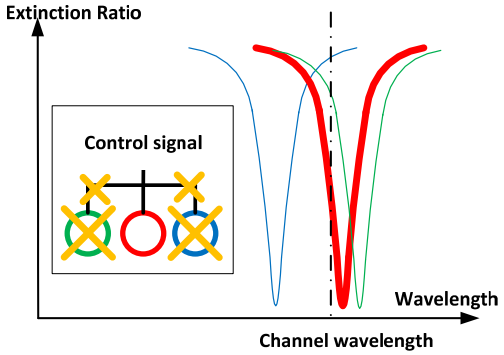

Fig5. Spare rings for random variation (3 rings)

B. *Spare Ring Parameters*

In this section, we discuss several parameters in designing the spare rings to compensate for both inter-core random variation and intra-core systematic variation.

Fully random uncorrelated variation causes each ring to shift independently. To compensate for random variation, all the spare rings would be designed to have the same resonant wavelength, which matches the channel wavelength. Fig. 5 shows an example of three spare rings for one modulator (inset). After fabrication, we measure the resonant wavelength of each ring, keep the one whose resonant wavelength is closest to the channel wavelength, then eliminate the other two. Based on our simulations, choosing 1 of 3 rings can decrease the variation to 42% of the original, and fabricating 10 can further decrease it to 14%.

Systematic variation causes the rings of a given cluster to shift together. The resonant wavelengths of the spare rings are designed to be slightly different from each other. Fig. 6 shows an example of 5 rings for each channel. The resonant wavelengths of the rings are evenly spaced. The wavelength gap between two adjacent rings would be selected based on the magnitude of the systematic variation. Note that the previous ring duplication scheme is simply a special case of this scheme where the wavelength spacing is equal to zero. Both random and systematic variations happen together during fabrication, so the amount of designed-in wavelength spacing among the rings should be selected based on the magnitude of each component.

Fig. 7 shows the effectiveness of using 10 rings, or 9 spares, with each scheme. The variation in resonant wavelength is 0.29nm per-cluster, and 0.87nm per ring [11]. Wavelength deviation in the figure is normalized to the original system without any spare rings. According to the results, properly spaced spare rings can decrease the variation to 20% of the original. This in turn could reduce the tuning power to 20% of the original.

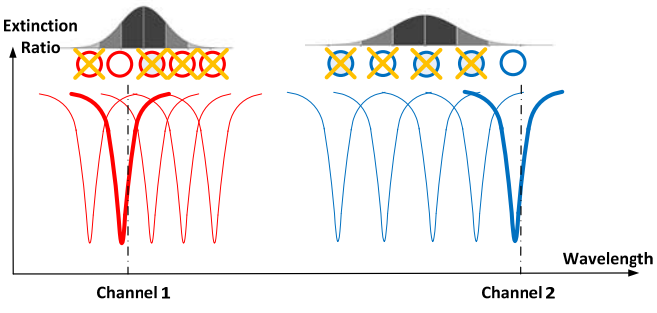

Fig6. Spare ring placement for systematic variation (5 rings)

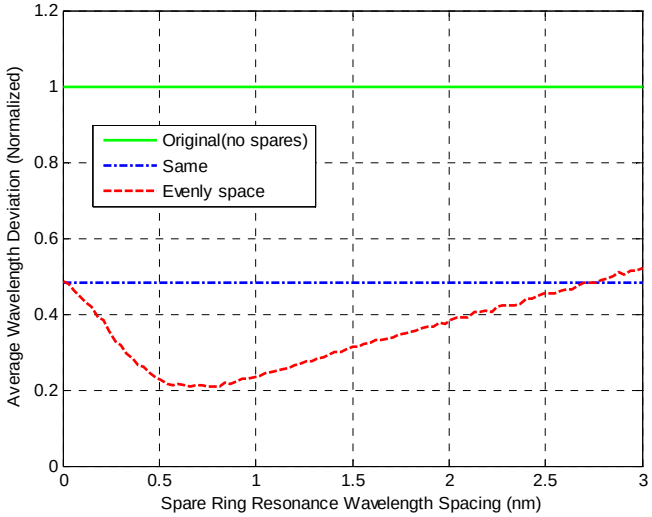

Fig.7. Effectiveness of spare placement schemes (10 rings)

V. Post Fabrication Reconfiguration II : Transceiver

A. *Effect of Reducing the Number of Working Rings*

Although spare-ring-based reconfiguration can decrease the effective process variation significantly, the total tuning power may still exceed the total tuning power budget. According to Fig. 4, if the overall variation is 0.87nm, the original anticipated tuning power is over 300W. Using spare rings limits effective variation to 0.17nm, and, according to Fig. 4, decreases the tuning power to about 30W. But this is still 3 times the original tuning power budget of 10W.

The next level of optimization is to entirely eliminate groups of rings that are collectively outside the acceptable variation tolerances. While doing so eliminates some links from the network, an improved connection scheme can tolerate these changes.

Fig. 8 shows the tuning power as a function of eliminated rings. If we can eliminate half of the worst quality rings, the total required tuning power could be decreased to 20% of the

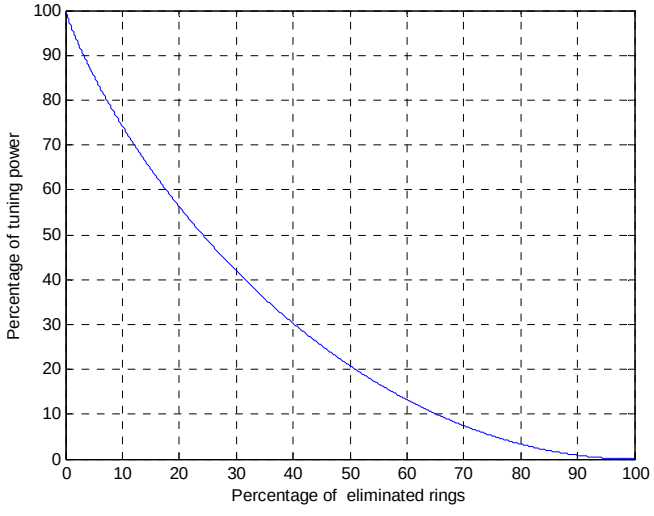

Fig8. Tuning power Percentage w.r.t eliminated rings

original (6W), bringing it within the total tuning power budget.

B. *Converting Modulators to Transceivers*

Bandwidth of a Corona-like system is very sensitive to the total number of working rings, and especially vulnerable to detector failure. This architecture has 64 ring resonators on each data channel, including 63 modulators and 1 detector.

Let's analyze the relationship between detector *n* (belonging to cluster *n*) and modulator *m* (belonging to cluster *m*). If modulator *m* is eliminated, one channel connection from cluster *m* to cluster *n* is lost, but all other clusters can still communicate with *n*. On the other hand, eliminating detector *n* would leave no more detectors on this channel, rendering it useless.

In order to solve this problem, we propose the addition of a detector to ever receiver in the system, effectively converting them from transmitters to transceivers. Fig. 9 shows two possible ring transceiver structures. The left one is a modulator with a Ge photo diode, a structure already proposed as part of the arbitration scheme in Corona. The right one is a modulator with additional drop waveguide and detector at the end of it. In this structure, light resonated into the ring will be coupled into the extra waveguide instead of simply resonating with the ring, and then detected by detector.

Adding this additional detection capability significantly changes the overall system connectivity. As shown in Fig. 10, when transceivers are used, additional communication options are added to every channel.

Note that at any given time, still only one pair of rings can use a given channel. So replacing the modulator with transceivers cannot increase the total bandwidth of the system, just the connectivity. In corona, we have in each channel 63 different communication pairs (from each modulator to the only detector), whereas in the new structure we could have as many as C(64,2) = 2016 different communication pairs *per channel*.

Converting modulators to transceivers will increase the area and control complexity; therefore, using both transceivers and spare rings may permit fewer spare rings.

In Fig. 11 we show a Corona-style network, but replace the original modulators with transceivers. In this system, A can send data to B only through the third waveguide (the only one where B has detectors), but now A can send data to B through

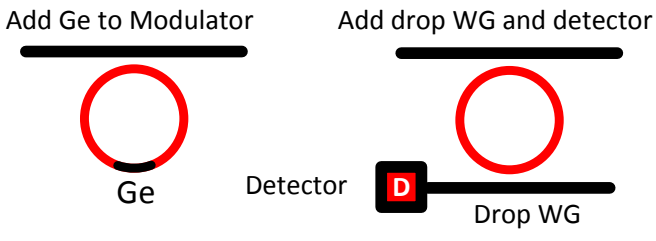

Fig9. Transceiver structure

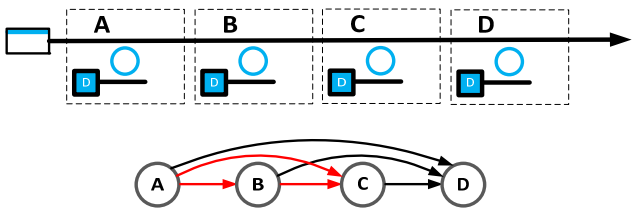

Fig10. Communication mechanism for 4 clusters using transceivers

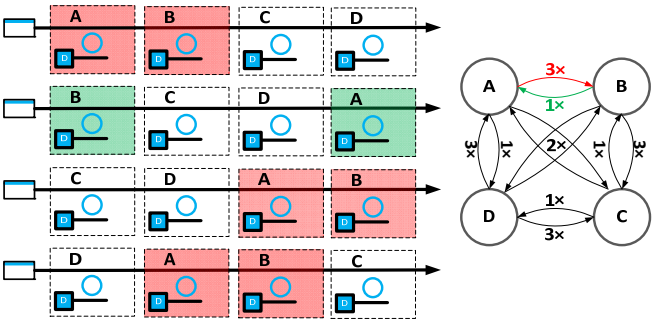

Fig11. Communication structure using transceivers for 4 clusters

3 different waveguides.

However, even with transceivers, B can send data to A also only through the second waveguide, because that is the only waveguide where B is upstream from A. This asymmetry will be exacerbated as the number of nodes increases.

To fix this asymmetry, recall that Corona uses 4 waveguide bundles for each cluster to increase bandwidth. Reversing the direction of half of the waveguides in these bundles will balance the connectivity: any given cluster now has 32 different ways to talk to any other cluster.

We analyze the throughput in two cases: the worst case (i.e. high congestion, all links saturated) and the best case (i.e. point-to-point throughput with no other traffic).

Assuming the throughput of one channel is T_C. In the worst case, every working ring is always attempting to take the channel to send data. For example, In Fig. 10, A, B and C are competing for the channel, and the throughput for each is at most $T_C/3$ (likely less due to arbitration overhead). If C is eliminated, the worst-case throughput from A to D and B to D will *increase* to $T_C/2$. In the new structure, since we have 6 communication pairs, the worst-case throughput of each pair is at most $T_C/6$. After C is eliminated, the throughput of each remaining pair will be $T_C/3$.

In the best case (i.e. no traffic or contention), the peak throughput of between any two clusters is defined by the total number of channels they can use. Again refer to the example in Fig. 10: the peak throughput of A, B and C to D is T_C. If C is eliminated, the peak throughput of A and B to D remains the same.

VI. Experimental Results

In this section we apply both proposed methods and observe the net improvement in yield and reliability. Fig. 12 shows the impact on the throughput when eliminating the worst rings for the proposed method. With the 50% of functioning rings, our proposed structure maintains 99.7% of the original worst-case throughput. From Fig. 8 we can see that this only needs 20% of the original tuning power, which is around 6W and within the original 10W tuning budget. Also note that the proposed method can use only 2.5% of the rings while still maintaining 50% of the original worst-case throughput. This could further reduce the tuning power to only 12mW (0.04% original).

Fig. 13 shows the system-wide average peak (no contention) throughput for Corona and our proposed structure as a function of operating rings. Results show that in our proposed method, only 17% of the rings are needed to match the peak

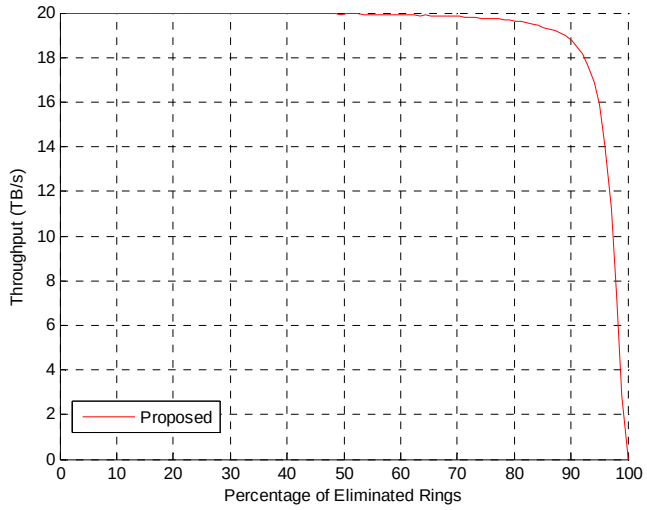

Fig12. Throughput w.r.t eliminated (worst) rings in worst case

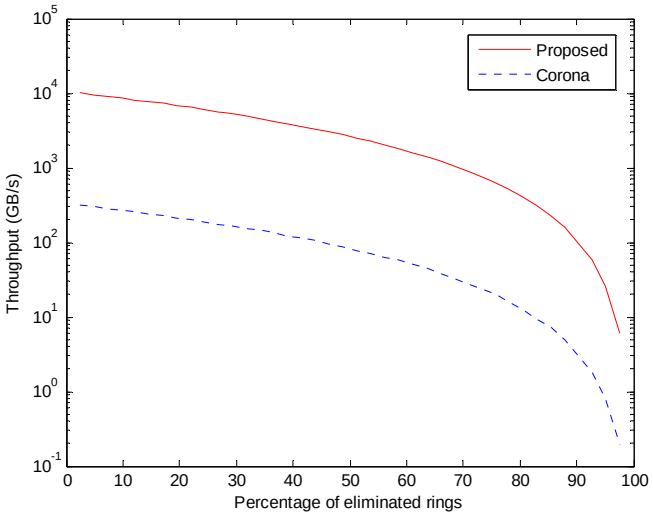

Fig13. Peak throughput w.r.t eliminated (worst) rings

throughput of Corona. This translates to a reduction in tuning power to 2.2% of the original.

VII. Conclusion

In this paper, we propose two post-fabrication reconfiguration methods to enable a reliable and power efficient optically-connected multi-core system. First, spare rings are proposed to reduce the effective variation and, in turn, the tuning power. Then, the connectivity of the system is enhanced by replacing the modulators with transceivers, which allows us to eliminate even more defective rings and cut the tuning power further, while still maintaining the original performance.

Experimental results show that only 12mW is needed to tune the rings in order to maintain 50% of the system's throughput, and maintaining 99.7% of throughput should be possible with as little as 6W of tuning power.

Acknowledgments

This research is partially supported by the Gigascale Systems Research Center (GSRC), one of six research centers funded under the Focus Center Research Program and the support by the Semiconductor Research Corporation.

Reference

[1] S.K. Moore, "Top 11 technologies of the decade; #5: Multicore CPUs," *IEEE Spectrum*, vol. 48, no. 1, pp. 40-42, Jan. 2011.

[2] G. T. Reed, et al., "Silicon optical modulators," *Nature Photonics*, July 2010.

[3] Q. Xu et al., "12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators," *Optcs Express*, vol. 15, no. 2, pp. 430–436, 2007.

[4] J. K. S. Poon et al., "Wide-range tuning of polymer microring resonators by the photobleaching of the CLD-1 chromophores," *Optical Letters*, vol. 29, no. 22, pp. 2584–2586, 2004.

[5] M. R. Watts et al., "Ultralow power silicon microdisk modulators and switches," in *IEEE Int. Conf. on Group IV Photonics*, pp. 4-6, Sep. 2008.

[6] D. Liang, et al., "Native-oxide-confined high-index-contrast lambda=1.15 μm strain-compensated InGaAs single quantum well ridge waveguide lasers," *Applied Physics Letters*, vol, 93, no. 16, pp. 161-108, 2008.

[7] D. Vantrease et al., "Corona: system implications of emerging nanophotonic technology," in *IEEE International Symposium on Computer Architecture*, pp. 153–164, 2008.

[8] Li Z., et al. "Reliability modeling and management of nanophotonic on-chip networks," *IEEE Trans. VLSI Systems*, no. 99, pp. 1-14, 2011.

[9] Q. Xu et al., "12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators," *Optics Express*, vol. 15, no. 2, pp. 430–436, 2007.

[10] M. A. Popovic, "Theory and design of high-index-contrast microphotonic circuits," *Ph.D. dissertation, MIT, Cambridge*, 2008.

[11] Li Z. "On-chip optical and electrical communication: analysis and design for system level performance and reliability," *Ph.D. dissertation, Tsinghua University, Beijing*, 2010.

[12] L. Youngs and S. Paramanandam, "Mapping and repairing embedded memory defects," *IEEE Design & Test*, pp. 18-24, January, l997.

[13] S. Shamshiri et al., "A cost analysis framework for multi-core systems with spares," in *IEEE Int. Test Conference*, Oct. 2008.

[14] S. Shamshiri and K.-T. Tim Cheng, "Yield and cost analysis of a reliable NoC," in *IEEE VLSI Test Symposium*, pp. 173-178, May 2009.

[15] I. Loi et al., "A low-overhead fault tolerance scheme for TSV-based 3D network on chip links," in *International Conference on Computer-Aided Design*, pp. 598-602, 2008.

[16] Q. Xu et al., "Micrometrescale silicon electro-optic modulator" *Nature*, vol. 435, pp. 325-327, May 2005.

[17] S. Manipatruni et al., "Wide temperature range operation of micrometerscale silicon electro-optic modulators," *Optical Letters*, vol. 33, no. 19, pp. 2185–2187, 2008.

[18] Po Dong, et al. "Broadly tunable high speed silicon micro-ring modulator," in *IEEE Photonics Society Summer Topical Meeting Series*, pp. 197-198, July 2010.

[19] Z. Li et al., "Global on-chip coordination at light speed," *IEEE Design & Test of Computers*, vol. 27, no. 4, pp. 54-67, July 2010.

[20] *International Technology Roadmap for Semiconductors*; http://www.itrs.net/, 2010.

[21] Q. Xu, D. Fattal, and R. G. Beausoleil, "Silicon microring resonators with 1.5-μm radius," *Optical Express* vol. 16, pp. 4309-4315, 2008.