Received January 28, 2019, accepted February 7, 2019, date of publication February 15, 2019, date of current version March 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2899662

A Novel Laser Source Supply Scheme for Optical Network on Chip

ZHUANGZHUANG LIU[1], HUAXI GU[1], KUN WANG[2], BOWEN ZHANG[1], AND KANG WANG[1]
[1]The State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China
[2]School of Computer Science, Xidian University, Xi'an 710071, China

Corresponding author: Huaxi Gu (hxgu@xidian.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61634004 and Grant 61472300, in part by the Fundamental Research Funds for the Central Universities under Grant JB170107 and Grant JB180309, in part by the China Postdoctoral Science Foundation 2018M633465, and in part by the Key Research and Development Plan of Shaanxi Province under Grant 2017ZDCXL-GY-05-01.

ABSTRACT With the increasing number of integrated processor cores on a single chip, the bandwidth, speed, and power consumption of electrical interconnect cannot meet the needs of researchers. As a new technology, optical interconnect has become an alternative scheme for electrical interconnection. However, the traditional supply strategy of laser power brings a lot of static power loss, which will become a factor that cannot be ignored in the development of optical interconnect. In this paper, we propose a dynamic laser source supply scheme based on the autoregressive integrated moving average model, which can reasonably adjust the laser source size by prediction. The simulation results show that our dynamic supply scheme can save more than 70% of the laser power compared with the static supply under the real application of the PARSEC flow benchmarks.

INDEX TERMS Optical network on chip, laser power, dynamic supply, ARIMA model.

I. INTRODUCTION

In order to meet the requirements of multi-core systems, integrating more and more cores on a single chip has become a means to improve its performance. Therefore, the interconnection architecture directly affects the performance of chip multiprocessors (CMPS) [1]. Traditional electrical interconnect cannot meet the high bandwidth, low delay and low power consumption requirements of Network on Chips. As a new technology, optical interconnect is being favored by more and more researchers because of its characteristics of high bandwidth and low power transmission over extended distances. These characteristics make nanophotonic an ideal solution for the challenges faced by interconnect designers [3].

Compared with electrical interconnect, optical interconnect has a significant improvement in performance, but it also faces some new challenges. The communication of optical network on chip needs laser source to provide energy. Firstly, the external laser source emits a certain number of wavelengths, which are coupled to the power waveguide by a fiber coupler and finally distributed to each node using a supply architecture. When the node has communication requirements, the energy on the power waveguide will be modulated into optical signal transmission. However, this static supply is not only poor in real-time, but also cause large static power consumption when the node has no communication demand. So, the static power loss of laser source is an important factor in the design of optical interconnect.

The associate editor coordinating the review of this manuscript and approving it for publication was Ting Wang.

A reasonable design of the laser power system can greatly reduce the power consumption of the Optical Network on Chip and improve the energy utilization efficiency. There have been many studies on laser source supply solutions, but most of them are designed for supply architecture. For example, Werner *et al.* [2] of the University of Manchester proposed an optical-electronic hybrid network Lego. The local communication using electrical links, and long-distance communication using optical links. A high-bandwidth link is divided into multiple low-bandwidth links to improve the utilization of optical links and reduce power consumption. However, this design increases the link utilization and reduces the communication delay, which also brings more crossover loss. Ohio University Kennedy and Kodi [4] proposed a CW/CCW supply architecture that reduces power consumption by decomposing crossbar while maintaining single-hop network communication. However, the architecture has high requirements for micro-rings, waveguide, and consumes a lot of on-chip resources. At the same time, there are some studies on the dynamic management of laser power. Zhou and Kodi [5] proposed an optical bandwidth scaling network based on prediction. By predicting channel utilization and buffer occupancy in the network, the optical bandwidth is dynamically scaled to reduce the static loss of the laser power. However, the number of laser power required in this design is too large and the scalability is relatively poor.

The research of dynamic laser power management is still not mature. In the design of [6], Wang *et al.* proposed a scheme that predicted network traffic through generalized linear model, and dynamically manage laser power based on network traffic. However, real network traffic has irregular, gusty characteristics, in which case the generalized linear model is no longer applicable. Therefore, this design is difficult to have a wide range of applications. In this paper, we propose a dynamic supply scheme based on ARIMA model. The ARIMA model can better describe the burstiness and self-similarity of network traffic, and it can better match the change of network traffic compared with the generalized linear model.

The rest of this paper is organized as follows. Section II describes the laser source supply architecture in Optical Network on Chips. The proposed laser source dynamic control scheme is presented in detail in Section III. The simulation results and is presented in Section IV. Finally, Section V concludes this paper.

II. LASER SOURCE SUPPLY ARCHITECTURE

The traditional laser power supply scheme is shown in figure 1. The external laser source distributes the laser power evenly to each node through the power waveguide to provide energy for routing node communication [7]. If the laser source provides a power of P, node 1 obtains $\frac{1}{n}$ of it, node 2 gets the remaining $\frac{1}{n-1}$, node 3 gets the remaining $\frac{1}{n-2}$, and so on, each node obtains a power of $\frac{1}{n}$P. In a moment, not all nodes in the network are communicating, and the power allocated to some of the nodes is not utilized. As a result, this power will be directly transmitted to the terminal of the power waveguide to be eliminated and caused a large part of the optical power will be wasted.

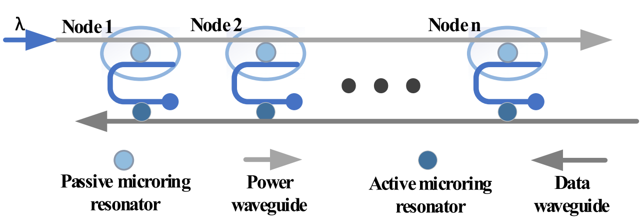


FIGURE 1. Traditional laser power supply scheme.

In this paper, the communication architecture uses a reservation-based Single Write Multi Read (SWMR) bus [8], as shown in Figure 2. The reserved channel ($CH_0 \ldots CH_{N-1}$) is used to establish communication, and the data waveguide ($DW_0 \ldots DW_{N-1}$) is responsible for the transmission of optical signals. All receivers are off by default. When the routing node attempts to send a data packet, Firstly, it broadcasts a flit through the reserved channel, which contains the destination information and the packet length information. Then, only the destination node receives the packet on the corresponding data channel, and the other nodes do not couple the laser energy. This avoids the expensive overhead of broadcasting packets on the data channel.

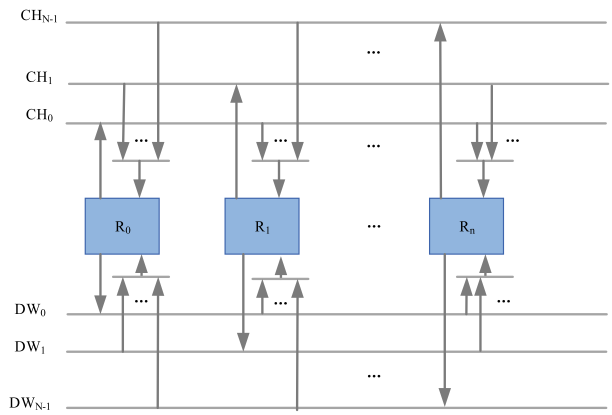


FIGURE 2. Reservation-based single write multi read bus.

We use the classic serpentine waveguide supply method [9], as shown in figure 3 is a 4 × 4 network layout. The external laser source emits a certain number of wavelengths, and then coupled into the power waveguide through a grating coupler. Finally, the laser power is evenly distributed to each node. If the node has communication requirement, we will modulate the wavelength to communicate with the destination node. If there is no communication requirement, the wavelength will be directly propagated to the end of the power waveguide and eliminated. We placed a power detector at the end of waveguide to measure the total optical power used at some points in the network. Assume that the power supplied by the laser source is P1 at a certain time, and the power detected by the power detector is P2. Using P1-P2 to indicate the power required in the network at this time.

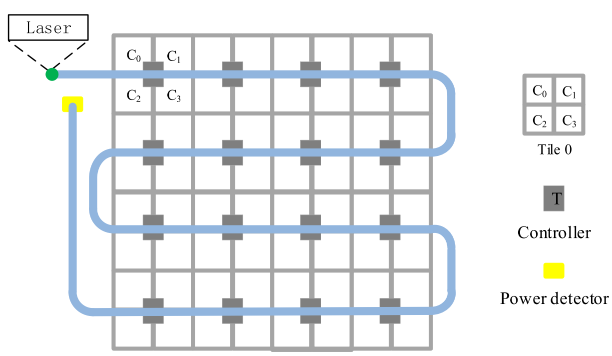


FIGURE 3. 4 × 4 network layout.

III. LASER SOURCE DYNAMIC CONTROL SCHEME

The amount of laser power that needs to be provided in the network at a time is directly proportional to the traffic in the network. Therefore, we adjust our laser power based on the distribution of traffic in the network. Network traffic is irregular and abrupt. We use the adaptive ARIMA model to predict the traffic in the network at the next moment based on the information of historical moments [11], and dynamically adjust the power of the laser source to reduce power consumption and improve energy utilization efficiency.

A. ARIMA MODEL

ARIMA contains three parts, AR, I, MA. AR represents the autoregressive model. I represents the single integer order, and the time series model must be a stationary sequence to establish the econometric model. Firstly, the roots of unit test should be carried out on the time series. If the time series is non-stationary, it will be transformed into stationary series by difference. After several times of difference, it will be called single integer of several orders. MA stands for moving average model. The ARIMA model is actually a combination of the AR model and the MA model.

Suppose $\{\mathrm{w_t} = 0, 1 \ldots\}$ represents a random time series, if there is a nonnegative integer d that satisfies:

$$\nabla^d w_t = x_t \tag{1}$$

Then the time series is a stationary time series, otherwise it needs to be smoothed.

We also definition polynomial $\varphi(B)$ and $\phi(B)$ as:

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_P B^P = 1 - \sum_{K=1}^{P} \varphi_K B^K \tag{2}$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_P B^q = 1 - \sum_{K=1}^{q} \phi_K B^K \tag{3}$$

where $\varphi(B)$ and $\phi(B)$ are relatively prime, $\varphi_p \phi_q \neq 0$, and $\{\alpha_i\}$ is white noise sequence

So $\{\omega_{\mathrm{t}}, t = 0, 1, 2, 3 \ldots\}$ should satisfy the difference equation:

$$\varphi(B)\, x_i = \phi(B)\, \alpha_i \tag{4}$$

where $|B| \leq 1$, $\{\omega_t, t = 0, 1, 2, 3 \ldots\}$ is an autoregressive differential moving average sequence, denoted as ARIMA (p, d, q) model, where d is the difference order.

Substituting Equation 1 into 4:

$$\varphi(1 - B)^{\mathrm{d}}\, \omega_t = \phi(B)\, \alpha_t \tag{5}$$

Equation 5 is called generalized autoregressive operator, and equation 5 can be converted into:

$$\varphi(B)\, \omega_t = \phi(B)\, \alpha_t \tag{6}$$

For a sample data set $\omega_t, t = 0, 1, 2, 3 \ldots\}$ that satisfies the ARIMA (p, d, q) model, take the natural logarithm and pass the d-order difference to obtain a stationary ARMA sequence. Determining regression parameters and smoothing parameters is the key to the ARIMA model. The ARIMA (p, d, q) model determines the values of the parameters p, q mainly based on the autocorrelation function (ACF) and the partial autocorrelation function (PACF). The optimum model is determined by Akaike information criterion (AIC) and Bayesian information criterion (SBC).

The AIC criterion function of ARIMA (p, d, q) model is defined as:

$$AIC = \mathrm{n} \ln\left(\sigma_g^2\right) + 2\,(p + q + 1) \tag{7}$$

The SBC criterion function of ARIMA (p, d, q) model is defined as:

$$SBC = \mathrm{n} \ln\left(\sigma_g^2\right) + 2\,(p + q + 1) \tag{8}$$

ARIMA (p, d, q) parameters are optimized according to the last AIC and SBC minimization principles. For the results calculated with different parameter models, the real data can also be used to analyze their similarity coefficient and fitting degree. If the similarity coefficient and fitting degree are the largest, the model is optimal.

B. ARIMA MODEL ANALYSIS

The ARIMA model process is shown in figure 4. The ARIMA model requires that the time series be stationary. Therefore, after obtaining the historical moment data, it is necessary to judge whether it is a stationary time series (flow data are generally non-stationary series), and then carry out differential treatment for the series. Generally, by a single difference series can be converted into a stationary time series, no more than twice. After the difference, the ADF roots of unit test sequence is the stationary time series. Then the autocorrelation function ACF and non-autocorrelation function PACF were calculated, and the model parameters p and q were

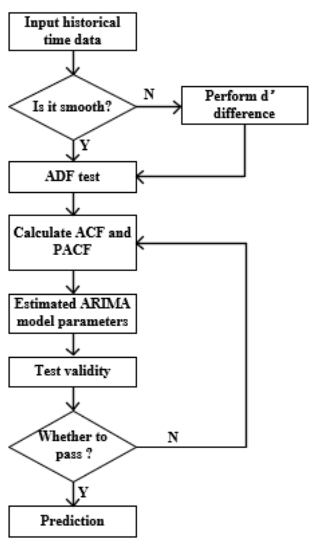
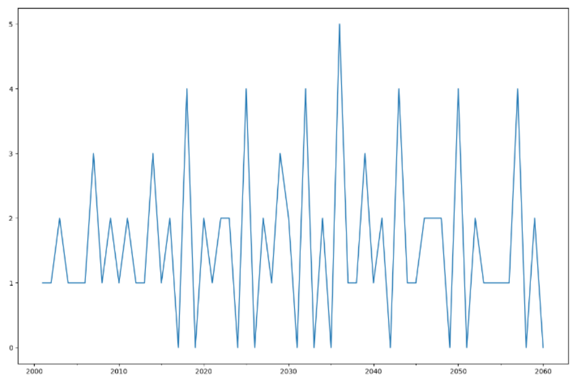


FIGURE 4. The flow graph of ARIMA.

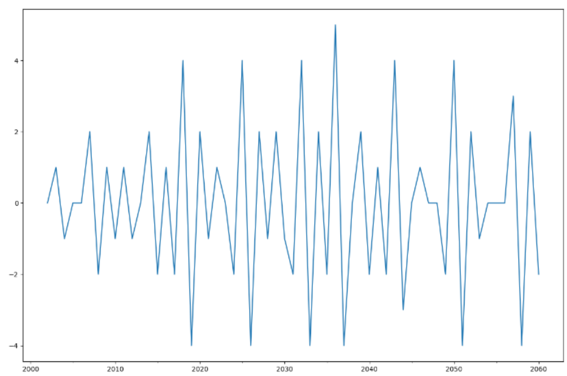
estimated by ACF and PACF graphs. Finally, check whether the model can meet the requirements of stationarity and reversibility, and whether the residual sequence of the model is white noise. If so, the model can be identified, and the model can be successfully established for prediction.

We demonstrate the modeling and prediction process through a set of data, and collect a kind of PARSEC traffic baseline data applied in the Network on Chip with the Netrace simulator.

Figure. 5(a) is the fitting of the original traffic data base on historical moment. It can be seen from the figure that the traffic model is non-stationary. After first difference, we transform it into a stationary time series, as shown in figure 5(b). Then, we analyzed ACF and PACF, as shown in figure 6. The p and q values of the model were determined, and the model was established for prediction. As shown in figure. 7(a), the orange curve is the predicted value, which achieves a good fitting effect compared with the real value 7(b).



(a)



(b)

FIGURE 5. (a) Original data model. (b) First difference stationary time series.

C. LASER SOURCE MANAGEMENT STRATEGY

As shown in figure 3, In the process of network operation, the photodetector placed at the tail also monitors the working conditions in the network at runtime. We build a global controller based on the collected sampling information of the historical moment, put N sets of sampling data into the queue to build an ARIMA model for prediction, as shown

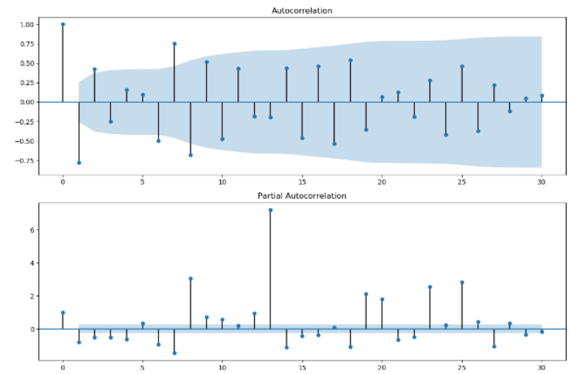
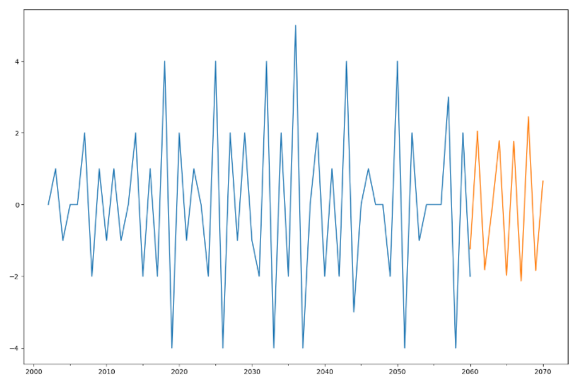
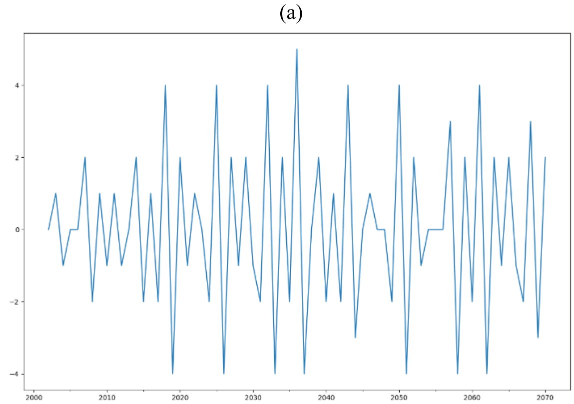


FIGURE 6. ACF and PACF graph.



(a)



(b)

FIGURE 7. (a) Prediction model timing sequence. (b) Actual model timing sequence.

in figure 8. We only need to change the corresponding pointer to conveniently schedule the predicted value to adjust the laser power. We update the data in the queue after each execution. The ARIMA model can predict 20 to 50 data information at a time, and we continuously update the sample data in the queue during the supply process, so there is no additional time overhead. Moreover, this implementation is light-weight and can improve overall performance with minimal resources.

If the photodetector can detect the laser power at a certain moment, it means that the supply of the laser source can meet the needs of the network and there is excess. If the photodetector does not detect the laser power at this time,

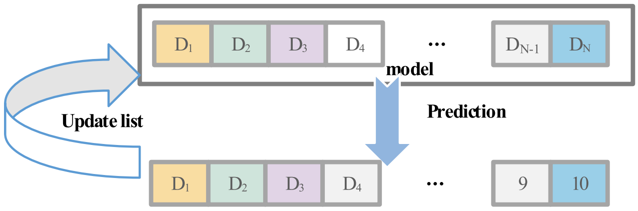


FIGURE 8. Predictive queue model.

it means that the supply of the laser source is just right or unable to meet the needs of the network. In order not to affect normal communication, some nodes can't get enough laser power to send packets. We feedback a message to the global controller and add a level of laser power based on the prediction at the next moment. Suppose the laser has n levels of P, 2P..., nP. When receiving the feedback information, we increase it to (k+1) P based on the predicted value kP at the next moment.

IV. SIMULATION AND ANALYSIS RESULTS

We simulated the proposed Optical Network on Chip laser power dynamic management scheme. Under the 64-core network, the ARIMA model is built using historical time data to predict network traffic for a period of time in the future. We tested our solution under the integrated traffic model, PARSEC traffic fluidanimate, blackschols, ×264 benchmarks. The sampling interval is set to 100 cycles. If there are n time slots in the simulation, the output power of the laser source in the i-th time slot is Pi. The instantaneous value of the output power of the laser source is constantly changing, and the average laser power is

$$p = \left(\sum_{i=1}^{n} p_i\right) / n \tag{9}$$

Figure 9 shows the average laser power comparison of the static flow and dynamic supply strategies for the integrated flow model uniform at different injection rates.

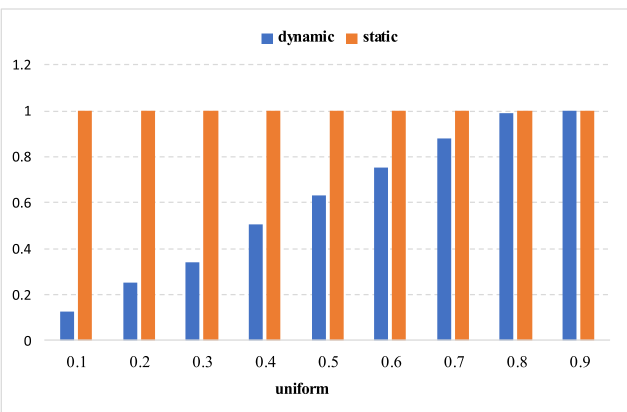


FIGURE 9. Normalized laser power under uniform traffic model.

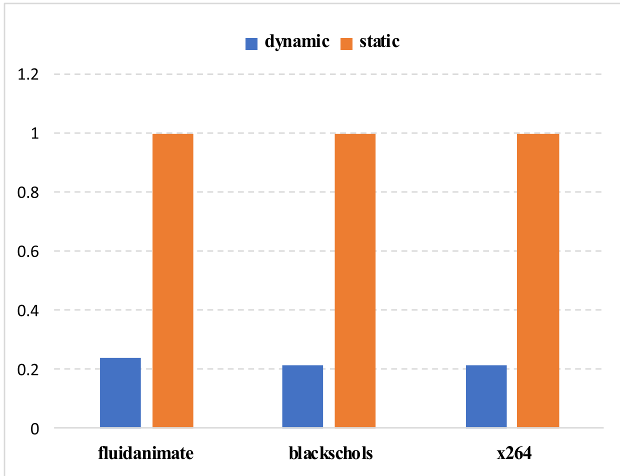


FIGURE 10. Normalized laser power under PARSEC benchmarks.

Figure 10 shows the average laser power comparison between the static supply and the dynamic supply strategy of the Netrace extraction fluidanimate, blackschols, and x264 on the 64-core optical Network on Chip. The static supply strategy is used as the benchmark.

Bienia *et al.* [10] have tested the PARSEC benchmark traffic model with 78.7% of the packets being single and sparse, which means that there will be only a few nodes working in the network at many times. If we use a static supply scheme, it will cause a large part of the waste. The simulation in this paper also verify the conclusion in [10] that the dynamic scheduling strategy can reduce the power consumption of the laser source by more than 70%. Through simulation, we can find that the smaller the traffic on the Network on Chip, the more the laser power saved by our proposed dynamic scheduling method. As the scale of the Network on Chip continues to expand in the future, we believe that the dynamic management of the laser source will greatly improve the performance of the optical Network on Chip.

V. CONCLUSIONS

Silicon photonic links is expected to replace electrical links in future many-core processors. However, the large laser power consumption in silicon photonic networks is limiting their development. In this paper, we proposed a novel laser source supply scheme for Optical Network on Chip. We use a dynamic supply scheme based on ARIMA model, which can reasonably adjust the laser source size by prediction. From the simulation results we can see that our dynamic supply scheme can save 70% of the laser power compared with the static supply under the real application of the PARSEC flow benchmarks.

REFERENCES

[1] S. Werner, J. Navaridas, and M. Luján, "A survey on optical network-on-chip architectures," *ACM Comput. Surveys (CSUR)*, vol. 50, no. 6, p. 89, Jan. 2018.

[2] S. Werner, J. Navaridas, and M. Luján, "Designing low-power, low-latency networks-on-chip by optimally combining electrical and optical links," in *Proc. IEEE Int. Symp. High Perform. Comput. Archit.*, Feb. 2017, pp. 265–276.

[3] C. Chen and A. Joshi, "Runtime management of laser power in silicon-photonic multibus NoC architecture," *IEEE J. Sel. Topics Quantum Electron.*, vol. 19, no. 2, Mar./Apr. 2013, Art. no. 3700713.

[4] M. Kennedy and A. K. Kodi, "Bandwidth adaptive nanophotonic crossbars with clockwise/counter-clockwise optical routing," in *28th Int. Conf. VLSI Des.*, Jan. 2015, pp. 123–128.

[5] L. Zhou and A. K. Kodi, "PROBE: Prediction-based optical bandwidth scaling for energy-efficient NoCs," in *Proc. 7th IEEE/ACM Int. Symp. Netw. Chip.*, Apr. 2013, pp. 1–8.

[6] X. Wang, Y. Gao, H. Gu, and K. Wang, "A laser power management method for on-chip photonic interconnect," in *Proc. 7th Int. Green Sustain. Comput. Conf. (IGSC)*, Nov. 2017, pp. 1–4.

[7] Y. Pan *et al.*, "Firefly: Illuminating future network-on-chip with nanophotonics," *ACM SIGARCH Comput. Archit. News*, vol. 37, no. 3, pp. 429–440, Jun. 2009.

[8] Z. Wang, H. Gu, Y. Yang, and B. Zhang, "Power allocation method for TDM-based optical network on chip," *IEEE Photon. Technol. Lett.*, vol. 25, no. 10, pp. 973–976, May 15, 2013.

[9] D. Vantrease *et al.*, "Corona: System implications of emerging nanophotonic technology," in *Proc. Int. Symp. Comput. Archit.*, Oct. 2008, pp. 153–164.

[10] C. Bienia, S. Kumar, J. P. Singh, and K. Li, "The PARSEC benchmark suite: Characterization and architectural implications," in *Proc. Int. Conf. Parallel Archit. Compilation Techn.*, Oct. 2008, pp. 72–81.

[11] H. Z. Moayed and M. A. Masnadi-Shirazi, "Arima model for network traffic prediction and anomaly detection," in *Proc. Int. Symp. Inf. Technol.*, Aug. 2008, pp. 1–6.

ZHUANGZHUANG LIU received the B.E. degree in electronic information engineering from the North University of China, in 2018. He is currently pursuing the M.E. degree in telecommunication and information systems with the State Key Laboratory of ISN, Xidian University. His main research interest includes network-on-chip and optical interconnect.

HUAXI GU received the Ph.D. degree in telecommunication and information system from Xidian University, Xi'an, Shaanxi, China, in 2005, where he is currently a Professor with the State Key Laboratory of ISN. He has authored or co-authored more than 100 publications in many international journals and conferences. His current research interests include interconnection networks, networks-on-chip and optical interconnect, and data center networks.

KUN WANG received the B.E. and M.E. degrees in computer science and technology from Xidian University, Xi'an, China, in 2003 and 2006, respectively, where she is currently a Lecturer with the Department of Computer Science. Her current research interests include high performance computing and cloud computing, and the network virtualization technology.

BOWEN ZHANG received the M.E. degree from the School of Telecommunications Engineering, Xidian University, in 2015, where he is currently pursuing the Ph.D. degree in telecommunication and information systems with the State Key Laboratory of ISN. His main research interests include optical network-on-chip, and artificial intelligence technology in ONoC and multiplex techniques in optical interconnected networks.

KANG WANG received the M.E. degree in communication engineering from Xidian University, in 2016, where he is currently pursuing the Ph.D. degree in telecommunication and information systems with the State Key Laboratory of ISN. His main research interests include on-chip optical networks for memory systems, on-chip floorplan optimization, and distributed artificial intelligence.