

# OPERON: Optical-electrical Power-efficient Route Synthesis for On-chip Signals

Derong Liu<sup>1</sup>, Zheng Zhao<sup>1</sup>, Zheng Wang<sup>2</sup>, Zhoufeng Ying<sup>1</sup>, Ray T. Chen<sup>1,2</sup>, and David Z. Pan<sup>1</sup>

<sup>1</sup>Department of Electrical and Computer Engineering

<sup>2</sup>Department of Materials Science and Engineering

<sup>1,2</sup>University of Texas at Austin, Austin, TX, USA

## ABSTRACT

As VLSI technology scales to deep sub-micron, optical interconnect becomes an attractive alternative for on-chip communication. The traditional optical routing works mainly optimize the path loss, and few works explicitly exploit the optical-electrical co-design of on-chip interconnects. To overcome these limitations, we present an efficient framework that directs the hybrid optical and electrical routes with a global view of power optimization. In this framework, on-chip signal bits are processed as hyper nets; the combination of optical and electrical routes are designed for hyper nets; then a formulation is given to find the appropriate solution of each hyper net and follows a speed-up algorithm; a min-cost max-flow network is utilized to reduce the consumed optical waveguides. Experimental results demonstrate the effectiveness of the proposed framework.

## 1. INTRODUCTION

As VLSI technology scales into deep sub-micron, interconnect delay becomes a bottleneck towards timing closure in comparison to cell delay. Based on this observation, research efforts have shifted to explore efficient interconnect to replace electrical wires. Due to the dominance of bandwidth-distance-power properties, optical communication based on chip-scale optical-electrical systems emerges as a promising alternative [1]. As recent fabrication techniques enable the on-chip integration of nano-scale devices with a high density, nanophotonic interconnects show great potential in unleashing the bandwidth limitation for memory access and processor communication.

With the increasing trend of on-chip communication density, Wavelength Division Multiplexing (WDM) exhibits the potential of offering high bandwidth with controllable overheads. During the transmission, signal quality can be affected by various losses at the receiving side. And EO/OE conversion power overheads should also be considered carefully [2]. Thus, it is desirable to propose an efficient optical-electrical co-design engine by offering optical and electrical connections for local and global communication, respectively. As shown in Figure 1, it consists of two layers: the bottom one for copper wires and the upper one for optical interconnects. The yellow arrows denote the electrical wires, and the blue arrows denote the optical interconnects for remote connections. Based on this infrastructure, an intelligent framework is essential to direct the distribution of electrical and

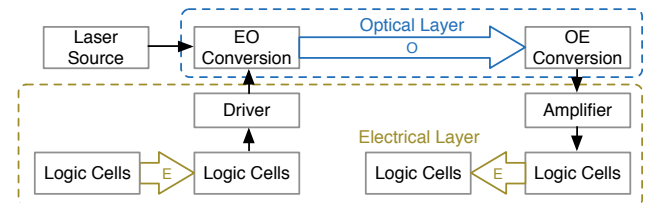


Figure 1: Block Diagram of Optical-Electrical On-Chip Design.

optical wires for on-chip communication.

There are a few works dealing with the integration of optical interconnects onto on-chip designs. Some works provide physical design of on-chip optical interconnect: Ding et al. [3] proposed the first optical router for low power consumption, and employed WDM architecture to reach high density/capacity during global routing, while ignoring the splitting loss [4]. As the first placement-and-routing tool for optical Network-on-Chip(NoC)s, Boos et al. [5] balanced the trade-off between propagation and crossing losses, but limited to two-pin connections. Besides, optical NoCs were also designed to enhance its resilience to physical variations, such as environmental temperature and manufacturing instability [4, 6]. To increase the potential bandwidth, communication parallelism was emphasized in [7] through a formal methodology. Very recently, with the proposal of LED-driven wires, an automatic place-and-route flow enabled the replacement of electrical wires with on-chip laser sources [8].

With the high-bandwidth demands of on-chip communication, it is desired that optical interconnects collaborate with electrical counterparts smoothly [9]. Very few of previous works provide a comprehensive routing flow for optical-electrical co-design of on-chip multi-pin signals. Generally, optical configurations are deployed for the distant connections. Nevertheless, signal transmission may suffer from non-negligible optical loss, resulting in the potential malfunction. And the splitting loss also aggravates the resulting loss, which plays an important role but is neglected in the previous optical physical design works. By introducing the reasonable optical-electrical configurations, we can make a better trade-off between the optical loss and power consumptions. Therefore, it is worthwhile to devise a routing flow where optical connections can collaborate with electrical wires seamlessly.

In this paper, we propose a routing synthesis flow for optical interconnects integrated with electrical wires. Our contributions are summarized as follows:

- With splitting loss into consideration, optical-electrical co-design routes are derived for power and loss optimization;
- Mathematical formulation targets at power minimization while satisfying the detection constraints, and a Lagrangian-Relaxation-based algorithm is presented for speed-up;
- Network-based algorithm assigns the optical connections for sharing WDMs with the control of distance.

The remainder of this paper is organized as follows. Section 2 presents the overview of our framework and adopted models.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

DAC'18, June 24-29, 2018, San Francisco, CA, USA

© 2018 ACM. ISBN 978-1-4503-5700-5/18/06...\$15.00

DOI: <https://doi.org/10.1145/3195970.3196084>

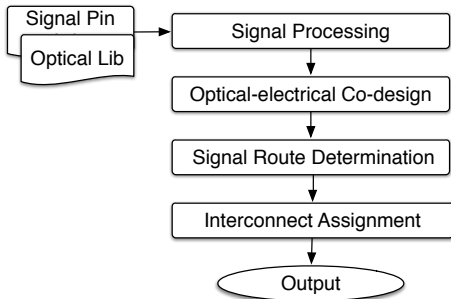


Figure 2: OPERON Flow.

Section 3 describes our routing procedure, presents a mathematical formulation to optimize power consumption while satisfying the detection constraints, and a Lagrangian-Relaxation method benefits the runtime. Section 4 presents the WDM assignment. Section 5 reports the experimental results, and followed by conclusion in Section 6.

## 2. PRELIMINARIES

In this section, we provide the overview of our proposed framework, and illustrate the adopted model and methodology, based on which a problem formulation is given.

### 2.1 Overall Flow

To provide a clear view of our framework, an outline is shown in Figure 2. Firstly, a processing step partitions signal groups into a set of hyper nets in a top-down manner while clustering the neighboring electrical pins from a bottom-up view. Secondly, we devise the optical-electrical co-design routes for each hyper net. With the acquired solution candidates, a mathematical formulation is given for solution determination; to avoid the potential runtime overheads, we adopt a fast algorithm for speed-up. Then the WDM placement is performed and an assignment procedure follows to allocate optical connections onto nearby WDMs. We finally acquire the electrical-optical co-design solutions.

### 2.2 Optical Device Model

With the advanced fabrication technologies, emerging optical devices have been promoting efficient on-chip communication. The WDM infrastructure offers multi-channel routing among physical locations, which is a promising alternative for high-speed data communication. Therefore, it provides great potential for routing on-chip parallel connections, without crosstalk issues between different channels.

In comparison to electrical wires, WDMs contribute to high bandwidth and low power consumption for data propagation. Nevertheless, the EO/OE conversion leads to non-negligible power consumption due to the extra driver and amplifier configuration, which are usually neglected during routing optimization. Hence we provide the optical power,  $p_o$ , by using WDMs,

$$p_o = p_{mod} \cdot n_{mod} + p_{det} \cdot n_{det}, \quad (1)$$

where  $n_{mod}, n_{det}$  represent the number of employed modulators and detectors, and  $p_{mod}, p_{det}$  represent the unit power cost of modulators and detectors [2].

Besides the power cost, the optical loss along the WDM should also be taken into careful consideration. As shown in Figure 3(a), the loss mainly consists of propagation loss, crossing loss, and splitting loss. The first two kinds of loss are respectively in proportion to the total WDM length and the number of crossing occurrences. Notably, the splitting loss has usually been neglected in the previous work, which, however, turns out to be one of the major sources of loss for on-chip optical routing [10]. The splitting loss happens whenever an input light source splits into

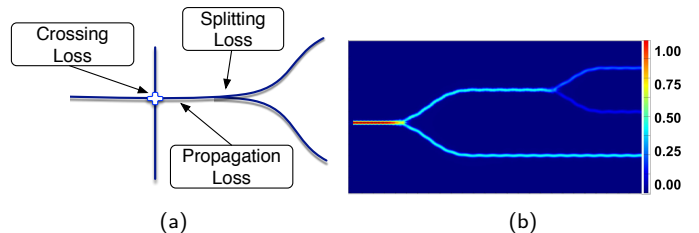


Figure 3: Optical model illustration. (a) Loss model for on-chip optical routing; (b) Simulation of the normalized power loss in Y-branch splitters.

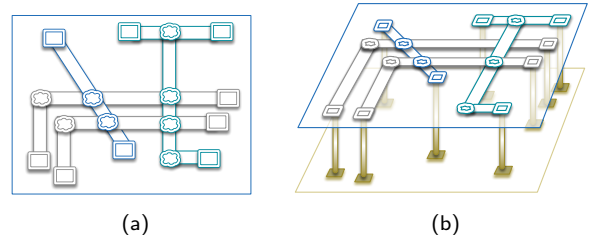


Figure 4: On-chip signal model. (a) Signal routing on 2D optical layer; (b) Signal routing on 3D optical-electrical architecture.

multiple light sinks. The expected splitting loss in  $dB$  for each splitting is calculated to be  $10 \cdot \sum \log(n_s)$ , where  $n_s$  is the number of splitting arms. As seen from Figure 3(b) that shows the power distribution of two cascaded 50-50 Y-branch splitters based on the simulation, each reduces the input light power into one half on the output sides. Thus, the loss is calculated as follows,

$$loss = \alpha \cdot WL + \beta \cdot n_x + 10 \cdot \sum \log(n_s), \quad (2)$$

where  $WL$  is the WDM length,  $n_x$  is the number of crossings, and  $n_s$  is the number of splitting arms.  $\alpha$  and  $\beta$  are the physical parameters for the propagation and crossing loss.

### 2.3 Proposed Signal Model

The on-chip integration of optical interconnect provides the opportunity for signals to be routed in parallel routes. In current industrial designs, the performance-critical signal bits are bound together for data communication between logic cells and memory interfaces, etc. Figure 4(a) gives an example of signal routing on a 2D optical layer, where each signal group is identified with one single color. For the grey signals, they are clustered into two hyper nets because the number of total bits exceeds the WDM capacity. Different from Manhattan routing based on electric wires, optical scheme allows routing in any direction, which can benefit the overall wire-length. With the EO/OE conversion deployment, the optical interconnects are coupled to the bottom electrical layer, as shown in Figure 4(b). This architecture offers us the flexibility of optical-electrical co-design for on-chip signals.

### 2.4 Problem Formulation

Based on the proposed flow and optical model discussed in the preceding section, we define the proposed optical-electrical route synthesis (OPERON) problem as follows:

**Problem 1 (OPERON).** *Given signals bundled in groups with the pin locations, our framework determines the routing topologies and optical-electrical configurations for each signal group so that the total power consumption can be optimized while the detection constraints are satisfied.*

## 3. ALGORITHMS

In this section, we will elaborate the procedure of signal routing through optical-electrical cooperation for power efficiency. With the constructed hyper nets, route candidates are generated with co-design optimization. A mathematical formulation guides the routing concurrently, which follows with a speed-up algorithm.

### 3.1 Signal Processing

Before deriving the route solutions, a processing procedure is required for creating the pseudo pins and hyper nets of the signal bits. Considering that the pin distances may vary significantly for one bit, it brings the necessity of constructing the pseudo pins to represent the neighboring electrical pins. Besides, signal bits can be bundled together as a hyper net whose pins are located in adjacent locations. Therefore, we employ the following clustering methodologies for hyper net construction.

#### 3.1.1 K-Means-based Clustering

Since the capacity, i.e., the number of allowable channels, of one WDM is limited according to the current fabrication technologies, we should determine how to cluster the bits to satisfy the capacity constraint. Thus, for a given signal group, we first check if the number of bits is above the capacity. If so, then we partition this group based on the K-Means strategy, which is widely used in clustering for its effectiveness [11].

Due to its top-down nature, the signal bits are divided into  $K$  clusters and  $K$  is the quotient of the total bit and the capacity value. Since K-Means itself cannot guarantee the cluster size during the solving process, we extend the K-Means strategy by checking the size in every iteration: if a cluster violates the capacity constraint, the additional bits will be assigned to the second closest one, and so on. As  $K$  clusters are adequate for accommodating all the bits, the neighboring bits can be located in one cluster with the decreasing distance variance. When the variance becomes lower than a given threshold, this iterative flow will stop. There may be a few empty clusters without any assigned bits, which will be removed afterward.

#### 3.1.2 Hyper Net Construction

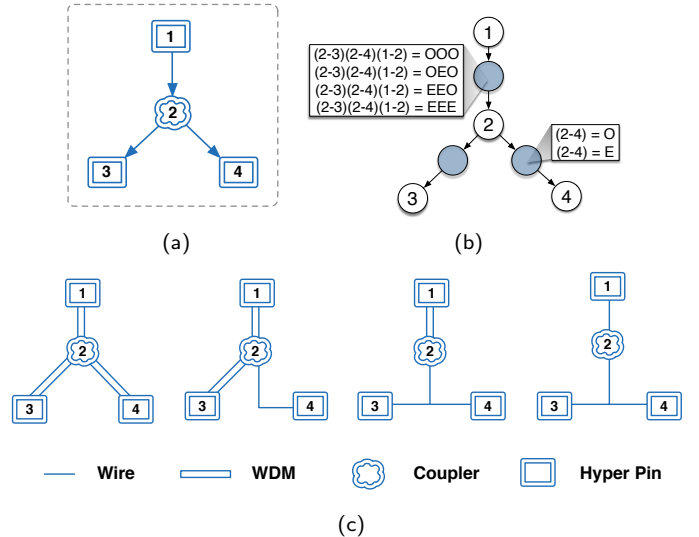
Based on the K-Means solution, we construct the hyper net to represent all the bits in a cluster. By replacing the set of individual nets with a single hyper net, the whole problem size can be reduced. We build up the pseudo pins for the hyper net and determine their corresponding electrical pins.

For one cluster containing a set of neighboring bits, each electrical pin itself is initialized as a hyper pin. Then we adopt a bottom-up clustering strategy for shrinking the hyper pins. During each iteration, a pair of hyper pins is selected with the minimum Euclidean distance. And we check if their physical distance is below the pre-specified threshold: if so, then these two hyper pins will be combined with the updated gravity center; otherwise, the clustering will return with the finalized set of hyper pins. As each hyper pin contains a set of electrical pins eventually, we should ensure the number of connections between the hyper pins. Based on the constructed hyper nets and pins, we perform routing design and synthesis in the following sections.

### 3.2 Optical-electrical Route Co-design

With the assistance of processing, we acquire a set of hyper nets with the corresponding pins for connection. Here we discuss how to combine optical and electrical design in a coordinated way. Besides supporting optical interconnects for distant connections as the previous works [4], our co-design methodology provides a more systematic analysis of power consumption and optical loss.

Before the optical-electrical development, the sets of optical route candidates are generated as the baselines for co-design processing. Here we choose to extend the Batched Iterated 1-Steiner (BI1S) algorithm which provides the flexibility of Steiner point selection. By sorting the Steiner points with the induced propagation and bending cost, we can acquire various baselines by visiting different points. Additionally, compared to Manhattan routing of electrical wires, optical interconnects are able to route in any direction, as shown in Figure 5. Thus, the rectilinear connections are not mandatory for optical interconnects, which offers



**Figure 5:** Optical-electrical co-design example. (a) Hyper net topology; (b) Dynamic programming based co-design scheme; (c) Corresponding optical-electrical candidates.

more feasible baseline topologies. After obtaining the baselines, we present the optical-electrical co-design scheme for power and loss optimization.

As the baseline is a tree topology, in essence, we take this advantage to develop a list of co-designs, each recorded with competitive optical loss and power cost. Taking Figure 5(a) as a baseline instance, we visit each connection between the hyper pins and Steiner points and decide which to employ optical interconnections. Inspired by the classic buffer insertion algorithm, we derive the promising co-designs in a bottom-up manner so that the trade-off between loss and power can be balanced.

The algorithm flow is illustrated in Figure 5(b). By traversing  $node_3$  and  $node_4$  in the bottom level, each interconnect can choose to route through optical WDMs or electrical wires, and the internal node records each solution with the specific power and optical loss. The optical power is calculated as in Eq. (1), while the electrical power is in proportional to the wire-length. For the optical loss, we are able to calculate the exact propagation and splitting loss and approximate the crossing loss based on the optical baselines. If the optical interconnect between  $node_2$  and  $node_4$  causes both higher loss and power costs compared to that between  $node_2$  and  $node_3$ , then the former candidate turns to be an inferior solution to be pruned beforehand, and we come to the upper level for further judgments. The internal node between  $node_1$  and  $node_2$  accumulates all the possible solutions, and the resulting power cost and optical loss are recalculated. As a redundant solution is removed, there remain four solutions with different configurations. Figure 5(c) lists all the finalized solutions. It can be seen that the third candidate can save the OE conversion overheads by implementing the bottom branches through electrical wires. After traversing all the nodes, we acquire the optical-electrical solutions while eliminating the non-competitive alternatives. The runtime complexity of this procedure is within  $\mathcal{O}(|N_c||d|)$ , where  $|N_c|$  is the number of connections between the hyper pins and Steiner points, and  $|d|$  is the depth of the tree in Figure 5(b).

### 3.3 Mathematical Formulation

To obtain the optimal solution for each object, the mathematical formulation is given in Formula (3). The objective is to minimize the total power overheads for all the hyper nets  $H$ . The set of optical-electrical solution candidates is denoted as  $H_{sol}$ , which consists of both optical-electrical co-design and pure electrical routes. The first item,  $a_{ij}$ , represents the  $j$ -th solution candidate

for hyper net  $i$ , and  $p_{oe}(i, j)$  gives the corresponding power costs with O/E conversions, as described in Section 2.2. Additionally,  $a_{ie}$  represents the electrical route alternative of hyper net  $i$  through electrical wires, as the fourth candidate in Figure 5(c), and  $p_e(i)$  represents the power consumption of  $a_{ie}$ .

$$\min \sum_{(i,j) \in H_{sol}} p_{oe}(i, j) \cdot a_{ij} + \sum_{i \in H} p_e(i) \cdot a_{ie} \quad (3a)$$

$$\text{s.t.} \quad \sum_{(i,j) \in H_{sol}} a_{ij} + a_{ie} = 1, \quad \forall i \in H, \quad (3b)$$

$$\sum_{(m,n) \in H_{sol}} l_x(i, j, m, n, p) \cdot a_{ij} \cdot a_{mn} + l_s(i, j, p) \cdot a_{ij} + l_{spl}(i, j, p) \cdot a_{ij} \leq l_m, \quad \forall p \in P(a_{ij}), \quad i \in H, \quad (3c)$$

$$a_{ij}, a_{ie} \text{ is binary, } \quad \forall i \in H, \quad \forall j. \quad (3d)$$

Meanwhile, constraint (3b) denotes that only one solution can be selected for each hyper net: if no appropriate optical-electrical co-design route is found, then this hyper net will be handled with electrical wires. Also, based on the optical constraints, the light intensity should be strong enough to be detected at the receiver. Hence, constraint (3c) guarantees that the total source-to-sink loss on path  $p$  should be lower than the maximum loss,  $l_m$ , and  $p$  corresponds to one source-to-sink path in a co-design candidate  $a_{ij}$ .  $l_x(i, j, m, n, p)$  refers to the crossing loss on path  $p$  resulting from the intersections between candidates  $a_{ij}$  and  $a_{mn}$ , while  $l_s(i, j, p)$  and  $l_{spl}(i, j, p)$  are the propagation loss and splitting loss of path  $p$ , respectively. Finally, with the constraints of both  $a_{ij}$  and  $a_{ie}$  as binary variables, it is seen that this quadratic programming problem can be solved through Integer Linear Programming (ILP). Due to the existing terms of  $a_{ie}$ s, a feasible solution can be guaranteed for all the hyper nets.

Nevertheless, solving an ILP would lead to prohibitive runtime overheads. Thus the speed-up technique is adopted to condense the solution space without sacrificing the performance. To reduce the number of variables, we can remove those crossing variables belonging to the pair of hyper nets with non-overlapped bounding boxes. By this setting, we can control the search space without performance degradation.

### 3.4 Lagrangian Relaxation-based Algorithm

Since solving the ILP even with the reduced variables would still be time-consuming, we re-formulate this problem in a more efficient way. Therefore, we propose a Lagrangian Relaxation-based (LR) approach, which relaxes the constraint (3c) into the objective function. The relaxed formula is shown in Formula (4) with the Lagrangian Multiplier (LM)  $\lambda_p$  for each path  $p$ . Since  $l_m$  is a constant value, we remove it in Formula (4a).

$$\begin{aligned} \min \quad & \sum_{(i,j) \in H_{sol}} p_{oe}(i, j) \cdot a_{ij} + \sum_{i \in H} p_e(i) \cdot a_{ie} \\ & + \sum_{p \in P(H_{sol})} \sum_{(m,n) \in H_{sol}} \lambda_p \cdot l_x(i, j, m, n, p) \cdot a_{ij} \cdot a_{mn} \\ & + \sum_{p \in P(H_{sol})} \lambda_p \cdot (l_s(i, j, p) + l_{spl}(i, j, p)) \cdot a_{ij} \quad (4a) \\ \text{s.t.} \quad & (3b), (3d). \end{aligned}$$

To resolve an LR-based algorithm, we update the LMs iteratively to explore the solution. Meanwhile, the quadratic terms can be linearized based on the last iteration, which empirically works well [12, 13]. Thus we employ this approximation method to capture the quadratic terms:

$$a_{mn} \cdot a_{ij} \approx a'_{mn} \cdot a_{ij} + a_{mn} \cdot a'_{ij}. \quad (5)$$

By substituting the linearization terms, it is seen that the routing optimization becomes a weighted sum of  $a_{ij}$ s and  $a_{ie}$ s. With the fixed set of LMs and constraint (3b), we search for the solution in each iteration. The pseudocode of the LR-based algorithm is shown in Algorithm 1. Initially, the LMs are set to a value proportional to the  $p_e$  (line 1) while the power cost is calculated for the candidate solutions (line 2). And the propagation and splitting loss,  $l_s(i, j, p)$ ,  $l_{spl}(i, j, p)$ , are also calculated for each optical-electrical route candidate (line 3). During the Lagrangian optimization, we traverse the hyper nets in each iteration and select the candidate with the best sum of weights, including both its inherent power and LM penalty costs (line 5). Based on the acquired solutions, we check the detection violations for each source-to-sink path (line 6). Then the LMs are updated at the end of each iteration according to the violations (line 7). To ensure the algorithm convergence, we update the LMs through a sub-gradient methodology. This procedure continues until a convergence is reached. The converging criteria are set as follows: the decrease in both power costs and violations reach a pre-defined ratio, or the iteration number is over 10.

---

#### Algorithm 1 LR-based Algorithm

---

**Input:** A set of hyper nets with candidates  $a_{ij}, a_{ie}$ .

- 1: Initialize LMs  $\lambda_p$ s;
  - 2: Calculate the power costs for  $a_{ij}, a_{ie}$ s;
  - 3: Calculate the loss  $l_s(i, j, p), l_{spl}(i, j, p)$  for  $a_{ij}$ s;
  - 4: **while** no converge **do**
  - 5:     Select the candidate with the best weight;
  - 6:     Calculate violation values for paths;
  - 7:     Update  $\lambda_p$ s based on violations;
  - 8: **end while**
- 

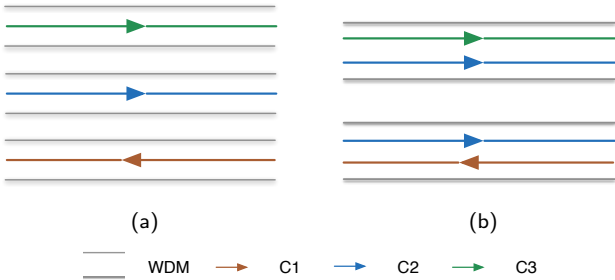
## 4. WDM ASSIGNMENT

After the last procedure, each hyper net has obtained its topology consisting of point-to-point connections. Similar to electrical wires, the distribution of optical connections should also be controlled. The WDM offers multi-channels for parallel routing, but its capacity is constrained. Thus, the assignment is able to utilize WDMs efficiently without disturbing the previous result. In this section, we describe the WDM placement and the assignment procedure through a network model.

### 4.1 WDM Placement

Before the assignment of connections, we perform the WDM placement to initialize their locations. An intuitive way is to place one WDM for each connection. However, this would lead to an unnecessarily high volume of WDMs for which can be shared by the hyper nets propagating in parallel; Also, we control the distance to be above  $dis_l$  between two nearby WDMs to avoid crosstalk. Thus, the placement procedure not only controls the number of WDMs, but also conforms to the distance bound.

Since the placement of vertical and horizontal WDMs follow the same way, for brevity we only discuss the horizontal WDMs. Initially, the horizontal connections are collected and sorted in an ascending order of their  $y$ -coordinates. Then we traverse each connection sequentially to determine the WDMs' locations. The first WDM is placed with the same location as the first connection, and recognized as the current WDM. For the next connection, we check whether the current WDM has enough capacity. Also, their distance should be below the distance  $dis_u$  to avoid causing the disturbances. If both of conditions are met, the connection will be assigned to the current WDM; otherwise, we place an additional WDM and turn to the next connection. After visiting all the connections, an adequate number of WDMs will be obtained for assignment. As the placement enables the combination of multiple connections in cases, the number of WDMs is



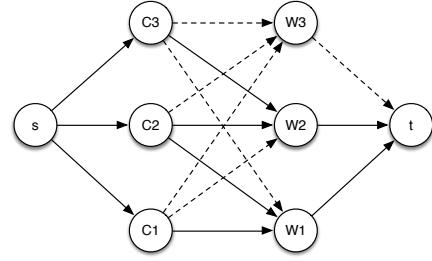
**Figure 6:** Example of WDM Placement. (a) Initial WDM placement for three connections; (b) WDM placement after the assignment.

already well controlled. It shall be noted that after the placement there may exist some congested regions where two nearby WDMs are within  $dis_u$  distance. To avoid these cases, we check those violating regions and adjust the WDMs’ locations in a one-by-one way for legalization.

## 4.2 Network-flow Based Assignment

After initializing the WDM placement, we will assign the optical connections while removing the idle WDMs. Previously, a number of WDMs are utilized to accommodate the signal bits in a sequential manner, which does not exploit the WDMs in a global view. This brings out the necessity of re-assigning the connections concurrently. As we observe in Figure 6(a), for three connections marked in different colors, we adopt three WDMs if each connection contains 20 bits and the capacity is 32 as set in [4]. Nevertheless, by re-assigning the connections, we can save one WDM while satisfying the capacities as shown in Figure 6(b). To reach this objective, we further present a min-cost max-flow network to re-allocate the connections. Due to its uni-modular property, the assignment solution can be acquired directly without any approximation or rounding methodologies.

Figure 7 illustrates how the network flow model resolves the assignment problem. Since the horizontal and vertical connections are re-assigned independently with the same method, we just list the horizontal connections. Given the sets of connections and WDMs, a directed graph  $G$  is constructed as follows: There are four types of vertices contained in  $G$ :  $V_s$  and  $V_t$  represent the pseudo starting and ending nodes, respectively;  $V_C$  and  $V_W$  represent the connections of hyper nets and the WDMs which have already been placed. Also, there are three types of edges in  $G$ :  $\{V_s \rightarrow V_C\}$ ,  $\{V_C \rightarrow V_W\}$ , and  $\{V_W \rightarrow V_t\}$ . The edges of the first type ensure that all the connections should be assigned to the WDMs; the edges of the second type determine which WDM will be chosen for assignment; and the edges of the third type guarantee that the WDM capacity constraint should be satisfied. Notably, in order not to disturb the routing results obtained in Section 3, we only allow each connection to connect with its neighboring WDMs, and the distance should be within  $dis_u$ , as shown in Figure 7. Meanwhile, the costs of edges are defined as follows: the cost from  $V_s$  to  $V_C$  is set to 0; the cost from  $V_C$  to  $V_W$  is the perpendicular distance between the WDM and the connection’s current location; and the cost from  $V_W$  to  $V_t$  is the WDM usage cost. The capacities of edges are also defined: the capacities from  $V_W$  to  $V_t$  are the maximum allowable capacity, and the capacities of other edges are the number of passing nets through the connection for flow accommodation. Since our target is to reduce the costs of WDMs, we prefer to apply higher costs to the edges from  $V_W$  to  $V_t$ . Thus, we normalize the costs of edges from  $V_C$  to  $V_W$  so that the WDMs’ usages are emphasized. After solving the example in Figure 6, the edges with flows are shown as solid lines in Figure 7. With the network model, three connections share two WDMs. It is seen that both the number of flow edges and vertices are proportional to the number of connections  $|C|$ , and the runtime complexity is within  $\mathcal{O}(|C|^2)$ .



**Figure 7:** Example of min-cost max-flow assignment.

## 5. EXPERIMENTAL RESULTS

We implemented the OPERON framework in C++, and tested it on a Linux machine with eight 3.3GHz CPUs. Meanwhile, we selected GUROBI [15] as our ILP solver, and open source graph library LEMON [16] as our min-cost max-flow network solver. For the optical parameters, we set the values of  $\alpha, \beta$  to 1.5 dB/cm, 0.52 dB, with the same optical settings in [5]; and the power consumptions of modulators and detectors are set to 0.511 pJ/bit and 0.374 pJ/bit [2]. With these parameters, we derive the test cases from the industrial benchmarks, by up-scaling the dimension into centimeter scale, and employ the signal processing for the hyper net generation. The details of each test case are listed in the left part of Table 1. The column “#Net” gives the overall number of signal bits, while the column “#HNet” and “#HPin” correspond to the number of hyper nets and hyper pins, respectively. Since our work targets at optical-electrical co-design, we focus on the power consumption caused by both optical and electrical routes, as listed in column “Power”. The optical power is calculated in Eq. (1), and the electrical power is estimated based on its dynamic power:

$$p_e = \gamma \cdot f \cdot V^2 \cdot Cap, \quad (6)$$

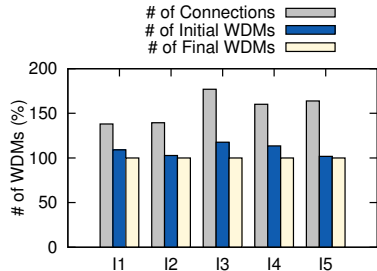
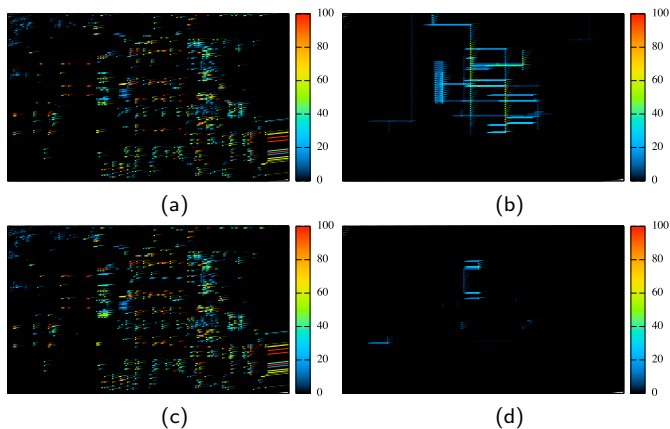
where  $\gamma, f, V, Cap$  denote the switching factor, system frequency, voltage level, and wire capacitance in proportional to the wire-length. Due to the signals’ performance-critical nature, the wire-lengths of electrical wires are estimated based on the Rectilinear Steiner Minimum Tree (RSMT) with the parameters in [2, 4].

To show the effectiveness, we implemented the similar GLOW framework for optical designs [4], and the electrical design based on Streak [14] for comparison. From the experimental results, it is shown that the utilization of optical interconnects consumes the power costs about one-third of the counterpart caused by electrical wires. This proves the high efficiency of optical propagation for distant communications, consistent with the optical inherent characteristics. Nevertheless, in order to satisfy the loss constraint, a few hyper nets cannot be routed on the optical layer and the residual nets have to be completed through electrical wires, resulting in additional power consumptions. To deal with this condition, we observe that after the employment of the optical-electrical paradigm, the overall power overheads are reduced by 14.0%. This is mainly because of the decreasing number of electrical routes, and the adjustment of EO/OE conversion also helps. However, due to the complexity of dealing with ILP formulation, the runtime overhead is significant, especially for the large benchmarks. Thus, the effectiveness of the proposed speed-up algorithm is shown, with the slightly worse performance but much shorter runtime.

To demonstrate the effectiveness of the WDM assignment, we compare the number of connections before the placement, and the WDMs before and after the assignment through the normalization in Figure 8. It is shown that the placement is able to reduce the usage of WDMs greatly for all the cases, but still, suffers from the sub-optimality due to its heuristic style. With the integration of the network flow algorithm, we observe that the number of WDMs can be further reduced by 8.9% on average.

**Table 1:** Performance Comparisons among Different Designs

Bench	#Net	#HNet	#HPin	Electrical [14]	Optical [4]	OPERON (ILP)		OPERON (LR)	
				Power	Power	Power	CPU(s)	Power	CPU(s)
I1	2660	356	1306	20.50	4.92	4.79	> 3000	4.88	2.1
I2	1782	837	1701	50.79	14.48	12.39	> 3000	12.77	5.0
I3	5072	168	336	17.96	2.70	2.49	4.4	2.57	0.9
I4	3224	403	1474	21.51	5.70	5.45	341.3	5.62	2.4
I5	1994	933	1897	54.21	18.40	14.61	> 3000	15.22	8.8
average	-	-	-	32.99	9.24	7.95	> 1869.1	8.21	3.8
ratio	-	-	-	<b>3.565</b>	<b>1.000</b>	<b>0.860</b>	-	<b>0.889</b>	-

**Figure 8:** Comparison of WDMs for optical connections before the placement, before the assignment and after the assignment.**Figure 9:** Power consumption distribution of I2. (a) Optical power in GLOW; (b) Electrical power in GLOW; (c) Optical power in OPERON; (d) Electrical power in OPERON.

Finally, we perform the experiments to measure the normalized power hotspots on both optical and electrical layers for GLOW and OPERON, as shown in Figure 9. Figure 9(a) and Figure 9(b) provide the power distributions of GLOW, while Figure 9(c) and Figure 9(d) provide the distributions of OPERON. We observe that the hotspots in Figure 9(a) and Figure 9(c) are distributed in a very similar manner. It is reasonable because they employ the similar amounts of EO/OE conversion overheads. By comparison, from the electrical layers shown in Figure 9(b) and Figure 9(d), the hotspots are alleviated greatly in OPERON. Consider that a higher flexibility is allowed through optical-electrical co-design procedure, much fewer electrical wires are allocated on the electrical layer in Figure 9(d), consistent with our motivation. Therefore, the hotspot comparison demonstrates the power efficiency of the proposed OPERON framework.

## 6. CONCLUSION

In this paper, we have proposed a set of algorithms for optical-electrical co-design of on-chip signals to optimize power consumptions. First, we adopt the clustering strategy to generate the sets of hyper nets and hyper pins, and construct the baseline topologies with optical interconnects. Based on the baseline, we develop the co-design solution set for the minimization of total loss and power costs. Then a mathematical formulation targets

at power optimization while satisfying the detection constraints, and follows a fast flow to reach a close performance with the ILP solution. Finally, a network flow model is adopted to utilize WDMs sufficiently. The results show that our route synthesis engine is able to provide the optical-electrical solution with legality and power efficiency. We believe that our work can open up opportunities for optical-electrical co-design research.

## Acknowledgement

The authors would like to thank Bei Yu from CUHK for helpful comments and discussions. This work is supported by the Multi-disciplinary University Research Initiative (MURI) program through the Air Force Office of Scientific Research (AFOSR), contract No.FA 9550-17-1-0071, monitored by Dr. Gernot S. Pomrenke.

## 7. REFERENCES

- [1] C. Sun, M. T. Wade, Y. Lee, J. S. Orcutt, L. Alloatti, M. S. Georgas, A. S. Waterman, J. M. Shainline, R. R. Avizienis, S. Lin *et al.*, "Single-chip microprocessor that communicates directly using light," *Nature*, 2015.
- [2] C. Sun, M. Wade, M. Georgas, S. Lin, L. Alloatti, B. Moss, R. Kumar, A. Atabaki, F. Pavanello, R. Ram *et al.*, "A 45nm SOI monolithic photonics chip-to-chip link with bit-statistics-based resonant microring thermal tuning," in *Proc. VLSIC*, 2015.
- [3] D. Ding, Y. Zhang, H. Huang, R. T. Chen, and D. Z. Pan, "O-Router: an optical routing framework for low power on-chip silicon nano-photonics integration," in *Proc. DAC*, 2009.
- [4] D. Ding, B. Yu, and D. Z. Pan, "GLOW: A global router for low-power thermal-reliable interconnect synthesis using photonic wavelength multiplexing," in *Proc. ASPDAC*, 2012.
- [5] A. Boos, L. Ramini, U. Schlichtmann, and D. Bertozzi, "Proton: An automatic place-and-route tool for optical networks-on-chip," in *Proc. ICCAD*, 2013.
- [6] M. Mohamed, Z. Li, X. Chen, L. Shang, A. Mickelson, M. Vachharajani, and Y. Sun, "Power-efficient variation-aware photonic on-chip network management," in *Proc. ISLPED*, 2010.
- [7] A. Peano, L. Ramini, M. Gavaneli, M. Nonato, and D. Bertozzi, "Design technology for fault-free and maximally-parallel wavelength-routed optical networks-on-chip," in *Proc. ICCAD*, 2016.
- [8] T. Krishna, A. Balachandran, S. B. Chiah, L. Zhang, B. Wang, C. Wang, K. L. E. Kian, J. Michel, and L.-S. Peh, "Automatic place-and-route of emerging LED-driven wires within a monolithically-integrated cmos-III-V process," in *Proc. DATE*, 2017.
- [9] W. Bogaerts, R. Baets, P. Dumon, V. Wiaux, S. Beckx, D. Taillaert, B. Luyssaert, J. Van Campenhout, P. Bienstman, and D. Van Thourhout, "Nanophotonic waveguides in silicon-on-insulator fabricated with CMOS technology," *Journal of Lightwave Technology*, 2005.
- [10] Z. Zhao, Z. Wang, Z. Ying, S. Dhar, R. T. Chen, and D. Z. Pan, "Logic synthesis for energy-efficient photonic integrated circuits," in *Proc. ASPDAC*, 2018.
- [11] S. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory (TIT)*, 1982.
- [12] B. Yu, D. Liu, S. Chowdhury, and D. Z. Pan, "TILA: Timing-driven incremental layer assignment," in *Proc. ICCAD*, 2015.
- [13] D. Liu, B. Yu, S. Chowdhury, and D. Z. Pan, "TILA-S: Timing-driven incremental layer assignment avoiding slew violations," *IEEE TCAD*, 2018.
- [14] D. Liu, V. Livramento, S. Chowdhury, D. Ding, H. Vo, A. Sharma, and D. Z. Pan, "Streak: Synergistic topology generation and route synthesis for on-chip performance-critical signal groups," in *Proc. DAC*, 2017.
- [15] Gurobi Optimization Inc., "Gurobi optimizer reference manual," <http://www.gurobi.com>, 2014.
- [16] "LEMON," <http://lemon.cs.elte.hu/trac/lemon>.