Energy-Optimal Routing on VCSEL-Based Interconnected Networks

Ilias Gravalos, Apostolos Siokis, Panagiotis Kokkinos, and Emmanouel A. Varvarigos

***Abstract*—Optical interconnection networks are being used in systems on chip, supercomputers, and datacenters, fueling exascale computing, big data, and artificial intelligence applications. The vertical cavity surface emitting laser (VCSEL) is a popular, mature, and cost-effective photonic transmitter technology that enables energy proportionality by allowing the links' data rate and the associated power consumption to be adjusted. Our work assumes VCSEL-based optical interconnects and presents intelligent centralized and distributed mechanisms to jointly and optimally select the routes, the flow sizes, and the transmission powers needed to serve a given input traffic load, minimize the consumed energy, and optimize performance. For this purpose, we use a detailed VCSEL energy model and formulate the energy minimization problem as a constrained nonlinear multicommodity optimization problem, which is solved optimally with the proposed approaches. The simulation results, carried out under a variety of scenarios, show the efficiency of these methods in terms of throughput and energy consumption.**

***Index Terms*—Energy aware flow control; Energy aware routing; Green networks; Network optimization; Optical interconnections; Optical networks; VCSEL.**

I. Introduction

Toward meeting the challenges raised by exascale computing, big data, and artificial intelligence applications, we are witnessing a rapid increase in the performance requested from systems at various scales [1]. As a result, systems on chip (SoCs) with many cores have emerged, supercomputers for high-performance computing (HPC) are becoming bigger, hosting millions of cores [2], and hyperscale datacenters are being built and deployed all over the world [3].

Naturally, an increase in size comes with an increase in complexity, bandwidth requirements, and energy consumption. The interconnection network, i.e., the communication fabric interconnecting the various components and systems, plays a prominent role in serving these demands, along with networking issues like the technologies used (optical versus electrical), the topology, the routing, the flow control (traffic engineering) algorithms, and other issues [4]. Reference [5] reports that datacenters are constrained by their interconnection networks instead of their computational power, with communication threatening to be the bottleneck of the entire system. Also, a large number of cores on a single chip may lead to the appearance of network hot spots, which in turn degrade the overall system's performance and lifetime [6].

Network interconnections contribute increasingly over 25% of the energy consumption of these systems [7–9], with a large percentage allocated to the communication links and the associated resources (30% in datacenters [10] and 65% in supercomputers [11]). This increase in energy consumption is not only due to the respective traffic increase [3] that puts a greater load in the network, but also due to the energy-efficient practices and technologies applied in other parts of these systems [12], making the networks' energy profile increasingly critical.

Optical technology can provide the flexibility required in the new generation of interconnection networks, offering high bandwidth, low latency, energy efficiency, and longer reach [2,5,13,14]. Figure 1 illustrates the vision for the application of optics in "in-the-box" networks such as datacenters and supercomputers: cabling of racks via active optical cables, opto-chips, and router chips with embedded photonic transmitters/receivers, coupled to boards with integrated optical waveguides. Although the replacement of copper with optics reduces the intrinsic energy consumption, the solution is not a panacea, for two reasons. First, the amount of potential solutions for energy-efficient networks is affected by all system levels, ranging from the physical layer to the applied algorithms and the served applications. Second, energy consumption at fiber-based transmissions is directly influenced by the transmission data rate on each link, and since the expectation of fiber networks is mainly the bandwidth increase, a respective increase in energy consumption is expected as well. Therefore, it is essential to enable the intelligent operation of network interconnections and to make optical interconnection networks' energy proportional based on their usage [15].

In this work, we propose centralized and distributed routing and flow control algorithms for optical interconnection networks utilizing vertical cavity surface emitting laser (VCSEL) photonic technology. The VCSEL is a popular, mature, and cost-effective photonic transmitter technology [16] that enables energy proportionality by allowing the

Manuscript received March 30, 2017; revised July 5, 2017; accepted July 6, 2017; published September 14, 2017 (Doc. ID 291837).

I. Gravalos (e-mail: grabalos@ceid.upatras.gr) and A. Siokis are with the Computer Engineering and Informatics Department, University of Patras, Patras, Greece.

A. Siokis, P. Kokkinos, and E. A. Varvarigos are with the Computer Technology Institute and Press "Diophantus," Patras, Greece.

E. A. Varvarigos is also with the School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece.

https://doi.org/10.1364/JOCN.9.000833

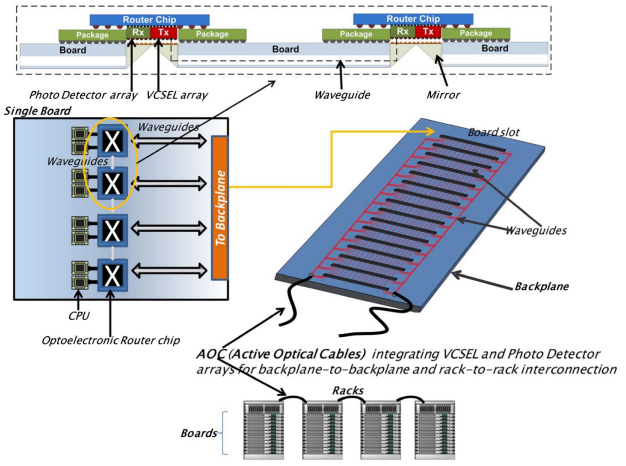

Fig. 1. Illustration of the application of photonics in "in-the-box" networks: active optical cables, optical printed circuit boards, opto-electronic chips. Rx, receiver; Tx, transmitter.

links' data rate and the associated power consumption to be adjusted based on the flow [17]. VCSELs are easily coupled to fiber and are close to achieving bit rates up to 100 Gbits/s on error-free transmissions for distances up to 1 km [18,19]. They are deployed in active optical cables for rack-to-rack interconnections, and for board-to-board, chip-to-chip, and inside-chip communication.

In this context, we aim to minimize the energy dissipation on VCSEL interconnected networks, applying suitable routing and flow control algorithms. Using a detailed VCESL energy model [17], in which the energy dissipation at each link is not linear with respect to the data rate, we convert the energy minimization problem to a constrained nonlinear multicommodity optimization problem, for which we propose two approaches to obtain the optimal solution. In this work we focus on the part of energy consumption that depends on the data rate used, and thus our approach focuses on the relevant elements, while we neglect components that are involved with static consumption. In particular, we assume that some components can be passive (such as multiplexers/demultiplexers) or not required for "in-the-box" networks (such as amplifiers). The OptiMal EnerGy Aware (OMEGA) scheme, based on the optimal routing and flow deviation concept [20], distributes traffic flows among several paths of an optical interconnection network in order to achieve the minimum aggregate flow at each link, while it also decides on a suitable bit rate for each VCSEL transmitter. Hence, the traffic fluctuates among the first derivative minimum (energy) paths, until the optimal solution is reached. In this way the total energy consumption is minimized, while the energy consumption is also balanced in the network links, reducing the risk for any heat hot spots (e.g., on chip-to-chip interconnects). We also propose the Distributed/Decomposable ENergy Aware RoutIng Optimisation (DENARIO) approach based on the alternating direction method of multipliers (ADMM) [21], using a set of auxiliary variables in order to decouple flow variables per links and benefit from the parallelization provided by the ADMM. A number of simulations are performed evaluating the OMEGA scheme, under a variety of realistic HPC-application traffic patterns and network topologies. These showcase OMEGA's energy efficiency and higher throughput in comparison to other energy- or not-aware routing algorithms and load balancing methods. In the current work we assume the same VCSEL technology is utilized to establish the network links.

The remainder of this paper is organized as follows. Section II reports on previous works related to network interconnects, routing algorithms, and energy-efficient practices. In Section III we give a detailed description of the considered energy model used for VCSELs. In Section IV we formulate the problem of energy-aware optimal routing, and we present the respective centralized and distributed schemes for its solution. In Section V we describe the simulation setup and present the results obtained. Finally, Section VI concludes our work.

II. Previous Work

The power consumption of interconnection networks in SoCs, supercomputers, and datacenters has received increased attention lately. Methods and technologies for power management along with energy-aware traffic engineering schemes have been identified as important means for achieving energy-efficient networks.

In particular, resource consolidation approaches consolidate the network traffic on fewer links and devices, then power off the idle devices and links. The authors of [11] investigate and propose self-regulating power-aware interconnection networks that turn their links on/off in response to changes in traffic in a distributed fashion for on-chip networks as well as for chip-to-chip networks. In [22], the authors propose for supercomputers the addition of hardware support for on/off control of links in software and the use of adaptive runtime systems to manage them. In [23] the authors present a mechanism that dynamically adjusts the available network bandwidth by switching links on and off according to the traffic requirements using a fat-tree interconnection network. The original underlying routing algorithm is maintained, at the expense of a slight latency increase for low loads. These approaches are effective compared to traditional networks, where it is evident that the energy consumption is dominated by the operation of network devices (even when they do not transmit any traffic), while the energy consumption remains rather constant with traffic load [24]. However, as will be shown later, in VCSEL technology the energy consumption increases progressively with the data rate (load) on the link.

"Proportional computing" refers to the concept of energy consumption in proportion to resource utilization, i.e., an idle or underutilized component or device should consume less energy [15]. Proportional computing methods can be classified into two main categories, namely, (a) dynamic voltage and frequency scaling (DVFS) and (b) adaptive link rate (ALR) [8]. The DVFS techniques are usually applied to server processors and to multicore SoCs [6], to scale the power supply and frequency of the processor(s) according to the system's load. When applied to networks, this

approach allows DVFS or dynamic voltage scaling (DVS) links to work in a discrete range of frequencies and supply voltages, which leads to different levels of power consumption in response to their traffic utilization [25,26]. The exact scaling decisions can be based on past network utilization to predict future traffic [25]. Reference [27] investigates hybrid techniques based on both DVS and on/off links. The idea is to shut down DVS links when traffic drops to very low levels. The ALR methods aim to reduce the energy states of the network interfaces by scaling down a communication link's data rate as a function of the network link loads [8]. However, slowing the communicational fabric down should be performed carefully and based on the demands of user applications; otherwise, it may lead to bottlenecks, thereby limiting the overall system performance. The methodologies presented in this work can also be applied with these ALR-capable transmission technologies, though we focus here on optical interconnects and VCSEL-based links.

Also, routing algorithms have a significant impact on the latency, throughput, reliability, and energy efficiency of interconnection networks. In SoC networks [networks on chip (NoCs)], the selection of the most optimal path to route the packets in the NoC minimizes the traffic flow on the network, which in turn results in energy efficiency, in addition to avoiding congestion and deadlocks [6]. Routing algorithms can be classified into static and adaptive/dynamic routing, while hybrid approaches have also been proposed for networks, and network interconnects in particular. Reference [26] compares DVS with the use of dynamic routing in non-DVS links, showing that as long as the network provides enough bandwidth to meet the needs of the application, an adaptively routed network can improve latency with the same power consumption.

In this work, we consider a VCSEL-based optical interconnection network. Optical technology is currently an active issue in interconnection networks of various scales: (inside) datacenters [5,28,29], exascale systems [2,4,30], and on-chip communications [13]. Reference [30] reports that up to two-thirds of the cost of an exascale interconnect is expected to be in the optical links. VCSELs are even considered for use in energy-efficient optical network units for access networks [31]. Generally, VCSELs are used for short-range communication ($\leq$1 km [18,19]). Their reach decreases with high data rates, and as a result VCSEL arrays with multimode fibers (MMFs) can be used to increase the provided capacity. Also, commodity VCSELs are incompatible with WDM technology, which is preferable for longer distances [29]. Most related works on VCSEL-based interconnects [17,32–34] focus on the physical properties or the design of the respective systems to achieve high error-free data rates with low energy dissipation without, however, considering the required optimization logic. The authors of Refs. [33,34] demonstrate efficient experimental performance of VCSELs with high data rates and low consumption. The authors of Refs. [17,32] consider VCSEL-based opto-electronic links and explore the energy savings achieved by scaling down the supply voltage of the link components when the required rate is less than the maximum link rate supported.

This is one of the first works (to our knowledge) on VCSEL-based optical interconnects to propose intelligent mechanisms that jointly and optimally select the route, the flow size, and the transmission power based on the input traffic load, minimizing the consumed energy and optimizing performance. We use a detailed VCSEL energy model to account for its energy dissipation under various transmission configurations. The proposed mechanisms can actually be applied in all scales of systems (SoCs, supercomputers, datacenters) where VCSEL technology is utilized, taking advantage of its widespread use [16]. Also, the power control is a preferable method, since the energy dissipation on each link (considering VCSELs and photodiodes) can be accurately obtained and depends on the respective data rate (increasing progressively with it). In other technologies and respective mechanisms that utilize the on/off approach [11], there is a sudden jump in consumption as soon as the device is turned on, and after that it is rather constant with traffic. However, the large amount of external traffic and the frequent data exchange in all links of HPCs and datacenter networks limit the deployment of on/off approaches that require the concentration of traffic on a few links. As a result, our approach is more appropriate, optimally balancing network traffic on VCSEL-based network interconnects.

III. VCSEL Energy Consumption Model

As a fundamental part of the architecture of optical interconnected systems, an opto-electronic link consists of the transmitter, the receiver, and the optical channel. In the case of a passive channel, the total energy consumption depends on the transmitter and the receiver. In particular, the energy-absorbing components that operate at the transmitter include a *laser source* that we assume to be implemented by a VCSEL, whose operation is to convert 0s and 1s into low and high intensities, respectively, and a *VCSEL driver* that modulates the driving current to the VCSEL, based on the input bit patterns. At the receiver, the power consumption is due to the *photodetector* that converts the optical bit stream back into electrical current signals and is implemented by a photodiode, the *transimpedance amplifier* (*TIA*) that converts these signals to amplified voltage signals, and finally the *clock and data recovery circuit* (*CDR*) [17,32]. An optical link with the aforementioned components is depicted in Fig. 2.

Consequently, the corresponding total power consumed by an opto-electronic link is given by [17,32]

$$P_L = P_V + P_{VD} + P_{PH} + P_{\text{TIA}} + P_{\text{CDR}}, \tag{1}$$

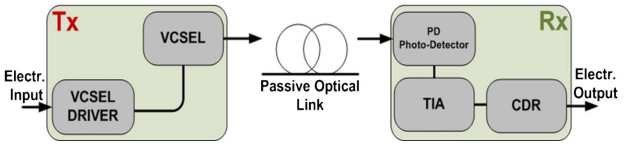

Fig. 2. Architecture of a VCSEL-based opto-electronic link. Tx, transmitter; Rx, receiver.

which adds the energy consumption of the respective components composing a link. Next, we analyze the individual power consumption of each component:

(a) The total power consumed by the VCSEL depends on a current threshold I_t, above which it can be stimulated and emit light. The light intensity then depends on the modulation current I_d that is fed by the driver. Thus, the power consumption of a VCSEL is

$$P_V = (I_t + \varepsilon I_d)(V_t + V_d + V_{dd} - V_{tn}), \tag{2}$$

where ε is the switching factor, V_{dd} is the supply voltage, V_t is the threshold voltage, and V_d is the voltage drop on the resistance, while the remainder of the subtraction corresponds to the minimum source drain voltage required for the gate to ensure saturation point.

(b) The power consumed by the VCSEL driver is dynamic and accrues to the charging of the inverter chain for each transmission. A VCSEL driver consists of a set of cascaded inverters, each of which has a size γ times larger than the previous one. The total power dissipated at the driver stages can be modeled as

$$P_{VD} = \varepsilon C_{VD} V_{dd}^2 T_{LR}, \tag{3}$$

where T_{LR} is the transmission link bit rate and C_{VD} is the total VCSEL drive capacitance of the inverters (sum of the input and output capacitance), given by

$$C_{VD} = C_L - C_{\text{in}} + \sum_{\gamma=0}^{k-1} (C_{\text{in}} + C_{\text{out}})\varepsilon^{\gamma}, \tag{4}$$

where C_L is the load capacitance of the inverter chain and C_{in} and C_{out} are the input and output capacitances of the minimum sized inverters, respectively.

(c) The photodetector at the receiver is responsible for converting the optical signal into a photon current. In order to assure successful detection, the photodetector must operate at the minimum receiver sensitivity power R_{min}, which is proportional to the transmitted bit rate. Therefore, the VCSEL power consumption depends on the receiver's requirements for a given T_{LR} and V_{dd}. Considering that the photodetector's power dissipation P_{PH} (<1 mW) is much lower than that of the other components, we will consider it negligible and will ignore it.

(d) The total power dissipated at the TIA can be calculated as

$$P_{\text{TIA}} = I_{\text{bias}} V_{dd} + I_d V_{dd} + \varepsilon(\alpha\beta)^2 R_f. \tag{5}$$

However, the dissipation in the TIA is dominated by the first term of the right-hand side, and hence we simplify the equation to

$$P_{\text{TIA}} = I_{\text{bias}} V_{dd}, \tag{6}$$

where I_{bias} is the bias current of the internal amplifier and $T_{LR\,\text{max}}$ is the maximum bit rate that assures the correct functionality of the TIA.

(e) Finally, the power consumed by the CDR unit is given by

$$P_{\text{CDR}} = \varepsilon C_{\text{CDR}} V_{dd}^2 T_{LR}, \tag{7}$$

where C_{CDR} is the capacitance of the CDR unit.

IV. Energy-Aware Routing in VCSEL Networks

The total power consumption at each opto-electronic link depends directly on the transmission bit rate at which the individual components operate. In particular, the supply voltage at each of the aforementioned components of a link can be restricted proportionally when the transmission bit rate required on the link is less than its capacity [17]. We will see that significant energy savings can be obtained by properly scaling the transmission bit rate on each link of a path, while satisfying the network traffic demands in terms of delay and throughput constraints.

In what follows, we assume the existence of a central controller that enables dynamic power management of all VCSEL links in the network (similar hypotheses were made in [17,32]), based on the decisions made in the presented methodologies. This implies that the electrical part of the opto-electronic router chips (such as in [13]) allows the reconfiguration of the characteristics of the embedded optical links, as described in Section III. We also assume that these mechanisms are only executed prior to system startup, to optimize its performance, and that the input source, the destination traffic matrix, does not change during the system's operation.

A. Energy-Aware Optimal Routing Problem

Considering the (energy) cost function of Eq. (1) and a given set of traffic flows $F = \{F_{s,d}\}$ comprising $F_{s,d}$ flow units (e.g., bits/s) from an origin node s to a destination node d, our objective is to find the set of paths $P_{s,d}$ (and associated traffic flows) that should be used for routing this traffic F in order to minimize the total energy consumption. The idea is that by distributing the total (s,d) traffic among several paths from s to d, and jointly optimizing power consumption for all source destination pairs (s,d) in the network, the amount of total flow per link is reduced, along with the required bit rate at each VCSEL, decreasing the total power dissipation. Hence the problem can be formulated as

$$\begin{aligned} &\text{minimize} \sum_{(i,j)} P_L^{ij}(T_{LR}^{ij}) \\ &\text{subject to} \sum_{p \in P_{s,d}} f_p = F_{s,d} \end{aligned}, \tag{8}$$

where T_{LR}^{ij} is the total flow in bits/s on link (i,j), $P_L^{ij}(T_{LR}^{ij})$ is the power consumption on link (i,j), and f_p is the flow of path p. For the total transmission rate T_{LR}^{ij} of all the flows on link (i,j), we have

$$T_{LR}^{ij} = \sum_{\substack{\text{all } p \text{ s.t.} \\ (i,j)\in p}} f_p. \tag{9}$$

We assume that this transmission bit rate T_{LR}^{ij} between two nodes i, j is generated by configurable and identical VCSELs, in which case the consumption *function* $P_L^{ij}(T_{LR}^{ij})$ is the same for all links (i,j), given by Eq. (1), but its value depends on the transmission rate on that link. Thus by substituting Eqs. (9) into Eq. (8), we obtain

$$P(F) = \sum_{(i,j)} P_L^{ij}\left(\sum_{\substack{\text{all } p \text{ s.t.} \\ (i,j)\in p}} f_p\right),$$

where F is a vector containing the flows f_p.

Furthermore, by obtaining the optimal bit rates for each link, we can adjust the supply voltage of a VCSEL to the minimum required for successful communication between two nodes, thus saving significant transmission energy. We assume that each node is equipped with one VCSEL per link and that the received flows are aggregated before transmitting them to the next link in their path. As mentioned in Refs. [17,32], the supply voltage varies in proportion to the bit rate variations. In particular, with a bit rate of 5 Gbits/s, a supply voltage of 0.9 V will be required. If the bit rate is doubled to 10 Gbits/s, the supply voltage must also be doubled. Hence, we further simplify Eq. (1) by substituting the supply voltage as follows:

$$V_{dd} = \frac{V_{dd(\max)}}{T_{LR(\max)}} * T_{LR}^{ij}, \tag{10}$$

where $T_{LR(\max)}$ is the maximum transmission rate of the VCSEL and $V_{dd(\max)}$ is the respective supply voltage. Substituting Eqs. (2), (3), (5), (6), (7), and (10) into Eq. (1), the cost function can be rewritten as

$$\begin{aligned} P_L^{ij}(T_{LR}^{ij}) = {} & (I_t + \varepsilon I_d)\left(V_t + V_d + \frac{V_{dd(\max)}}{T_{LR(\max)}} T_{LR}^{ij} - V_{tn}\right) \\ & + \varepsilon C_{VD}\left(\frac{V_{dd(\max)}}{T_{LR(\max)}}\right)^2 (T_{LR}^{ij})^3 + I_{\text{bias}} \frac{V_{dd(\max)}}{T_{LR(\max)}} * T_{LR}^{ij} \\ & + \varepsilon C_{\text{CDR}}\left(\frac{V_{dd(\max)}}{T_{LR(\max)}}\right)^2 (T_{LR}^{ij})^3. \end{aligned} \tag{11}$$

B. OptiMal EnerGy Aware Routing Algorithm (OMEGA)

To obtain the optimal solution for the multicommodity problem formulated in Eq. (8), we rely on the following general condition for optimality:

Lemma 1: If $f{:}R^n \to R$ is a differential convex function on R^n and X is a convex subset of R^n, then $x^* \in X$ is an optimal solution of the general minimization problem in the form of Eq. (8) if and only if $\nabla f(x^*)^T(x - x^*) \geq 0$, $\forall\, x \in X$. For proof we refer the reader to [20].

As one can see, Eq. (11) is a convex monotonically increasing function of T_{LR}^{ij} that increases sharply as T_{LR}^{ij} approaches the maximum capacity of link (i,j), since the second derivatives P_{ij}'' exist and are positive in $[0, T_{LR(\max)}]$. The partial derivative of P is given by

$$\frac{\partial P(F)}{\partial f_p} = \sum_{(i,j)\in p} (P_L^{ij})', \tag{12}$$

where the derivatives $(P_L^{ij})'$ are evaluated at the aggregated flows corresponding to F. Defining the first derivative $(P_L^{ij})'$ of the link energy cost with respect to its flow T_{LR}^{ij} as the (*first derivative*) *energy length* of link (i,j), Eq. (12) provides the *energy length* of that path p. Hence Lemma 1 can be applied to Eq. (8), which provides that

$$\frac{\partial P(F)}{\partial f_p}(f_p - f_p^*) \geq 0.$$

This, along with the requirement of $f_p^* > 0$, implies that the necessary and sufficient condition for optimality is that only paths of minimum first derivative length must have positive flows.

Interestingly, the proposed scheme can also be applied assuming the use of the on/off approach and of respective technologies. In this case, the first derivative cost on each link (from which the minimum path is obtained) will correspond to the start-up energy consumption of a device that operates at a specific data rate. Hence, the proposed solution will forward the traffic into minimum length (energy) first derivative paths, that is, to paths with the minimum number of devices required to serve the traffic demands, while powering off the rest of the devices.

C. Feasible Solution

Provided the abovementioned induction, we observe that, for each (s,d) pair, if the flow traverses a non-optimal path, then a portion of the corresponding flow could be redirected to the minimum first derivative path in order to come closer to the optimal solution. This can be shown by observing that if F is a feasible set path flow and ΔF is a corresponding portion shift, then the scalar β given by $G(\beta) = P(F + \beta\Delta F)$ has a first derivative around $\beta = 0$:

$$G'(\beta) = \sum_{\substack{\text{for all} \\ s,d\text{pairs}}} \sum_{p\in P_{s,d}} \frac{\partial P(F)}{\partial f_p} \Delta F.$$

Moreover, if ΔF is positive for minimum paths and negative for all other paths while maintaining flow conservation for each (s,d) pair, we will obtain

$$G'(\beta)|_{\beta=0} < 0,$$

which implies that the objective function can be reduced by a shift in direction ΔF.

However, since the path costs depend on flow values, the minimum path length generally changes after each flow redirection. Hence, the formulated problem can be optimally solved by methods such as the Frank–Wolfe (flow deviation) or the steepest descent. In this case, given an initial (non-optimal) vector F of flow allocations for all (s,d) pairs, the optimal solution can be obtained by iteratively shifting portions of flow β along the minimum paths, obtaining new values as

$$F = F + \beta(F^* - F),$$

where $\beta \in [0,1]$. The iterations continue until further flow redirections cannot improve the overall cost of Eq. (8). The value of β can be obtained by estimating the second-order Taylor approximation of $G(\beta) = P(F + \beta(F^* - F))$ around $\beta = 0$, deriving

$$\beta = \min\left[1, -\frac{\sum_{(i,j)}(T_{LR}^{ij*} - T_{LR}^{ij})P'_{ij}}{\sum_{(i,j)}(T_{LR}^{ij*} - T_{LR}^{ij})^2 P''_{ij}}\right],$$

where P'_{ij} and P''_{ij} (the first and second derivatives of P_L^{ij}, estimated at f_{ij}), are given by

$$P_L^{ij} = (I_t + \varepsilon I_d)\frac{V_{dd(\max)}}{T_{LR(\max)}} + 3\varepsilon C_{VD}\left(\frac{V_{dd(\max)}}{T_{LR(\max)}}\right)^2 (T_{LR}^{ij})^2 + I_{\text{bias}}\frac{V_{dd(\max)}}{T_{LR(\max)}} + 3\varepsilon C_{\text{CDR}}\left(\frac{V_{dd(\max)}}{T_{LR(\max)}}\right)^2 (T_{LR}^{ij})^2,$$

$$P''_{ij} = 6\varepsilon C_{VD}\left(\frac{V_{dd(\max)}}{T_{LR(\max)}}\right)^2 T_{LR}^{ij} + 6\varepsilon C_{\text{CDR}}\left(\frac{V_{dd(\max)}}{T_{LR(\max)}}\right)^2 T_{LR}^{ij}.$$

Apparently β cannot take large values because the optimality constraint will be violated (in other words, the upper part of the β estimation is a good approximation only for a small enough value of β). The power minimization algorithm at each step tries to shift a portion β of the flow from a non-shortest energy length path to the shortest one for each communication pair (s,d). In this way the flow is balanced between the links of a given topology, thus obtaining the minimum transmission rates (and minimum energy consumption) with respect to the traffic entering the network.

D. Decomposable ENergy Aware RoutIng Optimization (DENARIO)

Another attractive approach for solving convex optimization problems is the alternating direction method of multipliers (ADMM). This method is preferred for providing a level of parallelization toward the variables' estimation. However, in its basic version, this attribute requires decoupled variables, which is not the case for the problem as formulated in Eq. (8), due to the respective constraints. This can be clearly seen in the following analytical expression of Eq. (8):

$$\begin{aligned} &\text{minimize} \sum_{(i,j)} P_L^{ij}\left(\sum_{\substack{\text{for all}\\ s,d\text{pairs}}}\sum_{p\in P_{s,d}} \delta_p^{i,j} f_p\right) \\ &\text{subject to} \sum_{(i,j)} P_L^{ij}\sum_{p\in P_{s,d}} \delta_p^{i,j} f_p = F_{s,d} \qquad \begin{matrix}\forall\, i = s,\\ s \in (s,d)\end{matrix} \end{aligned}, \tag{13}$$

where

$$\delta_p^{ij} = \begin{cases} 1, & \text{if link } (i,j) \text{ belongs to path } p \\ 0, & \text{otherwise} \end{cases}.$$

Let us here substitute $\delta_p^{ij} f_p$ with f_p^{ij} [meaning the flow that traverses link (i,j) of path p] in order to lighten the notation without loss of generality. Notice that according to the constraint, the total flow for each source destination pair equals the sum of flows traversing the links directly attached to the source node.

Hence, to achieve primal splitting of the respective ADMM, we introduce a set of auxiliary variables z_p^{ij} and modify Eq. (13) as follows:

$$\begin{aligned} &\text{minimize} \sum_{(i,j)} P_L^{ij}\left(\sum_{\substack{\text{for all}\\ s,d\text{pairs}}}\sum_{p\in P_{s,d}} f_p^{ij}\right) \\ &\text{subject to} \sum_{p\in P_{s,d}} f_p^{i,j} - \frac{F_{s,d}}{m} = z_p^{i,j} \quad \forall (i,j),(s,d), \\ &\sum_{(i,j)} z_p^{ij} = 0, \end{aligned} \tag{14}$$

where m is the number of (i,j) pairs.

Both Eqs. (13) and (14) can obtain the same optimal solution over variables f_p^{ij} and thus approach the same global optima. To apply the ADMM, we use the following augmented Lagrangian formula:

$$L_y(f,z,\lambda) = \sum_{(i,j)}\left(P_L^{ij}\left(\sum_{\substack{\text{for all}\\ s,d\text{pairs}}}\sum_{p\in P_{s,d}} f_p^{i,j}\right) - \left\langle \lambda_p^{ij}, \sum_{p\in P_{s,d}} f_p^{i,j} - \frac{F_{s,d}}{m} - z_p^{ij}\right\rangle + \frac{y}{2}\left\| \sum_{p\in P_{s,d}} f_p^{i,j} - \frac{F_{s,d}}{m} - z_p^{ij}\right\|^2\right),$$

where $\langle *, * \rangle$ is the inner product, λ_p^{ij} is the Lagrangian multiplier, and y is a penalty parameter. Thus the respective ADMM estimates the f, z, and λ values through the iterations

$$(z)^{v+1} = \underset{z}{\text{argmin}}\left(\frac{y}{2}\sum_{(i,j)}\left\| \sum_{p\in P_{s,d}} (f_p^{ij})^v - \frac{F_{s,d}}{m} - z_p^{ij} - \frac{(\lambda_p^{ij})^v}{y}\right\|^2\right),$$

$$(f_p^{ij})^{v+1} = \underset{z}{\text{argmin}}\left(P_L^{ij}\left(\sum_{\substack{\text{for all}\\ s,d\text{pairs}}}\sum_{p\in P_{s,d}} f_p^{i,j}\right) + \frac{y}{2}\left\| \sum_{p\in P_{s,d}} f_p^{i,j} - \frac{F_{s,d}}{m} - (z_p^{ij})^{v+1} - \frac{(\lambda_p^{ij})^v}{y}\right\|^2\right),$$

$$(\lambda_p^{ij})^{v+1} = (\lambda_p^{ij})^v - y\left(\sum_{p\in P_{s,d}} (f_p^{i,j})^{v+1} - \frac{F_{s,d}}{m} - (z_p^{ij})^{v+1}\right).$$

Since z is a projection on the hyperplane $\sum_{(i,j)} z_p^{ij} = 0$, for each iteration $v+1$, the z values on every link can be calculated as follows:

$$(z_p^{ij})^{v+1} = -\frac{1}{m}\left(\sum_{(i,j)}\sum_{p\in P_{s,d}} (f_p^{ij})^v - \frac{F_{s,d}}{m} - \frac{(\lambda_p^{ij})^v}{y}\right) + \left(\sum_{p\in P_{s,d}} (f_p^{ij})^v - \frac{F_{s,d}}{m} - \frac{(\lambda_p^{ij})^v}{y}\right).$$

Hence, f_p^{ij} and λ_p^{ij} can be calculated in parallel for each (i,j) and z_p^{ij} as obtained from the last equation, the calculation step that decouples across the components of z [21]. Thus, the respective variables can be estimated distributedly at each node i of link (i,j). A suitable mechanism must

be applied to disseminate values calculated at step ν to proceed in step $\nu + 1$. It is important to mention that the execution time can further decrease by observing the following: $f_p^{ij} = f_p^{ef}$. In other words, the flows on links regarding the same path are equal. Hence, values f_p^{ij} and z_p^{ij} can be estimated once for each p.

V. Performance Evaluation

In order to assess the performance of the OMEGA scheme, we performed a number of simulation experiments. Also, the simulation environment was implemented in OMNET++. We use as evaluation criteria the network energy dissipation, the traffic losses, and the standard deviation of the links' load over varying traffic loads. We define total energy dissipation as the cumulative energy of all the network's links, as estimated by Eq. (1), which is necessary for each link to transmit at appropriate data rates in order to satisfy the traffic demands. The total power dissipation is metered in watts with respect to the network's traffic load given in Gbits/s. Traffic losses may occur when the total flows from all the paths traversing a link exceed the maximum link capacity. The load deviation is used as a metric of the load balancing that each routing algorithm achieves and shows how loads range between links. We would like to mention here that in our simulations we did not evaluate DENARIO, as the savings obtained with it would be the same as those obtained with OMEGA, since both approaches find an optimal solution based on flow deviation techniques. Our description of DENARIO aims at showing that the considered optimization problem can also be solved in a distributed manner. Real field implementation and measurements would be required to compare OMEGA's and DENARIO's different methods of operation.

In our performance results we opted to simulate links with single VCSELs for the sake of clarity. However, the proposed schemes can also be applied to topologies where larger capacity links are established utilizing arrays of VCSELs. In this case two scenarios can be considered: (a) the fibers are utilized as a single link, and (b) each fiber is utilized independently. In the former case the load is equally distributed among the VCSELs. Otherwise, the aggregate dissipation would be non-optimal, as described in Subsection IV.B. The latter case requires that the VCSELs in each link are independently deployed. Again, the proposed techniques can be applied with no adaptation.

A. Network Topologies

In our simulations we consider two network topologies: (a) $N = 16$ nodes placed on a mesh 4×4 topology and (b) a larger network with $N = 30$ nodes randomly connected by 37 bidirectional links. Mesh-like architectures are popular in several HPC systems, such as IBM, CRAY supercomputers, and Fujitsu's K-computer. For example, 16 nodes in a CRAY system [35] correspond to four blades, each hosting four nodes (with a single rack hosting 24 blades or 96 nodes). Each node hosts processing elements and routes traffic through proprietary CRAY router chips. In our case, we assume that each node's routing element is equipped with VCSEL transmitters and photodetectors that are assigned to the corresponding communication links. Such opto-electronic routing elements have been presented in the past [13]. The maximum link capacity considered matches the maximum VCSEL data rate, at 10 Gbits/s. Consequently, if the total flow assigned at a link exceeds the maximum capacity, the respective link dissipation will be estimated according to the maximum bit rate, since the surplus flow is assumed to be lost.

B. Traffic Patterns

The traffic patterns we used in the simulation scenarios correspond to two HPC applications, FFTW [36] and SuperLU [37], and one synthetic pattern, bit complement. The two HPC application profiles were obtained by using the integrated performance monitoring tool that profiles performance aspects and resource utilization of a parallel program, maintaining low overhead. The FFTW is an implementation of the discrete Fourier transform, and its behavior closely resembles uniform random traffic (URT), where each node communicates with all the other nodes (equally likely) using one-to-one communication (not broadcasting). Hence, the generated traffic L_n of node n is equally distributed among the individual flows of the corresponding communication pairs. Thus, every node n sends L_n/N units of traffic to every other node. The SuperLU is a general purpose library for the direct solution of large, sparse, nonsymmetric systems of linear equations on high-performance machines. The SuperLU is data intensive only locally (the majority of the traffic is destined to adjacent nodes). Finally, the bit complement is a permutation traffic pattern in which each source sends all of its traffic to a single destination (computed by complementing the bits of the source address). It is a traffic pattern that, among others, is typically relied on to demonstrate poor performance regarding throughput and delay [38].

C. Alternative Approaches

In order to further assess the performance of the proposed OMEGA scheme over VCSEL-based optical interconnects, we performed comparisons against four well-known routing algorithms: (a) The basic shortest path algorithm (BSP), i.e., the minimum hop routing algorithm, which is energy agnostic (in the sense that no voltage scaling is performed based on the link load). (b) An energy-aware variation of it, called energy shortest path (ESP). ESP determines the routing paths in terms of minimum hops, similar to BSP, but it also adjusts the voltage V_{dd} on each link to a suitable value with respect to instantaneous aggregate bit rates, as explained earlier in Section III, in order to further lower energy consumption. Since link loads depend on the routing algorithm, BSP and ESP accrue the same link loads; however, energy savings are expected using ESP as a result of voltage scaling due to bit rate variations. In BSP,

the value of V_{dd} is fixed and set to 1.8 V, which is the appropriate value to assure successful transmission at data rates of 10 Gbits/s. We also used two algorithms performing load balancing: (c) Valiant's algorithm [38] and (d) a shortest-path load balancing algorithm (LB). In Valiant, every packet sent from some source to some destination node is first sent from the source to a randomly chosen intermediate terminal node, then from the latter to the destination. This is an effective way to randomize routing for worst case traffic patterns, as it converts a traffic pattern into an "average" traffic pattern through randomization. For mesh/torus topologies, both the intermediate and destination nodes are reached through shortest path/dimension order routing [38]. Thus, in terms of flows, regardless of the original traffic pattern, each one of the two phases of Valiant's algorithm appears to be URT, eventually leading to twice the link loads of URT. The LB algorithm follows a similar approach, but load balances traffic using only the shortest paths (in terms of minimum hops) for the communication pairs. The voltage scaling is applied in those algorithms as well, as in ESP.

D. Simulation Results

OMEGA optimally load balances traffic among the networks' links, keeping their load low, thus requiring small data rates and correspondingly less transmission power. As a result, the network resources are utilized properly in order to serve the requested load, while achieving reduced energy consumption in comparison to the other methods evaluated.

Figure 3 presents the results of the algorithms' evaluation for the FFTW and URT traffic scenarios in the 4×4 mesh topology. We observed that the energy-aware approaches clearly outperform BSP and Valiant's algorithm. Also, OMEGA achieves significant energy savings of almost 18.5% compared to ESP, even when the traffic load increases, leading to the saturation of most links [at around 5 Gbits/s for the ESP algorithm, as shown in Fig. 3(b)]. Also, OMEGA achieves 70% energy savings compared to Valiant's algorithm. The LB approach performs well at light loads (starting to experience losses when traffic loads exceed 7 Gbits/s), but OMEGA still achieves 7% better energy savings. In addition, using OMEGA the load deviation of the links from the total average link load is much lower than with the other algorithms considered [Fig. 3(c)], thus achieving zero flow losses [Fig. 3(b)] and better utilizing the available network resources. The losses on links also explain why ESP and LB outperform the proposed method in some occasions. OMEGA manages to deliver more traffic, consequently consuming more energy.

OMEGA's optimal load balancing among several available paths for each communication pair reduces the bandwidth utilization of each link, even when the total traffic load of the network is high. This equilibrium is illustrated in Fig. 4, assuming a mesh network topology, which exhibits the link loads at the point where the BSP algorithm starts saturating. In the figure we observe that OMEGA achieves better load balancing. On the other hand, the use of BSP results in unbalanced link loads that reach the links' maximum capacity, leading to bottlenecks and probably heat hot spots, assuming on-chip communications.

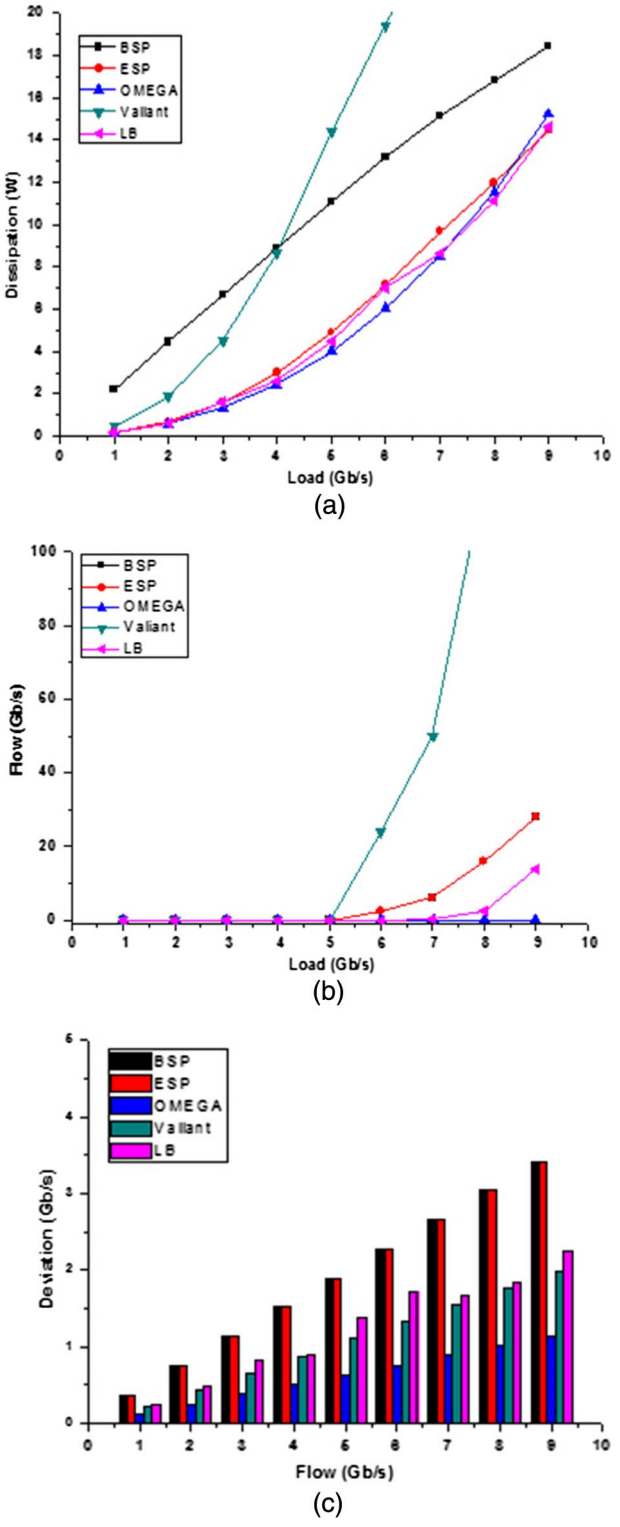

Fig. 3. (a) Network energy dissipation, (b) network flow losses, and (c) standard link load deviation of the 4×4 mesh topology over FFTW/URT.

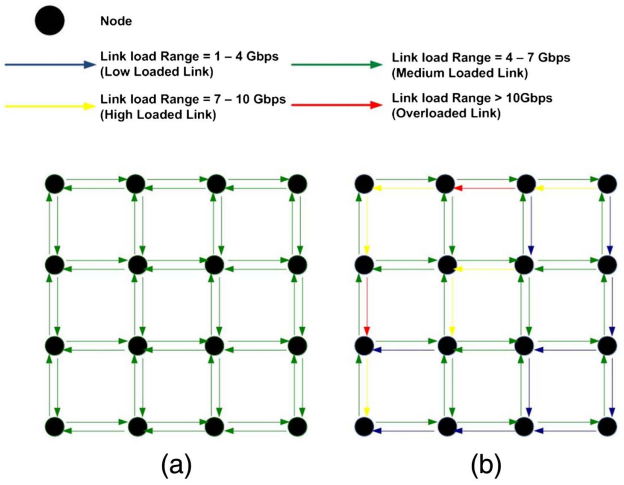

Fig. 4. Link loads for (a) OMEGA and (b) shortest path.

Similar results are obtained for the SuperLU traffic pattern depicted in Fig. 5. The somewhat increased locality of traffic compared to FFTW/URT lends itself to minimum hop routing. OMEGA routing performs marginally better than the LB algorithm in terms of energy consumption. Since both algorithms in this case use shortest paths, the small differences in deviation from the mean value of the demanded channel bandwidths in Fig. 5(c) indicate that channel loads are marginally better balanced in OMEGA.

For the bit complement traffic pattern, all the shortest path algorithms examined, using load balancing or not, perform poorly. BSP and ESP saturate the network for injection bandwidth of 1.66 Gbits/s, while LB saturates it for 3.22 Gbits/s. Valiant's algorithm exhibits identical behavior to that in Fig. 5, saturating the network for a load of 5 Gbits/s. The OMEGA scheme for 5 Gbit/s traffic yields 30% energy savings compared to Valiant's algorithm.

Similar results were obtained for the random network topology. Despite the sparse network connectivity, the proposed OMEGA algorithm can achieve less energy consumption and better link utilization than the other algorithms considered, since the load is optimally distributed and, as a result, links become saturated much more slowly. In particular, the obtained results showed energy savings of up to 76%, 20.5%, and 11% compared to BSP, ESP, and LB, respectively.

Regarding latency issues, our approach does not directly consider delay performance but focuses instead on energy aspects. Latency, however, is also reduced indirectly, because in our method the network load is distributed along several paths, and moreover, along the shortest paths between an origin–destination pair. As a result, the links of the topology do not saturate with respect to the rate [as shown in Figs. 3(b), 4, and 5(b)], reducing queuing delays, and traffic can be served with low latency. This notion of the throughput—delay trade-off complies perfectly with the comprehensive analysis provided in Ref. [20].

To sum up, simple energy-aware shortest path routing strategies achieve relatively low energy losses for low loads, but they do not perform well in terms of throughput for certain traffic patterns and tend to saturate the network early. On the other hand, load balancing traffic throughout the network topology (as in Valiant's algorithm) achieves good performance for adversarial traffic patterns where minimum hop routing performs poorly, disrupting any locality of the traffic. Another important observation is that such "blind" traffic load balancing all over the network is also prohibitive from an energy consumption perspective for VCSEL-based optical interconnection networks, since it keep all links loaded. Thus, the OMEGA routing and power optimization algorithm, while respecting the constraint that all generated traffic must reach the correct destinations, yields the optimal results in terms of both energy consumption and throughput.

VI. Conclusion

We considered a VCSEL-based optical interconnected network supporting the dynamic reconfiguration of the bit rate and energy footprint of its links, and proposed a centralized, namely OMEGA, and a distributed optimal routing and power control strategy. We used a detailed VCSEL energy model as a cost function to formulate the respective optimization problem and to run our simulations. Our results for a mesh network, with FFTW and SuperLU traffic patterns, showed for OMEGA energy consumption improvements up to 63.5% and 18.5%, respectively, compared to basic and energy-aware shortest path

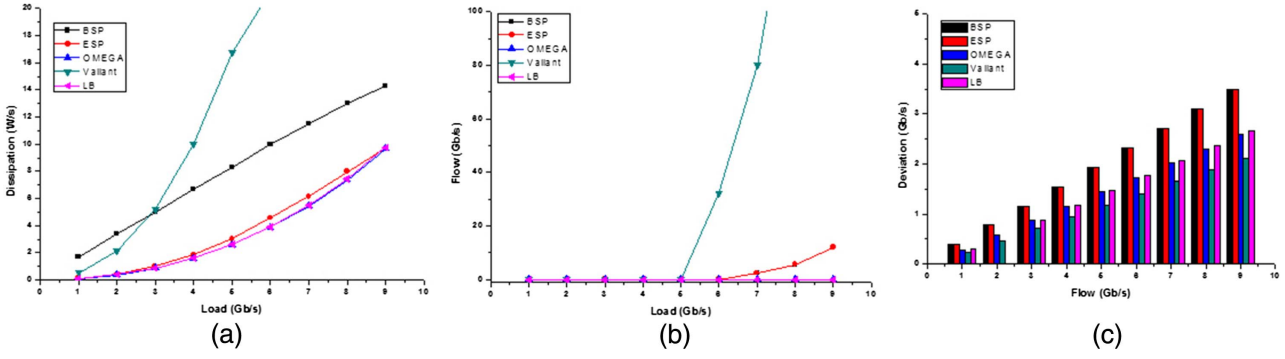

Fig. 5. Total network (a) energy dissipation, (b) flow losses, and (c) standard link load deviation of the 4 × 4 mesh topology over SuperLU.

algorithms and up to 70% compared to an energy-aware version of Valiant's algorithm, while achieving higher throughput than all of them. The respective improvements in energy consumption were 76% and 20.5% for a small network with randomly interconnected nodes.

References

[1] J. Escudero-Sahuquillo and P. J. Garcia, "High-performance interconnection networks in the exascale and big-data era," *J. Supercomput.*, vol. 72, no. 12, pp. 4415–4417, 2016.

[2] K. Ishii, T. Inoue, and S. Namiki, "Toward exa-scale optical circuit switch interconnect networks for future datacenter/HPC," *Proc. SPIE*, vol. 10131, 1013105, 2017.

[3] http://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf.

[4] R. Trobec, R. Vasiljević, M. Tomašević, V. Milutinović, R. Beivide, and M. Valero, "Interconnection networks in peta-scale computer systems: A survey," *ACM Comput. Surveys*, vol. 49, no. 3, 44, 2016.

[5] K. Bergman, "Optically interconnected high performance data centers," in *European Conf. and Exhibitionon on Optical Communication (ECOC)*, 2010.

[6] A. Abbas, M. Ali, A. Fayyaz, A. Ghosh, A. Kalra, S. U. Khan, M. U. S. Khan, T. De Menezes, S. Pattanayak, A. Sanyal, and S. Usman, "A survey on energy-efficient methodologies and architectures of network-on-chip," *Comput. Electr. Eng.*, vol. 40, no. 8, pp. 333–347, 2014.

[7] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 1, pp. 732–794, 2016.

[8] K. Bilal, S. U. Khan, S. A. Madani, K. Hayat, M. I. Khan, N. Min-Allah, J. Kolodziej, L. Wang, S. Zeadally, and D. Chen, "A survey on green communications using adaptive link rate," *Cluster Comput.*, vol. 16, no. 3, pp. 575–589, 2013.

[9] M. Buckler, W. Burleson, and G. Sadowski, "Low-power networks-on-chip: Progress and remaining challenges," in *IEEE Int. Symp. on Low Power Electronics and Design*, 2013.

[10] D. Kliazovich, P. Bouvry, Y. Audzevich, and S. U. Khan, "GreenCloud: A packet-level simulator of energy-aware cloud computing data centers," in *IEEE GLOBECOM*, 2010.

[11] V. Soteriou and L.-S. Peh, "Exploring the design space of self-regulating power-aware on/off interconnection networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 3, pp. 393–408, 2007.

[12] http://newscenter.lbl.gov/2016/06/27/data-centers-continue-proliferate-energy-use-plateaus/.

[13] K. Hasharoni, S. Benjamin, A. Geron, G. Katz, S. Stepanov, N. Margalit, and M. Mesh, "A high end routing platform for core and edge applications based on chip to chip optical interconnect," in *Optical Fiber Communication Conf.*, 2013.

[14] M. A. Taubenblatt, "Optical interconnects for high-performance computing," *J. Lightwave Technol.*, vol. 30, no. 4, pp. 448–457, 2012.

[15] D. Abts, M. R. Marty, P. M. Wells, P. Klausler, and H. Liu, "Energy proportional datacenter networks," in *Int. Symp. on Computer Architecture*, 2010, pp. 338–347.

[16] http://www.marketsandmarkets.com/PressReleases/vcsel.asp.

[17] X. Chen, L.-S. Peh, G.-Y. Wei, Y.-K. Huang, and P. Prucnal, "Exploring the design space of power-aware opto-electronic networked systems," in *Int. Symp. on High-Performance Computer Architecture*, 2005, pp. 120–131.

[18] K. L. Chi, Y. X. Shi, X. N. Chen, J. J. Chen, Y. J. Yang, J. R. Kropp, N. Ledentsov, M. Agustin, N. N. Ledentsov, G. Stepniak, and J. P. Turkiewicz, "Single-mode 850-nm VCSELs for 54-Gb/s on-off keying transmission over 1-km multi-mode fiber," *IEEE Photon. Technol. Lett.*, vol. 28, no. 12, pp. 1367–1370, 2016.

[19] D. M. Kuchta, "High speed VCSEL links using equalization," in *European Conf. on Optical Communications (ECOC)*, 2016.

[20] D. P. Bertsekas and R. G. Galager, *Data Networks*. Prentice-Hall, 1987.

[21] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[22] E. Totoni, N. Jain, and L. V. Kale, "Toward runtime power management of exascale networks by on/off control of links," in *Parallel and Distributed Processing Symp. Workshops & PhD Forum (IPDPSW)*, 2013.

[23] M. Alonso, S. Coll, J.-M. Martínez, V. Santonja, P. Lopez, and J. Duato, "Dynamic power saving in fat-tree interconnection networks using on/off links," in *Int. Conf. on Parallel and Distributed Processing*, 2006, p. 299.

[24] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsiang, and S. Wright, "Power awareness in network design and routing," in *IEEE INFOCOM*, 2008, pp. 457–465.

[25] L. Shang, L. Peh, and N. Jha, "Dynamic voltage scaling with links for power optimization of interconnection networks," in *Int. Symp. on High-Performance Computer Architecture (HPCA-9)*, 2003, pp. 79–90.

[26] J. Stine and N. Carter, "Comparing adaptive routing and dynamic voltage scaling for link power reduction," *IEEE Comput. Archit. Lett.*, vol. 3, p. 4, 2004.

[27] E. J. Kim, K. H. Yum, G. M. Link, N. Vijaykrishnan, M. Kandemir, M. J. Irwin, M. Yousif, and C. R. Das, "Energy optimization techniques in cluster interconnects," in *Int. Symp. on Low Power Electronics and Design*, 2003, pp. 459–464.

[28] K. Christodoulopoulos, K. Kontodimas, K. Yiannopoulos, and E. Varvarigos, "Bandwidth allocation in the NEPHELE hybrid optical interconnect," in *Int. Conf. Transparent Optical Networks (ICTON)*, 2016.

[29] A. Vahdat, H. Liu, X. Zhao, and C. Johnson, "The emerging optical data center," in *Optical Fiber Communication Conf.*, 2011.

[30] S. Rumley, S. Nikolova, R. Hendry, Q. Li, D. Calhoun, and K. Bergman, "Silicon photonics for exascale systems," *J. Lightwave Technol.*, vol. 33, no. 3, pp. 547–562, 2015.

[31] E. Wong, M. Mueller, P. I. Dias, C. A. Chan, and M. C. Amann, "Energy saving strategies for VCSEL ONUs," in *Optical Fiber Communication Conf.*, 2012.

[32] A. K. Kodi and A. Louri, "Energy-efficient and bandwidth-reconfigurable photonic networks for high-performance computing (HPC) systems," *IEEE J. Sel. Top. Quantum Electron.*, vol. 17, no. 2, pp. 384–395, 2011.

[33] J. A. Lott, P. Moser, A. Payusov, S. Blokhin, P. Wolf, G. Larisch, N. N. Ledentsov, and D. Bimberg, "Energy efficient 850 nm VCSELs operating error-free at 25 Gb/s over multimode optical fiber up to 600 m," in *IEEE Optical Interconnects Conf.*, 2012.

[34] K. L. Chi, J. L. Yen, J. M. Wun, J. W. Jiang, I. C. Lu, J. Chen, Y. J. Yang, and J. W. Shi, "Strong wavelength detuning of 850 nm vertical-cavity surface-emitting lasers for high-speed 40 <)Gbit/s) and low-energy consumption operation," *IEEE J. Sel. Top. Quantum Electron.*, vol. 21, no. 6, pp. 470–479, 2015.

[35] A. Bland, R. Kendall, D. Kothe, J. Rogers, and G. Shipman, "Jaguar: The world's most powerful computer," in *Cray User Group Conf.*, 2009.

[36] M. Frigo and S. G. Johnson, "The design and implementation of FFTW3," *Proc. IEEE*, vol. 93, no. 2, pp. 216–231, 2005.

[37] X. Li and J. W. Demmel, "SuperLU_DIST: A scalable distributed-memory sparse direct solver for unsymmetric linear systems," *ACM Trans. Math. Softw.*, vol. 29, no. 2, pp. 110–140, 2003.

[38] W. J. Dally and B. Towles, *Principles and Practices of Interconnection Networks*. Morgan Kaufmann, 2004.